Matérn 模型：穿越统计学、数值分析与机器学习之旅

Emilio Porcu[1], **Moreno Bevilacqua**、**Robert Schaback** 和 **Chris J. Oates**

arXiv:2303.02759v1 [math.ST] 5 Mar 2023

摘要。半个多世纪以来，Matérn 模型一直是空间统计学的基石。近年来，Matérn 模型更是成为数值分析、逼近理论、计算统计学、机器学习和概率论等多个领域的核心。本文将以 Matérn 模型为基础，展开一场跨越这些领域的探索。首先，我们探讨 Matérn 模型对空间统计学中估计与预测的重要性，并阐明它与其他受其深刻影响的学科之间的联系。随后，我们将 Matérn 模型置于大数据与可扩展计算的研究背景中，讨论随机偏微分方程方法、Vecchia似然近似，以及近期在贝叶斯计算中的应用。最后，我们回顾近期进展，包括 Matérn 模型的灵活替代方案，并从估计、预测、屏蔽效应、计算和Sobolev正则性等方面比较其表现。

关键词：逼近理论、紧支撑、协方差、核、克里金、机器学习、最大似然、再生核希尔伯特空间、空间统计学、Sobole夫空间。

1. 引言

本文有两个目的：一方面，我们跨越多个学科，对 Matérn 模型进行全景式介绍。另一方面，本文对 Matérn 模型在多个学科中的作用提出建设性批评，同时讨论替代模型或更一般的模型，以及它们与统计建模、估计、预测、计算统计学、数值分析和机器学习诸多方面的相关性。

Guttorp 和 Gneiting [70] 对 Matérn 模型的历史沿革进行了介绍。Matérn 模型——也称为 Matérn 协方差函数，或 Matérn 核，具体名称取决于语境——通常被认为源自 Matérn [108], ，但在科学文献的不同分支中也有其他称谓。Matérn 模型应用广泛，不可能在此涵盖其所有多样化的应用；我们的综述聚焦于其中一些特别值得关注且具有重要意义的应用。具体而言，我们旨在涵盖

1. 统计学中使用 Matérn 模型进行估计和预测，重点关注最大似然估计、克里金预测及相关的屏蔽效应；2. Matérn 模型的应用：a) 计算统计学，包括随机微分方程（SDE）和随机偏微分方程（SPDE）方法、似然近似、偏微分方程（PDE）的推断以及 Stein 方法；b) 统计建模，包括非标准情形，例如无法假设各向同性和平稳性，或对方向和曲线进行建模；

Emilio Porcu 是哈利法大学数学系教授，阿布扎比，阿联酋（电子邮箱：*emilio.porcu@bku.ac.ae*）。Moreno Bevilacqua 是阿道夫·伊瓦涅斯大学统计系教授（电子邮箱：*moreno.bevilacqua@uai.cl*）。Robert Schaback 是德国哥廷根大学数学系教授（电子邮箱：*schaback@math.uni-goettingen.de*）。Chris J. Oates 是英国纽卡斯尔大学数学、统计学与物理学院教授（电子邮箱：*chris.oates@ncl.ac.uk*）。[1]通讯作者。

c) 逼近理论和数值分析，其中 Matérn 模型用于构造基于核的插值函数；d) 机器学习，其中 Matérn 模型是高斯过程建模相关文献的核心；以及 e) 概率论，其中 Matérn 模型启发了若干基于相关随机过程样本路径性质的研究成果，并与某些类别随机微分方程的解相结合；

3. 与近期灵活的 Matérn 模型替代方案进行比较，重点关注

a) 具有紧支撑或多项式衰减等有趣特征的增强模型；b) 渐近估计精度、模型设定错误下的预测以及筛选效应；c) 在近似理论、计算统计学和机器学习中使用某些类别的紧支撑核所带来的影响。

本文的创新之处在于，它首次以 Matérn 模型为视角，广泛审视了科学文献。具体而言，我们并未试图全面综述一般的协方差函数。近期的综述对从空间到时空的协方差模型进行了相当全面的梳理[126], , 以及多元协方差函数[60], 和基于协方差的球面与流形建模 [122]。此外，尽管 Matérn 模型在科学图景中有许多引人入胜的应用，我们也不可能面面俱到。因此，我们将重点限定在方法论和理论问题上，并希望这些问题能与广泛使用 Matérn 模型的各个学科相关。

1.1 设定与记号

全文中，粗体字母表示向量和矩阵，转置算子记为 $\top$。令 $d\in\mathbb{N}$，并令 $Z=\{Z(\boldsymbol{x})\ \ \boldsymbol{x}\in\mathbb{R}^d\}$ 为一个实值高斯随机场，其均值为零，协方差函数为 $K:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$，定义为 $K(\boldsymbol{x},\boldsymbol{y}):=\mathrm{Cov}(Z(\boldsymbol{x})\ Z(\boldsymbol{y}))$。协方差函数是对称且正定的。本文中，正定一词的含义是

$$\sum_{i=1}^{n}\sum_{j=1}^{n}c_iK(\boldsymbol{x}_i,\boldsymbol{x}_j)c_j\geq 0 \tag{1}$$

对所有 $c_i\in\mathbb{R}$、所有 $n\in\mathbb{N}$ 和所有 $\boldsymbol{x}_i\in\mathbb{R}^d$均成立。如果上述不等式严格成立，则称 K 为严格正定。

每个对称正定函数 $K:\ \mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ 定义了平移函数 $K(\boldsymbol{x},\ \cdot)$，定义在 $\mathbb{R}^d$ 上，对所有 $\boldsymbol{x}\in\mathbb{R}^d$ 均成立。此外，可以定义两个平移函数之间的内积:

$$\langle K(\boldsymbol{x},\cdot),K(\boldsymbol{y},\cdot)\rangle_{\mathcal{H}(\mathcal{K})}:=K(\boldsymbol{x},\boldsymbol{y}),\ \ \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d, \tag{2}$$

就 K 本身而言。该定义可推广至所有平移函数的线性组合，并通过完备化生成一个定义在 $\mathbb{R}^d$上的函数希尔伯特空间 $\mathcal{H}(\mathcal{K})$。该空间称为 K的原生空间。请注意，该希尔伯特空间允许连续的点求值 $\delta_{\boldsymbol{x}}:\ f\mapsto f(\boldsymbol{x})$，其依据是重现公式

$$f(\boldsymbol{x})=\langle f,K(\boldsymbol{x},\cdot)\rangle_{\mathcal{H}(K)},\ \boldsymbol{x}\in\mathbb{R}^d,\ \ f\in\mathcal{H}(K) \tag{3}$$

这由（2）可得。于是，$\mathcal{H}(K)$ 称为再生核希尔伯特空间（RKHS），其核为 K。特别地，平移函数 $K(\boldsymbol{x},\cdot)$ 属于 $\mathcal{H}(K)$，构成该空间的完备化，并且是狄拉克泛函 $\delta_{\boldsymbol{x}}$ 的 Riesz 表示元。它们是机器学习、数值分析和逼近理论的核心，因为（2）使得抽象空间 $\mathcal{H}(K)$ 中的内积可以利用核显式计算，这就是所谓的核技巧。更多细节请参见第6.1节和 [166]。对于正定且ˆ平稳的核 K，其傅里叶变换 K 可用于改写希尔伯特空间 $\mathcal{H}(K)$上的内积（2），形式如下:

$$\langle f,g\rangle_{\mathcal{H}(K)}=\int_{\mathbb{R}^d}\frac{\hat{f}(\boldsymbol{\omega})\overline{\hat{g}(\boldsymbol{\omega})}}{\hat{K}(\boldsymbol{\omega})}\mathrm{d}\boldsymbol{\omega},\ \ f,g\in\mathcal{H}(K), \tag{4}$$

相差一个常数因子。这里，$\overline{g}$ 表示函数 g 的复共轭。注意，K 的谱会对 $\mathcal{H}(K)$ 中函数的谱施加惩罚。粗略地说，希尔伯特空间 $\mathcal{H}(K)$ 由满足 $\hat{}f/\sqrt{\hat{}K}$ 在 $\mathbb{R}^d$ 上平方可积的函数 f 组成。希尔伯特空间 $\mathcal{H}(K)$ 与协方差函数为 K 的高斯过程样本路径之间微妙的联系，将在本文多处出现，例如第 2、4.4、6.3 和 7.1 节。从这个意义上说，核是连接确定性模型与概率模型的重要纽带。

如果严格正定核 K 满足 $K(\boldsymbol{x},\boldsymbol{y})\equiv K(\boldsymbol{x}-\boldsymbol{y})$，则称其为平稳核。根据博赫纳定理[28]，$K$ 是正且有界测度 F的傅里叶变换，即

$$K(\boldsymbol{x}-\boldsymbol{y})=\int_{\mathbb{R}^d}\mathrm{e}^{\mathrm{i}(\boldsymbol{x}-\boldsymbol{y},\boldsymbol{\omega})}F(\mathrm{d}\boldsymbol{\omega}),\qquad \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d.$$

这里，$(\cdot,\cdot)$ 是 $\mathbb{R}^d$ 中的内积，i 是单位复数。当 K绝对可积时，可以进行傅里叶反演，此时我们称傅里叶ˆ ˆ变换 K 为其谱密度。注意，K 非负且可积。此外，本文大部分内容假设协方差函数 K具有平稳性和各向同性，因此

$$\mathrm{Cov}(Z(\boldsymbol{x}),Z(\boldsymbol{y}))=K(\boldsymbol{x}-\boldsymbol{y})=\sigma^2\varphi(\|\boldsymbol{x}-\boldsymbol{y}\|), \tag{5}$$

其中 $\boldsymbol{x}$、$\boldsymbol{y} \in \mathbb{R}^d$ 和 $\|\cdot\|$ 表示欧几里得距离。这里，我们假设 φ 连续，且 $\varphi(0) = 1$。下文中，我们将 φ 等同地称为一个函数或相关函数；后者是 $\varphi(\|\cdot\|)$ 的简写。因此，对于所有 $\boldsymbol{x} \in \mathbb{R}^d$，参数 $\sigma^2 > 0$ 是 $Z(\boldsymbol{x})$的方差。令 Φ_d 表示这类函数φ 所构成的类，它们通过恒等式 (5) $i.e.$ Φ_d 导出协方差函数 K。这是定义在 $\mathbb{R}^d$上的连续各向同性相关函数类。根据 Schoenberg[142], ，这类函数具有如下精确的积分表示：

$$\varphi(x) = \int_0^\infty \Omega_d(rx) F_d(\mathrm{d}r), \qquad x \geq 0, \tag{6}$$

其中 F_d 是一个概率测度，且

$$\Omega_d(x) = \Gamma(d/2) \left(\frac{2}{x}\right)^{d/2-1} J_{d/2-1}(x), \qquad x \geq 0, \tag{7}$$

其中 $\Gamma(\cdot)$ 是伽马函数，J_ν 是阶数为 $\nu > 0$ [118, 的第一类贝塞尔函数，公式为 10.2.2]。对于类 Φ_d 中的一个成员 φ，我们可以利用其 $\varphi(\|\boldsymbol{x} - \boldsymbol{y}\|)$ 的 d 元傅里叶变换仍具有各向同性这一性质，因此可将其化简为一个标量积分公式

$$\widehat{\varphi}(z) = \frac{z^{1-d/2}}{(2\pi)^{d/2}} \int_0^\infty u^{d/2} J_{d/2-1}(uz) \varphi(u) \mathrm{d}u, \; z \geq 0, \tag{8}$$

定义其 d 元各向同性谱密度，并假设该积分存在。若省略分母 $(2\pi)^{d/2}$，同一公式也适用于逆径向傅里叶变换。全文中，我们用 Φ_∞ 表示$\bigcap_{d\geq 1} \Phi_d$，即在每个 d 维欧几里得空间上诱导正定径向函数的函数类 φ。因此，当且仅当 $\varphi(\|\cdot\|)$ 是 $\mathbb{R}^d$ 中的相关函数时，$\varphi \in \Phi_d$ 成立。

2. MATÉRN 模型

Matérn 模型，$\mathcal{M}_{\nu,\ \alpha}$，定义为 [148]

$$\mathcal{M}_{\nu,\alpha}(x) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{x}{\alpha}\right)^\nu \mathcal{K}_\nu \left(\frac{x}{\alpha}\right), \qquad x \geq 0, \tag{9}$$

其中，$\alpha > 0$ 为尺度参数，$\nu > 0$ 为光滑性参数，$\mathcal{K}_\nu$ 为第二类修正贝塞尔函数，其阶数为 ν [2, 9。6.22]。可以验证，$\mathcal{M}_{\nu,\ \alpha}(0) = 1$，因此 (9) 是一个相关函数。Stein [148, p48] 中的论证表明，$\mathcal{M}_{\nu,\ \alpha}$ 属于类 Φ_∞。函数 $\sigma^2 \mathcal{M}_{\nu,\ \alpha}$ 将称为 Matérn 协方差函数，而 $\sigma^2 > 0$ 表示相应高斯随机场的方差。

Matérn 类的重要性源于参数 ν，它控制相关高斯场样本路径的可微性。具体而言，对于任意正整数 k，定义在 $\mathbb{R}^d$ 上、具有 Matérn 相关函数的高斯场 Z 的样本路径当且仅当 $\nu > k$ 时，均方可微 k 阶（沿任意方向）。此外，经过重新缩放的 Matérn 相关函数会在 $\nu \to \infty$ 时收敛到高斯核或平方指数核，即

$$\mathcal{M}_{\nu,\alpha/(2\sqrt{\nu})}(x) \xrightarrow[\nu\to\infty]{} \exp(-x^2/\alpha^2), \qquad x \geq 0, \tag{10}$$

其在 $\mathbb{R}^d$ 的任意紧集上一致收敛。因此，有时也采用参数化形式 $\mathcal{M}_{\nu,\ \alpha/(2\sqrt{\nu})}$ [169]。

当 $\nu = k + 1/2$ 时，对于 k 为非负整数的情形，Matérn 相关函数可简化为负指数相关函数与一个 k 阶多项式的乘积。例如，$\mathcal{M}_{1/2,\ 1}(x) = \exp(-x)$ 以及 $\mathcal{M}_{3/2,\ 1}(x) = \exp(-x)(1+x)$。一般而言，

$$\mathcal{M}_{k+1/2,1}(x) = \exp(-x) \sum_{i=0}^{k} \frac{(k+i)!}{2k!} \binom{k}{i} (2x)^{k-i} \tag{11}$$

对于 $k \in \mathbb{N}_0$。这种简单的 Matérn 相关函数代数形式无疑促成了 Matérn 模型的广泛流行。

现在，我们可以详细探讨 Matérn 模型的诸多面向。第 3 节讨论最大似然估计、克里金预测和屏蔽效应；第 4 节探讨 Matérn 模型的随机偏微分方程刻画。第 5 节讨论作为更复杂模型构建模块的 Matérn 模型；第 6 节通过 Matérn 模型的视角审视科学图景，重点关注数值分析、概率论和机器学习。第 7 节介绍近期提出的 Matérn 模型替代方案和推广形式；第 8 节从估计、预测和屏蔽效应的角度比较这些替代模型。

3. 使用 Matérn 模型进行估计和预测

令 $D \subset \mathbb{R}^d$ 为 $\mathbb{R}^d$ 的子集。考虑 D 中的 $X_n = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ 个（互不相同的）位置，在这些位置观测高斯随机场 Z 的值 $\boldsymbol{Z}_n = (Z(\boldsymbol{x}_1), \ldots\ldots, Z(\boldsymbol{x}_n))^\top$；该高斯随机场定义于第 1.1 节。一个重要问题是，如何在未观测位置 $\boldsymbol{x}_0 \in D \setminus X_n$ 处预测值 $Z(\boldsymbol{x}_0)$。于是，对于 $Z(\boldsymbol{x}_0)$，一个尤其自然的预测器是

$$\widehat{Z}_n = \boldsymbol{c}_n^\top \boldsymbol{R}_n^{-1} \boldsymbol{Z}_n \tag{12}$$

其中向量 $[\boldsymbol{c}_n]_i = K(\boldsymbol{x}_0, \boldsymbol{x}_i)$ 以及核矩阵$[\boldsymbol{R}_n]_{i,\ j} = K(\boldsymbol{x}_i, \boldsymbol{x}_j)$。预测量 (12) 可以从多个角度加以论证。经典观点认为，(12) 是 $Z(\boldsymbol{x}_0)$ 的最佳线性无偏预测量（BLUP），通常称为$Z(\boldsymbol{x}_0)$ [44] 的简单克里金预测量。从现代视角来看，随着无偏估计的作用日益受到质疑，我们可以依据其他最优性性质来论证这一选择，包括：

1. 它是以实现值 $Z(\boldsymbol{x}_0)$ 为条件的期望 $\boldsymbol{Z}_n$；2. 它是基于数据集 $Z(\boldsymbol{x}_0)$ 对 $\boldsymbol{Z}_n$作出的最优估计（即贝叶斯决策），其损失函数为平方误差，参见 [116, 第 13 节3]；3. 它给出在 $\boldsymbol{x}_0$处对数据进行插值的最小再生核希尔伯特空间范数插值函数，参见第 6.1节；4. 它是基于 $Z(\boldsymbol{x}_0)$ 从$\boldsymbol{Z}_n$ 出发进行逼近的算法，按照基于信息的复杂性[116, 第 10 节2] 以及逼近理论（参见第6.1节），该算法使最坏情形误差最小。

不一一列举。若假设 Z 具有平稳各向同性协方差函数$\sigma^2\mathcal{M}_{\nu,\,\alpha}$，我们便可以在 Matérn 模型这一自然框架下，研究（12）的性能。如何为参数 σ、α、ν 选择合适的取值是一个关键问题，我们将在第3.1节首先讨论，随后在第12节研究（3.2）的性能。第3.3节将讨论如何将 Matérn 模型直接推广至更一般的定义域，例如流形和图。

3.1 使用最大似然进行估计

在这种情形下，最大似然（ML）及类似的估计方法颇受欢迎，这是因为已有实用的数值计算方法（包括基于梯度的方法），且最大似然有经典理论作为基础。另一方面，使用最大似然隐含着统计模型设定正确这一前提，而这必须针对具体情况逐一判断。为限定讨论范围，下面我们将聚焦于最大似然估计。我们的目标是了解何时能够根据数据对 Matérn 模型的参数进行一致估计，并了解极大似然估计量的渐近分布。为此，回顾一下，高斯对数似然函数为

$$\mathcal{L}_n(\boldsymbol{\theta}) = -\frac{1}{2}\left(\log(|\sigma^2\boldsymbol{R}_n)|) + \frac{1}{\sigma^2}\boldsymbol{Z}_n^\top\boldsymbol{R}_n^{-1}\boldsymbol{Z}_n\right), \tag{13}$$

加上一个常数项后，其中 $\boldsymbol{\theta}=$（ν、α、σ^2）。最大似然估计量定义为

$$\widehat{\boldsymbol{\theta}}_n = \underset{\boldsymbol{\theta}\in\mathbb{R}^3_+}{\operatorname{argmax}}\,\mathcal{L}_n(\boldsymbol{\theta}). \tag{14}$$

方差参数的最大似然估计可以用闭式形式计算为 $\widehat{\sigma}_n^2 = \boldsymbol{Z}_n^\top\boldsymbol{R}_n^{-1}\boldsymbol{Z}_n/n$；将这一表达式代入（13）后，数值问题便转化为对所谓的 集中似然 关于$\mathbb{R}^2_+$进行优化。然而，最大化对数（集中）似然需要求解一个非线性优化问题，因此必须使用数值方法；见第4.3节。

最大似然估计的性能主要在两种不同的渐近极限下进行了研究。在固定域渐近理论下，采样域 D 有界且采样位置集合 X_n 在 D中变得越来越稠密。在增长域渐近理论下，定义域 D 随观测数据的数量 n 增加而扩大，任意两个采样位置之间的距离均有一个正的下界。Zhang 和 Zimmerman [180]指出，极大似然估计量的性能在这两种框架下可能大不相同，下面将对此进行讨论。

3.1.1 增长域渐近理论。Mardia 和 Marshall [107] 利用增长域渐近理论证明，在温和的正则性条件下，最大似然估计量是强一致的，即对于真实参数 $\boldsymbol{\psi}_0$，$\widehat{\boldsymbol{\theta}}_n \xrightarrow{a.s.} \boldsymbol{\theta}_0$。此外，他们还证明最大似然估计量服从渐近正态分布，即

$$\boldsymbol{F}^{1/2}(\boldsymbol{\theta}_0)(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}) \tag{15}$$

其中 $\boldsymbol{F}(\boldsymbol{\theta}) = -E[\mathcal{L}_n''(\boldsymbol{\theta})]$ 是 Fisher 信息矩阵，其元素为

$$F(\boldsymbol{\theta})_{i,j} = \frac{1}{2}\operatorname{tr}\left(\frac{\mathrm{d}\boldsymbol{\Sigma}_n}{\mathrm{d}\boldsymbol{\theta}_i}\boldsymbol{\Sigma}_n^{-1}\frac{\mathrm{d}\boldsymbol{\Sigma}_n}{\mathrm{d}\boldsymbol{\theta}_j}\boldsymbol{\Sigma}_n^{-1}\right),$$

以及 $\boldsymbol{\Sigma}_n = \sigma^2\boldsymbol{R}_n$。尽管我们关注的是 Matérn 模型，但值得注意的是，只要满足特定的正则性条件，这类渐近结果便适用于任何参数化相关函数。这些条件以相关矩阵及其导数的特征值条件来表述 [107]，不过，这些条件通常可能不容易验证（例如，参见 Shaby 和 Ruppert [144], 关于指数情形的研究）。一般来说，只要采样区域的空间范围远大于随机场的依赖范围，增域渐近理论就能非常准确地描述最大似然估计的行为[180, 144, 84]。

3.1.2 固定域渐近理论。Zhang[179] 研究了固定域渐近条件下 Matérn 模型的最大似然估计，并证明当光滑性参数 ν 已知且固定时，参数 σ^2和 α 均无法一致估计，其中 $d=1,\ 2,\ 3$。相反，只有参数

$$\text{micro}_{\mathcal{M}} = \sigma^2/\alpha^{2\nu}, \tag{16}$$

有时称为微遍历参数 [180, 148], 可以被一致估计。这是相应的两个高斯测度等价性的结果，我们将其记为 $P(\sigma_i^2\mathcal{M}_{\nu,\,\alpha_i})$，其中 $i=0,\ 1$。特别是，对于任意有界无限集 $D\subset\mathbb{R}^d$, $d=1,\ 2,\ 3$, $P(\sigma_0^2\mathcal{M}_{\nu,\,\alpha_0})$ 与 $P(\sigma_1^2\mathcal{M}_{\nu,\,\alpha_1})$ 在 $Z(\boldsymbol{x})$ $\boldsymbol{x}\in D$ 的样本路径上等价，当且仅当

$$\sigma_0^2/\alpha_0^{2\nu} = \sigma_1^2/\alpha_1^{2\nu}. \tag{17}$$

相比之下，对于 $d \geq 5$，Anderes [8] 证明了具有不同 Matérn 协方差函数的两个高斯测度相互正交，因此，在这种情况下，在固定域渐近条件下，所有参数都可以一致估计。情形 $d = 4$ 最近由 Bolin 和 Kirchner 进行了研究 [31]。

对于固定且已知的光滑度参数 ν，有关微遍历参数最大似然估计的渐近结果，可见 Zhang [179]，以及之后的 Kaufman 和 Shaby [84]。特别地，对于定义在有界无限集 $D \subset \mathbb{R}^d$, $d = 1,2,3$上的零均值高斯场，其 Matérn 协方差函数为 $\sigma_0^2\mathcal{M}_{\nu,\alpha_0}$，微遍历参数的最大似然估计量 $\hat{}\sigma_n^2/\hat{}\alpha_n^{2\nu}$是强相合的，$i.e.$,

$$\hat{\sigma}_n^2/\hat{\alpha}_n^{2\nu} \xrightarrow{a.s.} \sigma_0^2/\alpha_0^{2\nu},$$

其渐近分布为

$$\sqrt{n}(\hat{\sigma}_n^2/\hat{\alpha}_n^{2\nu} - \sigma_0^2/\alpha_0^{2\nu}) \xrightarrow{d} \mathcal{N}(0, 2(\sigma_0^2/\alpha_0^{2\nu})^2).$$

一般而言，当随机场的依赖范围相对于采样区域的空间范围较大时，固定域渐近理论能够非常准确地描述微遍历参数最大似然估计 [84]的行为。关于将这些结果推广到 Z 含有高斯误差的情形，可见 Tang 等人 [156]，；而有关 Matérn 模型时空版本的结果，可见 Ip 和 Li [77] 以及 Faouzi 等人 [55]。最后需要指出的是，在假设光滑度参数未知的情况下，微遍历参数的有效估计仍是一个尚未解决的问题；Loh 等人 [105]在这一方向上取得了一些有希望的结果。

3.2 预测与屏蔽效应

Matérn 类内高斯测度的等价性会对在未观测位置 $\boldsymbol{x}_0 \in D \setminus X_n$ 预测 $Z(\boldsymbol{x}_0)$ 产生影响；下面将讨论这些影响。下文假设 ν 已知且固定，并考虑 σ 和 α 设定错误的情形。也就是说，我们假设 Z 是一个具有 Matérn 协方差 $\sigma_0^2\mathcal{M}_{\nu,\alpha_0}$ 的高斯场，并考察使用 Matérn 模型 $\sigma_1^2\mathcal{M}_{\nu,\alpha_1}$ 时预测器 (12) 的表现。由于数据生成过程的真实参数 σ_0 和 α_0 通常未知，这种情形很常见。我们的理论框架采用固定域渐近。

首先请注意，(12) 不取决于σ_1的值，但取决于参数 α_1 的值（以及参数 ν，不过该参数是固定的）。将使用记号 $\boldsymbol{c}_n(\alpha_1)$ 和$\boldsymbol{R}_n(\alpha_1)$来强调这种依赖关系。在与真实模型 $\sigma_0^2\mathcal{M}_{\nu,\alpha_0}$ 相关联的高斯测度 $P(\sigma_0^2\mathcal{M}_{\nu,\alpha_0})$ 下，均方

$\hat{}$预测器 $Z_n(\alpha_1)$ 的误差为

$$\begin{aligned}
&\mathrm{VAR}_{\alpha_0,\sigma_0^2}\left[\widehat{Z}_n(\alpha_1) - Z(\boldsymbol{x}_0)\right] \\
&= \sigma_0^2\Big(1 - 2\boldsymbol{c}_n(\alpha_1)^\top \boldsymbol{R}_n(\alpha_1)^{-1}\boldsymbol{c}_n(\alpha_0) \\
&\qquad + \boldsymbol{c}_n(\alpha_1)^\top \boldsymbol{R}_n(\alpha_1)^{-1}\boldsymbol{R}_n(\alpha_0)\boldsymbol{R}_n(\alpha_1)^{-1}\boldsymbol{c}_n(\alpha_1)\Big),
\end{aligned}$$

而如果不存在模型错设，那么前面的表达式可简化为

$$\begin{aligned}
(18)\quad &\mathrm{VAR}_{\alpha_0,\sigma_0^2}\left[\widehat{Z}_n(\alpha_0) - Z(\boldsymbol{x}_0)\right] \\
&= \sigma_0^2\big(1 - \boldsymbol{c}_n(\alpha_0)^\top \boldsymbol{R}_n^{-1}(\alpha_0)\boldsymbol{c}_n(\alpha_0)\big).
\end{aligned}$$

在正则性条件下，对于固定域渐近，Stein [146] 证明，当两个高斯测度$P(\sigma_i^2\mathcal{M}_{\nu,\alpha_i})$，$i = 0,1$ 等价时，渐近有效预测和预测方差的渐近正确估计均成立，即 (17)。具体而言，

$$(19)\qquad \frac{\mathrm{VAR}_{\sigma_0^2,\alpha_0}\left[\widehat{Z}_n(\alpha_1) - Z(\boldsymbol{x}_0)\right]}{\mathrm{VAR}_{\sigma_0^2,\alpha_0}\left[\widehat{Z}_n(\alpha_0) - Z(\boldsymbol{x}_0)\right]} \xrightarrow{a.s.} 1$$

and

$$(20)\qquad \frac{\mathrm{VAR}_{\sigma_1^2,\alpha_1}\left[\widehat{Z}_n(\alpha_1) - Z(\boldsymbol{x}_0)\right]}{\mathrm{VAR}_{\sigma_0^2,\alpha_0}\left[\widehat{Z}_n(\alpha_1) - Z(\boldsymbol{x}_0)\right]} \xrightarrow{a.s.} 1.$$

(19)的含义是，在真实模型下，如果使用 ν 的正确值，那么 α_1 取任何值都将具有渐近效率。(20)的含义更强，它保证在均方误差意义下，在正确模型和设定错误的模型下使用设定错误的预测器是渐近等价的。需要注意的是，这类结果没有考虑与设定错误模型的协方差参数相关的不确定性。Kaufman 和 Shaby [84] 表明，通过使用方差$\hat{}\sigma_n^2 = \boldsymbol{Z}_n^\top \boldsymbol{R}_n^{-1}(\alpha_1)\boldsymbol{Z}_n/n$ 的极大似然估计量代替 σ_1^2，(20)仍然成立。

基于时空 [77] 和二元 [14] 版本的 Matérn 模型，两个高斯测度等价的条件也已被建立。接下来，我们考虑预测的一个具有实际重要性的方面，即所谓的屏蔽效应。

屏蔽效应。该屏蔽效应指的是这样一种现象：预测器（12）几乎完全取决于那些距离待预测值最近的观测值 [149]。因此，在大型数据集的情况下，屏蔽效应是一种重要工具，可用于减轻计算（12）所带来的负担。这一问题一直是地统计学中的重要研究课题 [109, 110, 111, 41]。事实上，巴黎高等矿业学院地统计学学院的 Matheron [109, 110], 提出了对屏蔽效应的首次正式定义

屏蔽效应指的是这样一种情形：距离待预测值较远的观测值所获得的克里金权重为零。Matheron 的定义与实数轴上的马尔可夫性质直接相关；当采用指数模型进行克里金时，就会出现这种性质（事实上，$\mathcal{M}_{1/2,\ \alpha}$）。

M. Stein [148, 149, 151, 152] 采用了屏蔽效应的另一种定义，现描述如下。设 Z 是定义在 $\mathbb{R}^d$ 上的均方连续、零均值且弱平稳的高斯随机场。令$e(X_n)$ 表示基于$\boldsymbol{Z}_n$对 $Z(\boldsymbol{x}_0)$ 进行预测时，预测器（12）的误差。将考虑观测位置集合 X_n 的两种选择。为此，设 F_ϵ, N_ϵ 为由 $\epsilon>0$ 索引的集合，其中 N_ϵ 包含距离待预测值最近的观测值，而 F_ϵ 包含距离最远的观测值。于是 Stein [149] 指出，当

$$\lim_{\epsilon\downarrow 0}\frac{\mathbb{E}\, e(N_\epsilon\cup F_\epsilon)^2}{\mathbb{E}\, e(N_\epsilon)^2}=1. \tag{21}$$

关于这一定义含义的深入讨论可见 Porcu 等人 [128], 的研究，其中报告了固定定义域渐近理论与增长域渐近理论之间的显著差异。

采样点 X_n 的空间配置决定了屏蔽效应是否成立。Porcu 等人 [128] 将规则方案定义为满足以下条件的方案：$F_\epsilon=\{\epsilon(\boldsymbol{x}_0+j)\}$，其中 $j\in\mathbb{Z}^d$ 和 N_ϵ 是将 F_ϵ 限制在某个固定区域后的结果，且 $\boldsymbol{x}_0$ 位于该区域内部，并假设 $\boldsymbol{x}_0\notin\mathbb{Z}^d$。对于规则方案，当谱 K 在无穷远处沿每个方向均具有共同的变异指数 [的正则变化时，Stein [149] 证明了（21）*[27]*，引自 128]。然而，对于时空过程而言，这一条件可能并不实用，因为空间坐标和时间坐标的可微性不一定相同。为解决这一问题，我们转而考虑非规则方案：对于 $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$，它们是 $\mathbb{R}^d$ 中互不相同的非零元素；$\boldsymbol{y}_1,\ldots,\boldsymbol{y}_N$ 是 $\mathbb{R}^d$ 中互不相同的元素；当 $\boldsymbol{x}_0=\boldsymbol{0}\in\mathbb{R}^d$ 和 $\boldsymbol{y}_0\in\mathbb{R}^d$ 均为非零时，有 $N_\epsilon=\{\epsilon\boldsymbol{x}_1,\ldots,\epsilon\boldsymbol{x}_n\}$ 和 $F_\epsilon=\{\boldsymbol{y}_0+\epsilon\boldsymbol{y}_1,\ldots,\boldsymbol{y}_0+\epsilon\boldsymbol{y}_N\}$。在 128] 中，这被称为 Stein假设 [

$$\forall R>0,\quad \lim_{\|\boldsymbol{\omega}\|\to\infty}\sup_{\|\boldsymbol{\tau}\|<R}\left|\frac{\widehat{K}(\boldsymbol{\omega}+\boldsymbol{\tau})}{\widehat{K}(\boldsymbol{\omega})}-1\right|=0, \tag{22}$$

在该设定下给出了屏蔽效应的一个充分条件^（在对 K 和 N_ϵ施加一些温和的附加条件下）；对于某些特定设计方案 N_ϵ [151], 该条件可在维度 $d=1$ 和$d=2$ 下针对均方连续但不可微的随机场进行验证。对于 $K=\mathcal{M}_{\alpha,\ \nu}$，Matérn 模型的谱有一个简单的表达式 [2, 11。4.44]:

$$\widehat{\mathcal{M}}_{\nu,\alpha}(z)=\frac{\Gamma(\nu+d/2)}{\pi^{d/2}\Gamma(\nu)}\frac{\alpha^d}{(1+\alpha^2z^2)^{\nu+d/2}},\ z\geq 0, \tag{23}$$

由此可验证（22）。

因此，在规则方案和不规则方案下，均可确立 Matérn 模型的屏蔽效应，从而证明对预测器采用“局部”近似是合理的（12）。

3.3 流形和图上的 Matérn 模型

设 M 为一般流形。一个实际问题是，能否将 Matérn 相关函数（9）与定义在该流形上的适当度量 g复合，从而在 M上保持正定性。对于球面，测地距离是一种自然的度量；它是球壳上任意两点之间圆弧的长度。对于这种度量，$(x,\ y)\mapsto\mathcal{M}_{\nu,\ \alpha}(g(x,y))$ 只有在 $0<\nu\leq 1/2$ [63]时才是相关函数。Alegría 等人[3], 强调了这一限制，并提出了 $\mathcal{F}$ 族；该模型适用于球面，且在均方可微性和分形维数方面具有与 Matérn 函数相同的性质。Li 等人 [98]研究了其他一般流形上的 Matérn 函数。Guinness 和 Fuentes [69] 提出用谱展开定义一个模仿 Matérn 模型的协方差函数，但 Lindgren 等人 [101] 批评这一构造有误，因为该谱展开无法重现 Matérn 模型的相同性质。

遗憾的是，Matérn 模型在任何非平坦空间上的适用性有限，这一问题似乎也延伸到了更抽象的情形。Anderes 等人 [9] 通过一种优雅的等距嵌入论证 $0<\nu\leq 1/2$ 证明，当输入空间是具有欧几里得边的图时，必须满足这一限制。Menegatto 等人 [112] 通过一个更一般的论证证明，对于赋有测地度量的一般拟度量空间，也存在同样的限制。Bolin 和 Kirchner [30] 所做的重要工作提出了一种在度量图上一次可微的模型。合理的结论是，要将 Matérn 模型推广到一般流形，通常需要采用某种形式的随机偏微分方程方法；我们将在第 4.2 节讨论这一方法。

4. 计算统计学中的 Matérn 模型

本节探讨 Matérn 模型与计算统计学的交互，首先介绍 Matérn 模型的数值实现方法（第 4.1、4.2 和 4.3 节），然后转而讨论如何利用 Matérn 模型促进数值计算本身（第 4.4 节）。

4.1 作为随机微分方程的实现

Matérn 模型可以表示为随机微分方程的状态空间表示，从而能够采用信号处理文献中的高效计算技术进行模拟、估计和预测。事实上，

聚焦于维度 $d=1$，并令

$$\mathbf{Z}(x)=(Z,\mathrm{d}Z/\mathrm{d}x,\ldots,\mathrm{d}^kZ/\mathrm{d}x^k),$$

Matérn 模型 $\mathcal{M}_{\nu,\ \alpha}$ 在 $\nu=k+1/2$ 的条件下可表示为

$$\mathrm{d}\mathbf{Z}=\begin{pmatrix}0 & 1 & & \\ \vdots & \vdots & & \\ & 0 & 1 & \\ -a_0 & -a_1 & \ldots & -a_{k-1}\end{pmatrix}\mathbf{Z}\,\mathrm{d}x+\begin{pmatrix}0\\ \vdots\\ 0\\ 1\end{pmatrix}\mathrm{d}\mathcal{W}$$

其中 $a_i={}_{k+1}\mathrm{C}_i\cdot\alpha^{-k-1+i}$，.C. 为二项式系数，而 $\mathcal{W}(x)$ 表示定义在 $x\in\mathbb{R}$ [73]上的零均值白噪声过程。状态空间形式的优势在于，只需对数据进行一次遍历，即可完成估计和预测，计算成本为线性$O(n)$，并可使用熟悉的卡尔曼更新方程，如 Sarkka 等人[135] 所述；Hennig 等人在第 6 章中对此作了更详细的阐述 [75]。Sarkka 等人 [135],还给出了高维空间中的类似表示，包括时空版本的 Matérn 模型。不过需要注意的是，这些方法仅在时间步数方面保持线性复杂度；其复杂度相对于空间网格大小呈立方增长。随机偏微分方程方法可以解决这一问题，接下来我们将对此进行讨论。

4.2 作为随机偏微分方程的实现

Matérn 模型持续广受欢迎的一个重要原因，是有高效且可扩展的数值方法可用于模拟，这在很大程度上要归功于 Lindgren 等人[102]。这些作者考虑了随机偏微分方程

$$(\alpha^{-2}-\Delta)^{\gamma/2}Z(\boldsymbol{x})=\mathcal{W}(\boldsymbol{x}),\qquad \boldsymbol{x}\in\mathbb{R}^d, \tag{24}$$

其中 $\alpha>0$, Δ 是拉普拉斯算子，$\mathcal{W}$ 是定义在 $\mathbb{R}^d$ 上的高斯白噪声，因此 $\mathrm{Cov}(\mathcal{W}(A_1)\ \mathcal{W}(A_2))=|A_1\cap A_2|$，其中 A_i 是 $\mathbb{R}^d$ 的子集，$i=1,2$，且 $|\cdot|$是体积积分。Whittle [167] 和 Whittle [168]证明了，方程（24）的解是一个具有 Matérn 协方差$\sigma^2\mathcal{M}_{\nu,\alpha}$ 的高斯场，其参数为 α（与之前相同）以及

$$\sigma^2=\frac{\Gamma(\nu)\alpha^{2\nu}}{\Gamma(\nu+d/2)(4\pi)^{d/2}},\qquad \nu=\gamma-d/2.$$

这一视角带来了两点启示：第一，为随机偏微分方程数值近似而开发的工具可以用于 Matérn 模型；第二，我们可以清晰地沿此路径将 Matérn 模型的定义推广到任何可定义相应随机偏微分方程的（平面或非平面）流形上。（例如，Jansson 等人[78] 采用这一视角，将 Matérn 模型推广到球面 $\mathbb{S}^d$。）

为了对（24）提供一种便于计算的近似，Lindgren 等人 [102] 考虑了（24）的弱解以及对该弱解的近似使用定义域 $\Omega\subset\mathbb{R}^d$（具体而言，采用了 *Galerkin* 近似，使用有限元基函数）。由此，作者建立了一条将随机场 Z 近似为具有 *Gauss–Markov*结构且具有稀疏精度矩阵的随机场的正式途径。稀疏矩阵代数使得能够快速模拟 Matérn 随机场的实现，并快速计算似然（13）（但无法快速计算似然的梯度）。

定义域 Ω 的选择会引入边界效应，必须谨慎地加以缓解。Khristenko 等人 [87],Brown 等人 [37] 为 γ是整数的情形提供了一种解决方案；Bolin 和 Kirchner [30]则研究了非整数情形。Cameletti 等人 [39]将基于随机偏微分方程的 Matérn 场扩展到了时空情形，随后 Bakka 等人 [15], Clarotto 等人 [42],也进行了相关研究，而 Bolin 和 Wallin [33]则探讨了多元 Matérn 情形。Lang 和 Pereira [92]提出了基于流形上 Galerkin 方法的替代近似方案。Pereira 等人 [121]提出了一种有趣的方法，使得能够在流形上处理海量数据。这些文献的研究兴趣有两个方面。一方面，与高斯随机场有限维表示相关的技术问题本身就极具趣味性；另一方面，这一群作者的研究动力实际上是提供高效计算工具。现有的相关软件包（例如 R-INLA、inlabru 和 rSPDE）便是明证，相关综述可参见 Lindgren 等人 [101]。

Sanz-Alonso 和 Yang [133] 试图解释 Matérn 模型数值近似中精度与可扩展性之间的权衡。回顾一下，在 SPDE方法中 [102]，Z 式（24）中的内容通过一个高斯过程进行数值近似

$$Z_\delta(\boldsymbol{x})=\sum_{k=1}^{n_\delta}\omega_k\epsilon_k(\boldsymbol{x}),\qquad \boldsymbol{x}\in\Omega, \tag{25}$$

其中 ϵ_k 是有限元基函数，向量 $\boldsymbol{\omega}=(\omega_1,\ \ldots\ldots,\ \omega_{n_\delta})^\top$服从均值为零且具有稀疏精度矩阵的多元高斯分布。近似 Z_δ 的精度取决于：(a) 有限元基函数的紧支撑；(b) 定义域 Ωf所导致的边界效应；以及 (c) 决定 δ 的基数的网格宽度 n_δ（见25）。早期的大多数文献都研究了（25），其中 n_δ 与所建模数据集的样本量 n 成正比。Sanz-Alonso 和 Yang [133] 采用固定域渐近方法来解释何时 $n_\delta\ll n$ 可能是一种合理的策略。为此，他们研究了高斯过程回归，并在贝叶斯收缩率框架下开展工作。他们的结果为特定的

缩放关系提供了依据： n_δ 与 $n_\delta = o(n)$ ，前提是光滑度 ν 足够高。

Sanz-Alonso 和 Yang 最近采用了另一种研究随机偏微分方程和 Gauss–Markov随机场的路径 [134]，他们采用基于图的随机偏微分方程离散化方法。这种方法非常适用于处理离散且非结构化的点云，例如机器学习任务中的点云数据；在这类任务中，数据属于一个隐式定义的低维流形。这种方法的第二个优点是，无需对定义域进行显式三角剖分。

4.3 近似似然与 Matérn 模型

使用最大似然估计（14）来估计 Matérn 模型的参数时，需要进行数值优化。虽然可以使用通用优化算法，但通常更好的做法是先构造一个精确但计算成本低的似然近似，以便随后更轻松地对其进行最大化。事实上，在处理大数据集时，近似似然至关重要，因为计算（13）需要求相关矩阵的逆和行列式，通常通过乔列斯基分解来完成，其计算复杂度为 $O(n^3)$，存储成本为 (n^2)。

或许最成功的近似方法是 Vecchia方法[163]，，它近年来 [（包括 147, 48, 49, 67, 47]）受到了极大的关注。Vecchia近似可用于任意相关模型，其基本思想是用高斯条件分布的乘积替代（13），其中每个条件分布只涉及一小部分数据。该近似要求对数据进行排序，并指定用于条件化的“先前”数据数量 m 。通常， m 越大，近似的精度越高，但计算成本也越高；而排序方式会影响近似的精度[67]。Vecchia方法为精度矩阵的 Cholesky 因子提供了一种稀疏近似，因此可在 $O(nm^3)$ 时间内计算近似似然，存储成本为 $O(nm^2)$ 。更多详情请参阅 Katzfuss 和 Guinness [83]近期发表的综述。Vecchia似然可视为一类更一般估计方法中的一个特例，这类方法称为准似然或复合似然 [103, 162]，已广泛用于采用 Matérn 模型 [52, 26, 13]的高斯场估计。

缓解最大似然估计计算负担的另一种方法是协方差截断 [59]。其基本思想是将 Matérn 模型与一个具有紧支撑的相关函数相乘，从而得到一个具有紧支撑的“修正”Matérn 模型。这会使相应的协方差矩阵具有稀疏性，因此可以利用稀疏矩阵算法，高效地计算 Cholesky 分解 [59]。然而，一些作者 [25, 23] 认为，鉴于近期有人提出了灵活的紧支撑模型，且 Matérn 模型是其特例，因此截断方法可能已经过时；参见第 8节。本文不打算全面综述似然近似，因此请读者参阅 Sun 等人 [154] 和 Heaton 等人 [74] 了解详情。

4.4 用于贝叶斯计算的 Matérn 模型

过去十年来，人们对使用核方法求解偏微分方程的兴趣日益浓厚。考虑一个系统

$$\begin{aligned} \mathcal{A}u &= f \qquad \text{in } \Omega \\ \mathcal{B}u &= g \qquad \text{on } \partial\Omega \end{aligned}$$

由一个涉及以下内容的微分方程指定： $\mathcal{A}$ 和 f，以及由以下内容指定的初始条件或边界条件： $\mathcal{B}$ 和 g。至少可以追溯到 Fasshauer[57] 在确定性情形下的工作，后来 Cockayne 等人等作者又从贝叶斯视角对其进行了重新诠释 [43]，；我们可以通过将 $u : \Omega \to \mathbb{R}$ 建模为 u 先验上的一个高斯随机场，并对该随机场进行条件化，使其在位置 $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} \subset \Omega$处满足微分方程，并在位置 $\{\boldsymbol{x}_{m+1}, \ldots, \boldsymbol{x}_n\} \subset \partial\Omega$处满足边界条件，从而近似求得强解。该过程的条件均值与对称配置法一致，该方法由 Fasshauer 引入 [57]，，我们将在第6.1节再作介绍；条件方差则为解提供概率不确定性量化，表达因计算预算有限而残留的不确定性。要实现这些方法，需要使用样本路径具有足够正则性的高斯过程，从而使对导数 $\mathcal{A}u$ 进行条件化的操作定义良好。另一方面，假设过高的光滑性可能会导致不确定性量化过于自信。因此，需要一种光滑性可调、能够适应具体微分方程的核。Matérn 类满足这一要求，但并非只有它能够做到这一点；我们将在第7节继续讨论这一问题。

贝叶斯统计学界近期受到广泛关注的一个特定偏微分方程是 Stein 方程，对于该方程，$\mathcal{A}u = c + p^{-1}\nabla \cdot (p\nabla u)$ ，其中 p是所关注后验分布的概率密度函数，f 是我们希望计算其后验期望的函数，而 c 是一个常数。如果 Stein 方程存在解，那么 c 必然是我们所要计算的后验期望值。这促成了若干数值求解 Stein 方程的研究工作，因为与先对 p 进行近似（例如使用马尔可夫链蒙特卡罗），再利用 p 的近似值来估计所关注期望相比，这是一种更直接的替代方法。在这一背景下，通常会使用核方法

，[117, 145] 尤其是，核的光滑性应比函数 f 高两个阶，因为 Stein 方程是一个二阶偏微分方程。在黎曼流形上对 Stein 方程的推广由 [19], 进行了研究；他们倡导使用具有可定制光滑性的核，以复现流形上函数的索об列夫空间，例如 Matérn 模型（在流形上的推广形式）。Matérn 模型与索об列夫空间之间的联系将在第 6.1节阐述。

5. 使用 Matérn 进行灵活建模

有人可能会认为，Matérn 模型对于许多统计应用而言灵活性不足，因为它仅适用于平稳、各向同性的高斯标量值随机场。然而，Matérn 模型也是许多更复杂模型的重要构建模块，其中一些将在下文介绍。这方面的文献十分丰富，因此本文的讨论必然较为简略；本节的扩展版本见补充材料的附录A 。

5.1 标量值随机场

首先讨论基于 Matérn 模型构建的标量值随机场模型。需要注意的是，可以轻而易举地在 Matérn 模型中引入非零均值函数，也可以将核相加或相乘，以获得表达能力可能更强的核；这两点我们都不再赘述。

为放宽 Matérn 模型的各向同性假设，[7] 可以考虑能够纳入空间依赖发展过程中优先方向的尺度混合。另一方面，时空模型需要特殊处理；Gneiting [62], Zas-tavnyi 和 Porcu [178]介绍了 Matérn 核的非可分离版本。

Paciorek 和 Schervish [119], 以参数化方式放宽了平稳性假设，随后 Roininen 等人 [132]又以非参数化方式放宽了这一假设。Wilson 等人 [170],提出了输入扭曲方法，通过神经网络变换输入，再将其送入 Matérn 模型，力求在参数模型的计算可处理性与非参数模型的灵活性之间取得平衡。

高斯假设可以通过输出扭曲来放宽，即采用形式为 $Z(\boldsymbol{x}) = w(Z(\boldsymbol{x}))$ 的变换，其中 $w(\cdot)$ 是从 $\mathbb{R}^d$ 到 $\mathbb{R}^d$的非线性映射。通常，当 Z 的协方差函数为 Matérn 型时，Z 的协方差函数并非 Matérn 型；但如果 w 足够光滑，那么 Z 的光滑性属性就会传递到 Z。关于是否存在协方差函数属于 Matérn 类的非高斯过程这一问题，Åberg 和 Podgórski 给出了肯定的回答。Yan 和 Genton 提出了具有 Matérn 协方差函数的变换高斯随机场。Bolin 随后与 Wallin 为非高斯 Matérn 场提供了基于随机偏微分方程的构造。Palacios 和 Steel、Xua 和 Genton、Bevilacqua 等人以及 Morales-Navarrete 等人提供了具有协方差 $g(\mathcal{M}_{\nu,\,\alpha})$ 的一般非高斯场，其中 $g(\cdot)$ 是能够保持 Matérn 模型正定性和光滑性属性的适当函数。

Matérn 模型的一项重要扩展近来备受关注：将其推广到那些经典光滑性概念无法明确定义的空间上的随机场。例如，Anderes 等人 [9] 研究了具有欧几里得边的图，并为其配备图上的测地距离或电阻度量。Menegatto 等人 [112] 通过考察拟度量空间，对这一设定作了推广。Bolin等人 [32]采用了不同的方法，在度量图上构建具有相应协方差结构的随机场。Tang 和 Zimmerman[155] 以及 Porcu 等人 [127]研究了具有欧几里得边的图上的 Matérn 模型时空版本。这些研究显著拓展了 Matérn 模型的适用范围。

Matérn 协方差函数随距离呈指数衰减，这可能不适合对涉及长记忆的过程进行建模。人们已开发出若干方法来修正 Matérn 相7。

[68] 考虑了定义在格点上的高斯随机场 $\mathbb{Z}^d$，其协方差函数是 Matérn 协方差在 $\mathbb{Z}^d$ 上的限制。所得谱是与 Matérn 协方差相关的谱密度的平滑版本。对于这种特定情形，随机偏微分方程近似可能会高估尺度，α。然而，尚不清楚这一结论如何推广到在 $\mathbb{R}^d$ 中连续索引的高斯场。

5.2 向量值随机场

与多元空间建模相关的方法层出不穷，读者可参阅 Genton和Kleiber [60]。这里，各向同性协方差函数 $\boldsymbol{K}: [0,\infty) \to \mathbb{R}^{p\times p}$ 是矩阵值的。对角线上的元素 K_{ii} 称为自协方差函数，而元素 K_{ij}、$i \neq j$ 称为交叉协方差函数。Gneiting 等人 [64] 提出了一种多元 Matérn 模型

$$K_{ij}(x) = \sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{M}_{\nu_{ij},\alpha_{ij}}(x), \qquad x \geq 0, \tag{26}$$

其中，σ_{ii}^2 是 $\mathbb{R}^p$中多元随机场第 i个分量 Z_i的方差，而ρ_{ij} 是同位

相关系数。为确保正定性，参数ν_{ij}、α_{ij}和ρ_{ij}受到一些限制，而且对同位相关系数ρ_{ij}的限制通常相当严格。最后这一点促使研究者探索替代方法，读者可参阅 Apanasovich 等人 [11]，以及较新的 Emery 等人[54]。Allard 等人[6] 提出了多变量时空 Matérn 结构的扩展，Porcu 等人[125]则通过一种技术方法进行了扩展。Kleiber 和 Nychka [88]提出了多变量非平稳 Matérn 函数。Alegría 等人 [4] 研究了具有凹陷效应的多元 Matérn 模型；时空协方差模型中的“凹陷”指的是 Cov($Z(\boldsymbol{x},t)$ $Z(\boldsymbol{x}',t')$) 大于 Cov(($Z(\boldsymbol{x},t)$ $Z(\boldsymbol{x}',t)$) 的情形，这需要特别的数学处理。

Dey 等人 [51], 最近研究了图上的多元 Matérn 建模。他们通过拼接提出了一类多变量图形高斯过程。拼接是一种基于图构造多元协方差函数的方法，并能确保变量之间在过程层面具有条件独立性。与 Matérn 模型结合时，拼接会生成一个多元高斯过程，其单变量分量均为 Matérn高斯过程，且符合图模型所规定的过程层面条件独立性。拼接能够大幅提升计算效率并减少参数维度。Li 等人 [97], 最近提出了一种巧妙的高斯过程构造方法，其中涉及 Matérn 协方差函数。他们考虑了一个乘积空间，该空间由 d维欧几里得空间与一个可用于索引组标签的抽象集合构成。

5.3 方向、形状与曲线

Matérn 模型在方向过程研究中发挥着重要作用。Banerjee 等人 [18] 对方向有限差分过程和方向导数过程的概念进行了形式化，尤其关注 Matérn 模型。Matérn 模型也应用于形状分析。Banerjee 和 Gelfand [16] 引入了贝叶斯边界分析，通过对边界进行“边界分析”来测量与曲线相关的空间梯度；Halder 等人 [71]进一步发展了这一方法。Matérn 模型的光滑性非常适合这一框架。Quick 等人 [131] 提出了使用 Matérn 模型对时间梯度进行建模的方法。与这些方法相关的是，Matérn 模型的光滑度参数 ν 在 Halder 等人 [71], 近期发表的论文中发挥着核心作用；他们通过分析随机曲面，解释感兴趣响应变量中的潜在依赖关系。

这只是对 Matérn 模型在统计应用中的简要介绍，但它的影响远不止于统计学；接下来，我们将探讨 Matérn 模型对相关领域的重要性。

6. 统计学之外的 Matérn 模型

本节将探讨 Matérn 模型对数值分析和逼近理论（第6.1节）、机器学习（第 6.2节）以及概率理-论（第6.3节）的影响。

6.1 数值分析与逼近理论

这里考虑的问题是：根据定义域 $D \subset \mathbb{R}^d$ 上一组互不相同的数据位置 $X_n = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ 处给定的数据值 $y_i = f(\boldsymbol{x}_i)$，重建实值函数 f。与第3.1节从统计学角度进行的阐述不同，从数值分析的角度来看，这些数据并不被假定以任何方式随机产生。尽管如此，我们此前从统计学角度加以阐释的许多数学表达式，也会出现在数值任务的求解过程中。数据向量 $\boldsymbol{Z}_n$ 被重新解释为$\boldsymbol{Z}_n = (f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_n))^\top$，而任务是近似未知函数 f 在未采样位置 $\boldsymbol{x} \in D \setminus X_n$ 处的值 $f(\boldsymbol{x})$。一种自然的解法是求一个最小范数插值函数

$$s_{f,X_n,K} = \operatorname*{arg\,min}_{s \in \mathcal{H}(K)} \|s\|_{\mathcal{H}(K)} \quad \text{s.t.} \quad \begin{aligned} s(\boldsymbol{x}_i) &= f(\boldsymbol{x}_i), \\ i &= 1, \ldots, n, \end{aligned}$$

我们回想一下，这正是第3节中提到的第三个最优性性质。因此，再次利用核矩阵$\boldsymbol{R}_n = [K(\boldsymbol{x}_i, \boldsymbol{x}_j)]_{i,j=1}^n$，求解系统 $\boldsymbol{R}_n \boldsymbol{c}_n = \boldsymbol{Z}_n$，得到一个固定的系数向量 $\boldsymbol{c}_n$，它确定了一个线性组合

$$s_{f,X_n,K}(\boldsymbol{x}) = \sum_{i=1}^{n} c_i K(\boldsymbol{x}_i, \boldsymbol{x}), \qquad \boldsymbol{x} \in D,$$

属于以下平移函数张成空间的平移函数 $K(\boldsymbol{x}_i, \cdot)$。根据再生公式（3）和（2），这一点很容易得出。设定后，上述公式与（12）相同，而所得的值$s_{f,X_n,K}(\boldsymbol{x})$ 可解释为对 $f(\boldsymbol{x})$的数值近似。对数似然函数（13）也可等价地视为对插值函数范数的惩罚，因为$\|s_{f,X_n,K}\|^2_{\mathcal{H}(K)} = \boldsymbol{Z}_n^\top \boldsymbol{R}_n^{-1} \boldsymbol{Z}_n$.

第3 节中的第四个最优性原理在这里对应于这样一个事实：误差泛函 $\epsilon_{\boldsymbol{x}}: f \mapsto f(\boldsymbol{x}) - s_{f,X_n,K}(\boldsymbol{x})$ 在对偶空间 $\mathcal{H}(K)^*$ 中的范数，在所有使用相同数据 $\boldsymbol{Z}_n$的 $\mathcal{H}(K)$ 上的线性重构算法中取最小值。关键工具是幂函数P_{K,X_n}，它对所有$\boldsymbol{x} \in D$ 定义为

$$\begin{aligned} &P_{K,X_n}(\boldsymbol{x}) \\ &= \sup\left\{ f(\boldsymbol{x}) : f \in \mathcal{H}(K),\, f(X_n) = 0,\, \|f\|_{\mathcal{H}(K)} \le 1 \right\} \end{aligned}$$

它具有性质 $P_{K,X_n}(\boldsymbol{x}) = \|\epsilon_{\boldsymbol{x}}\|_{\mathcal{H}^*(K)}$，并导出如下形式的最优误差界

$$|f(\boldsymbol{x}) - s_{f,X_n,K}(\boldsymbol{x})| \le P_{K,X_n}(\boldsymbol{x}) \|f\|_{\mathcal{H}_K}.$$

对所有 $\boldsymbol{x} \in D$ 以及 $f \in \mathcal{H}(K)$。它可以利用基于 $X_n \cup \{\boldsymbol{x}\}$ 的核矩阵进行数值计算，但这里省略细节。值得注意的是，幂函数恰好等于 克里金方差 [141], 的平方根，后者给出了在 $\boldsymbol{x}$ 处、给定数据位置 X_n 和核 K 时克里金误差的方差。

因此，在此情境下，对逼近误差的分析可归结为对幂函数的分析，进而归结为对空间 $\mathcal{H}$（K）的分析。根据（4）和（23），由 Matérn核 $\mathcal{M}_{\nu,\,1}$ 生成的再生核希尔伯特空间，其内积为

$$\langle f, g\rangle_{\mathcal{H}(\mathcal{M}_{\nu,1})} = \int_{\mathbb{R}^d} \frac{\hat{f}(\boldsymbol{\omega})\overline{\hat{g}(\boldsymbol{\omega})}}{(1+\|\boldsymbol{\omega}\|^2)^{\nu+d/2}}\mathrm{d}\boldsymbol{\omega} \tag{27}$$

差一个常数因子，这与经典索об列夫空间 $W_2^{\nu+d/2}$（$\mathbb{R}^d$）的内积一致。根据 Sobole夫嵌入定理，只要满足 $\nu > 0$，该空间中的元素便是定义良好的连续函数。这个空间是偏微分方程数学分析的典型框架；我们在第4.4节探讨过这一联系。总而言之，使用 Matérn核，可以根据给定的采样数据，对 Sobole夫空间中的函数进行最优恢复。利用导数数据进行推广恢复，可得到用于求解 Sobole夫空间中偏微分方程的无网格数值方法，其中包括基于 Wu 的偏微分方程导数数据的对称配置方法 [172], ，它与 Schaback 所述的方法具有类似的希尔伯特空间最优性性质[137]。使用 Matérn核的一个重要动因是，偏微分方程理论通常意味着解属于 Sobole夫空间。另一方面，也有充分理由用多调和函数替代 Matérn核 [138, 50]。

关于使用核进行确定性恢复问题的许多其他结果，可参见Wendland [166], ；相关应用见 Schaback和Wendland [139] ，而结合基本理论的 MATLAB程序见Fasshauer和McCourt [58]。

在数值分析和逼近理论中，Matérn 核及其他核通常采用较大的光滑度参数，因为它们旨在解决插值问题，而非回归问题。Nar-cowich 等人 [114] 证明，收敛速率取决于提供数据的函数 f 与核二者光滑度中的较小者；错误设定的 Matérn 核（其光滑度参数 ν 相对于 f的光滑度取值过大）所产生的误差，其收敛速率与正确设定 ν时所能达到的速率相同。另一方面，Tuo 和 Wang [159] 证明，在相同的设定下，预测误差对设计点的空间填充特性更加敏感。特别是，要达到最优收敛速率，还要求控制实验设计的准均匀性。

当然，在数值分析和逼近理论中使用核，需要估计核参数。量 σ 不会出现在相关矩阵 $\boldsymbol{R}_n$中，但尺度参数 α 会显著影响插值函数的误差。关于尺度估计已有大量文献，其中部分研究借鉴了最大似然等统计概念（参见第 3 节中的参考文献）。另一方面，第 7.3节介绍的多调和核等 Matérn 模型的特定替代方案，凭借插值函数独立于所用尺度参数取值这一显著性质，无需进行尺度估计。参见 Wendland [166] 和第 7.3节。

6.2 机器学习

核方法是机器学习研究的一个重要分支，其中核通常用于解决各种监督学习和无监督学习任务。与第6.1节的插值情形相比，机器学习中的数据通常带有噪声，因此必须指定似然或损失函数。

Matérn 模型通常便于核方法的分析；例如，Tuo等人 [160] 给出了充分条件，使Matérn核岭回归在 L_2 范数和再生核希尔伯特空间范数下的收敛速率超过标准极小极大速率。然而，数据中的噪声会给光滑度参数的选择带来重大挑战，例如Matérn 模型中的 ν。Karvonen[81] 证明，在固定定义域渐近情形下，ν 的最大似然估计在渐近意义下不会低估真实值；也就是说，如果真实回归函数具有 Sobolev光滑性 $\nu_0 + d/2$，那么光滑度参数估计在渐近意义下不会小于 $\nu_0 + d/2$，但这本身并不足以成为使用最大似然[82]的有力动机。由于这些额外的挑战，标准做法是在开发方法时尽可能保持核的一般性，并尽可能利用数据和模型选择准则学习出适当的核形式。然而，近期针对非欧几里得数据的机器学习方法依赖于SPDE方法，因此Matérn模型及相关模型正被明确地加以使用。

随着研究人员试图分析的数据类型日益多样且结构化，人们对灵活的高斯过程模型的需求也不断增长。这类模型定义在流形和离散的基于图的域等非欧几里得定义域上。在高斯过程框架下，Borovitskiy 等人 [35] 提出不对随机偏微分方程（24）进行数值求解，而是采用高斯过程模型的有限秩逼近。具体而言，他们考虑对黎曼流形 M 上作适当调整后的式（

其对应的 Matérn 模型可展开为如下级数：

$$\sum_{n=0}^{\infty}\left(\frac{2\nu}{\alpha^2}+\lambda_n\right)^{-\nu-d/2} f_n(\boldsymbol{x})f_n(\boldsymbol{x}'), \qquad \boldsymbol{x},\boldsymbol{x}'\in M$$

其中 $\{\lambda_n\}_{n=0}^{\infty}$ 和 $\{f_n\}_{n=0}^{\infty}$ 分别是拉普拉斯–贝尔特拉米算子的特征值序列和特征函数序列 $-\Delta_M$。作者提出，先数值求解拉普拉斯–贝尔特拉米算子的前几个特征函数$\{f_i\}_{i=0}^{n}$，然后使用一个有限秩高斯过程，其实现是这些特征函数的线性组合 $\{f_i\}_{i=0}^{n}$。尽管求解特征值问题可能比数值求解随机偏微分方程更困难，但作者认为，在同一流形上需要解决多项任务的场景中，缓存特征函数可以降低成本。Borovitskiy 等人巧妙地将这一方法扩展到无向图[34], ，并直接影响了定义在神经网络上的高斯过程[80],以及高斯过程的路径条件化[171], 、人工智能中的仿真智能[93] 和图上跨时间的核方法扩展 [115]。其他应用包括神经信息系统中的汤普森采样 [161], 、机器人学中的贝叶斯优化 [79],以及度量空间上的高斯过程回归[90]。

6.3 概率论与随机过程

从概率论和随机过程的角度来看，Matérn 模型已得到深入研究。从正则性的角度出发，Scheuerer [140] 总结了具有 Matérn 协方差函数的高斯随机场的性质；当且仅当 $\nu > k$ 时，样本路径在均方意义下可微 k 次。在相同条件下，样本路径的（局部）索об列夫空间指数恒等于 k。此外，具有 Matérn 协方差的高斯随机场的分形维数恒等于 $\min(\nu, d)$，其中 d 是定义该随机场的欧几里得空间的维数。对于具有 Matérn 协方差的非高斯随机场，Kent[86]研究了其连续性性质。

Matérn 模型的其他若干性质也已得到研究。Kelbert 等人 [85] 研究了随机分数阶热方程背景下 Matérn 模型中的分数阶随机场；另见 Leonenko 等人[96]。Leonenko 等人[95]研究了定义在嵌入于 $\mathbb{R}^d$中的单位球上的随机场，其协方差函数是 Matérn 模型在有限范围上的限制。Leonenko 和 Malyarenko [94]研究了具有等价类 Matérn 协方差函数的张量值随机场。Terdik [157]考察了具有 Matérn 协方差函数的非高斯随机场的角谱。近期的一项研究[158] 揭示了 Matérn 模型与一类随机过程的某些拉普拉斯 ARMA 表示之间有趣的联系。Lilly 等人[99] 表明，Matérn过程是分数阶 Brownian 运动的阻尼版本。Lim 和 Teo [100] 研究了具有广义 Matérn 协方差的随机场，该协方差由含有两个分数阶的分数阶随机微分方程的解得到，使作者能够推导出相关随机场的样本路径性质。Angulo 等人[10]通过随机 Helmholtz 方程给出了 Matérn随机场的时空扩展。

据该领域的重要贡献者 N. Leonenko[1], 所述，Matérn 模型的重要性基于对偶定理[72, 定理 1] ，该定理给出了对称随机向量的某些特征函数类别与其密度之间的显式关系。具体而言，与 Matérn 模型相关联的谱密度本身就是一个协方差函数，称为柯西协方差函数或反多二次协方差函数，它可以用于参数化相关高斯随机场的赫斯特效应。

至此，我们透过 Matérn 模型的视角，完成了对科学图景的探索。现在，我们将目光转向未来，探讨 Matérn 模型有望实现的改进。

7. Matérn 模型的改进

本节介绍了Matérn 模型的改进；这些协方差函数至少在一定程度上保留了 Matérn 模型的局部性质，同时又提供了额外的特性和功能。这里我们先逐一介绍这些模型，关于其特性的深入评论则留待第8节。

7.1 紧支撑模型

紧支撑协方差模型的历史悠久，可以追溯到 Askey [12], ，他提出了核

$$\mathcal{A}_{\mu,\beta}(x)=\left(1-\frac{x}{\beta}\right)_+^{\mu}, \qquad x\geq 0, \tag{28}$$

其中 β 和 μ 均严格为正，并且 $(x)_+ = \max(0, x)$ 是截断幂函数。该研究表明，对于所有 $\beta > 0$ ，$\mathcal{A}_{\mu,\beta}$ 属于 Φ_d 当且仅当 $\mu \geq (d+1)/2$。显然，映射 $\boldsymbol{x} \mapsto \mathcal{A}_{\mu,\beta}(\|\boldsymbol{x}\|)$ 在半径为 β 、嵌入于 $\mathbb{R}^d$ 的球上具有紧支撑。因此，当相关状态 $\boldsymbol{x}_i$ 和 $\boldsymbol{x}_j$ 满足 $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \geq \beta$ 时，协方差矩阵中会出现精确的零项；这种稀疏性的计算优势将在第 8.5 节中进一步讨论。

[1]个人通信，2023年1月。

Matheron 的升维与降维 [110] 方法被 Wendland [165] 应用于 Askey函数，从而获得了具有更高阶光滑性的紧支撑协方差函数；这些函数关于 $\|\boldsymbol{x}\|$ 是截断多项式。该策略无法在偶数空间维度中生成整数阶索об列夫空间，这一问题后来由 Schaback[136] 解决，他识别出了“缺失的”Wendland函数。Gneiting [62]提出了 Wendland函数的统一观点。Zastavnyi [175] 给出了一般类的充要条件，该类同时涵盖普通的和缺失的 Wendland函数。Buhmann [38] 推广了 Wendland函数，并给出了充分的参数条件，使这一新类在给定 d时属于 Φ_d。这些被称为 Buhmann函数的函数随后由 Zastavnyi[176] 研究，之后又由 Zastavnyi 和 Porcu [177], 、Porcu 等人 [129]以及 Faouzi 等人 [56]进行了研究。Hubbert[76] 以及 Chernih 和 Hubbert [40]研究了 Wendland函数的其他表示形式和性质。此外，人们还发展了 Wendland函数在多变量 [124, 46]、时空 [123] 和非平稳过程 [89]中的扩展。

接下来将进行更具技术性的讨论，我们将在其中介绍另外两类具有紧支撑的相关函数，并分别在第8节讨论。

1. 广义 Wendland ($\mathcal{GW}$) 族 [61, 176]包含具有紧支撑的相关函数；与 Matérn 模型一样，这些相关函数允许对底层高斯随机场的光滑性进行连续参数化。$\mathcal{GW}_{\kappa,\ \mu,\ \beta}$ 模型通过恒等式依赖于参数 $\kappa \geq 0$ 和 $\mu,\ \beta > 0$

(29) $$\mathcal{GW}_{\kappa,\mu,\beta}(x) = \frac{\Gamma(\kappa)\Gamma(2\kappa+\mu+1)}{\Gamma(2\kappa)\Gamma(\kappa+\mu+1)2^{\mu+1}}\mathcal{A}_{\kappa+\mu,\beta^2}\left(x^2\right) \times {}_2F_1\left(\frac{\mu}{2}, \frac{\mu+1}{2}; \kappa+\mu+1; \mathcal{A}_{1,\beta^2}\left(x^2\right)\right),$$

其中，为使 $\mathcal{GW}_{\kappa}$、$_{\mu}$、$_{\beta}$ 属于类 Φ_d，需要满足 $\mu \geq (d+1)/2+\kappa$；而 ${}_2F_1(a,b,c,)$ 是高斯超几何函数 [2]。$\mathcal{GW}_{\kappa}$、$_{\mu}$、$_{\beta}$ 模型的样本路径在任意方向上均为 k 次均方可微，当且仅当 $\kappa > k-1/2$ [61]，因此 κ 在该模型中起到光滑度参数的作用。当 $\kappa = k \in \mathbb{N}$ 时，$\mathcal{GW}_{k}$、$_{\mu}$、$_{\beta}$ 可分解为 Askey 函数 $\mathcal{A}_{\mu+k}$ 与一个 k 次多项式 $_{\beta}$ 的乘积。该模型包含 Wendland函数（$\kappa = k$，为正整数）以及缺失的 Wendland函数（$\kappa = k+1/2$）。Bevilacqua 等人 [25] 的定理 1(3) 意味着由 $\mathcal{GW}_{\kappa/2-(d+1)/4,\mu,\beta}$ 诱导的再生核希尔伯特空间，其中 $\kappa \geq (d+1/2)$，与索об列夫空间 $W_2^{\kappa}(\mathbb{R}^d)$ 范数等价。

2. 高斯超几何 ($\mathcal{GH}$) 族 [53] 定义为

(30) $$\mathcal{GH}_{\kappa,\delta,\gamma,\beta}(x) = \frac{\Gamma(\delta-d/2)\Gamma(\gamma-d/2)}{\Gamma(\delta-\kappa+\gamma-d/2)\Gamma(\kappa-d/2)} \times \mathcal{A}_{\delta-\kappa+\gamma-d/2+1,\beta^2}\left(x^2\right) \times {}_2F_1\left(\delta-\kappa; \gamma-\kappa; \delta-\kappa+\gamma-d/2; \mathcal{A}_{1,\beta^2}(x^2)\right).$$

该模型有四个参数，并且对于每个正的 β，只要 $\kappa > d/2$，它就属于类 Φ_d，其中

$$2(\delta-\kappa)(\gamma-\kappa) \geq \kappa, \quad \text{and} \quad 2(\delta+\gamma) \geq 6\kappa+1.$$

当且仅当 $\kappa > (k+d)/2$时，$\mathcal{GH}_{\kappa,\ \delta,\ \gamma,\ \beta}$ 模型的样本路径沿任意方向均可进行 $\lceil k/2 \rceil$ 次均方可微。因此，参数 κ 也控制着该模型样本的光滑性。

.....的重要性将在第 $\mathcal{GW}$ 和 $\mathcal{GH}$ 节讨论8。

7.2 具有多项式衰减的模型

具有多项式衰减的相关模型，例如广义柯西[65] 或 Dagum 模型 [22]，在对具有长期依赖性的数据进行建模时可能很有用。然而，使用这些相关模型时，将无法控制样本路径的可微性，而这是 Matérn 模型的一项关键性质。Ma 和 Bhadra[106] 最近提出了 Matérn 类的一种修改形式，使其能够实现多项式衰减，同时保持传统 Matérn 模型的局部性质。与该模型相关的相关函数为

(31) $$\mathcal{CH}_{\nu,\eta,\beta}(x) = \frac{\Gamma(\nu+\eta)}{\Gamma(\nu)}\mathcal{U}\left(\eta, 1-\nu, \nu\left(\frac{x}{\beta}\right)^2\right),\ x \geq 0,$$

其中 $\mathcal{U}$ 是第二类合流超几何函数 [2]。这里 $\nu > 0$ 控制原点附近的均方可微性，与 Matérn 模型中的情况类似，而$\eta > 0$ 控制尾部的厚重程度。构造（31）基于涉及逆伽马分布的尺度混合（经过重新参数化的）Matérn 模型。Ma 和 Bhadra [106] 指出，这一类模型特别适用于大距离占主导的外推问题。

7.3 多调和核

我们对 Matérn 模型的改进综述以多调和核作结，其定义如下：

$$H_{\nu,d}(x) := \left\{ \begin{matrix} x^{2\nu-d}\log x & \text{for } 2\nu-d\in 2\mathbb{Z} \\ x^{2\nu-d} & \text{else} \end{matrix} \right\} \tag{32}$$

差一个符号 $(-1)^{\lfloor \nu-d/2\rfloor+1}$。作为 $x=\|\boldsymbol{x}\|$、$\boldsymbol{x}\in\mathbb{R}^d$ 的函数，Matérn 核 $\mathcal{M}_{\nu-d/2}$、1 以 x 的偶次幂开头，后面接着 $H_{\nu,\,d}$；从这个意义上说，这两个模型彼此相关。忽略一个常数因子，$H_{\nu,\,d}(\|x\|)$ 在 $\mathbb{R}^d$ 上的广义傅里叶变换为 $\|\omega\|^{-2\nu}$，因此尺度参数也不过是另一个常数因子。这使得基于多调和函数的插值与尺度无关。对照（23），可以看出它与傅里叶空间中的 $\mathcal{M}_{\nu-d/2,\,\alpha}$ 的联系。Stein [150] 给出了多调和核与 Matérn 模型之间的形式化联系；多调和核也称为幂律协方差函数。多调和核是 $\lfloor \nu-d/2\rfloor+1$ 阶条件正定的；技术性定义见 Wendland [166]。多调和核生成的不是希尔伯特空间，而是 Beppo–Levi 空间。此类空间与索об列夫空间有相似之处，但为实现预测（第 3 节）和插值（第 6.1 节），还必须添加一个多项式空间；参见 Wendland [166]。一般而言，多调和核作为分数阶高斯场中的协方差出现，其中包括布朗运动的某些形式 [104，定理 3.3]。

接下来，我们将重点讨论对 Matérn 模型进行此类改进是否有必要。

8. Matérn 模型的改进是否有用?

本节最后将对 Matérn 模型及第 7 节介绍的模型改进版本进行批判性评述。

8.1 Matérn 模型的严格推广

Matérn 模型无法实现紧支撑、远距离处的空穴效应（正值与负值之间的振荡），也无法产生适用于建模长程依赖的缓慢衰减尾部。第 7 节中的大多数改进都旨在解决这类问题；此处我们将说明如何将 $\mathcal{GW}$、$\mathcal{GH}$ 和 $\mathcal{CH}$ 模型视为 Matérn 模型的严格推广。

Bevilacqua 等人 [23] 已证明，Matérn 模型是经过重新尺度化的 $\widetilde{\mathcal{GW}}$ 模型的一个极限情形。具体而言，他们考虑了定义为 $\mathcal{GW}$ 的模型

$$\widetilde{\mathcal{GW}}_{\kappa,\mu,\beta}(x) = \mathcal{GW}_{\kappa,\mu,\beta\left(\frac{\Gamma(\mu+2\kappa+1)}{\Gamma(\mu)}\right)^{\frac{1}{1+2\kappa}}}(x), \qquad x\geq 0,$$

并证明了

$$\lim_{\mu\to\infty} \widetilde{\mathcal{GW}}_{\kappa,\mu,\beta}(x) = \mathcal{M}_{\kappa+1/2,\beta}(x), \quad \kappa\geq 0,$$

在集合 $x\in(0,\ \infty)$ 上一致收敛。参数 μ 因此可以实现在紧支撑模型与全局支撑模型之间切换；它既可以固定，以确保得到稀疏相关矩阵，也可以根据数据集进行估计。不过，这种等价性仅适用于 Matérn 模型中大于或等于 $1/2$ 的光滑度参数，因此并未涵盖光滑度参数的整个取值范围。令人遗憾的是，分形维数 [是衡量时间序列和空间数据样本路径粗糙度的一种常用指标；当光滑度参数介于0 和1之间时， 66] 可完全通过 Matérn 模型进行参数化。因此，$\mathcal{GW}$（或 $\mathcal{GW}$）模型无法对随机场的分形维数进行完全参数化。此类问题可通过 $\mathcal{GH}$ 模型解决，该模型将 $\mathcal{GW}$ 模型作为一个特例 [53]：

$$\mathcal{GH}_{\frac{d+1}{2}+\nu,\frac{d+\mu+1}{2}+\nu,\frac{d+\mu}{2}+1+\nu,\beta}(x) = \mathcal{GW}_{\nu,\mu,\beta}(x)$$

当 β、δ 和 γ 以使得 $\beta/\sqrt{4\delta\gamma}$ 趋于 $\alpha>0$ 的方式趋于无穷大时，$\mathcal{GH}$ 模型（30）一致收敛于 Matérn 模型 $\mathcal{M}_{\kappa-d/2,\ \alpha}(x)$，并且在这种情况下涵盖了 Matérn 模型光滑度参数的完整取值范围。

Matérn 模型也可作为 $\mathcal{CH}$ 模型的一个特殊极限情形出现。具体而言，Ma 和 Bhadra [106] 表明

$$\lim_{\eta\to\infty} \mathcal{CH}_{\nu,\eta,2\sqrt{\nu(\eta+1)}\beta}(x) = \mathcal{M}_{\nu,\beta}(x),$$

该收敛在任意紧集上一致成立。

Matheron 的转带算子 [109] 可应用于相关函数，在保持核的正定性的同时产生空洞效应。Schoenberg 的一个论证证明，对于 $\mathbb{R}^d$ 中的各向同性相关函数，相关值不能小于 $-1/d$ [143]。由于 Matérn 模型对所有 d 都是有效模型，这意味着将转带法应用于 Matérn 模型不会产生任何空穴效应。另一方面，$\mathcal{GW}$ 和 $\mathcal{GH}$ 模型允许出现这种效应。

8.2 改进模型的估计

Matérn 模型的最大似然估计已得到充分研究；此处我们将讨论，针对 Matérn 模型的改进，能在多大程度上获得类似结果。

在增长域渐近理论的背景下，$\mathcal{GW}$ 和 $\mathcal{CH}$ 模型的参数可以使用最大似然法进行一致估计，且其相应的渐近分布已知；参见第3.1.1节。

在固定域渐近的背景下，与经典的 Matérn 模型类似，这些模型的参数

无法被一致估计。例如，Bevilacqua 等人[25] 表明，协方差模型 $\sigma^2\mathcal{GW}_{\kappa,\ \mu,\ \beta}$ 的微遍历参数（假设 κ 和 μ 已知）为 $\text{micro}_{\mathcal{GW}}=\sigma^2/\beta^{2\kappa+1}$. 此外，他们还证明，对于定义在有界无限集 $D\subset\mathbb{R}^d(d=1,\ 2,\ 3)$ 上的零均值高斯场，其协方差模型为 $\sigma_0^2\mathcal{GW}_{\kappa,\ \mu,\ \beta_0}$，微遍历参数的极大似然估计量 $\hat{}\sigma_n^2/\hat{}\beta_n^{2\kappa+1}$ 是强一致的，$i.e.$,

$$\hat{\sigma}_n^2/\hat{\beta}_n^{2\kappa+1}\xrightarrow{a.s.}\sigma_0^2/\beta_0^{2\kappa+1}.$$

此外，对于 $\mu>(d+1)/2+\kappa+3$，其渐近分布由下式给出

$$\sqrt{n}(\hat{\sigma}_n^2/\hat{\beta}_n^{2\kappa+1}-\sigma_0^2/\beta_0^{2\kappa+1})\xrightarrow{d}\mathcal{N}(0,2(\sigma_0^2/\beta_0^{2\kappa+1})^2).$$

目前尚无针对所提出的 $\mathcal{GH}$ 模型的类似结果。

类似地，Ma 和 Bhadra [106] 表明，协方差模型 $\sigma^2\mathcal{CH}_{\nu,\eta,\beta}$的微遍历参数，在假设 ν 已知的情况下，表示为

$$\text{micro}_{\mathcal{CH}}=(\sigma^2\Gamma(\nu+\eta))/(\beta^{2\nu}\Gamma(\eta)).$$

此外，他们证明，对于定义在有界无限集 $D\subset\mathbb{R}^d$ ($d=1,2,3$)上的零均值高斯场，具有协方差模型 $\sigma_0^2\mathcal{CH}_{\nu,\eta_0,\beta_0}$，微遍历参数的最大似然估计量$(\hat{}\sigma_n^2/\hat{}\beta_n^{2\nu})$ $(\Gamma(\nu+\hat{}\eta_n)/\Gamma(\hat{}\eta_n))$ 是强一致的，$i.e.$,

$$\frac{\hat{\sigma}_n^2(\Gamma(\nu+\hat{\eta}_n)}{\hat{\beta}_n^{2\nu}\Gamma(\hat{\eta}_n)}\xrightarrow{a.s.}\frac{\sigma_0^2\Gamma(\nu+\eta_0)}{\beta_0^{2\nu}\Gamma(\eta_0)}$$

并且，如果 $\eta_0>d/2$，其渐近分布由下式给出

$$\frac{\hat{\sigma}_n^2(\Gamma(\nu+\hat{\eta}_n)}{\hat{\beta}_n^{2\nu}\Gamma(\hat{\eta}_n)}-\frac{\sigma_0^2\Gamma(\nu+\eta_0)}{\beta_0^{2\nu}\Gamma(\eta_0)}\xrightarrow{d}\mathcal{N}\left(0,2\left(\frac{\sigma_0^2\Gamma(\nu+\eta_0)}{\beta_0^{2\nu}\Gamma(\eta_0)}\right)^2\right).$$

这些结果总体上支持在这些增强版 Matérn 模型中使用最大似然代入估计；预测性能问题将在下文讨论。

8.3 使用增强模型进行预测

如果两个高斯测度是等价的，那么相关的预测和均方误差在渐近意义下是相同的（参见第 3.2 节）。为此，近期研究试图建立 Matérn 模型与 Matérn 模型改进形式的高斯测度之间的等价性。Bevilacqua 等人 [25] 考虑 $\sigma_1^2\mathcal{GW}_{\kappa,\mu,\beta}$ 模型，并证明对于给定的 $\sigma_1\geq 0$、$\nu\geq 1/2$ 和 $\kappa\geq 0$，如果 $\nu=\kappa+1/2$、$\mu>d+\kappa+1/2$ 和

$$\sigma_0^2\alpha^{-2\nu}=\left(\frac{\Gamma(2\kappa+\mu+1)}{\Gamma(\mu)}\right)\sigma_1^2\beta^{-(1+2\kappa)},\tag{33}$$

那么，在 () 的路径上，对于 ，$P\ \sigma_0^2\mathcal{M}_{\nu\,\alpha}$ $P\ \sigma_1^2\mathcal{GW}_{\kappa\,\mu\,\beta}$ $d=(,)$ 与 (,,)是等价的，其中 ,,。因此，使用具有紧支撑的 $\mathcal{GW}$ 模型所作的预测，与使用 Matérn 模型所作的预测在渐近意义下是相同的。同样，Ma 和 Bhadra [106] 证明，对于给定的 $\eta\geq d/2$ 和 $\nu\geq 0$，如果

$$\sigma_0^2\alpha^{-2\nu}=\left(\frac{\Gamma(\nu+\eta)}{\Gamma(\eta)}\right)\sigma_1^2\left(\frac{\beta^2}{2}\right)^{-\nu},\tag{34}$$

那么 $P\ \sigma_0^2\mathcal{M}_{\nu\,\alpha}$ $P\ \sigma_1^2\mathcal{CH}_{\nu\,\eta\,\beta}$ $d=$ （, ）等价于（, , ），对于, , , 在（ ）的路径上。因此，使用具有多项式尾部衰减的 $\mathcal{GW}$ 模型进行的预测，与使用 Matérn 模型进行的预测渐近相同。

如果关注的是预测器（12），而不是相关高斯随机场所带来的预测不确定性，那么值得注意的是，Matérn 模型的平稳性假设或许并非必需。Stein等人 [153] 指出，在适当的参数条件下，可以在 Matérn 模型中考虑 $\alpha=0$，这等价于使用$H_{\nu,\ d}$ 中的多调和核进行预测（32）。该研究中的定理1表明，若 $d\leq 3$ 且参数 ν 满足其中的条件（2）（或 $d=1$），则在有界区域上，无法区分 $\alpha>0$ 与 $\alpha=0$。上述观察反映出，像第6.1节那样，使用多调和核进行预测与尺度无关。这源于傅里叶变换的齐次性，因此在此情境下无需进行尺度估计。

8.4 增强模型下的筛选效应

筛选效应也适用于增强版的 Matérn 模型。对于规则方案，Porcu 等人 [128] 中的定理1表明，当 $\mu>(d+1)/2+\kappa$时，$\mathcal{GW}$ 模型具有渐近筛选效应。这个条件并不苛刻，因为要使 $\mathcal{GW}_{\kappa}$、$_{\mu}$、$_{\beta}$ 属于类 Φ_d，本来就要求满足 $\mu\geq(d+1)/2+\nu$。对于不规则方案，情况则更为复杂。例如，对于 $d=1$中的不可微场，Stein [151] 中的定理1与 Porcu 等人[128] 中的定理1共同说明，只要满足 $\mu>1$，Askey 模型 $\mathcal{GW}_0$、$_{\mu}$、$_{\beta}$ 就具有筛选效应。对于 $d=2$，Stein [151] 中的定理2表明，只要满足 $\mu>3/2$，Askey 模型就具有筛选效应。$\mathcal{GW}$模型满足 Porcu 等人 [128], 中式（1.3）的 Stein条件，进而可以验证 Stein假设（22）。

Porcu 等人 [128] 的数值实验表明，与标准 Matérn 模型相比，在采用紧支撑的增强模型下，屏蔽效应甚至更强。这可以带来计算优势，我们将在下文讨论。

8.5 可扩展计算

统计精度与计算可扩展性之间旷日持久的较量，催生了试图应对这一棘手权衡的方法。下文将特别聚焦于 Matérn 模型背景下的这一权衡。诸如基于预测过程[17] 和固定秩克里金法的方法 [45], 等通用方法将不在讨论范围之内；感兴趣的读者可参阅 Sun 等人的综述[154]。

与 Matérn 模型相关的计算复杂度，主要取决于输入空间的维数 d、必须估计的核参数数量 p，以及数据量 n。下面将依次探讨这些挑战。

首先，我们考虑 d 较大带来的挑战。在机器学习中，当针对高维输入空间 [169] 执行高斯过程回归时，经常会遇到这一情况。由于 Matérn 模型 $\mathcal{M}_{\nu,\alpha}$ 可重现一个索об列夫空间，且范数等价（参见第 6.1 节），并且要使该空间中的元素逐点有定义，就必须满足 $\nu > d/2$，由此可知，当 d 趋于无穷时，ν 必须趋于无穷，因此 Matérn 模型会退化为高斯模型 (10)。在此极限下，一些增强模型的灵活性也会丧失；$\mathcal{GW}_{\kappa,\mu,\beta}$ 模型中的条件 $\mu \geq (d+1)/2 + \kappa$ 迫使参数 μ 随着 d 趋于无穷，而这又会迫使 $\mathcal{GW}_{\kappa,\mu,\beta}$ 趋近于 $\mathcal{M}_{\nu,\alpha}$。从这个角度看，类 $\mathcal{GH}_{\kappa,\delta,\gamma,\beta}$ 似乎更适合用于较大的 d。另外值得一提的是，对于 $d \geq 5$，所有具有 Matérn 协方差函数的高斯测度彼此正交 [8]。当把 Matérn 模型视为表达先验信念的先验分布时，这对高斯过程回归具有哲学层面的影响，因为核参数的微小变化会导致先验分布的整个支撑集发生改变。

与高输入维度相关联的 d 挑战是模型中会出现大量参数 p。多元 Matérn 模型的问题在于，不仅 p 会随 d呈指数增长，而且模型有效性条件还对（26）中的同位相关系数ρ_{ij} 施加了严格限制。Emery 等人 [54] 表明，当 $p = 3$ 时，这类限制就已经变得极其严格。其他多元协方差函数也有类似情况，包括 Daley 等人提出的多元 $\mathcal{GW}$ 模型[46]。

最后，我们考虑数据量 n 很大的情况，这会导致与预测器（(n^3) 计算和 $O(n^2)$存储成本相关联的开销（12）。针对 Matérn 模型，人们提出了多种方法来降低这些成本，其中许多方法利用了协方差（$\mathbf{\Sigma}_n$）或精度矩阵（$\mathbf{\Sigma}_n^{-1}$），或其 Cholesky 因子（ch($\mathbf{\Sigma}_n^{-1}$)）的（近似）稀疏性：

- 直接利用 Matérn 模型协方差矩阵稀疏性的方法 $\mathbf{\Sigma}_n$ 是第7节中 Matérn 模型的增强版本。这些方法在（估计的）紧支撑相对于采样区域的空间范围较小时可能很有用，此时近似结果极为稀疏；下文将通过实证研究探讨这一点。

- 与 Matérn 模型相关联的精度矩阵 $\mathbf{\Sigma}_n^{-1}$通常是非稀疏的（$d = 1$ 和 $\nu = 0.5$的情形除外），但实际上，矩阵元素通常相对接近0，也就是说，$\mathbf{\Sigma}_n^{-1}$是准稀疏的。因此，用稀疏矩阵近似 $\mathbf{\Sigma}_n^{-1}$可能是一种有效策略。第4.2节介绍的随机偏微分方程（SPDE）方法就是这种策略的一个典型实例。这种方法也可以从数值线性代数的结果得到解释：这些结果表明，如果矩阵元素呈现衰减特性，那么其逆矩阵的元素也会呈现类似（且更快）的衰减行为[20]。
- Vecchia 近似[163] 及其扩展[（例如 48, 67, 83, 47] ）意味着对 ch($\mathbf{\Sigma}_n^{-1}$) 进行稀疏近似，并且通常应用于 Matérn 模型，尽管它们也可以应用于任何协方差模型。这些方法的一个潜在局限是，它们依赖于变量的排序以及条件集合的选择，而后者决定了 Cholesky 稀疏模式 [见 67]。

对与 Matérn 模型改进相关的矩阵的稀疏性进行数值研究是很有启发性的；为此，我们聚焦于 $\mathcal{GW}_{\kappa,\mu,\beta}$ 模型。通过该模型，我们可以从具有半径为

$$C = \beta \left(\frac{\Gamma(\mu + 2\kappa + 1)}{\Gamma(\mu)} \right)^{\frac{1}{1+2\kappa}}$$

的紧支撑模型切换到 Matérn 模型，只需增大 μ 参数即可。在我们的实验中，我们报告了 $\mathbf{\Sigma}_n$ 的稀疏性以及 $\mathbf{\Sigma}_n^{-1}$和 ch($\mathbf{\Sigma}_n^{-1}$) 的准稀疏性；后者定义为上三角矩阵中绝对值低于任意小常数 ϵ的数值所占的百分比，在本例中我们设 $\epsilon = 1.e-8$。

经验评估考虑了 $n = 1$、156 和$n = 4$、489 个位置点，位于 $[0,1]^2$,，其中各点分别以 0.03 和 0.015 的间距排列在规则网格上。对于 $\nu = 0$ 1、2，我们设置 β，使 Matérn 模型的实际范围等于 0.15（β =分别为 0.050、0.0316、0.0253），并考虑增加 $\mu = 1.5+\kappa$、4、8、16、32、120，∞（其中$\mathcal{GW}_{\kappa}$、∞、β为 Matérn 模型 $\mathcal{M}_{\kappa+1/2,\ \beta}$）。

结果见表1。对于较低的值 $\mu = 1.5$，2.5，3.5 和 $\nu = 0$，1，2，协方差矩阵高度稀疏，而稀疏性会随着

$\kappa = 2$				
$\boldsymbol{\Sigma}_n^{-1}$		$\mathrm{ch}(\boldsymbol{\Sigma}_n^{-1})$		
1156	4489	1156	4489	
2.90	1.31	5.63	6.12	
10.4	11.9	15.0	22.6	
42.8	61.0	47.3	64.8	
51.7	73.0	54.6	74.2	
52.3	75.7	54.8	77.2	
52.4	75.7	54.1	76.5	
52.4	75.4	53.6	74.6	

$\kappa = 2$					
$\boldsymbol{\Sigma}_n$		$\boldsymbol{\Sigma}_n^{-1}$		$\mathrm{ch}(\boldsymbol{\Sigma}_n^{-1})$	
1156	4489	1156	4489	1156	4489
71.4	0	0	0	0	
66.4	0	0	0	0	
29.6	0.71	2.73	2.13	6.90	
1.29	13.7	25.8	14.8	27.9	
0	25.5	23.5	24.8	42.7	
0	26.0	10.9	25.9	46.4	
0	26.0	18.2	25.9	46.6	

矩阵，那么 $\boldsymbol{\Sigma}_n^{-1}$或 $\mathrm{ch}(\boldsymbol{\Sigma}_n^{-1})$ 往往具有很强的准稀疏性。

我们重复了相同的实验，但将 Matérn 模型的实际作用范围设为 0.4。这使得$\beta = 0.133, 0.084, 0.067$ 分别对应 $\nu = 0, 1, 2$ 。结果见表 2。结论与前一种设定相同，但在这种情况下， $\boldsymbol{\Sigma}_n$ 的稀疏性较低， $\boldsymbol{\Sigma}_n^{-1}$的准稀疏性也较低，$\mathrm{ch}(\boldsymbol{\Sigma}_n^{-1}$亦如此。

这些数值实验凸显了一个明显的权衡：当固定 β 和 ν时，随着 μ 增大，即从紧支撑 Matérn 模型转向全局支撑 Matérn 模型， $\boldsymbol{\Sigma}_n^{-1}$（或 $\mathrm{ch}(\boldsymbol{\Sigma}_n^{-1})$）的（准）稀疏性与$\boldsymbol{\Sigma}_n$ 之间存在权衡。具体而言，当 $\mu \to \infty$ （Matérn 模型）时， $\boldsymbol{\Sigma}_n^{-1}$具有很强的准稀疏性，而 $\boldsymbol{\Sigma}_n$ 是稠密的。相反，当 μ 较小时， $\boldsymbol{\Sigma}_n^{-1}$不具备准稀疏性，但 $\boldsymbol{\Sigma}_n$ 具有很强的稀疏性。这似乎表明，稀疏精度矩阵近似对于 Matérn 模型应当效果良好，但在处理表现出短程紧支撑依赖的数据时可能会遇到问题。在这种情况下，更好的方法应当是利用 $\boldsymbol{\Sigma}_n$的稀疏性，这可通过增强版 Matérn 模型实现。

9. 结论

自 Matérn 模型问世以来，它产生了深远影响，并持续广泛应用于各类科学学科及其他领域。Matérn 模型最初的研究动机源于其在空间插值中的灵活性，如今，关于该模型的替代版本和增强版本也已有丰富的研究成果。尤其是，随机偏微分方程及相关方法使人们能够在相当一般的定义域上定义 Matérn 模型的类比模型，并采用精度矩阵的稀疏近似；与此同时，近期具有紧支撑的增强模型研究进展，有助于通过对协方差矩阵进行稀疏近似来实现可扩展计算，而且非常适用于具有短尺度依赖的过程。近来，人们对这些增强模型的理论和经验性质进行了积极研究。另一方面，仍有一些具有实际重要性的理论问题尚待解决，例如有限样本量下的参数估计，以及参数估计对相关预测性能的影响。

我们目前对 Matérn 模型的理解，是不同学科的科学家与从业者相互交流的成果；我们希望，这种多学科视角能进一步加深对 Matérn 模型的认识。

参考文献

[1] Åberg, S. 和 Podgórski, K.（2011）。一类非高斯二阶随机场。Extremes，14(2)：187–222。

[2] Abramowitz, M. 和 Stegun, I. A.，编（1970年）。《数学函数手册》。Dover，纽约。[3] Alegría, A.、Cuevas-Pacheco, F.、Diggle, P. 和 Porcu, E.（2021a）。F族协方差函数：一种用于球面随机场建模的 Matérn 类比。Spatial Statistics，43：100512。[4] Alegría, A.、Emery, X. 和 Porcu, E.（2021b）。带交叉凹陷的双变量 Matérn 协方差：用于对区域化变量建模。Spatial Statistics，41：100491。[5] Alegría, A.、Ramírez, F. 和 Porcu, E.（2023年）。空间随机场各向同性协方差函数的混合参数类。arXiv预印本 arXiv:2301.05602。[6] Allard, D.、Clarotto, L. 和 Emery, X.（2022）。用于多元时空数据的完全不可分离Gneiting协方差函数。Spatial Statistics，52：100706。[7] Allard, D.、Senoussi, R. 和 Porcu, E.（2016年）。空间数据的各向异性模型。Mathematical Geosciences，48(3)：305–328。[8] Anderes, E.（2010年）。高斯随机场中尺度与方差的一致分离。《统计年鉴》，38(2)：870–893。[9] Anderes, E.、Møller, J. 和 Rasmussen, J. G.（2020年）。图及其边上的各向同性协方差函数。《统计年鉴》，48(4)：2478–2503。[10] Angulo, J.、Kelbert, M. Y.、Leonenko, N. 和 Ruiz-Medina, M. D.（2008年）。与随机分数阶Helmholtz方程和热方程相关的时空随机场。Stochastic Environmental Research and Risk Assessment，22(1)：3–13。[11] Apanasovich, T. V.、Genton, M. G. 和 Sun, Y.（2012年）。适用于任意数量分量多元随机场的有效Matérn类互协方差函数。美国统计学会会刊，107(497)：180–193。[12] Askey, R.（1973年）。径向特征函数。技术报告，研究中心，威斯康星大学。[13] Bachoc, F.、Bevilacqua, M. 和 Velandia, D.（2019年）。固定域渐近条件下高斯过程的复合似然估计。多元分析杂志，174：104534。[14] Bachoc, F.、Porcu, E.、Bevilacqua, M.、Furrer, R. 和 Faouzi, T.（2022）。多元地统计学中的渐近等价预测。伯努利，28(4)：2518–2545。[15] Bakka, H.、Krainski, E.、Bolin, D.、Rue, H. 和 Lindgren, F.（2020年）。Matérn场向时空的基于扩散的推广。arXiv预印本 arXiv:2006.04917。[16] Banerjee, S. 和 Gelfand, A. E.（2006年）。贝叶斯边界分析：空间过程模型下的曲线梯度评估。美国统计学会会刊，101(476)：1487–1501。[17] Banerjee, S.、Gelfand, A. E.、Finley, A. O. 和 Sang, H.（2008年）。大型空间数据集的高斯预测过程模型。Journal of the Royal Statistical Society: Series B (Statistical Methodology)，70(4)：825–848。[18] Banerjee, S.、Gelfand, A. E. 和 Sirmans, C. F.（2003年）。空间过程模型下的方向变化率。美国统计学会会刊，98(464)：946–954。[19] Barp, A.、Oates, C. J.、Porcu, E. 和 Girolami, M.（2022）。一种黎曼–Stein核方法。伯努利，28(4)：2181–2208。[20] Benzi, M.（2016年）。矩阵计算中的局部化：理论与应用。载于 Benzi, M. 和 Simoncini, V.（编），《利用矩阵计算中的隐藏结构：算法与应用：意大利切特拉罗，2015》，第211–317页。Springer国际出版。[21] Berg, C.（2008年）。Stieltjes–Pick–Bernstein–Schoenberg及其与完全单调性的联系。载于 Mateu, S. 和 Porcu, E.（编），《正定函数：从Schoenberg到时空挑战》，第15–45页。西班牙：数学系，海梅一世大学，卡斯特利翁-德拉普拉纳。

[22] Berg, C.、Mateu, J. 和 Porcu, E.（2008年）。各向同性相关函数的Dagum族。伯努利，14(4)：1134–1149。[23] Bevilacqua, M.、Caamaño-Carrillo, C. 和 Porcu, E.（2022）。在空间统计学中统一紧支撑协方差函数与Matérn协方差函数。多元分析杂志，189：104949。[24] Bevilacqua, M.、Caamaño-Carrillo, C.、Arellano-Valle, R. B. 和 Morales-Oñate, V.（2021）。使用（偏斜）t过程进行非高斯地统计建模。Scandinavian Journal of Statistics，48(1)：212–245。[25] Bevilacqua, M.、Faouzi, T.、Furrer, R. 和 Porcu, E.（2019年）。固定域渐近条件下使用广义Wendland函数进行估计和预测。《统计年鉴》，47(2)：828–856。[26] Bevilacqua, M. 和 Gaetan, C.（2015年）。比较基于空间高斯随机场成对数据的复合似然方法。《统计与计算》，25(5)：877–892。[27] Bingham, N. H.、Goldie, C. M. 和 Teugels, J.（1987年）。正则变差。《数学及其应用百科全书》，27：145–156。[28] Bochner, S.（1955年）。《调和分析与概率论》。《加州数学科学专著》。加利福尼亚大学出版社。[29] Bolin, D.（2014年）。由非高斯噪声驱动的空间Matérn场。Scandinavian Journal of Statistics，41(3)：557–579。[30] Bolin, D. 和 Kirchner, K.（2020年）。具有一般光滑性的高斯随机场的有理SPDE方法。Journal of Computational and Graphical Statistics，29(2)：274–285。[31] Bolin, D. 和 Kirchner, K.（2022）。具有分数阶协方差算子的高斯随机场的测度等价与渐近最优线性预测。arXiv预印本 arXiv:2101.07860v2。[32] Bolin, D.、Simas, A. B. 和 Wallin, J.（2022）。度量图上的高斯Whittle-Matérn场。[33] Bolin, D. 和 Wallin, J.（2016年）。多元G型Matérn随机偏微分方程随机场。arXiv预印本 arXiv:1606.08298。[34] Borovitskiy, V.、Azangulov, I.、Terenin, A.、Mostowsky, P.、Deisenroth, M. 和 Durrande, N.（2021）。图上的Matérn高斯过程。载于 Banerjee, A. 和 Fukumizu, K.（编），《第24届人工智能与统计国际会议论文集》，《机器学习研究论文集》第130卷，第2593–2601页。机器学习研究论文集。[35] Borovitskiy, V.、Terenin, A.、Mostowsky, P. 等（2020年）。黎曼流形上的Matérn高斯过程。Advances in Neural Information Processing Systems，33：12426–12437。[36] Bourotte, M.、Allard, D. 和 Porcu, E.（2016年）。用于多元时空数据的灵活不可分离交叉协方差函数类。Spatial Statistics，18：125–146。[37] Brown, R. D.、Bardsley, J. M. 和 Cui, T.（2020年）。贝叶斯反问题中用于建模Whittle–Matérn先验的半变异函数方法。《反问题》，36(5)：055006。[38] Buhmann, M.（2001年）。一种具有紧支撑的新型径向基函数类。《计算数学》，70(233)：307–318。[39] Cameletti, M.、Lindgren, F.、Simpson, D. 和 Rue, H.（2013）。通过SPDE方法对颗粒物浓度进行时空建模。《AStA统计分析进展》，97(2)：109–131。[40] Chernih, A. 和 Hubbert, S.（2014年）。广义Wendland函数的闭式表示及其性质。《逼近理论杂志》，177：17–33。[41] Chilès, J. 和 Delfiner, P.（2012年）。《地统计学：空间不确定性建模》。Wiley，纽约。[42] Clarotto, L.、Allard, D.、Romary, T. 和 Desassis, N.（2022）。具有平流和扩散的时空数据集的SPDE方法：一种无矩阵方法。载于第15届国际会议

计算统计与方法学ERCIM工作组（CMStatistics 2022）。[43] Cockayne, J., Oates, C. J., Sullivan, T. J., and Girolami, M.（2019年）。贝叶斯概率数值方法。SIAM评论，61(4):756–789。[44] Cressie, N.（1990年）。克里金的起源。数学地质学，22(3):239–252。[45] Cressie, N. and Johannesson, G.（2008年）。超大型空间数据集的固定秩克里金。Journal of the Royal Statistical Society: Series B (Statistical Methodology)，70(1):209–226。[46] Daley, D. J., Porcu, E., and Bevilacqua, M.（2015年）。多元随机场的紧支撑协方差函数类。随机环境研究与风险评估，29(4):1249–1263。[47] Datta, A.（2022）。空间统计中的最近邻稀疏Cholesky矩阵。WIREs计算统计学，14(5):e1574。[48] Datta, A., Banerjee, S., Finley, A. O., and Gelfand, A. E.（2016年）。大型地统计数据集的分层最近邻高斯过程模型。美国统计学会会刊，111(514):800–812。[49] Datta, A., Banerjee, S., Finley, A. O., Hamm, N. A., and Schaap, M.（2016年）。用于大型时空数据的非可分动态最近邻高斯过程模型及其在颗粒物分析中的应用。《应用统计年鉴》，10(3):1286。[50] Davydov, O. and Schaback, R.（2019年）。Sobole夫空间中的最优模板。IMA数值分析杂志，39:398–422。[51] Dey, D., Datta, A., and Banerjee, S.（2022）。高度多元空间数据的图高斯过程模型。Biometrika，109(4):993–1014。[52] Eidsvik, J., Shaby, B., Reich, B., Wheeler, M., and Niemi, J.（2013）。采用分块复合似然的空间模型估计与预测。Journal of Computational and Graphical Statistics，即将发表。[53] Emery, X. and Alegría, A.（2022）。用于欧几里得空间中二阶平稳随机场建模的高斯超几何协方差核：其紧支撑、性质与谱表示。随机环境研究与风险评估，36:2819—-2834。[54] Emery, X., Porcu, E., and White, P.（2022）。多元Matérn共区域化模型的新有效性条件及其在勘探地球化学中的应用。数学地球科学，页码1–26。[55] Faouzi, T., Porcu, E., and Bevilacqua, M.（2022）。固定域渐近下采用紧支撑协方差函数的时空估计与预测。Statistica Sinica，32:1–17。[56] Faouzi, T., Porcu, E., Bevilacqua, M., and Kondrashuk, I.（2020年）。Zastavnyi算子与正定径向函数。统计与概率快报，157:108620。[57] Fasshauer, G.（1997）。通过径向基函数配置法求解偏微分方程。载于LeMéhauté, A., Rabut, C., and Schumaker, L.（编），《曲面拟合与多分辨率方法》，页131–138。范德堡大学出版社，田纳西州纳什维尔。[58] Fasshauer, G. and McCourt, M.（2015年）。使用MATLAB的基于核的近似方法，《跨学科数学科学》第19卷。世界科学出版社，新加坡。[59] Furrer, R., Genton, M. G., a nd Nychka, D.（2006）。大型空间数据集插值的协方差截断。Journal of Computational and Graphical Statistics，15(3):502–523。[60] Genton, M. G. and Kleiber, W.（2015年）。多元地统计学的互协方差函数。统计科学，30(2):147–163。

[61] Gneiting, T.（2002a）。紧支撑相关函数。多元分析杂志，83:493–508。[62] Gneiting, T.（2002b）。时空数据的不可分离平稳协方差函数。J. Am. Statist. Ass.，97:590–600。[63] Gneiting, T.（2013）。球面上的严格与非严格正定函数。伯努利，19(4):1327–1349。[64] Gneiting, T., Kleiber, W., and Schlather, M.（2010）。多元随机场的Matérn交叉协方差函数。美国统计学会会刊，105:1167–1177。[65] Gneiting, T. and Schlather, M.（2004）。分离分形维数与赫斯特效应的随机模型。SIAM评论，46(2):269–282。[66] Gneiting, T., Ševˇcíková, H., and Percival, D. B.（2012）。分形维数估计量：评估时间序列与空间数据的粗糙度。统计科学，27(2):247 – 277。[67] Guinness, J.（2018）。用于改进高斯过程近似的置换与分组方法。技术计量学，60(4):415–429。[68] Guinness, J.（2022）。网格上的Matérn协方差逆。Biometrika，109(2):535–541。[69] Guinness, J. and Fuentes, M.（2016年）。球面上的各向同性协方差函数：一些性质与建模考量。多元分析杂志，143:143–152。[70] Guttorp, P. and Gneiting, T.（2006）。概率与统计学史研究XLIX：关于Matérn相关函数族。Biometrika，93(4):989–995。[71] Halder, A., Banerjee, S., and Dey, D. K.（2023年）。采用空间曲率过程的贝叶斯建模。美国统计学会会刊，?(just-accepted):1–27。[72] Harrar, S. W., Seneta, E., and Gupta, A. K.（2006）。矩阵变元t分布与矩阵变元vg分布之间的对偶性。多元分析杂志，97(6):1467–1475。[73] Hartikainen, J. and Särkkä, S.（2010）。时间高斯过程回归模型的卡尔曼滤波与平滑解。载于2010年IEEE机器学习信号处理国际研讨会，页379–384。电气电子工程师学会。[74] Heaton, M. J., Datta, A., Finley, A. O., Furrer, R., Guinness, J., Guhaniyogi, R., Gerber, F., Gramacy, R. B., Hammerling, D., Katzfuss, M., Lindgren, F., Nychka, D. W., Sun, F., and Zammit-Mangion, A.（2019年）。分析大型空间数据的方法案例研究竞赛。农业、生物与环境统计学杂志，24:398–425。[75] Hennig, P., Osborne, M. A., and Kersting, H. P.（2022）。概率数值学：作为机器学习的计算。剑桥大学出版社。[76] Hubbert, S.（2012）。一类紧支撑径向基函数的闭式表示。计算数学进展，36(1):115–136。[77] Ip, H. L. and Li, W. K.（2017）。关于时空随机场的若干Matérn协方差函数。Statistica Sinica，27:805–822。[78] Jansson, E., Kovács, M., and Lang, A.（2022）。球面Whittle–Matérn高斯随机场的曲面有限元近似。SIAM科学计算杂志，44(2):A825–A842。[79] Jaquier, N., Borovitskiy, V., Smolensky, A., Terenin, A., Asfour, T., and Rozo, L.（2022）。利用黎曼Matérn核进行几何感知贝叶斯机器人学优化。载于机器人学习会议，页794–805。机器学习研究论文集。[80] Jensen, K., Kao, T.-C., Tripodi, M., and Hennequin, G.（2020年）。用于发现神经数据中非欧几里得潜在结构的流形高斯过程潜变量模型。Advances in Neural Information Processing Systems，33:22580–22592。[81] Karvonen, T.（2022）。高斯过程回归中光滑性参数估计的渐近界。arXiv预印本arXiv:2203.05400。[82] Karvonen, T. 和 Oates, C. J.（2022）。高斯过程回归中的最大似然估计是病态的。机器学习研究杂志。即将发表。 [83] Katzfuss, M. 和 Guinness, J.（2021）。高斯过程 Vecchia 近似的通用框架。统计科学，36(1):124 – 141。 [84] Kaufman, C. G. 和 Shaby, B. A.（2013）。范围参数在地统计学估计与预测中的作用。Biometrika，100:473–484。[85] Kelbert, M. Y.、Leonenko, N. N. 和 Ruiz-Medina, M.（2005）。与随机分数阶热方程相关的分数阶随机场。应用概率进展，37(1):108–133。[86] Kent, J. T.（1989）。随机场的连续性性质。概率年刊，页码 1432–1440。 [87] Khristenko, U.、Scarabosio, L.、Swierczynski, P.、Ullmann, E. 和 Wohlmuth, B.（2019）。对基于偏微分方程的 Whittle-Matérn 随机场采样中的边界效应进行分析。SIAM/ASA 不确定性量化杂志，7(3):948–974。 [88] Kleiber, W. 和 Nychka, D.（2012）。多元空间过程的非平稳建模。多元分析杂志，112:76–91。[89] Kleiber, W. 和 Porcu, E.（2015）。非平稳矩阵协方差：紧支撑、长程依赖与拟算术构造。随机环境研究与风险评估，29(1):193–204。[90] Koepernik, P. 和 Pfaff, F.（2021）。度量空间中高斯过程回归的一致性。J. Mach. Learn. Res.，22:244–1。 [91] Laga, I. 和 Kleiber, W.（2017）。修正 Matérn 过程。Stat，6(1):241–247。[92] Lang, A. 和 Pereira, M.（2021）。紧黎曼流形上高斯随机场的 Galerkin–Chebyshev 近似。arXiv 预印本 arXiv:2107.02667。[93] Lavin, A. 、 Zenil, H. 、 Paige, B. 、 Krakauer, D. 、Gottschlich, J.、Mattson, T.、Anandkumar, A.、Choudry, S.、Rocki, K.、Baydin, A. G. 等（2021）。模拟智能：迈向新一代科学方法。arXiv 预印本 arXiv:2112.03235。 [94] Leonenko, N. 和 Malyarenko, A.（2017）。Matérn 类张量值随机场及其他。统计物理杂志，168(6):1276–1301。[95] Leonenko, N.、Malyarenko, A. 和 Olenko, A.（2022）。关于球内随机场的谱理论。Theory of Probability and Mathematical Statistics，107:61–76。[96] Leonenko, N.、Ruiz-Medina, M. D. 和 Taqqu, M. S.（2011）。分数阶椭圆、双曲和抛物随机场。Electronic Journal of Probability，16:1134–1172。 [97] Li, D.、Jones, A.、Banerjee, S. 和 Engelhardt, B. E.（2021a）。多组高斯过程。arXiv 预印本 arXiv:2110.08411。 [98] Li, D.、Tang, W. 和 Banerjee, S.（2021b）。紧黎曼流形上具有 Matérn 协变图的高斯过程推断。arXiv 预印本 arXiv:2104.03529。 [99] Lilly, J. M.、Sykulski, A. M.、Early, J. J. 和 Olhede, S. C.（2017）。分数布朗运动、Matérn 过程与湍流扩散的随机建模。地球物理学非线性过程，24(3):481–514。 [100] Lim, S. 和 Teo, L. P.（2009）。广义 Whittle-Matérn 随机场：相关涨落的模型。物理学杂志A：数学与理论，42(10):105202。[101] Lindgren, F.、Bolin, D. 和 Rue, H.（2022）。高斯场与非高斯场的随机偏微分方程方法：十年历程，仍在继续。空间统计，页码 100599。 [102] Lindgren, F.、Rue, H. 和 Lindstroem, J.（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。英国皇家统计学会会刊B辑

, 73:423 –498.

[103] Lindsay, B.（1988年）。复合似然方法。《当代数学》，80：221–239。[104] Lodhia, A., Sheffield, S., Sun, X., 和 Watson, S. S.（2016年）。分数阶高斯场：综述。《概率调查》，13(none)：1–56。[105] Loh, W.-L., Sun, S., 和 Wen, J.（2021）。关于固定域渐近、参数估计以及具有 Matérn 协方差函数的各向同性高斯随机场。《统计年鉴》，49(6)：3127–3152。[106] Ma, P. 和 Bhadra, A.（2022）。超越 Matérn：一类可解释的合流超几何协方差函数。《美国统计学会会刊》，页码 1–14。[107] Mardia, K. V. 和 Marshall, J.（1984）。空间回归中残差协方差模型的最大似然估计。Biometrika，71：135–146。[108] Matérn, B.（1986）。《空间变异：随机模型及其在森林调查和其他抽样调查问题中的应用》。Springer，海德堡，第2版。[109] Matheron, G.（1963）。《应用地质统计学论》。第二卷：克里金法。法国地质与矿业研究局回忆录，第24号。法国地质与矿业研究局出版社，巴黎。[110] Matheron, G.（1965）。《区域化变量及其估计》。Masson，巴黎。[111] Matheron, G.（1971）。《区域化变量理论及其应用》。地质统计学中心，巴黎矿业学院，法国枫丹白露。[112] Menegatto, V., Oliveira, C., 和 Porcu, E.（2020）。Gneiting类、半度量空间与等距嵌入。《构造性数学分析》，3(2)：85–95。[113] Morales-Navarrete, D., Bevilacqua, M., Caamaño-Carrillo, C., 和 Castro, L.（2022）。点参照空间计数数据建模：一种泊松过程方法。《美国统计学会会刊》，0(0)：1–25。[114] Narcowich, F. J., Ward, J. D., 和 Wendland, H.（2006）。通过径向基函数进行散乱数据插值的索伯列夫误差估计与伯恩斯坦不等式。《构造性逼近》，24(2)：175–186。[115] Nikitin, A. V., John, S., Solin, A., 和 Kaski, S.（2022）。基于随机偏微分方程的不可分离时空图核。载于《人工智能与统计学国际会议》，页码 10640–10660。机器学习研究论文集。[116] Novak, E. 和 Wo´zniakowski, H.（2008年）。《多元问题的可处理性：泛函的标准信息》，第2卷。欧洲数学会。[117] Oates, C. J., Girolami, M., 和 Chopin, N.（2017）。蒙特卡罗积分的控制泛函。Journal of the Royal Statistical Society: Series B (Statistical Methodology)，79(3)：695–718。[118] Olver, F. W., Lozier, D. M., Boisvert, R. F., 和 Clark, C. W.（2010年）。《NIST数学函数手册》。剑桥大学出版社，剑桥。[119] Paciorek, C. J. 和 Schervish, M. J.（2006）。使用一类新的非平稳协方差函数进行空间建模。《环境计量学：国际环境计量学会官方期刊》，17(5)：483–506。[120] Palacios, M. B. 和 Steel, M. F. J.（2006）。非高斯贝叶斯地质统计建模。《美国统计学会会刊》，101：604–618。[121] Pereira, M., Desassis, N., 和 Allard, D.（2022）。黎曼流形上大数据集的地统计学：一种无矩阵方法。arXiv预印本 arXiv:2208.12501。[122] Porcu, E., Alegría, A., 和 Furrer, R.（2018）。空间全局且随时间演变的数据建模。《国际统计评论》，86：344–377。[123] Porcu, E., Bevilacqua, M., 和 Genton, M. G. (2020a). 具有动态紧支撑的不可分离时空协方差函数。Statistica Sinica, 30:719–739。[124] Porcu, E., Daley, D. J., Buhmann, M., 和 Bevilacqua, M. (2013). 用于多变量地统计学的紧支撑径向基函数。随机环境研究与风险评估, 27(4):909–922。[125] Porcu, E., Emery, X., 和 Mery, N. (2022a). 空间各向同性且时间对称的矩阵值协方差函数的判据与表征。Comp. Appl. Math., 41(223)。[126] Porcu, E., Furrer, R., 和 Nychka, D. (2021). 时空协方差函数的30年。Wiley Interdisciplinary Reviews: Computational Statistics, 13(2):e1512。[127] Porcu, E., White, P. A., 和 Genton, M. G. (2022b). 跨时间网络上的不可分离时空平稳协方差函数。arXiv预印本 arXiv:2208.03359。[128] Porcu, E., Zastavnyi, V., Bevilacqua, M., 和 Emery, X. (2020b). 具有紧支撑的协方差的Stein假设与屏蔽效应。《电子统计学杂志》, 14(2):2510–2528。[129] Porcu, E., Zastavyi, P., 和 Bevilacqua, M. (2016). Buhmann协方差函数、其紧支撑及其光滑性。arXiv预印本 arXiv:1606.09527。[130] Posa, D. (2023年). 各向同性协方差函数的特殊类别。随机环境研究与风险评估，即将发表。[131] Quick, H., Banerjee, S., 和 Carlin, B. P. (2013). 加州区域汇总哮喘住院数据的时间梯度建模。《应用统计年鉴》, 7(1):154。[132] Roininen, L., Girolami, M., Lasanen, S., 和 Markkanen, M. (2019年). 用于贝叶斯反演的Matérn场超先验。《逆问题与成像》, 13(1)。[133] Sanz-Alonso, D. 和 Yang, R. (2022a). 高斯过程的有限元表示：平衡数值精度与统计精度。SIAM/ASA Journal on Uncertainty Quantification, 10(4):1323–1349。[134] Sanz-Alonso, D. 和 Yang, R. (2022b). Matérn场的随机偏微分方程方法：图表示。统计科学, 37(4):519–540。[135] Sarkka, S., Solin, A., 和 Hartikainen, J. (2013). 通过无限维贝叶斯滤波与平滑进行时空学习：从卡尔曼滤波视角看高斯过程回归。《IEEE信号处理杂志》, 30(4):51–61。[136] Schaback, R. (2011). 缺失的Wendland函数。Advances in Computational Mathematics, 34(1):67–81。[137] Schaback, R. (2015). 用于比较所有线性偏微分方程求解器的计算工具。Advances of Computational Mathematics, 41:333–355。[138] Schaback, R. (2017). 节点无网格方法的误差分析。载于 Griebel, M. 和 Schweitzer, M.，编者，《偏微分方程无网格方法 VIII》，计算科学与工程讲义第115卷，第117–143页。Springer。[139] Schaback, R. 和 Wendland, H. (2006年). 核技术：从机器学习到无网格方法。《数值学报》, 15:543–639。[140] Scheuerer, M. (2010年). 一般二阶随机场样本路径的正则性。Stochastic Processes and their Applications, 120(10):1879–1897。[141] Scheuerer, M., Schaback, R., 和 Schlather, M. (2013). 空间数据插值——随机问题还是确定性问题? 《欧洲应用数学杂志》, 24(4):601–629。[142] Schoenberg, I.-J. (1938). 度量空间与完全单调函数。《数学年刊》, 39(4):811–841。

[143] Schoenberg, I. J.（1942年）。关于正定函数球面。Duke Math. Journal, 9:96–108。[144] Shaby, B. 和 Ruppert, D. (2012年). 截尾协方差：贝叶斯估计与渐近理论。Journal of Computational and Graphical Statistics, 21:433–452。[145] South, L. F., Karvonen, T., Nemeth, C., Girolami, M., 和 Oates, C. J. (2022). 源自Sard方法的半精确控制泛函。Biometrika, 109(2):351–367。[146] Stein, M. (1988年). 使用误设协方差函数对随机场进行渐近有效预测。The Annals of Statistics, 16:55–63。[147] Stein, M. (2005). 常规监测数据的统计方法。Journal of the Royal Statistical Society B, 67:667–687。[148] Stein, M. L. (1999年).《空间数据插值：克里金的一些理论》。Springer，纽约。[149] Stein, M. L. (2002年). 克里金中的屏蔽效应。The Annals of Statistics, 30(1):298–323。[150] Stein, M. L. (2004年). 某些非平稳随机场的高斯测度的等价性。Journal of Statistical Planning and Inference, 123(1):1–11。[151] Stein, M. L. (2011). 2010年Rietz讲座：屏蔽效应何时成立？The Annals of Statistics, 39(6):2795–2819。[152] Stein, M. L. (2015年). 屏蔽效应何时不成立？空间统计, 11:65–80。[153] Stein, M. L., Chi, Z., 和 Welty, L. J. (2004年). 大型空间数据集的似然近似。《皇家统计学会杂志：B辑（统计方法学）》，66(2):275–296。[154] Sun, Y., Li, B., 和 Genton, M. G. (2012 年). 大数据集的地统计学。载于《自然事件时空建模的进展与挑战》，第55–77页。Springer。[155] Tang, J. 和 Zimmerman, D. (2020年). 网络上的时空协方差模型及其在河流中的应用。arXiv预印本 arXiv:2009.14745。[156] Tang, W., Zhang, L., 和 Banerjee, S. (2021). 高斯空间过程模型中块金效应的可识别性与一致性。Journal of the Royal Statistical Society: Series B (Statistical Methodology), 83(5):1044–1070。[157] Terdik, G. (2015年). 非高斯各向同性场的角谱。Brazilian Journal of Probability and Statistics, 29(4):833–865。[158] Terdik, G. (2022年). 高维空间中拉普拉斯ARMA场的时空协方差函数。Theory of Probability and Mathematical Statistics, 107:111–132。[159] Tuo, R. 和 Wang, W. (2020年). 具有各向同性Matérn相关函数的克里金预测：稳健性与试验设计。J. Mach. Learn. Res., 21:187–1。[160] Tuo, R., Wang, Y., 和 Jeff Wu, C. (2020年). 用于计算机模型校准的Matérn型核岭回归改进收敛速率。SIAM/ASA Journal on Uncertainty Quantification, 8(4):1522–1547。[161] Vakili, S., Moss, H., Artemev, A., Dutordoir, V., 和 Picheny, V. (2021). 使用稀疏高斯过程模型的可扩展汤普森采样。Advances in Neural Information Processing Systems, 34:5631–5643。[162] Varin, C., Reid, N., 和 Firth, D. (2011). 复合似然方法概述。Statistica Sinica, 21:5–42。[163] Vecchia, A. V. (1988年). 连续空间过程的估计与模型识别。Journal of the Royal Statistical Society. Series B (Methodological), 第297–312页。[164] Wallin, J. 和 Bolin, D. (2015年). 使用非高斯Matérn场进行地统计建模。Scandinavian Journal of Statistics, 42(3):872–890。[165] Wendland, H. (1995年). 最小次数的分段多项式、正定且紧支撑的径向函数。Advances in Computational Mathematics, 4:389–396。[166] Wendland, H.（2004年）。散乱数据逼近，第17卷。剑桥大学出版社。[167] Whittle, P.（1953年）。多重平稳时间序列分析。英国皇家统计学会会刊B辑，15(1)：125–139。[168] Whittle, P.（1963年）。多维随机过程。国际统计学会通报，40(2)：974–994。[169] Williams, C. K. 和 Rasmussen, C. E.（2006年）。用于机器学习的高斯过程，第2卷。麻省理工学院出版社，剑桥，马萨诸塞州。[170] Wilson, A. G.、Hu, Z.、Salakhutdinov, R. 和 Xing, E. P.（2016年）。深度核学习。载于《人工智能与统计学》，第370–378页。机器学习研究论文集。[171] Wilson, J. T.、Borovitskiy, V.、Terenin, A.、Mostowsky, P. 和 Deisenroth, M. P.（2021年）。高斯过程的路径条件化。J. Mach. Learn. Res.，22：105–1。[172] Wu, Z.（1992年）。利用径向基函数对散乱数据进行Hermite–Birkhoff插值。《逼近理论及其应用》，8/2：1–10。[173] Xua, G. 和 Genton, M. G.（2017年）。Tukey g-and-h随机场。美国统计学会会刊，112：1236–1249。[174] Yan, Y. 和 Genton, M. G.（2018年）。具有Matérn协方差函数的变换高斯随机场数据的高斯似然推断。《环境计量学》，29(5-6)：e2458。[175] Zastavnyi, V.（2002年）。正定径向函数与样条函数。《数学报告》，66：213–216。[176] Zastavnyi, V.（2006年）。关于Buhmann函数的一些性质。《乌克兰数学杂志》，58(8)。[177] Zastavnyi, V. 和 Porcu, E.（2017年）。关于某些径向函数的正定性。《罗巴切夫斯基数学杂志》，38(2)：386–394。[178] Zastavnyi, V. P. 和 Porcu, E.（2011年）。Gneiting类时空协方差函数的刻画定理。Bernoulli，17(1)：456–465。[179] Zhang, H.（2004年）。基于模型的地质统计学中的不一致估计与渐近等价插值。美国统计学会会刊，99：250–261。[180] Zhang, H. 和 Zimmerman, D.（2005）。趋向调和空间统计学中的两个渐近框架。Biometrika，92：921–936。

补充材料

附录A：非常规情境下的 Matérn 建模：扩展版

有人可能会认为，Matérn 类仅限于取标量值且平稳、各向同性的随机场。尽管这一点属实，但同样确实的是，Matérn 类是构建更为复杂情形的基础模块。我们将在全文中逐一列出这些情形。

1. 标量值随机场。

a) 各向异性。如果空间维度沿某些优先方向延伸，那么各向同性就不再是空间建模中的现实假设。文献中探讨了多种各向异性 [7], , 并表明 Matérn 类可以与专门设计的变形组合，从而将优先方向纳入空间依赖性的考量。b) 非平稳性。Matérn 核已被 [119] 用于构建非平稳模型。考虑一组以均值为索引的高斯分布，其中均值为 $\boldsymbol{x}$ 的分布具有协方差矩阵 $\Sigma_{\boldsymbol{x}}$。设

$$Q_{\boldsymbol{x},\boldsymbol{y}} = (\boldsymbol{x}-\boldsymbol{y})^\top \left(\frac{\Sigma_{\boldsymbol{x}}+\Sigma_{\boldsymbol{y}}}{2}\right)^{-1} (\boldsymbol{x}-\boldsymbol{y})$$

and

$$K(\boldsymbol{x},\boldsymbol{y}) = \left|\Sigma_{\boldsymbol{x}}\right|^{1/4}\left|\Sigma_{\boldsymbol{y}}\right|^{1/4}\left|\frac{\Sigma_{\boldsymbol{x}}+\Sigma_{\boldsymbol{y}}}{2}\right|^{-1/2} \mathcal{M}_{\nu,\alpha}\left(\sqrt{Q_{\boldsymbol{x},\boldsymbol{y}}}\right).$$

那么， K 是正定的。Roininen等人 [132] 最近将这一方法推广了。另一种方法是对 Matérn 协方差函数的输入进行扭曲，以此引入非平稳性，如 $K(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{M}_{\nu,\alpha}(\|\boldsymbol{w}(\boldsymbol{x}) - \boldsymbol{w}(\boldsymbol{y})\|)$，其中微分同胚 $\boldsymbol{w}: \mathbb{R}^d \to \mathbb{R}^d$ 本身也可以参数化。Wilson等人 [170] 探讨了将 $\boldsymbol{w}$ 参数化为深度神经网络的情形。

c) 图与拟度量空间。[9] 考虑配备欧几里得边的图，并在图上采用测地距离或电阻度量。他们证明，只要满足 $0 < \nu \leq 1/2$，就可以将 $\mathcal{M}_{\nu,\alpha}$ 与图上的电阻度量结合使用。最近，[112] 通过考虑拟度量空间，对这一设定进行了推广。显然，在这种情况下也存在类似的限制。近期， [32] 采用了不同的方法，在度量图上构造具有相应协方差结构的随机场。 [155]和 [127]研究了空间为具有欧几里得边的图时，Matérn 类的时空版本。

d) 时空。对于时空高斯随机场 $\{Z(\boldsymbol{x},t) \quad \boldsymbol{x} \in \mathbb{R}^{d+1}, t \in \mathbb{R}\}$，一个典型的二阶假设是协方差在空间上各向同性、在时间上平稳。也就是说，

$$\operatorname{Cov}\left(Z(\boldsymbol{x},t), Z(\boldsymbol{y},t')\right) = K\left(\|\boldsymbol{x}-\boldsymbol{y}\|, |t-t'|\right) \tag{35}$$

对所有 $(\boldsymbol{x},t),(\boldsymbol{y},t') \in \mathbb{R}^d \times \mathbb{R}$均成立。Matérn 函数已被用作构造此类结构的基本构件。特别值得关注的是非可分协方差函数，它们允许空间与时间之间存在相互作用；在这一背景下， [62] 和 [178] 证明了

$$K(x,u) = \frac{\sigma^2}{\psi(u^2)}\mathcal{M}_{\nu,\alpha}\left(\frac{x}{\psi(u^2)}\right), \; x,u \geq 0, \tag{36}$$

生成一个类型为（35）的有效时空协方差函数。这里， ψ 是一个具有完全单调导数的严格正函数 [21]。

e) 非高斯场。通过考虑变换 $w: \mathbb{R} \to \mathbb{R}$，可以将 Matérn 模型 $\tilde{Z}$ 作为一系列非高斯模型的基础 $Z(\boldsymbol{x}) = w(\tilde{Z}(\boldsymbol{x}))$。然而， Z 的协方差函数通常并非 Matérn 协方差函数。关于是否存在协方差函数仍属于 Matérn 类的非高斯过程这一问题，[1]给出了肯定的回答。 [174] 提出了具有 Matérn 协方差函数的变换高斯随机场。 [29] 随后， [164] 为非高斯 Matérn 场提供了基于随机偏微分方程的构造。通过对高斯场（的独立重复样本）进行变换，可以得到一类具有核 $g(\mathcal{M}_{\nu,\alpha})$的广义非高斯场，其中 $g(\cdot)$ 是保持正定性的适当函数（例如参见 [120, 173, 24, 113]）。

f) 高斯场混合。可以将现实世界现象建模为叠加 Z_1+Z_2，其中 Z_1可能被选用来捕捉总体趋势（例如，对于一些固定函数ϕ_1，.....，ϕ_p，以及一些系数 $\boldsymbol{c}\in\mathbb{R}^p$，$Z_1(\boldsymbol{x})=c_1\phi_1(\boldsymbol{x})+\cdots+c_p\phi_p(\boldsymbol{x})$），而 Z_2可以是 Matérn 模型，用来捕捉过程的光滑程度。这类过程出现在贝叶斯线性回归中，此时系数 $\boldsymbol{c}$ 被视为随机变量，这是由认识不确定性导致的，例如 $\boldsymbol{c}\sim\mathcal{N}(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$，并且需要联合推断 $\boldsymbol{c}$和 Z_2。由此得到的混合 Z_1+Z_2 又是一个高斯随机场。

g) 具有修正尾部的 Matérn 模型。 Matérn 协方差函数随距离呈指数衰减。在存在长记忆的情况下，这可能会成为一个缺点。某些协方差函数类能够为空间数据中的长记忆与短记忆建立索引。广义 Cauchy[65] 和 Dagum [22] 是其中的典型例子；这些族既能对分形维数建立索引，也能对长记忆（亦称赫斯特效应）建立索引。然而，这些族无法像 Matérn 类那样对光滑性进行参数化。这一难题一直困扰着许多科学家。下面我们将介绍一些旨在修正 Matérn 模型尾部的方法，同时保持（a）正定性和（b）局部行为；后者又与均方可微性和 Sobolev 空间参数化相关。下述每项研究的动机各不相同，全文将对此加以说明。

(a) 具有周期性尾部的 Matérn 模型。有人$\widetilde{\ }$提出了对 Matérn 模型谱密度的修改方案，作者为 [91]。其主要思想是提出一个各向同性的谱密度类，$M_{\nu,\ \alpha,\ \sigma^2}$，该类$\widehat{\ }$通过以下恒等式与 Matérn 族 $M_{\nu,\ \alpha,\ \sigma^2}$ 相关联

$$\widetilde{M}_{\nu,\alpha,\sigma^2}(z)=\left(b^2+z^2\right)^{\xi}\widehat{\mathcal{M}}_{\nu,\alpha,\sigma^2}(z), \tag{37}$$

其中 $b\geq 0$ 和 $\xi<\nu$。 b 是一个额外的范围参数，而 ξ 与相应过程的光滑性有关。更确切地说，当且仅当$\nu-\xi>k$时，随机场是 k 次均方可微的。在极限情形 $\xi\to 0$下，可恢复传统 Matérn 模型。

谱密度中的这一尺度会导致谱的众数发生偏移，因此特别适用于获得具有强周期性的过程。与 (37) 相关联的协方差函数没有已知的显式表达式，因此处理该模型时应采用谱域中的统计方法。

(b) 混合模型。Matérn 模型的另一种推广利用了这样一个事实：它可以表示为高斯核相对于逆伽马型概率密度函数的尺度混合

$$\mathcal{M}_{\nu,\alpha}(x)=\int_0^{\infty}\exp(-ux^2)\pi_{IG}(u;\nu,1/(4\alpha^2))\mathrm{d}u \tag{38}$$

。[5] 利用上述恒等式构造混合模型，这些模型既保留了传统 Matérn 模型的局部性质，又能在较大距离下呈现更加灵活的行为。一种可能的混合构造称为 Matérn-Cauchy，记作 $\mathcal{MC}_{\nu_1,\nu_2,\ \alpha,\xi}$，其构造方式如下：

$$\mathcal{MC}_{\nu_1,\nu_2,\alpha,\xi}(x)=\int_0^{\xi}\exp(-ux^2)\pi_G(u;\nu_1/2,\alpha)\mathrm{d}u+\int_{\xi}^{\infty}\exp(-ux^2)\pi_{IG}(u;\nu_2,1/(4\alpha^2))\mathrm{d}u. \tag{39}$$

此处，π_G 是采用形状-率参数化的伽马概率密度函数，π_{IG} 是式（41）中的密度，$\alpha>0$ 是范围参数，$\nu_1>0$ 控制协方差的多项式衰减率，$\nu_2>0$ 表示均方可微性的阶数，而 $\xi\geq 0$ 是一个附加参数，用于平衡 Matérn 和柯西对总协方差函数的贡献。当 $\xi\to 0$时，混合模型趋于 Matérn 协方差。该模型的闭式表达式见[5]。请注意，该协方差函数在任意维度下均为正定，是模型（31）的一个自然竞争者。

文献 [5] 通过将式（38）中的高斯核替换为高斯核之差，构造了另一种混合模型

$$\left(a\exp(-ubx^2)-\exp(-ux^2)\right)(a-1)^{-1},$$

其中 a 和 b 满足条件 $1<b<a^{2/d}$，以便在 $\mathbb{R}^d$[130]中获得正定核。所得模型具有 Matérn 类型的局部行为，并允许在较远距离处出现负相关（空穴效应）。参数 a 和 b 控制空穴效应的尖锐度。当 a任意大时，即可恢复传统 Matérn 模型。[5]中给出了代数闭式表达式。

2. 向量值随机场。
令 $\{\boldsymbol{Z}(\boldsymbol{x})\ \ \boldsymbol{x}\in\mathbb{R}^d\}\subset\mathbb{R}^p$ 为一个 p元随机场，其各向同性协方差映射为 $\boldsymbol{K}\colon\mathbb{R}^d\to\mathbb{R}^{p\times p}$，其元素 K_{ij} 定义如下

$$K_{ij}(\boldsymbol{x})=\operatorname{cov}\left(Z_i(\mathbf{0}),Z_j(\boldsymbol{x})\right),\ \mathbf{0},\boldsymbol{x}\in\mathbb{R}^d.$$

具体而言，K_{ii} 和 K_{ij} 分别称为自协方差函数和交叉协方差函数。

a) 多元空间。关于多元空间建模，已有大量相关方法，读者可参阅 [60]。[64] 提出了一种如下形式的多元协方差结构

$$K_{ij}(\boldsymbol{x})=\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{M}_{\nu_{ij},\alpha_{ij}}(\|\boldsymbol{x}\|),\ \boldsymbol{x}\in\mathbb{R}^d,$$

其中，σ_{ii}^2 是 $\boldsymbol{Z}$中 Z_i 的方差，而 ρ_{ij} 是同位相关系数。为保持正定性，参数 ν_{ij},α_{ij} 和 ρ_{ij} 受到限制，而对同位相关系数 ρ_{ij} 的限制往往十分严格。这促使研究者探索其他方法以放宽参数限制，读者可参阅 [11]，以及较新的 [54]。

b) 多元时空。通过考虑 $\{\boldsymbol{Z}(\boldsymbol{x},t)\ \ \boldsymbol{x}\in\mathbb{R}^d,t\in\mathbb{R}\}\subset\mathbb{R}^p$，可将上述情形推广到时空领域；这是一个 p元随机场，其各向同性协方差映射为 $\boldsymbol{K}\colon\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}^{p\times p}$，其元素 K_{ij} 定义如下

$$K_{ij}(\boldsymbol{x},u)=\operatorname{cov}\left(Z_i(\mathbf{0},t),Z_j(\boldsymbol{x},t+u)\right)$$

对于$\mathbf{0}$，$\boldsymbol{x}\in\mathbb{R}^d$、$t$，$u\in\mathbb{R}$。[36] 考虑以下类型的映射 K_{ij}

$$K_{ij}(\boldsymbol{x},u)=\frac{\sigma_{ii}\sigma_{jj}\rho_{ij}}{\psi(u^2)^{d/2}}\mathcal{M}_{\nu_{ij},\alpha_{ij}\psi(u^2)}(\|\boldsymbol{x}\|)$$

对于 $\boldsymbol{x}\in\mathbb{R}^d$、$u\in\mathbb{R}$ 以及适当的正值连续函数 ψ。[6] 近期对此设定进行了推广；[125]则通过技术方法加以推广：两项工作都采用了这样的思路，即将映射 ψ（逐点）替换为映射 $\boldsymbol{\psi}$，其元素连续且严格为正 ψ_{ij}。

c) 多元非平稳。[88] 推导出一类矩阵值协方差函数，其中直接协方差和交叉协方差函数均属于 Matérn 类。Matérn 类的参数允许随位置变化，从而产生局部方差、局部范围、局部几何各向异性和局部光滑度。定义$\Sigma_{i,\ \boldsymbol{x}}=\Sigma_i(\boldsymbol{x})$：$\mathbb{R}^d\to\mathbb{R}^{d\times d}$ 并假设 $\Sigma_i(\boldsymbol{x})$ 对所有 i 和所有 $\boldsymbol{x}\in\mathbb{R}^d$均为正定。令

$$Q_{ij;\boldsymbol{x},\boldsymbol{y}}=(\boldsymbol{x}-\boldsymbol{y})^\top\left(\frac{\Sigma_{i,\boldsymbol{x}}+\Sigma_{j,\boldsymbol{y}}}{2}\right)^{-1}(\boldsymbol{x}-\boldsymbol{y})$$

and

$$K_{ij}(\boldsymbol{x},\boldsymbol{y})=\rho_{ij}\sigma_{i,\boldsymbol{x}}\sigma_{j,\boldsymbol{y}}\mathcal{M}_{\nu,\alpha}\left(\sqrt{Q_{ij;\boldsymbol{x},\boldsymbol{y}}}\right).$$

那么，$\boldsymbol{K}(\cdot,\ \cdot)=[K_{ij}(\cdot,\ \cdot)]_{i,\ j=1}^p$ 是正定的。

d) 带 Dimple 的多元 Matérn。在二元空间背景下，矩阵值 Matérn 协方差函数的每个元素都具有如（38）所示的尺度混合表示。[4] 考虑对这种混合进行修改，以得到 Matérn 模型的一种推广，其形式为

$$\widetilde{K}_{ij}(x)=\sigma_{ii}\sigma_{jj}\rho_{ij}\int_0^\infty\exp(-ux^2)g_{ij}(u;\xi)\times\pi_{IG}(u;\nu_{ij},1/(4\alpha_{ij}^2))\mathrm{d}u, \tag{40}$$

其中 $g_{ii}(u;\ \xi)=1$，且 $g_{ij}(u;\ \xi)=1\{u\le\xi\}-1\{u\ge\xi\}$ 适用于 $i\neq j$，其中 $1\{\cdot\}$为指示函数，ξ 为非负参数；此外，π_{IG} 是逆伽马随机变量的概率密度函数，即

$$\pi_{IG}(z;a,b)=\frac{b^a}{\Gamma(a)}z^{-a-1}\exp(-b/z),\ z>0. \tag{41}$$

请注意，矩阵值协方差的对角元素并未改变，因此 Matérn 模型颇具吸引力的局部属性得以保留。这一构造只会影响交叉协方差。~事实上，[4] 证明了 $K_{12}(x)$ 并非关于 x单调递减的函数。更确切地说，交叉协方差可以在严格大于零的距离处达到最大值。[4]将这一性质称为交叉凹陷。参数 ξ 控制交叉凹陷的强度。显然，传统二元Matérn模型是这一构造的一个极限情形（$\xi\to\infty$）。[4]给出了（40）的闭式表达式。~此外，对于 $\nu_{12}=1/2+n$，$n\in\mathbb{N}$，$K_{12}(x)$ 可以用误差函数和指数函数表示。

3. 方向、形状与曲线。Matérn 模型是研究方向过程的核心。 [18] 形式化了方向有限差分过程和方向导数过程的概念，并特别关注 Matérn 协方差函数。在数据或空间随机效应服从平稳高斯过程（具有 Matérn 协方差）模型的假设下，他们给出了完整的分布理论结果。[16] 引入贝叶斯边界分析，以衡量与边界分析边界相关的曲线梯度。Matérn 模型的光滑性已证明在此类框架中十分有效。提出了使用 Matérn 模型对时间梯度进行建模的方法。[131]。