arXiv:2601.11437v1 [stat.CO] 16 Jan 2026

通过稳定光滑度优化实现精确 Matérn 协方差估计的 Fisher 评分法

Yiping Hong*北京理工大学数学与统计学院、北京理工大学唐山研究院；Sameh Abdulah、Marc G. Genton 和 Ying Sun，阿卜杜拉国王科技大学计算机、电气与数学科学和工程部

2026年1月19日

摘要

具有 Matérn 协方差函数的高斯随机场（GRF）已成为建模空间过程的强大框架，因为它们能够灵活地捕捉空间场的不同特征。然而，由于数值不稳定性，使用最大似然估计（MLE）估计光滑度参数 ν 颇具挑战；MLE 需要基于 GRF 的完整协方差矩阵计算似然。此外，对于大型空间数据集，MLE 的计算成本仍高得难以承受。为应对这一挑战，我们提出 Fisher 回溯

*作者衷心感谢阿卜杜拉国王科技大学、中国国家自然科学基金（编号：12271286）、河北省自然科学基金（编号：A2025105010）以及北京理工大学青年学者学术启动计划（XSQD-6120220282）。

（Fisher-BT）方法。该方法将 Fisher 评分算法与回溯线搜索策略相结合，并对修正贝塞尔函数采用级数近似。借助 ExaGeoStat高性能计算框架，该方法能够高效地对空间数据集进行最大似然估计。与无导数优化方法相比，我们提出的方法不仅减少了迭代次数、加快了收敛速度，还提高了 ν 估计的数值稳定性。通过模拟以及使用覆盖密西西比河流域的土壤水分数据集开展的真实数据分析，我们表明，所提出的 Fisher-BT 方法能够达到与现有方法相当的精度，同时在计算效率和数值稳定性方面显著优于 BOBYQA 和 Nelder–Mead 等无导数算法。

关键词：回溯线搜索；ExaGeoStat；高斯随机场；高性能计算；最大似然估计；空间统计

1 引言

高斯过程（GP）模型广泛应用于空间统计和机器学习，涉及环境科学和地球科学等多个领域（Abdulah 等人 2024）、能源科学（Koncewicz 等人 2023）、生物学（Barac 等人 2019）和林业（Finley 等人 2017）。在各种协方差结构中，Matérn 协方差函数因能够灵活地捕捉空间数据的不同光滑度水平，而在地统计学中尤为突出（Stein 2012）。然而，实践中通常将 Matérn 协方差模型中的光滑度参数 ν 视为固定值，这主要是因为基于似然的估计存在计算上的挑战。例如，Diggle & Ribeiro (2007年)指出，由于对数似然曲面存在脊线或平台形状，Matérn 协方差模型中三个参数的估计非常困难，因此建议从离散候选集中选择 ν。文献中最常见的光滑度参数取值为 $\nu = 0.5$，此时模型简化为指数协方差模型。

然而，正如 De Oliveira & Han (2022年) 所展示的，地统计数据可能包含大量关于光滑度参数的信息 ν，该参数反映了底层随机场的均方可微性。 Hong等人(2021) 表明，不准确的ν 可能导致空间预测效率低下，并且预测误差估计不准确。因此，获得 ν 的数值稳定且精确的估计，对于推断和预测都至关重要。

为了计算精确的最大似然估计量（MLE），即使用 Matérn 模型的完整协方差矩阵评估似然，已有多个成熟的软件包可供使用，包括 Fortran 程序 `MLMATERN` （Pardo-Igúzquiza 等人 2009年）以及 R软件包，例如 `fields` （Nychka 等人 2021年）、`geostatsp` （Brown 2015年）、 `geoR`（Ribeiro Jr 等人 2003年）、 `RandomFields` （Schlather 等人 2015年）以及 `ExaGeoStatR` （Abdulah 等人 2023年）。这些软件包通常依赖无导数方法，例如模式搜索

（Hooke& Jeeves 1961）、Nelder–Mead（Nelder & Mead 1965）和 BOBYQA（Powell 2009），这些方法忽略了 Matérn协方差模型的对数似然函数相对于所有参数都是无穷可微的这一事实。尽管 `fields` 软件包使用了基于导数的方法，例如 BFGS（Broyden 1970），但该软件包并不涉及对光滑度参数 ν的极大似然估计。 此外，BFGS 方法没有充分利用似然函数的统计结构，例如 Fisher信息。

Fisher评分法是一种基于导数的优化方法，通过利用期望 Fisher信息更新参数，已成功应用于近似最大似然估计场景（Geoga 等人 2020年、Guinness 2021）。然而，据我们所知，该方法尚未用于 Matérn协方差模型的精确最大似然估计。一个主要障碍是计算 $\partial_\nu \mathcal{K}_\nu$时存在数值不稳定性，其中 $\mathcal{K}_\nu$ 是阶数为 ν的第二类修正贝塞尔函数。这些导数的有限差分近似往往会导致显著的舍入误差和不可靠的海森矩阵估计（Geoga 等人 2023年）。尽管 Geoga 等人(2023年) 针对这些导数提出了稳定级数近似，但将其用于精确最大似然估计计算时，仅限于 $n = 512$等小型数据集，其中 n 为空间观测的数量。

对于较大的空间数据集，精确最大似然估计计算受限于求逆一个 $n \times n$ 协方差矩阵所需的计算复杂度，这需要 $O(n^3)$ 次运算和$O(n^2)$ 内存。为了解决这一问题， `ExaGeoStat` 框架（Abdulah 等人 2018年、2023年）利用了先进的高性能并行稠密线性代数库，例如`Chameleon` （Agullo 等人 2012年），以加速对数似然评估中的矩阵运算。这使得利用并行与分布式计算，能够对多达数十万至数百万个观测值的合成数据集和真实数据集进行精确最大似然估计计算。然而，该框架的优化过程是无导数的，因此未能利用似然的 Fisher

信息结构。在本研究中，我们将这一精确最大似然估计计算框架整合进来，使我们提出的基于导数的方法能够在大规模系统上运行。

在本研究中，我们提出 Fisher 回溯（Fisher-BT）方法，这是一种针对 Matérn协方差模型的稳健、高效的精确最大似然估计计算方法，结合了 Fisher评分法、回溯线搜索和 Nelder-Mead回退方法。借助高性能的 `ExaGeoStat` 框架，我们提出的方法能够对大型数据集进行精确推断。我们的方法利用级数近似稳定地计算导数，从而使 Fisher评分法收敛；与无导数优化方法相比，还显著减少了迭代次数。通过全面的模拟以及对密西西比河流域土壤水分的真实数据分析，我们证明 Fisher-BT 在精度方面表现出竞争力，同时在计算速度和数值稳定性方面显著优于现有方法，如 `ExaGeoStat`/BOBYQA 和 `ExaGeoStat`/Nelder-Mead，尤其是在 ν取极端值时。

本文其余部分安排如下。第2 节介绍极大似然估计优化框架，并说明所提出的 Fisher-BT 算法。第3 节报告数值模拟结果，第4 节则展示该方法在真实世界土壤湿度数据集上的应用。最后，第5 节总结全文并展开讨论。

2 方法论

本节介绍在 Matérn 协方差模型下计算精确最大似然估计所采用的方法框架。我们首先介绍所提出的 Fisher-BT优化策略，该策略结合了 Fisher评分法、回溯线搜索和 Nelder–Mead 备用方法，以确保效率与稳定性兼具。随后，我们详细说明

对数似然、其梯度和费舍尔信息矩阵的计算方法，包括对光滑度参数 ν 的导数进行数值处理。

2.1 Fisher-BT 优化算法

我们考虑平稳 Matérn 协方差模型的精确最大似然计算。设 $\{Z(\boldsymbol{s}), \boldsymbol{s} \in \mathbb{R}^2\}$ 为零均值平稳高斯随机场，则两个位置 $\boldsymbol{s}_1$、$\boldsymbol{s}_2$ 之间的协方差仅与它们的距离 $\|\boldsymbol{s}_1 - \boldsymbol{s}_2\|$ 有关。我们假设：

$$\mathrm{Cov}(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)) = C(\|\boldsymbol{s}_1 - \boldsymbol{s}_2\|; \boldsymbol{\theta}),$$

其中，

$$C(h; \boldsymbol{\theta}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} \left(\frac{h}{\alpha}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{h}{\alpha}\right)$$

是参数为 $\boldsymbol{\theta} = (\sigma^2, \alpha, \nu)^\top$ 的 Matérn 协方差函数，$\mathcal{K}_{\nu}$ 是阶数为 ν 的第二类修正贝塞尔函数。考虑在 n 个位置 $\boldsymbol{s}_i$、$i = 1$、…、n 上观测到的空间数据 $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n))^\top$，则 $\boldsymbol{Z}$ 的联合分布为多元高斯分布 $N_n(0, \boldsymbol{\Sigma}(\boldsymbol{\theta}))$，其中 $[\boldsymbol{\Sigma}(\boldsymbol{\theta})]_{i,\,j} = C(\|\boldsymbol{s}_i - \boldsymbol{s}_j\|; \boldsymbol{\theta})$、$i$、$j = 1$、…、$n$ 为协方差矩阵。在这种情况下，$\boldsymbol{Z}$ 的对数似然函数为

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log\{\det(\boldsymbol{\Sigma}(\boldsymbol{\theta}))\} - \frac{1}{2}\boldsymbol{Z}^\top \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{Z}, \tag{1}$$

而 $\ell(\boldsymbol{\theta})$ 的最大值点定义为最大似然估计（MLE）。

ExaGeoStat 软件（Abdulah 等人 2018年）采用 BOBYQA 算法（Pow-ell 2009），这是一种无导数算法，要求优化空间为矩形。我们的目标是使用基于导数的方法（例如 Fisher 评分算法）加快精确最大似然估计计算，从而消除参数空间约束、减少迭代次数，进而加快计算速度。

当光滑度参数 ν 较大时，对数似然曲面在 ν 方向上趋于平坦，费舍尔信息矩阵的数值计算会变得不稳定，其元素在各次迭代中出现明显波动。在这种情况下，直接使用Fisher评分法可能会产生过于激进的更新，导致优化过程不稳定。为解决这一问题，我们提出 Fisher–BT算法，通过加入回溯线搜索并采用 Nelder–Mead 回退策略来增强稳健性。其核心思想是在导数信息可靠时加以利用，而当 ν 的值较大、妨碍稳定收敛时，则切换到 Nelder–Mead。所提出方法的整体流程总结于算法1中。

算法首先选择一个初始值 $\boldsymbol{\theta}^{(0)}$，因为较差的初始化可能会显著增加迭代次数，甚至导致 Fisher 评分法无法收敛。与无导数优化算法类似，我们首先假设真实参数 $\boldsymbol{\theta}$ 可能位于一个立方体内，其左下角向量为$\boldsymbol{\theta}^{(\ell)} = (\sigma^2_{(\ell)}, \alpha_{(\ell)}, \nu_{(\ell)})^\top$，右上角向量为 $\boldsymbol{\theta}^{(u)} = (\sigma^2_{(u)}, \alpha_{(u)}, \nu_{(u)})^\top$。然后，我们考虑表1 中给出的 L9 设计的九个初始值候选项，并选择对数似然值最高的参数作为初始值 $\boldsymbol{\theta}^{(0)}$。请注意，与 ExaGeoStat/BOBYQA 算法不同，立方体 $\Theta = [\sigma^2_{(\ell)}, \sigma^2_{(u)}] \times [\alpha_{(\ell)}, \alpha_{(u)}] \times [\nu_{(\ell)}, \nu_{(u)}]$仅用于选择初始值，因此最终的极大似然估计结果可能位于该立方体之外。当第 ν 3 节介绍的数值实验中真实参数较大时，就可能出现这种情况。

选定初始值后，算法使用回溯线搜索启动 Fisher 评分算法。为了得到增量向量 $\boldsymbol{\phi}^{(t)}$，我们不仅需要计算对数似然函数 $\ell(\boldsymbol{\theta})$，还需要计算其梯度和费舍尔信息矩阵。记作 $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^\top$，则偏导数的计算如下：

$$\ell_i(\boldsymbol{\theta}) := \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_i} = -\frac{1}{2}\mathrm{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_i) + \frac{1}{2}\boldsymbol{Z}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}, \tag{2}$$

算法1 我们提出的 Fisher-BT 算法

输入： 初始猜测向量 $\boldsymbol{\theta}^{(\ell)}$ 和 $\boldsymbol{\theta}^{(u)}$;

输出： 最大似然估计 $\hat{\boldsymbol{\theta}}_{MLE}$ 及其费舍尔信息矩阵 $I(\hat{\boldsymbol{\theta}}_{MLE})$.

使用 L9 设计确定初始参数 $\boldsymbol{\theta}^{(0)}$;

设置 $t = 0$;

执行一次费歇尔评分迭代：计算 $\ell(\boldsymbol{\theta}^{(0)})$, $\nabla\ell(\boldsymbol{\theta}^{(0)})$, $\boldsymbol{I}(\boldsymbol{\theta}^{(0)})$，然后计算 $\boldsymbol{\phi}^{(0)} \leftarrow \boldsymbol{I}(\boldsymbol{\theta}^{(0)})^{-1}\nabla\ell(\boldsymbol{\theta}^{(0)})$;

当 停止条件和切换条件均未满足**时执行**

检查 $\boldsymbol{\phi}^{(t)}$ 是否满足 Armijo条件;

当 不满足 Armijo条件**时**

$\boldsymbol{\phi}^{(t)} \leftarrow \boldsymbol{\phi}^{(t)}/2$;

检查更新后的 $\boldsymbol{\phi}^{(t)}$ 是否满足 Armijo条件;

结束循环

$\boldsymbol{\theta}^{(t+1)} \leftarrow \boldsymbol{\theta}^{(t)} + \boldsymbol{\phi}^{(t)}$;

$t \leftarrow t + 1$;

执行一次Fisher评分迭代：计算 $\ell(\boldsymbol{\theta}^{(t)})$、 $\nabla\ell(\boldsymbol{\theta}^{(t)})$、 $\boldsymbol{I}(\boldsymbol{\theta}^{(t)})$，然后计算 $\boldsymbol{\phi}^{(t)} \leftarrow \boldsymbol{I}(\boldsymbol{\theta}^{(t)})^{-1}\nabla\ell(\boldsymbol{\theta}^{(t)})$;

结束循环

如果未满足停止条件**则**

以 $\boldsymbol{\theta}^{(t)}$ 为初始值，并采用相同的停止条件执行Nelder-Mead优化;

将 $\boldsymbol{\theta}^{(t)}$ 更新为上一步的极大似然估计结果;

计算 $\boldsymbol{I}(\boldsymbol{\theta}^{(t)})$ ，作为费舍尔信息矩阵。

结束 if

$\hat{\boldsymbol{\theta}}_{MLE} \leftarrow \boldsymbol{\theta}^{(t)}$, $\boldsymbol{I}(\hat{\boldsymbol{\theta}})_{MLE} \leftarrow \boldsymbol{I}(\boldsymbol{\theta}^{(t)})$.

表1：Fisher 评分算法的初始值候选项

No.	θ_1	θ_2	θ_3
1	$\frac{1}{2}\sigma^2_{(\ell)}+\frac{1}{2}\sigma^2_{(u)}$	$\frac{1}{2}\alpha_{(\ell)}+\frac{1}{2}\alpha_{(u)}$	$\frac{1}{2}\nu_{(\ell)}+\frac{1}{2}\nu_{(u)}$
2	$\frac{1}{2}\sigma^2_{(\ell)}+\frac{1}{2}\sigma^2_{(u)}$	$\frac{5}{6}\alpha_{(\ell)}+\frac{1}{6}\alpha_{(u)}$	$\frac{5}{6}\nu_{(\ell)}+\frac{1}{6}\nu_{(u)}$
3	$\frac{1}{2}\sigma^2_{(\ell)}+\frac{1}{2}\sigma^2_{(u)}$	$\frac{1}{6}\alpha_{(\ell)}+\frac{5}{6}\alpha_{(u)}$	$\frac{1}{6}\nu_{(\ell)}+\frac{5}{6}\nu_{(u)}$
4	$\frac{5}{6}\sigma^2_{(\ell)}+\frac{1}{6}\sigma^2_{(u)}$	$\frac{1}{2}\alpha_{(\ell)}+\frac{1}{2}\alpha_{(u)}$	$\frac{5}{6}\nu_{(\ell)}+\frac{1}{6}\nu_{(u)}$
5	$\frac{5}{6}\sigma^2_{(\ell)}+\frac{1}{6}\sigma^2_{(u)}$	$\frac{5}{6}\alpha_{(\ell)}+\frac{1}{6}\alpha_{(u)}$	$\frac{1}{6}\nu_{(\ell)}+\frac{5}{6}\nu_{(u)}$
6	$\frac{5}{6}\sigma^2_{(\ell)}+\frac{1}{6}\sigma^2_{(u)}$	$\frac{1}{6}\alpha_{(\ell)}+\frac{5}{6}\alpha_{(u)}$	$\frac{1}{2}\nu_{(\ell)}+\frac{1}{2}\nu_{(u)}$
7	$\frac{1}{6}\sigma^2_{(\ell)}+\frac{5}{6}\sigma^2_{(u)}$	$\frac{1}{2}\alpha_{(\ell)}+\frac{1}{2}\alpha_{(u)}$	$\frac{1}{6}\nu_{(\ell)}+\frac{5}{6}\nu_{(u)}$
8	$\frac{1}{6}\sigma^2_{(\ell)}+\frac{5}{6}\sigma^2_{(u)}$	$\frac{5}{6}\alpha_{(\ell)}+\frac{1}{6}\alpha_{(u)}$	$\frac{1}{2}\nu_{(\ell)}+\frac{1}{2}\nu_{(u)}$
9	$\frac{1}{6}\sigma^2_{(\ell)}+\frac{5}{6}\sigma^2_{(u)}$	$\frac{1}{6}\alpha_{(\ell)}+\frac{5}{6}\alpha_{(u)}$	$\frac{5}{6}\nu_{(\ell)}+\frac{1}{6}\nu_{(u)}$

而 $\boldsymbol{I}(\boldsymbol{\theta})$ 的元素由下式计算

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{i,j}=\mathrm{E}_{\boldsymbol{\theta}}\left(-\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right)=\frac{1}{2}\mathrm{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_j), \tag{3}$$

其中 $[\boldsymbol{\Sigma}_i]_{j,\ k}=\frac{\partial}{\partial\theta_i}C(\|\boldsymbol{s}_j-\boldsymbol{s}_k\|;\boldsymbol{\theta})$ 是协方差矩阵相对于 θ_i 的逐元素偏导数。计算 $\ell(\boldsymbol{\theta})$, $\nabla\ell(\boldsymbol{\theta})$ 和 $\boldsymbol{I}(\boldsymbol{\theta})$ 的详细算法将在第 2.2 节介绍，其中涉及相对于ν 的导数项将使用与 Geoga 等人 (2023年) 类似的级数近似算法进行计算。

Armijo条件确保步长 s 满足

$$\ell(\boldsymbol{\theta}^{(t)}+s\boldsymbol{\phi}^{(t)})\geq\ell(\boldsymbol{\theta}^{(t)})+cs\nabla\ell(\boldsymbol{\theta}^{(t)})\cdot\boldsymbol{\phi}^{(t)}, \tag{4}$$

其中， c 是一个较小的正数。该条件要求优化过程在每次迭代中都取得足够的进展。回溯线搜索方法会检查$s=1$ 是否满足该条件，并将 s 缩小至 ρs ，重复一定的 $0<\rho<1$ ，直到满足该条件。在我们的过程中，取 $c=0.001$、 $\rho=0.5$。我们还放宽了 Armijo条件，

因此，当（4）式中的左侧减去右侧大于 -0.001 时，就认为满足该条件。这样可以避免在 ν 较大时步长变得过小。因此，算法 1 中的 Armijo条件为

$$\ell(\boldsymbol{\theta}^{(t)} + \boldsymbol{\phi}^{(t)}) \geq \ell(\boldsymbol{\theta}^{(t)}) + 0.001\nabla\ell(\boldsymbol{\theta}^{(t)}) \cdot \boldsymbol{\phi}^{(t)} - 0.001,$$

并将 $\boldsymbol{\phi}^{(t)}$ 更新为 $\boldsymbol{\phi}^{(t)}/2$，直到满足此条件。满足此条件后，Fisher评分法的下一个输入参数为 $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} + \boldsymbol{\phi}^{(t)}$，直至满足停止条件或切换条件。

我们将停止条件设为 $\|\nabla\ell(\boldsymbol{\theta}^{(t)})\|_2 \leq 0.001$。当此条件成立时，Fisher评分法收敛。然而，在某些情况下，例如 ν 过大或过小时，Fisher评分法可能无法收敛。因此，我们根据对数似然函数及其导数的调用次数设置切换条件。当程序尝试计算 $\ell(\boldsymbol{\theta})$ 超过 60 次（包括寻找初始值），或尝试计算 $\nabla\ell(\boldsymbol{\theta})$ 超过 20 次时，即满足切换条件，这意味着此时最好更换优化方法。因此，当停止条件尚未满足而切换条件已满足时，我们将切换到Nelder-Mead优化，并使用最后一个参数值 $\boldsymbol{\theta}^{(t)}$ 作为初始值。在 Nelder-Mead 方法中，若某个优化步骤使 $\ell(\boldsymbol{\theta})$ 的值变化小于 10^{-9}，我们便认为满足停止条件。
Nelder-Mead 方法计算最终估计值 $\boldsymbol{\theta}^{(t)}$，并在进行一次费舍尔信息矩阵计算后输出费舍尔信息矩阵 $\boldsymbol{I}(\boldsymbol{\theta}^{(t)})$ 。

2.2 似然计算算法

本节将介绍计算对数似然函数及其导数的详细算法。计算 $\ell(\boldsymbol{\theta})$ 的算法见算法 2，其主要涉及 Cholesky 分解。根据（2）和（3），为了计算 $\nabla\ell(\boldsymbol{\theta})$ 和 $\boldsymbol{I}(\boldsymbol{\theta})$，我们需要处理 $\boldsymbol{\Sigma}_i$ 和 $\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_i$，这两部分是计算中最耗时的部分。

对于导数矩阵 $\boldsymbol{\Sigma}_i$，注意到 $\boldsymbol{\Sigma}_{\sigma^2} = \frac{1}{\sigma^2}\boldsymbol{\Sigma}$。$\boldsymbol{\Sigma}_{\alpha}$可直接通过以下方式计算

$$\frac{\partial}{\partial\alpha}C(h;\boldsymbol{\theta}) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \cdot \frac{\sigma^2}{\alpha} \cdot \left(\frac{h}{\alpha}\right)^{\nu+1} K_{\nu-1}\left(\frac{h}{\alpha}\right).$$

算法 2 计算对数似然

输入： 参数 $\boldsymbol{\theta}$；空间数据 $\boldsymbol{Z}$ 在位置 $\boldsymbol{s}_i$, $i = 1, ..., n$处观测

输出： 对数似然函数 $\ell(\boldsymbol{\theta})$。

计算协方差矩阵 $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$。对 $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}^\top$进行 Cholesky 分解，其中 $\boldsymbol{L}$ 是下三角矩阵。计算 $\log\{\det(\boldsymbol{\Sigma})\} = 2 \cdot \sum_{i=1}^{n} \log(\ell_{ii})$，其中 ℓ_{ii} 是 i阶的对角项，属于 $\boldsymbol{L}$。计算 $\boldsymbol{Y} = \boldsymbol{L}^{-1}\boldsymbol{Z}$。计算 $\boldsymbol{Z}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{Z} = \boldsymbol{Y}^\top\boldsymbol{Y}$。根据（1）和前面的结果计算 $\ell(\boldsymbol{\theta})$ 。

导数矩阵 $\boldsymbol{\Sigma}_\nu$ 由 $x^\nu K_\nu(x)$ 的导数计算得出，采用 Geoga 等人 (2023年)提出的级数近似算法。该导数的详细计算步骤见补充材料。为计算迹$(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_j)$，我们采用以下迹恒等式。设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为实 $n \times n$矩阵，则：

$$\operatorname{trace}(\boldsymbol{A}\boldsymbol{B}) = \frac{1}{2}\left(\|\boldsymbol{A} + \boldsymbol{B}^\top\|_F^2 - \|\boldsymbol{A}\|_F^2 - \|\boldsymbol{B}\|_F^2\right), \tag{5}$$

其中$\|\cdot\|_F^2$表示 Frobenius 范数，仅需进行 $O(n^2)$ 次计算。算法 $\ell(\boldsymbol{\theta})$、 $\nabla\ell(\boldsymbol{\theta})$和 $\boldsymbol{I}(\boldsymbol{\theta})$的连续计算过程见算法 3。

3 模拟

本节展示了所提出的 Fisher-BT 方法在计算效率和估计精度方面的表现，并与采用

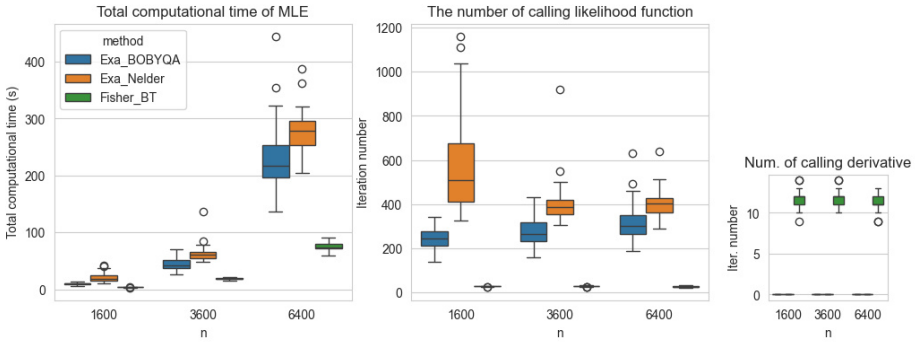
算法 3 计算对数似然、梯度和费舍尔信息矩阵**输入：** 参数 $\boldsymbol{\theta}$；在位置 $\boldsymbol{s}_i$、 $i = 1$、...、 n 上观测到的空间数据 $\boldsymbol{Z}$。

O输出：对数似然 $\ell(\boldsymbol{\theta})$、其梯度 $\nabla\ell(\boldsymbol{\theta})$以及 Fisher信息 $\boldsymbol{I}(\boldsymbol{\theta})$。步骤 1：计算 $\ell(\boldsymbol{\theta})$。计算协方差矩阵 $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$。进行 Cholesky 分解 $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}^\top$，其中 $\boldsymbol{L}$ 是下三角矩阵。计算 $\log\{\det(\boldsymbol{\Sigma})\} = 2 \cdot \sum_{i=1}^{n} \log(\ell_{ii})$，其中 ℓ_{ii} 是 $\boldsymbol{L}$的第 i 个对角项。计算 $\boldsymbol{Y} = \boldsymbol{L}^{-1}\boldsymbol{Z}$ 和 $\boldsymbol{Z}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{Z} = \boldsymbol{Y}^\top\boldsymbol{Y}$。利用公式 (1) 和前述结果计算 $l(\boldsymbol{\theta})$ 。步骤 2：计算 $\ell_1(\boldsymbol{\theta})$ 和$[\boldsymbol{I}(\boldsymbol{\theta})]_{1,1}$。计算 $\ell_1(\boldsymbol{\theta}) = \frac{1}{2\sigma^2}(\boldsymbol{Z}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{Z} - n)$ 和 $[\boldsymbol{I}(\boldsymbol{\theta})]_{1,1} = \frac{n}{2(\sigma^2)^2}$。步骤 3：计算 $\ell_2(\boldsymbol{\theta})$ 和 $[\boldsymbol{I}(\boldsymbol{\theta})]_{j,k}$，其中 $j, k \leq 2$。计算导数协方差矩阵 $\boldsymbol{\Sigma}_\alpha = \frac{\partial}{\partial\alpha}\boldsymbol{\Sigma}(\boldsymbol{\theta})$。计算 $\boldsymbol{A} \leftarrow \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\alpha$ 及其迹 $\text{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\alpha)$。计算$[\boldsymbol{I}(\boldsymbol{\theta})]_{1,2} = [\boldsymbol{I}(\boldsymbol{\theta})]_{2,1} = \frac{1}{2\sigma^2}\text{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\alpha)$。计算向量 $\boldsymbol{w} \leftarrow \boldsymbol{\Sigma}^{-1}\boldsymbol{Z}$ 和 $\boldsymbol{v} \leftarrow \boldsymbol{\Sigma}_\alpha\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}$，以及 $\boldsymbol{w}^\top\boldsymbol{v} = \boldsymbol{Z}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\alpha\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}$。根据公式 (2) 计算 $\ell_2(\boldsymbol{\theta})$ 、 $\boldsymbol{w}^\top\boldsymbol{v}$和 $\text{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\alpha)$。根据公式 (3) 和 (5) 计算$[\boldsymbol{I}(\boldsymbol{\theta})]_{2,2}$，其中 $\boldsymbol{A} = \boldsymbol{B} = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\alpha$。步骤 4：计算 $\ell_3(\boldsymbol{\theta})$ 和 $[\boldsymbol{I}(\boldsymbol{\theta})]_{j,k}$，其中 $j, k \leq 3$。计算导数协方差矩阵 $\boldsymbol{\Sigma}_\nu = \frac{\partial}{\partial\nu}\boldsymbol{\Sigma}(\boldsymbol{\theta})$。计算 $\boldsymbol{A} \leftarrow \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\nu$ 及其迹 $\text{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\nu)$。计算$[\boldsymbol{I}(\boldsymbol{\theta})]_{1,3} = [\boldsymbol{I}(\boldsymbol{\theta})]_{3,1} = \frac{1}{2\sigma^2}\text{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\nu)$。计算向量 $\boldsymbol{w} \leftarrow \boldsymbol{\Sigma}^{-1}\boldsymbol{Z}$ 和 $\boldsymbol{v} \leftarrow \boldsymbol{\Sigma}_\nu\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}$，以及 $\boldsymbol{w}^\top\boldsymbol{v} = \boldsymbol{Z}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\nu\boldsymbol{\Sigma}^{-1}\boldsymbol{Z}$。根据公式 (2) 计算 $\ell_3(\boldsymbol{\theta})$ 、 $\boldsymbol{w}^\top\boldsymbol{v}$和 $\text{trace}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\nu)$。根据公式 (3) 和 (5) 计算$[\boldsymbol{I}(\boldsymbol{\theta})]_{2,3} = [\boldsymbol{I}(\boldsymbol{\theta})]_{3,2}$，其中 $\boldsymbol{A} = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\alpha$, $\boldsymbol{B} = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\nu$。根据公式 (3) 和 (5) 计算$[\boldsymbol{I}(\boldsymbol{\theta})]_{3,3}$，其中 $\boldsymbol{A} = \boldsymbol{B} = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\nu$。

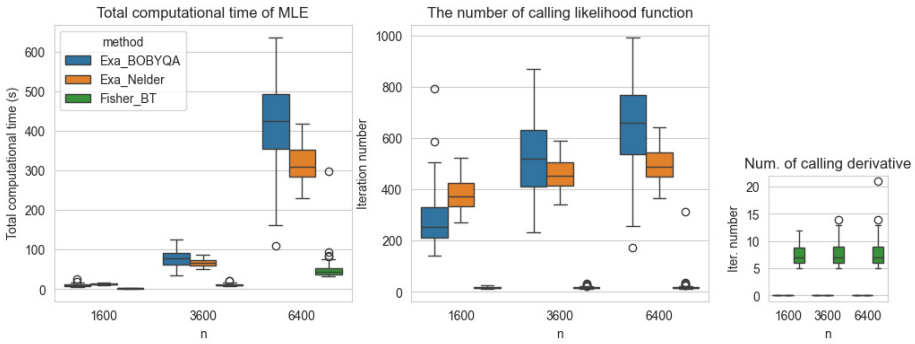
BOBYQA 和 Nelder–Mead 优化算法分别记为 ExaGeoStat/BOBYQA 和 ExaGeoStat/Nelder-Mead。数据由平稳高斯随机场生成，其协方差函数为 Matérn 协方差函数，真实参数值 (σ^2, α, ν) 设为 $(1, 0.1, 0.5)$，$(0.1, 0.1, 0.1)$，$(0.05, 0.05, 0.05)$，$(2, 0.8, 1)$ 和 $(1.5, 1.55, 1.3)$。这些设置包括一个 ν取中等值的情况、两个 ν相对较小的情况，以及两个 ν相对较大的情况。我们将数据规模设为$n = 1{,}600$、$3{,}600$ 和 $6{,}400$。对于每个样本量 n 和真实参数配置，我们从 Matérn 模型中生成 $M = 50$ 个随机实现，并使用前文所述的三种优化方法计算精确的最大似然估计（MLE）。在 ExaGeoStat/BOBYQA 方法中，优化范围为$(\sigma^2, \alpha, \nu) \in [0.01, 5] \times [0.01, 5] \times [0.01, 2]$；而在 ExaGeoStat/Nelder-Mead 方法中，优化初始值设为这些范围的中点，即 $(2.505, 2.505, 1.005)$。对于 Fisher-BT 方法，初始值按表 1 所示选取，其中 $\boldsymbol{\theta}^{(\ell)} = (0.01, 0.01, 0.01)$，$\boldsymbol{\theta}^{(u)} = (5, 5, 2)$ 是 ExaGeoStat/BOBYQA 优化范围的顶点，以确保初始值选取的公平性。

首先，我们在图中展示了三种方法收敛所需的计算时间和迭代次数1 和2。如图所示，对于两种优化算法，所提出的 Fisher-BT 算法相比 ExaGeoStat 具有显著的计算优势。唯一的例外是 $n = 3{,}600$, $\nu = 0.05$，在这种情况下，所提出的算法仍具有与 ExaGeoStat/BOBYQA 方法相当的计算性能。对于另外两种方法， ExaGeoStat/BOBYQA 方法在 ν较小或中等时更快，而 ExaGeoStat/Nelder-Mead 方法在 ν较大时更快，例如 $\nu = 1.0$ 或 1.3。请注意，BOBYQA 方法是 Abdulah 等人Abdulah 等人(2018年)所述 ExaGeoStat 优化框架中的默认优化方法。

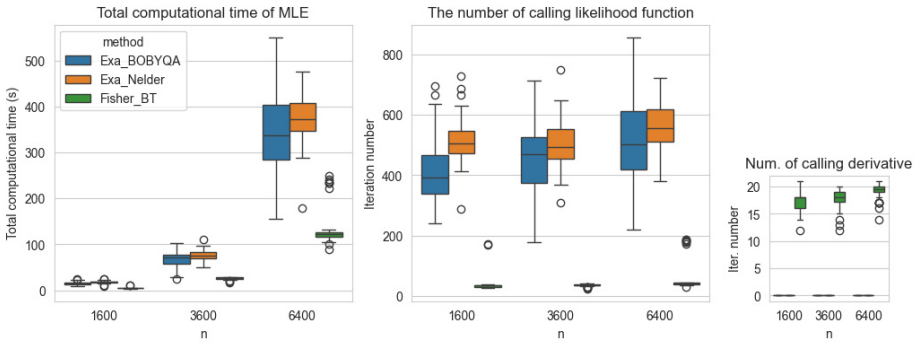
我们所提出方法在计算时间上的改进主要来自



(a) $(\sigma^2, \alpha, \nu) = (1.0, 0.1, 0.5)$

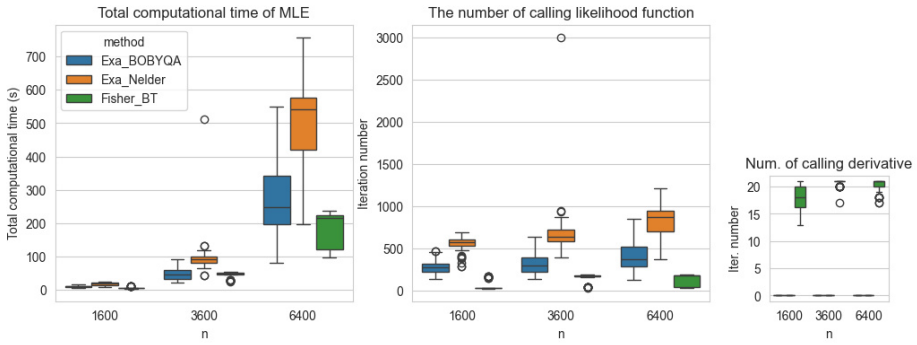


(b) $(\sigma^2, \alpha, \nu) = (2.0, 0.8, 1.0)$

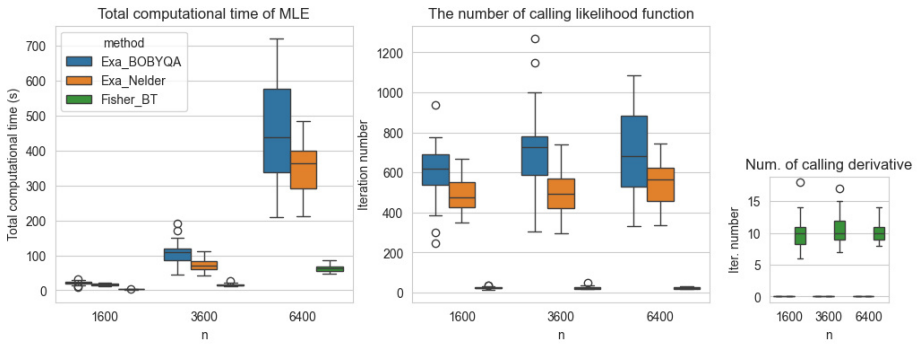


(c) $(\sigma^2, \alpha, \nu) = (0.1, 0.1, 0.1)$

图1：当 ν 的真实值适中时，不同最大似然估计算法的计算时间、迭代次数以及调用导数函数的次数。



(a) $(\sigma^2, \alpha, \nu) = (0.05, 0.05, 0.05)$

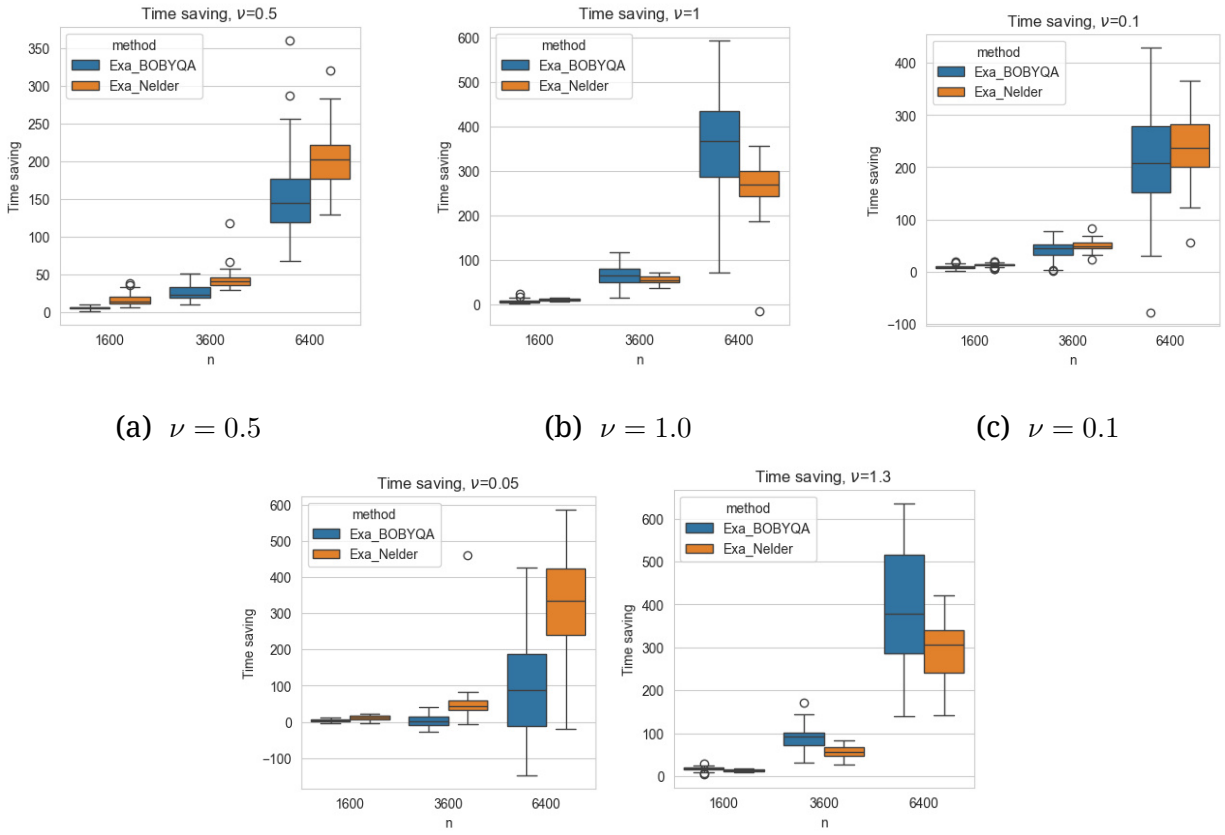


(b) $(\sigma^2, \alpha, \nu) = (1.5, 1.55, 1.3)$

图 2：当 ν 的真实值较小和较大时，不同最大似然估计算法的计算时间、迭代次数以及导数函数调用次数。

减少了似然评估次数。对于较大的 ν 值（例如 $\nu = 1.0$ 和 1.3），ExaGeoStat/Nelder–Mead 方法在大多数情况下所需的似然函数调用次数较少；而对于较小和中等的 ν 值（例如 $\nu = 0.05$、0.1 和 0.5），ExaGeoStat/BOBYQA 方法的调用次数较少。我们提出的算法以增加导数计算为代价，显著减少了似然函数调用次数。对于中等和较大的 ν值，在 ν较小的情况下，导数函数的调用次数较少。当 $\nu = 0.05$ 或 0.1 时，调用次数接近导数调用次数的上限，因此此时会用到 ExaGeoStat/Nelder-Mead 算法。即使在最坏情况下，我们提出的算法的计算时间仍与 ExaGeoStat中使用的无导数方法相当。我们还展示了.....之间的差异

我们提出的方法与两种方法之间的计算时间 ExaGeoStat方法的对比见图3-4，这表明所提出的方法总体上比这两种方法具有更高的计算效率。尽管在个别重复试验中（尤其是在 $\nu = 0.05$）计算时间并非最优，但我们提出的方法在大多数重复试验中都呈现出明显的节省计算时间趋势。



(a) $\nu = 0.5$ (b) $\nu = 1.0$ (c) $\nu = 0.1$

(d) $\nu = 0.05$ (e) $\nu = 1.3$

图3：所提出的方法与两种 ExaGeoStat方法之间计算时间节省的箱线图。

接下来，如图 5 和 6所示，我们考察通过三种方法计算得到的极大似然估计的估计性能。 ExaGeoStat/Nelder-Mead 方法可能会产生严重的离群估计值，尤其是在 ν 取较小值时。比如，当 $\sigma^2 = 0.5$时， σ^2 的最大估计值在 2.0×10^4 时达到 $n = 1$、600，以及 3.0×10^5 时达到 $n = 6$、400。

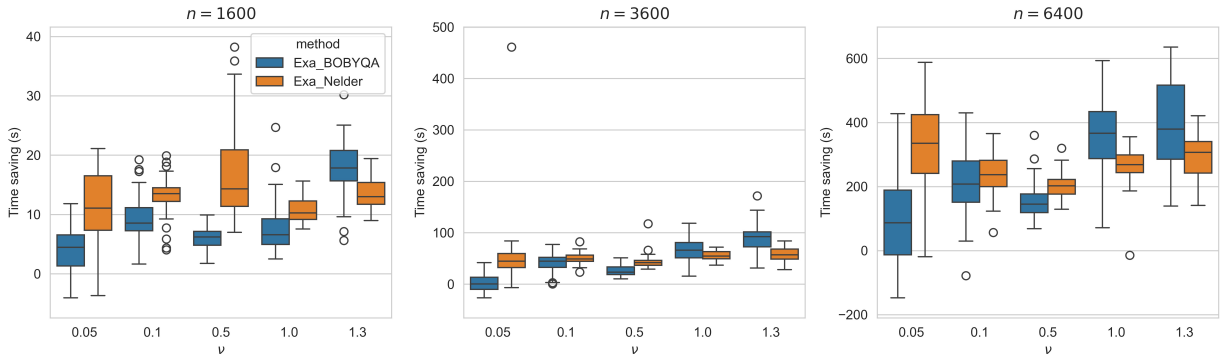
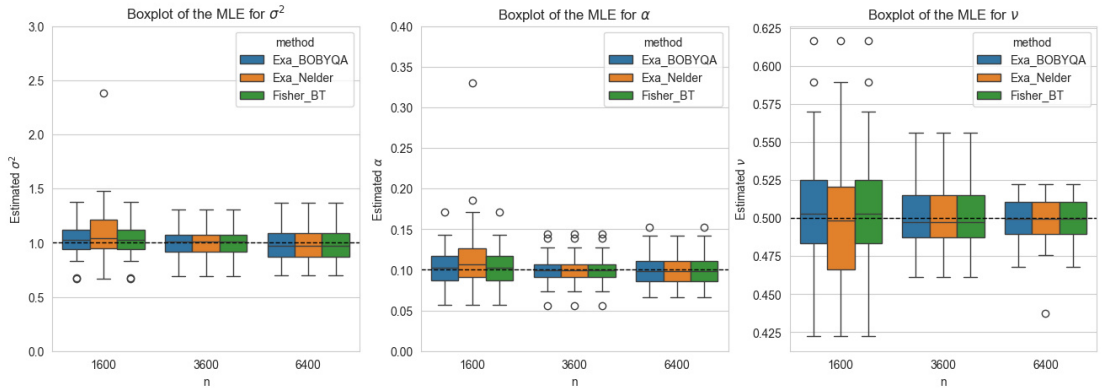


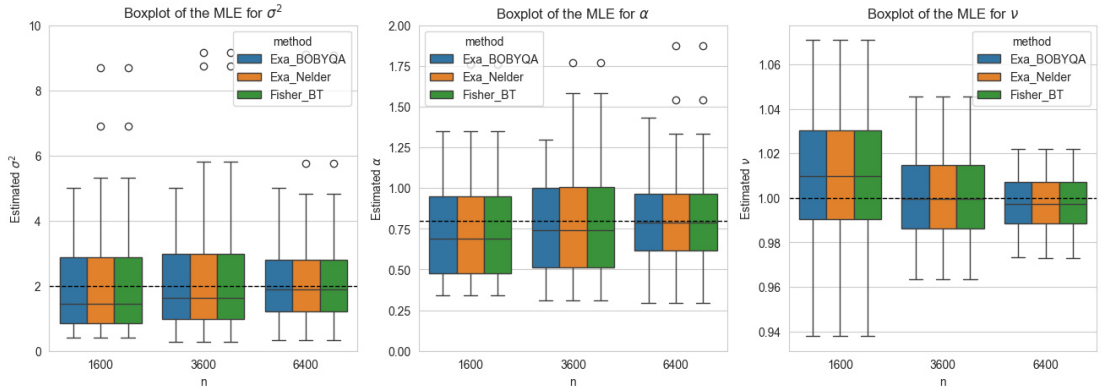
图4：相对于 n，所提方法与两种 ExaGeoStat方法之间计算时间节省的箱线图。

因此，我们忽略这些严重的离群值，以使箱线图清晰易读。在所有考察的情形下，Fisher-BT 算法都能得到方差最小且无偏性良好的极大似然估计。 ExaGeoStat/Nelder-Mead 算法在 $\nu = 1.0$ 时生成的箱线图与 Fisher-BT 算法的箱线图相似，并且在 $\nu = 1.3$时优于 ExaGeoStat/BOBYQA 算法。另一方面，当 $\nu = 0.05$、0.1、0.5 时，ExaGeoStat/BOBYQA 算法表现更好，因为与其 ExaGeoStat/Nelder-Mead 算法对应的箱线图相比，其箱线图的方差更小、离群值更少。因此，就估计性能而言， ExaGeoStat/BOBYQA 算法在较小和中等的 ν下表现更好，而 ExaGeoStat/Nelder-Mead 算法在较大的 ν下表现更佳。然而，我们的 Fisher-BT 算法能够针对所有考察的 ν给出精度良好的估计值，因此，我们所提出的算法还具有允许真实值 ν取更宽范围的优势。

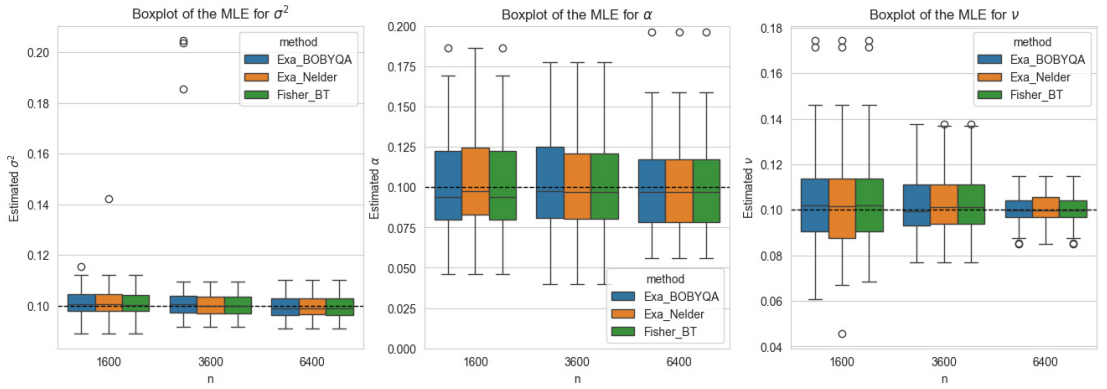
在图 5 和 6中， ν的估计方差随着 n 的增加而减小，但 σ^2 和 α 的该方差随着不同 n的变化并不显著。由于我们没有改变观测区域的大小，而是增加了观测位置的密度，这对应于固定区域渐近框架，因此根据Zhang (2004)，参数 σ^2 和 α 无法得到一致估计。然而， Zhang (2004)指出



(a) $(\sigma^2, \alpha, \nu) = (1.0, 0.1, 0.5)$



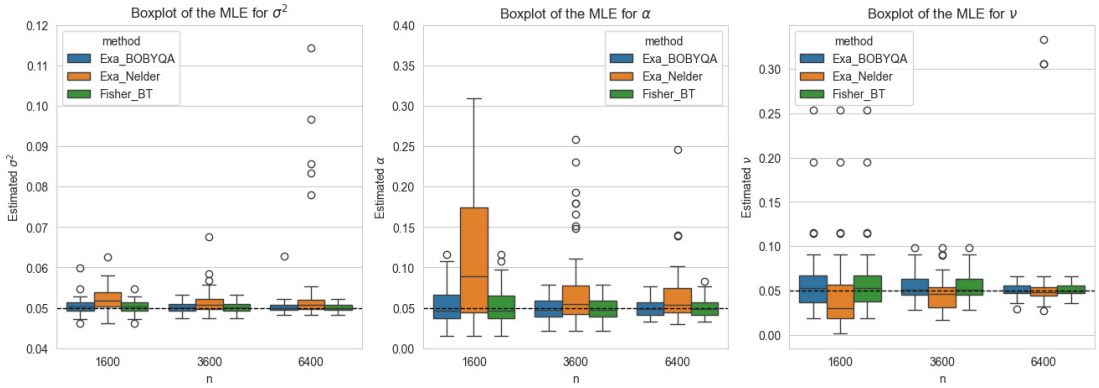
(b) $(\sigma^2, \alpha, \nu) = (2.0, 0.8, 1.0)$



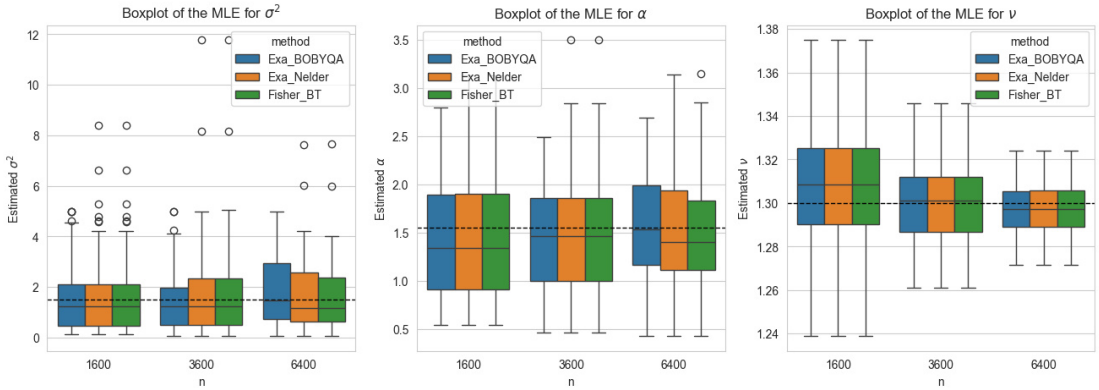
(c) $(\sigma^2, \alpha, \nu) = (0.1, 0.1, 0.1)$

图 5：真值处于中等水平时，不同最大似然估计算法所得估计值的箱线图

ν.



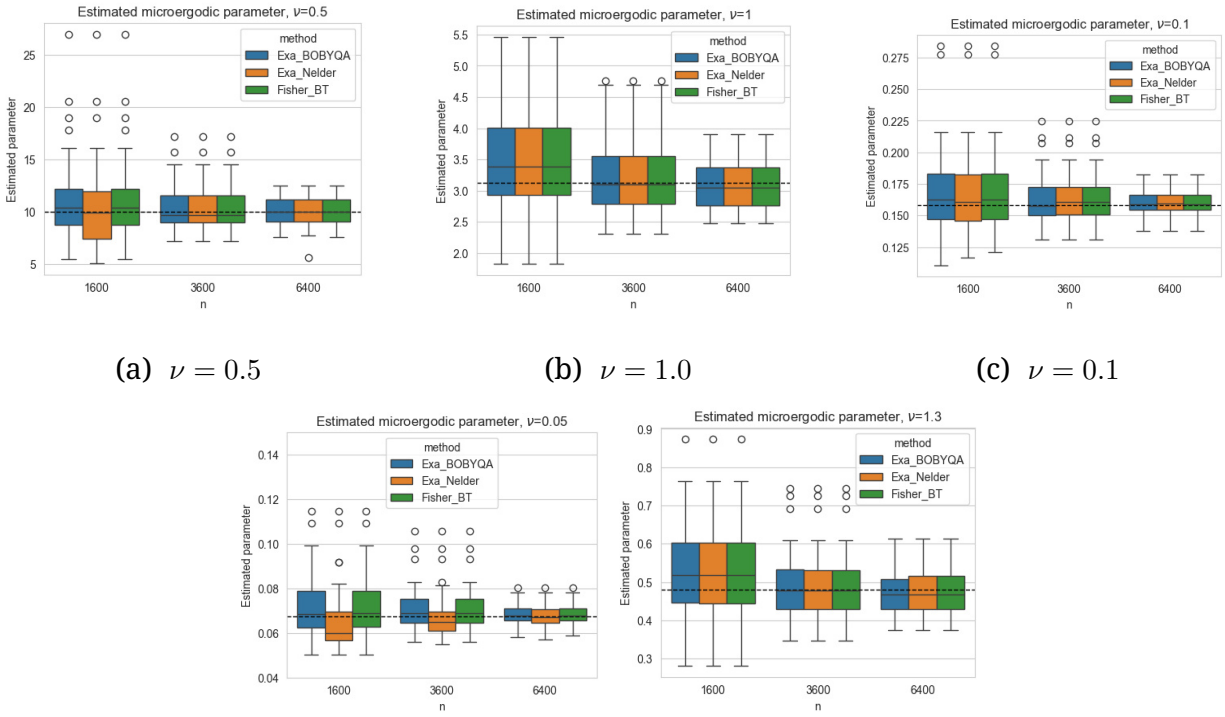
(a) $(\sigma^2, \alpha, \nu) = (0.05, 0.05, 0.05)$



(b) $(\sigma^2, \alpha, \nu) = (1.5, 1.55, 1.3)$

图6：当 ν 的真实值较小和较大时，不同最大似然估计算法所得估计值的箱线图。

微遍历参数 $\theta_m = \sigma^2\alpha^{-2\nu}$在固定域渐近下可以得到一致估计。因此，我们还展示了不同算法下 θ_m 的极大似然估计箱线图，如图 7所示。请注意，为了使图表更清晰，我们仅在 $\nu = 0.05$ 的情况下去除了六个离群值。图 7 显示，较大的 θ_m对应的极大似然估计方差较小，这与 Zhang (2004) 中固定 ν的渐近结果相似。 `ExaGeoStat`/Nelder-Mead 算法在 $\nu = 0.05$ 或 $\nu = 0.5$ 且 $n = 1$为 600时的表现不如另外两种算法；而在其他情况下，我们考察的方法对 θ_m 给出的估计值相近。



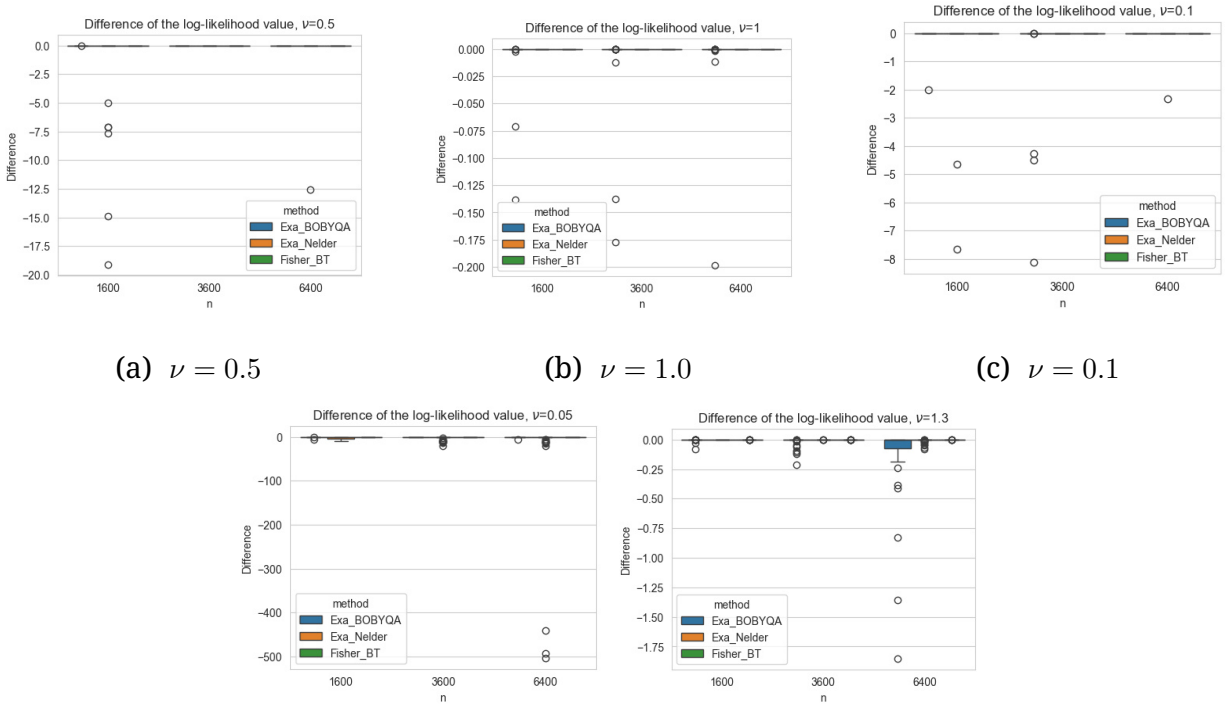
(a) $\nu = 0.5$ (b) $\nu = 1.0$ (c) $\nu = 0.1$

(d) $\nu = 0.05$ (e) $\nu = 1.3$

图7：不同最大似然估计算法所得微遍历参数估计的箱线图。

最后，为了评估优化算法的效率，我们比较了三种最大似然优化方法中，每种优化算法计算出的对数似然值与同一数据对应的最大对数似然值之间的差异，结果如图 8所示。我们忽略了该差异绝对值不大于 10^{-6}的情况。图 8 显示，当 ν 较小或适中时，例如 $\nu = 0.05$、0.1、0.5， ExaGeoStat/Nelder-Mead 方法更容易得到次优的对数似然值。然而，当 ν 较大时，例如 $\nu = 1.0$、1.3，ExaGeoStat/BOBYQA 方法在更多情况下可能得到次优结果。在大多数情况下，所提出的 Fisher-BT 方法在三种方法中取得了最优结果。即使在某些情况下 Fisher-BT 方法并非最优（这种情况出现在 $\nu = 1.3$时），其对数似然函数的损失也小于另外两种方法，因为最大的对数似然值损失为 5.4×10^{-5}。因此，我们提出的方法也优于

ExaGeoStat/BOBYQA 和 ExaGeoStat/Nelder-Mead 方法在此模拟中的表现。



(a) $\nu = 0.5$ (b) $\nu = 1.0$ (c) $\nu = 0.1$

(d) $\nu = 0.05$ (e) $\nu = 1.3$

图8：计算得到的对数似然值与其最大值之差的箱线图。

4 应用于土壤水分数据

我们考虑 Huang & Sun （2018年）提出的土壤水分数据集。该数据集由截至2004年1月1日美国密西西比盆地表层的高分辨率每日土壤水分数据组成。 Huang& Sun （2018年）使用经度和纬度作为协变量，通过线性模型对这些数据进行拟合，随后对残差进行带有一定平移的对数变换，从而获得高斯分布土壤水分残差数据集。我们将数据截断在 $[-84^\circ E, -80^\circ E] \times [34^\circ N, 42^\circ N]$的范围内，得到一个包含 $n_{\text{all}} = 542{,}629$ 个观测值的数据集。图 9展示了本文所考虑的土壤水分数据集。

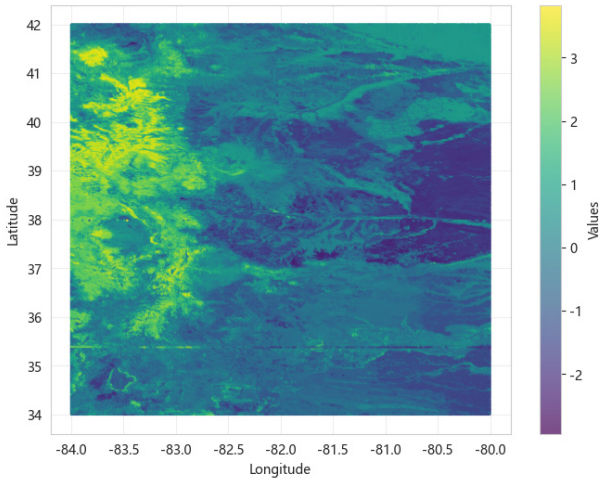


图9：我们真实案例研究中土壤湿度数据集的示意图。

为比较不同样本量下的计算精度和估计性能，我们考虑样本量分别为 $n = 3,600$、$14{,}400$、$32{,}400$和 $57,600$的子集。对于每个 n，我们均匀抽取10个样本量为 n 的随机子样本，然后评估计算时间、迭代次数和估计结果。我们选择 $(\sigma^2, \alpha, \nu) \in [0.01, 5] \times [0.01, 5] \times [0.01, 2]$ 作为 `ExaGeoStat`/BOBYQA 方法的搜索范围，并将下极值 $(0.01, 0.01, 0.01)$ 和上极值 $(5, 5, 2)$ 分别作为我们所提出方法的初始猜测向量，将中点 $(2.505, 2.505, 1.005)$ 作为 `ExaGeoStat`/Nelder-Mead 方法的初始值。图 10 展示了计算时间、迭代次数，以及所提出方法与其他方法之间的计算时间差。表 2 展示了不同 n下估计参数的均值和标准差。

如图 9所示，所提出的 Fisher-BT 方法在所有考察的情形下都缩短了计算时间。尽管该方法需要计算对数似然函数的导数，但其所需的对数似然函数评估次数仍显著较少。表 2 显示，我们考察的方法在所有情形下得出的估计值相近。事实上，不同方法对参数 σ^2、α、 ν的估计值几乎相同

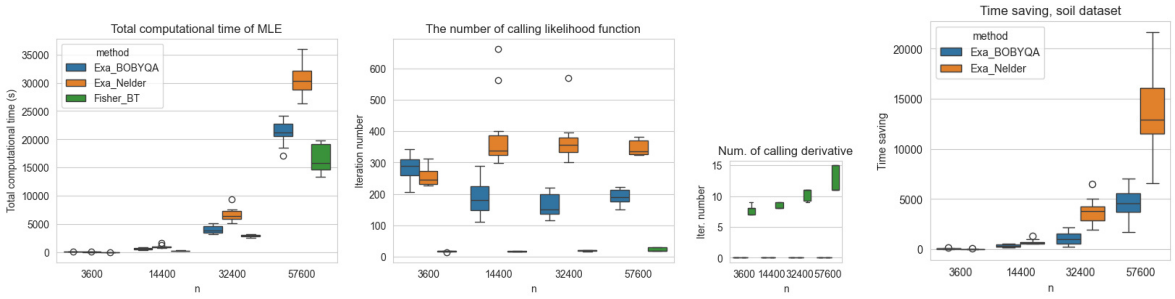


图10：土壤湿度数据集估计的计算时间、迭代次数和计算时间差的箱线图。

表2：土壤湿度数据集估计中不同样本量和方法的估计值均值（标准差）

n	σ^2			α			ν		
	Exa_BOBYQA	Exa_Nelder	Fisher_BT	Exa_BOBYQA	Exa_Nelder	Fisher_BT	Exa_BOBYQA	Exa_Nelder	Fisher_BT
3,600	1.5213	1.5214	1.5214	2.8993	2.8997	2.8997	0.2420	0.2420	0.2420
	(0.0488)	(0.0488)	(0.0488)	(0.3610)	(0.3612)	(0.3612)	(0.0108)	(0.0108)	(0.0108)
14,400	1.3366	1.3366	1.3366	1.3406	1.3406	1.3406	0.2773	0.2773	0.2773
	(0.0394)	(0.0394)	(0.0394)	(0.1307)	(0.1307)	(0.1307)	(0.0050)	(0.0050)	(0.0050)
32,400	1.2122	1.2122	1.2122	0.7553	0.7553	0.7553	0.3114	0.3114	0.3114
	(0.0128)	(0.0127)	(0.0127)	(0.0413)	(0.0413)	(0.0413)	(0.0053)	(0.0053)	(0.0053)
57,600	1.1307	1.1307	1.1307	0.5045	0.5045	0.5045	0.3397	0.3397	0.3397
	(0.0130)	(0.0130)	(0.0130)	(0.0321)	(0.0322)	(0.0322)	(0.0061)	(0.0061)	(0.0061)

最大差异分别仅为 3.34×10^{-4}、1.70×10^{-3} 和 2.62×10^{-6}。除两个 $n = 57$ 为 600 的样本外，所提出的方法在所有数据集上都取得了最大的对数似然值；对于这两个样本，其对数似然值分别仅比 `ExaGeoStat`/BOBYQA 方法低 1.35×10^{-10} 和 3.64×10^{-12}。总之，对于本研究考察的土壤水分子样本数据集，所提出的方法能够在实现有竞争力的估计精度的同时，显著缩短计算时间。

5 结论与讨论

本文提出了 Fisher-BT 优化算法，用于计算大规模空间数据集的平稳 Matérn 协方差模型的精确最大似然估计。我们将 Fisher评分法与回溯线搜索和 Nelder-Mead 方法相结合，为该算法引入了两阶段结构。Fisher评分法步骤通过减少对数似然函数的调用次数来加快收敛；Geoga 等人 (2023年) 的级数近似方法通过简化进一步稳定了收敛。当Fisher评分法无法收敛时，Nelder-Mead 步骤可作为保障机制，从而在更广泛的光滑度参数取值范围内获得更稳健的估计值 ν。借助ExaGeoStat 框架，我们的方法可在多核及众核系统上实现。模拟和真实数据应用表明，我们提出的方法通过减少对数似然函数的调用次数，在保持估计精度的同时缩短了精确最大似然估计的计算时间。模拟还表明，与 ν 相比，我们提出的方法对较小或较大的 ExaGeoStat/BOBYQA 和 ExaGeoStat/Nelder-Mead 方法更加稳健。

所提出的算法可以轻松推广到其他平稳协方差模型，例如包含块金项或几何各向异性的模型。未来，我们计划将该算法与矩阵近似技术相结合，例如分块低秩近似（Abdulah 等人 2018年），从而开发一种计算效率更高的方法，适用于具有 $n > 10^5$ 的超大空间数据集。

我们的方法在稳健性方面面临的主要障碍是第二类修正贝塞尔函数的计算精度。计算二阶导数的稳健性较差，且耗费更多计算时间，因此我们没有采用牛顿法进行优化。因此，如果能找到一种数值稳定性更高的算法来

计算 K_ν 该函数及其导数。我们提出的方法也可以与Sun & Stein （2016）提出的无偏估计方程法结合使用，从而消除因计算不稳定而要求固定光滑度参数的限制。总之，我们提出的 Fisher-BT 算法为精确最大似然估计计算提供了一种快速、准确且可靠的工具，并为评估 MLE 近似方法的计算精度确立了基准。

6 利益冲突声明

作者不存在利益冲突。

7 数据可用性声明

去标识化数据已在补充材料 C 中提供。

补充材料

A. 计算偏导数的算法：一份文档介绍了计算 $\boldsymbol{\Sigma}_\nu$ 的详细算法，即协方差矩阵关于光滑度参数的偏导数。

在本补充材料 A 中，我们介绍 $\boldsymbol{\Sigma}_\nu$的计算算法，该算法基于Geoga 等人 (2023年提出的级数近似算法。程序中，计算导数的条件如下：

1. 检查 $x = 0$。当 $x = 0$ 时，导数为零；

2. 如果不是，请检查 $x < 8.5$ 和 $|\nu - [\nu]| > 10^{-6}$，其中 $[\nu]$ 是 ν最接近的整数；
3. 如果不是，请检查 $x \geq 8.5$ 和 $x < 30$是否满足条件；

4. 如果不是，请检查 $x \geq 30$ 和 $|(\nu+0.5)-[\nu+0.5]| > 10^{-6}$是否满足条件；

5. 如果不是，使用滞后量 10^{-9} 的差分近似来计算导数。

以下是情况 2–4 的算法：

情况 2： $0 < x < 8.5$**，** ν **不是接近整数的值。**

在这种情况下， $x^{\nu}K_{\nu}(x)$ 的导数满足以下关系，其中 K_{ν} 是第二类修正贝塞尔函数，

$$K_{\nu}(x) = \sum_{k=0}^{\infty} \left(\frac{x}{2}\right)^{2k} \frac{1}{2 \cdot k!} \left\{ \Gamma(\nu) \left(\frac{x}{2}\right)^{-\nu} \frac{\Gamma(1-\nu)}{\Gamma(1+k-\nu)} + \Gamma(-\nu) \left(\frac{x}{2}\right)^{\nu} \frac{\Gamma(1+\nu)}{\Gamma(1+k+\nu)} \right\}.$$

令 $g_k(\nu) = \Gamma(1+\nu)/\Gamma(1+k+\nu)$，则

$$g_k(\nu) = \begin{cases} \prod_{j=1}^{k} (\nu+j)^{-1}, & k \geq 1, \\ 1, & k = 0. \end{cases} \qquad d_k(\nu) := \frac{\partial}{\partial \nu} \log(g_k(\nu)) = \begin{cases} -\sum_{j=1}^{k} (\nu+j)^{-1}, & k \geq 1, \\ 0, & k = 0. \end{cases}$$

通过直接计算，

$$\begin{aligned} \frac{\partial}{\partial \nu}[x^{\nu}K_{\nu}(x)] = \sum_{k=0}^{\infty} \left(\frac{x}{2}\right)^{2k} \frac{1}{2 \cdot k!} \Big[& 2^{\nu} \left\{ \log 2 + \psi(\nu) - d_k(-\nu) \right\} \Gamma(\nu) g_k(-\nu) \\ & + 2^{-\nu} x^{2\nu} \left\{ -\log 2 + 2\log(x) - \psi(-\nu) + d_k(\nu) \right\} \Gamma(-\nu) g_k(\nu) \Big], \end{aligned}$$

其中 $\psi(\nu) = \Gamma'(\nu)/\Gamma(\nu)$ 是双伽马函数。该级数用截至第 20 项的部分和近似。

情况 3： $8.5 \leq x < 30$。

当 $\nu \to \infty$ 时， $K_{\nu}(\nu x)$ 用以下公式近似

$$K_{\nu}(\nu x) \approx \sqrt{\frac{\pi}{2\nu}} \frac{\exp(-\nu\eta(x))}{(1+x^2)^{1/4}} \sum_{k=0}^{\infty} (-1)^k \nu^{-k} U_k(p(x)),$$

其中

$$\eta(x) = \sqrt{1+x^2} + \log\left(\frac{x}{1+\sqrt{1+x^2}}\right), \quad p(x) = (1+x^2)^{-1/2},$$

且 $U_k(p)$ 是 $3k$ 阶多项式，递归定义为 $U_0(p)\equiv 1$,

$$U_{k+1}(p)=\frac{1}{2}p^2(1-p^2)U_k'(p)+\frac{1}{8}\int_0^p(1-5t^2)U_k(t)dt.$$

令 $U_k(p)=\sum_{j=0}^{3k}c_j^{(k)}p^j$，则通过直接计算,

$$c_0^{(k+1)}=0;\quad c_1^{(k+1)}=c_0^{(k)}/8;\quad c_2^{(k+1)}=9c_1^{(k)}/16;\quad c_3^{(k+1)}=25c_2^{(k)}/24-5c_0^{(k)}/24;$$
$$c_j^{(k+1)}=\frac{1}{2}\left(j-1+\frac{1}{4j}\right)c_{j-1}^{(k)}-\frac{1}{2}\left(j-3+\frac{5}{4j}\right)c_{j-3}^{(k)},\quad j\in\{4,\dots,3k+1\};$$
$$c_j^{(k+1)}=-\frac{1}{2}\left(j-3+\frac{5}{4j}\right)c_{j-3}^{(k)},\quad j\in\{3k+2,3k+3\}.$$

通过直接计算,

$$\begin{aligned}&\frac{\partial}{\partial\nu}[x^\nu K_\nu(x)]\\&\approx\left[\log\nu+\log\left\{1+\sqrt{1+\left(\frac{x}{\nu}\right)^2}\right\}-\frac{1}{2\nu\left\{1+(x/\nu)^2\right\}}\right]g_\nu(x)\sum_{k=0}^{\infty}(-1)^k\nu^{-k}U_k(p(x/\nu))\\&-g_\nu(x)\sum_{k=0}^{\infty}(-1)^k k\nu^{-k-1}U_k(p(x/\nu))\\&+\frac{x^2}{\nu^3}\left\{1+(x/\nu)^2\right\}^{-3/2}g_\nu(x)\sum_{k=0}^{\infty}(-1)^k\nu^{-k}U_k'(p(x/\nu)),\end{aligned}$$

其中

$$g_\nu(x)=x^\nu\sqrt{\frac{\pi}{2\nu}}\frac{\exp(-\nu\eta(x/\nu))}{\{1+(x/\nu)^2\}^{1/4}}.$$

当 $8.5\le x<15$ 时，级数用截至第 12 项的部分和近似

当 $15\le x<30$ 时，则用截至第 8 项的部分和近似。

情形4： $x\ge 30$ **和** $\nu+0.5$ **不接近整数**。

当 $x\to\infty$、$K_\nu(x)$ 由以下内容近似时

$$K_\nu(x)=\sqrt{\frac{\pi}{2x}}e^{-x}\sum_{k=0}^{\infty}x^{-k}a_k(\nu),$$

其中

$$a_k(\nu)=\begin{cases}1, & k=0;\\ \frac{1}{8^k\Gamma(k+1)}\prod_{j=1}^{k}\{4\nu^2-(2j-1)^2\}, & k\ge 1.\end{cases}$$

通过直接计算，

$$\frac{\partial}{\partial \nu}[x^{\nu} K_{\nu}(x)] \approx \sqrt{\frac{\pi}{2}} x^{\nu-1/2} e^{-x} \sum_{k=0}^{\infty} x^{-k} \{\log x + b_k(\nu)\} a_k(\nu),$$

其中

$$b_k(\nu) = \begin{cases} 0, & k = 0; \\ \sum_{j=1}^{k} \frac{8\nu}{4\nu^2 - (2j-1)^2}, & k \geq 1. \end{cases}$$

当 $x \geq 30$ 时，该级数用截至第 5 项的部分和近似。

B. 代码： 一个 ZIP 文件，包含用于复现模拟和土壤水分应用结果的 C 和 Python 代码。

C. 土壤水分数据集： 一个 ZIP 文件，包含文章中使用的土壤湿度数据集以及用于抽样的代码。

参考文献

Abdulah, S., Baker, A. H., Bosilca, G., Cao, Q., Castruccio, S., Genton, M. G., Keyes, D. E., Khalid, Z., Ltaief, H., Song, Y. 等（2024），利用百亿亿级气候模拟器提升地球系统模型的输出质量，并节省其存储空间中的拍字节，载于 《SC24：高性能计算、网络、存储与分析国际会议》，IEEE，第 1–12 页。

Abdulah, S., Li, Y., Cao, J., Ltaief, H., Keyes, D. E., Genton, M. G. & Sun, Y.（2023年），《基于 ExaGeoStatR 的大规模环境数据科学》，*Environmetrics***34**(1)，e2770。

Abdulah, S., Ltaief, H., Sun, Y., Genton, M. G. & Keyes, D. E. (2018年), ‘Exa GeoStat：面向众核系统的地统计学高性能统一软件’, *IEEE*并行与分布式系统汇刊 **29**(12), 2771–2784。

Agullo, E., Augonnet, C., Dongarra, J., Ltaief, H., Namyst, R., Thibault, S. & Tomov, S. (2012年), 面向图形处理器的高性能线性代数软件混合化方法，载于‘GPU Computing Gems Jade Edition’，爱思唯尔，第473–484页。

Barac, D., Multerer, M. D. & Iber, D.（2019年），‘利用高斯过程从图像数据估计生物参数的全局优化’，《理论生物学杂志》**481**，233–248。

Brown, P. E.（2015年），‘基于模型的地统计学：简易方法’，《统计软件杂志》**63**，1–24。

Broyden, C. G.（1970年），‘一类双秩最小化算法的收敛性 1. 一般性考量’，《*IMA*应用数学杂志》**6**（1），76–90。

DeOliveira, V. & Han, Z.（2022年），‘关于高斯 Matérn 随机场中协方差参数的信息’，《农业、生物与环境统计学杂志》**27**（4），690–712。

Diggle, P. & Ribeiro, P.（2007年），基于模型的地统计学，施普林格。

Finley, A. O., Datta, A., Cook, B. C., Morton, D. C., Andersen, H. E. & Banerjee, S. (2017年), ‘将最近邻高斯过程应用于海量空间数据集：塔纳纳谷阿拉斯加的森林冠层高度预测’, *arXiv* 预印本 *arXiv:1702.00434***7**。

Geoga, C. J., Anitescu, M. & Stein, M.L. (2020年), ‘使用层次矩阵实现可扩展的高斯过程计算’, 《计算与图形统计学杂志》**29**(2), 227–237。

Geoga, C. J., Marin, O., Schanen, M. & Stein, M. L.（2023年），‘使用自动微分拟合 Matérn 光滑度参数’，《统计与计算》**33**(2), 48。

Guinness, J.（2021），‘通过对 Vecchia 近似进行 Fisher 评分法实现高斯过程学习’，《统计与计算》**31**(3), 25。

Hong, Y.、Abdulah, S.、Genton, M. G. 和 Sun, Y.（2021），“大型数据集空间预测近似方法的效率评估”，《空间统计》**43**，100517。

Hooke, R. 和 Jeeves, T. A.（1961），“数值与统计问题的‘直接搜索’解法”，《*ACM*学报》（*JACM*）**8**（2），212–229。

Huang, H. 和 Sun, Y.（2018），“大型空间数据集似然函数的分层低秩近似”，《计算与图形统计学杂志》**27**（1），110–118。**网址：** *https://doi.org/10.1080/10618600.2017.1356324*

Koncewicz, W.、Moździerz, M. 和 Brus, G.（2023），“一种用于评估烃重整过程中碳沉积的快速基于高斯过程的方法”，《国际氢能杂志》**48**（31），11666–11679。

Nelder, J. A. & Mead, R. (1965)，《函数最小化的单纯形法》，《计算机杂志》**7**(4), 308–313。

Nychka, D., Furrer, R., Paige, J. & Sain, S. (2021)，《fields：空间数据工具》。R软件包版本 16.3.1。

网址： *https://github.com/dnychka/fieldsRPackage*

Pardo-Igúzquiza, E., Mardia, K. V. & Chica-Olmo, M. (2009)，《MLMATERN：采用空间Matérn协方差模型进行最大似然推断的计算机程序》，《计算机与地球科学》**35**(6), 1139–1150。

Powell, M. J.（2009），《用于有界约束优化且无需导数的 BOBYQA 算法》，剑桥 *NA*报告 *NA2009/06* **26**，26–46。

Ribeiro Jr, P. J., Christensen, O. F. & Diggle, P. J.（2003），地统计软件 geoR 和 geoRglm，载于《DSC会议论文集》，第 2 卷，第 1–15 页。

Schlather, M., Malinowski, A., Menck, P.J., Oesting, M. 和 Strokorb, K.（2015年），《使用 RandomFields 程序包分析、模拟和预测多元随机场》，《统计软件杂志》**63**，1–25。

Stein, M. L.（2012年），空间数据插值：克里金的一些理论，Springer Science & Business Media。

Sun, Y. 和 Stein, M. L.（2016），《适用于大型空间数据集的统计与计算高效估计方程》，《计算与图形统计学杂志》**25(1)**，187–208。

Zhang, H.（2004），《基于模型的地统计学中的不一致估计与渐近等价插值》，《美国统计学会杂志》**99(465)**，250–261。