arXiv:2601.01345v1 [math.ST] 4 Jan 2026

协变量依赖 Copula 参数局部似然估计的一致渐近理论

Mathias N. Muia†

†数学与统计学系，南阿拉巴马大学，莫比尔，阿拉巴马州 36688，美国

✉mnmuia@southalabama.edu

摘要

条件 Copula 模型允许依赖结构随观测到的协变量变化，同时保持边缘行为与关联之间的区分。我们研究多元条件Copula模型中，针对平滑变化的 Copula 参数的核加权局部似然估计量的一致渐近行为。通过对经过适当变换的校准函数进行局部多项式近似，我们建立了局部对数似然、其得分函数及海森矩阵在紧协变量集上的一致收敛速率。这些结果进一步保证了局部最大似然估计量及诱导的Copula参数函数的一致性。该分析基于针对具有收缩邻域和多项式熵界的核索引类所采用的经验过程技术，为协变量依赖的 copula 模型中的全局一致性与稳定的局部优化提供了理论依据。

关键词： 条件 Copula；局部似然估计；一致性；经验过程理论；局部多项式方法；协变量依赖的依赖关系

1 引言

Copula为独立于边缘分布建模依赖结构提供了灵活的框架。根据Sklar定理 Sklar(1959年)，任何边缘分布连续的多元分布都存在唯一的Copula表示，从而可以对依赖关系进行研究

而不受边缘行为的影响。然而，在许多当代应用中，依赖关系并非恒定不变，且往往随可观测协变量而变化。这种由协变量驱动的依赖关系在金融、环境科学和生物医学研究等领域十分常见，在这些领域中，多个结果变量之间的关联会随时间、空间或人口特征而变化。

条件 Copula 模型通过将 Sklar 表示扩展到条件分布，旨在捕捉这种异质性（Patton, 2006年）。设 (X, Y) $=(X_1,\ldots,X_n,Y_1,\ldots,Y_s)$ 为一个随机向量，并将条件边缘分布定义为 $F_{k|\mathbf{Y}}(x_k \mid \mathrm{Y}=\mathrm{y}) = \mathbb{P}(X_k \le x_k \mid \mathrm{Y}=\mathrm{y})$，$k=1,\ldots,n$。给定 Y $=$ y 时 X 的条件联合分布为 $H_{\mathbf{X}|\mathbf{Y}}(x_1,\ldots,x_n \mid \mathrm{y}) = \mathbb{P}(X_1 \le x_1,\ldots,X_n \le x_n \mid \mathrm{Y}=\mathrm{y})$。当条件边缘分布连续时，存在唯一的 n维条件 Copula $C(\cdot \mid \mathrm{y})$，满足$H_{\mathbf{X}|\mathbf{Y}}(x_1,\ldots,x_n \mid \mathrm{y}) = C\Big(F_{1|\mathbf{Y}}(x_1 \mid \mathrm{y}) \ldots, F_{n|\mathbf{Y}}(x_n \mid \mathrm{y}) \mid \mathrm{y}\Big)$.。条件 Copula 与条件边缘分布共同完整地刻画了条件联合分布。

这种表述意味着，条件伪观测值 $U_{k|\mathbf{Y}} = F_{k|\mathbf{Y}}(X_k \mid \mathrm{Y}), k=1,\ldots,n$ 在 (0, 1) 上服从均匀分布，并满足 $(U_{1|\mathbf{Y}},\ldots,U_{n|\mathbf{Y}}) \mid \mathrm{Y}=\mathrm{y} \sim C(\cdot \mid \mathrm{y})$。因此，在给定协变量值 Y $=$y 时，条件 Copula 捕捉了 X 的全部相依结构。在参数化条件 Copula 模型中，协变量效应通常通过平滑参数函数 θ(y) 编码，从而得到形如$C(\cdot \mid \theta(\mathrm{y}))$ 的 Copula。常见示例包括带有参数化生成元的阿基米德族 McNeil 和 Nešlehová（2009年）；Genest 等（1995年）。

在这种设定下进行推断，需要估计未知的校准函数 $\theta(\cdot)$，同时遵循 Copula 族施加的约束。局部似然方法为此提供了一种自然的非参数策略。该方法通过处理变换后的校准函数 $\nu(\mathbf{y}) = \psi(\theta(\mathbf{y}))$，用多元多项式在局部对其进行近似，并最大化核加权条件对数似然。由此得到的估计量兼具灵活性与良好的渐近行为。

已有大量研究在各种假设下探讨了 Copula 模型的估计问题。早期研究考察了 Copula 密度及相关量的基于核的估计，以及其一致性与渐近正态性Gijbels 和 Mielniczuk (1990)。后续研究着重于减轻边界偏差并开发基于变换的修正，从而改善弱收敛性质Omelka 等人(2009)。此外，研究者还采用最小距离和基于秩的方法，在不指定边缘分布的情况下对 Copula 参数进行半参数估计Tsukahara (2005)。

近年来，关于协变量依赖的依赖结构的研究显著增多。研究者提出了具有未知校准函数的半参数条件 Copula 模型，并采用局部多项式伪似然方法进行估计Abegaz 等人(2012)。针对具有随协变量变化参数的参数条件 Copula，局部似然方法得出了逐点渐近偏差与方差表达式，并提出了基于数据驱动的模型选择和推断程序Acar etal.(2011b)。研究者还考察了条件 Copula 及其相关条件依赖度量的全非参数估计量Veraverbeke 等人 (2011)。

相关进展包括通过局部似然估计时变依赖参数的动态 Copula 模型，这类模型已应用于金融时间序列分析 Hafner 和 Reznikova (2010年)。对参数方法、半参数方法和边缘分布推断函数方法的比较表明，在边缘分布设定错误时，半参数方法具有较强的稳健性 Kim 等人 (2007年)。近期研究考察了 Copula 密度的非参数估计，尤其关注边界效应以及角点附近的无界行为，并提出了具有理论保证和实证证据支持的基于核的方法Muia etal.(2025a,b)。

尽管取得了这些进展，现有关于协变量依赖 Copula 模型的大多数渐近结果仍是逐点的，仅描述固定协变量值处的行为。这些结果未能解决若干具有重要理论和实践意义的问题。要确立全局一致性、确保数值优化程序的稳定性，并证明数据驱动的带宽选择合理，就需要在整个协变量空间上进行一致控制。得分与海森过程的一致收敛，对于分析局部极大值点并推导更高层次的渐近性质也起着关键作用。

本文通过为协变量依赖的 copula 模型中的局部似然估计量建立一致渐近理论，弥补了这些不足。尽管先前的研究给出了逐点偏差与方差展开 Acar et al. (2011b);Muia et al. (2025d)，我们证明了核加权局部对数似然及其梯度和海森矩阵在协变量空间的紧子集上一致收敛。为此，需要谨慎控制由协变量位置和局部多项式系数共同索引的经验过程，并处理核局部化所导致的收缩支撑集。

我们的主要目标是证明协变量依赖型 Copula 参数的局部似然估计量具有一致性。我们推导了局部

对数似然、得分函数和海森矩阵的一致收敛速率，进而证明校准函数 $\nu(\mathbf{y})$的局部最大似然估计量以及诱导的 Copula 参数函数 $\theta(\mathbf{y})$具有一致性。尽管这些结果为数据驱动的带宽选择提供了理论依据，但对所选带宽本身进行详细的渐近分析超出了本文的研究范围。

该分析依赖于经验过程技术，用于研究由协变量位置和局部多项式系数共同索引的函数类。主要困难源于多元核局部化：随着带宽减小，它会产生收缩邻域和发散包络函数。我们利用协变量空间的几何性质、多项式熵界以及经验过程的极大不等式来应对这些挑战。与非参数平滑中的通常情形一样，所得收敛速率取决于协变量维数 s，体现了人们熟知的维数灾难。

本文的贡献主要有三点。第一，在温和的正则性条件下，我们建立了核加权局部似然过程的一致随机等度连续性。第二，我们获得了明确的一致收敛速率，反映了带宽选择、样本量和协变量维数的影响。第三，我们为协变量依赖 Copula 模型中的一致推断和可靠的全局优化提供了理论基础。

本文其余部分的组织如下。第 2节介绍条件 Copula 模型和局部似然估计框架。第3节讨论实现问题，包括使用交叉验证准则进行带宽选择和 Copula 族选择。第 4节给出主要理论结果，包括局部对数似然、其得分函数及其海森矩阵的一致收敛速率，以及由此得到的一致性结论。第 5 节报告用于说明有限样本表现的模拟结果。附录包含补充背景材料和技术论证。

2 模型设定与局部似然估计

令 $(\mathbf{X},\ \mathbf{Y}) = (X_1,\ \ldots\ldots,\ X_n,\ Y_1,\ \ldots\ldots,\ Y_s)$ 为一个随机向量，其中 $\mathbf{Y} \in \mathbb{R}^s$。考虑一个参数 Copula 族 $\{C(\cdot \mid \theta) :\ \theta \in \Theta\}$ ，其对应的 Copula 密度为 $c(\cdot \mid \theta)$。我们假设，给定$\mathbf{Y} = \mathbf{y}$，随机向量 $\mathbf{X}$ 具有 Copula $C(\cdot \mid \theta(\mathbf{y}))$，其中 Copula参数$\theta(\mathbf{y})$ 是一个未知的光滑函数，其取值属于参数空间 $\Theta \subset \mathbb{R}^d$。我们的目标是通过用一个多元

次数为 p 的多项式。

由于 Copula 参数函数 $\theta(\mathbf{y})$必须取值于可容许参数空间Θ，我们引入严格单调的链接变换 $\nu(\mathbf{y}) = \psi(\theta(\mathbf{y}))$ $\theta(\mathbf{y}) = \psi^{-1}(\nu(\mathbf{y}))$，其中逆链接 $\psi^{-1}: \mathbb{R} \to \Theta$ 将实数轴映射到参数空间，从而确保$\theta(\mathbf{y}) \in \Theta$ 对所有$\mathbf{y}$均成立。变换后的函数 $\nu(\mathbf{y})$称为校准函数。这一变换至关重要，因为局部多项式建模通常是在$\mathbb{R}$上构建的，而 Copula 参数通常受到约束。例如，高斯藤参数满足 $\theta \in (-1,1)$，而对于 Clayton藤，通常假设$\theta \in (0, \infty)$。链接函数在整个估计过程中确保这些约束始终成立。

接下来，我们介绍校准函数的局部多项式近似。设$\{(X_{1i}, \ldots, X_{ni}, Y_{1i}, \ldots, Y_{si})\}_{i=1}^{N}$ 为一个随机样本，并假设校准函数$\nu(\cdot)$在内点$\mathbf{y} \in \mathbb{R}^s$处具有（$p+1$）阶连续可微性。对 $\nu(\mathbf{Y}_i)$在$\mathbf{y}$ 处作多元泰勒展开，得到

$$\nu(\mathbf{Y}_i) = \nu(\mathbf{y}) + \nabla\nu(\mathbf{y})^\top(\mathbf{Y}_i - \mathbf{y}) + \frac{1}{2}(\mathbf{Y}_i - \mathbf{y})^\top\nabla^2\nu(\mathbf{y})(\mathbf{Y}_i - \mathbf{y}) + \cdots + R_{p+1}(\mathbf{Y}_i - \mathbf{y}), \quad (2.1)$$

其中，$\nabla\nu(\mathbf{y})$ 和 $\nabla^2\nu(\mathbf{y})$ 分别表示 ν 在$\mathbf{y}$处的梯度和海森矩阵，余项满足 $R_{p+1}(\mathbf{Y}_i - \mathbf{y}) = O(\|\mathbf{Y}_i - \mathbf{y}\|^{p+1})$。为方便记号，我们用 $\nu(\mathbf{Y}_i) \approx \boldsymbol{\phi}_p(\mathbf{Y}_i - \mathbf{y})^\top\boldsymbol{\gamma}$, 近似 $\nu(\mathbf{Y}_i)$，其中 $\boldsymbol{\phi}_p(\cdot)$ 汇集了总次数不超过 p 的所有单项式，而 $\boldsymbol{\gamma}$ 表示相应的局部多项式系数向量。

基于这一近似，我们构造核加权局部对数似然。令$\mathbf{U}_i = (U_{1i}, \ldots, U_{ni})^\top$ 表示来自条件边缘分布的伪观测值，其中 $U_{ki} = F_{k|\mathbf{Y}}(X_{ki} \mid \mathbf{Y}_i)$。假设观测值之间条件独立，则数据点 $(\mathbf{U}_i \mid \mathbf{Y}_i)$ 对数似然中的贡献为其在$\mathbf{y}$ 邻域内的 $\log c(\mathbf{U}_i \mid \psi^{-1}(\nu(\mathbf{Y}_i)))$, 其中$c(\cdot \mid \cdot)$ 表示条件藤密度函数。由此得到的核加权局部对数似然为

$$L_N(\boldsymbol{\gamma}, \mathbf{y}, p, h) = \sum_{i=1}^{N} \log c\Big(\mathbf{U}_i \,\Big|\, \psi^{-1}\big(\boldsymbol{\phi}_p(\mathbf{Y}_i - \mathbf{y})^\top\boldsymbol{\gamma}\big)\Big)\, K_h(\mathbf{Y}_i - \mathbf{y}), \quad (2.2)$$

其中，$K_h(\mathbf{v}) = h^{-s}K(\mathbf{v}/h)$ 是经过重新缩放的多元核。

核权重 $K_h(\mathbf{Y}_i - \mathbf{y})$ 将估计局部化到协变量接近$\mathbf{y}$的观测值。更一般地，可以考虑带宽矩阵为 H 以下形式的多元核：$K_H(\mathbf{v}) = |H|^{-1}K(H^{-1}\mathbf{v})$；例如，参见 Wand and Jones (1995年)。为简单起见，我们采用

各向同性设定 $H = hI_s$，该设定在所有协变量维度上进行均匀平滑，同时保留标准归一化和渐近性质。 K的典型选择包括乘积 Epanechnikov核和多元高斯核。

为刻画（2.2）的最大化点，我们定义关于局部系数向量 $\boldsymbol{\gamma}$ 的得分向量和 Hessian矩阵。令 $\ell(\theta,\mathbf{u}) = \log c(\mathbf{u} \mid \theta)$ 表示对数Copula密度，并记 $(\cdot) = (\boldsymbol{\phi}_p(\mathbf{Y}_i - \mathbf{y})^\top \boldsymbol{\gamma})$。应用链式法则可得得分函数

$$\nabla_{\boldsymbol{\gamma}} L_N(\boldsymbol{\gamma}, \mathbf{y}) = \sum_{i=1}^{N} \ell'\Big(\psi^{-1}(\cdot), \mathbf{U}_i\Big)\, (\psi^{-1})'(\cdot)\, \boldsymbol{\phi}_p(\mathbf{Y}_i - \mathbf{y})\, K_h(\mathbf{Y}_i - \mathbf{y}),$$

以及海森矩阵

$$\nabla^2_{\boldsymbol{\gamma}} L_N(\boldsymbol{\gamma}, \mathbf{y}) = \sum_{i=1}^{N} \Big[\ell''\Big(\psi^{-1}(\cdot), \mathbf{U}_i\Big)\, \{(\psi^{-1})'(\cdot)\}^2 + \ell'\Big(\psi^{-1}(\cdot), \mathbf{U}_i\Big)\, (\psi^{-1})''(\cdot)\Big] \times \boldsymbol{\phi}_p(\mathbf{Y}_i - \mathbf{y}) \boldsymbol{\phi}_p(\mathbf{Y}_i - \mathbf{y})^\top K_h(\mathbf{Y}_i - \mathbf{y}).$$

局部最大似然估计量定义为

$$\hat{\boldsymbol{\gamma}}(\mathbf{y}) = \arg\max_{\boldsymbol{\gamma}} L_N(\boldsymbol{\gamma}, \mathbf{y}, p, h), \tag{2.3}$$

并通过求解估计方程 $\nabla_{\boldsymbol{\gamma}} L_N(\boldsymbol{\gamma}, \mathbf{y}) = 0$ 得到。实际应用中，可使用 Newton–Raphson算法进行此最大化，迭代更新形式为

$$\boldsymbol{\gamma}^{(m+1)} = \boldsymbol{\gamma}^{(m)} - \Big\{\nabla^2_{\boldsymbol{\gamma}} L_N(\boldsymbol{\gamma}^{(m)}, \mathbf{y})\Big\}^{-1} \nabla_{\boldsymbol{\gamma}} L_N(\boldsymbol{\gamma}^{(m)}, \mathbf{y}), \tag{2.4}$$

迭代直至收敛。

校准函数在 $\mathbf{y}$ 处的估计量为 $\hat{\nu}(\mathbf{y}) = \mathbf{e}_0^\top \hat{\boldsymbol{\gamma}}(\mathbf{y})$，其中 $\mathbf{e}_\alpha$表示在局部多项式展开中选取截距项（α^{th}阶）系数的单位向量。相应的Copula参数估计量为 $\hat{\theta}(\mathbf{y}) = \psi^{-1}(\hat{\nu}(\mathbf{y}))$。将 $\hat{\theta}(\mathbf{y})$ 代入 $C(\cdot \mid \theta(\mathbf{y}))$，即可得到协变量值 $\mathbf{y}$ 处的Copula函数估计量。

需要注意的是，上述最大化问题是在固定的评估点 $\mathbf{y}$ 处进行局部化的，因此只能在该点得到 Copula 参数的估计值。因此，为了恢复完整的校准函数 $\theta(\cdot)$，必须在覆盖协变量 $\mathbf{Y}$取值域的足够精细的点网格 $\mathbf{y}$ 上求解局部最大化问题。为完整起见，

局部似然估计量的其他渐近性质，包括逐点偏差与方差展开以及渐近正态性，将在附录中讨论。

本文的主要重点是确立这些估计量在评估点$\mathbf{y}$ 遍历紧集$\mathcal{Y}_0$时的一致渐近行为。

3 模型调优

协变量依赖的 copula 模型调优涉及两个相关的决策。第一个是选择局部似然估计所用的平滑带宽。第二个是选择合适的参数 Copula 族。我们采用局部似然推断中常用的交叉验证策略，并遵循Acar et al.(2011年b)提出的一般思路，并将其应用于多元协变量。

3.1 带宽选择

由于估计基于核加权局部 copula 似然，因此很自然地采用留一交叉验证来选择带宽。令 $\theta_h(\cdot)$ 表示使用带宽 h得到的 copula 参数函数估计量。对于每个观测值$(\mathbf{U}_i, \mathbf{Y}_i)$，使用除第 i个观测值以外的所有观测值计算留一估计量$\hat{}\theta_{h,-i}(\mathbf{Y}_i)$。然后通过最大化交叉验证局部似然来选择带宽

$$\mathrm{CVL}(h) = \sum_{i=1}^{N} \log c\Big(\mathbf{U}_i \mid \hat{\theta}_{h,-i}(\mathbf{Y}_i)\Big)\,. \tag{3.1}$$

选定的带宽 h_{cv} 是式（3.1）的最大化点。在协变量设计高度不均匀或局部变化较强的情况下，也可以采用可变带宽或最近邻带宽等自适应选择方式。

3.2 Copula 族选择

选择带宽并不能决定依赖关系的形式，这由 Copula 族决定。由于不同族之间的似然值无法直接比较，因此选择依据是预测性能，而非似然最大化。

令 $\mathcal{C} = \{C_q\colon\ q = 1, \ldots, Q\}$ 表示一个候选 copula 族的集合。对于每个族C_q，带宽选择程序会得出一个最优带宽 h_q^*，随后用它来

计算留一估计 $\hat{\theta}_{h_q^*}^{(-i)}(\mathbf{Y}_i)$。这些估计定义了如下形式的候选条件 Copula 模型： $C_q(\cdot \mid \hat{\theta}_{h_q^*}^{(-i)}(\mathbf{Y}_i))$。

预测精度通过条件预测进行评估：给定其余分量和协变量，预测每个伪观测。对于给定的族 C_q， U_{ki} 在给定$\mathbf{U}_{-k、i}$ 和$\mathbf{Y}_i$ 时的条件期望为

$$\widehat{\mathbb{E}}_q^{(-i)}(U_{ki} \mid \mathbf{U}_{-k,i}, \mathbf{Y}_i) = \int_0^1 u_k \, c_{q,k}\Big(u_k \mid \mathbf{U}_{-k,i}, \hat{\theta}_{h_q^*}^{(-i)}(\mathbf{Y}_i)\Big) \, du_k,$$

其中，$c_{q,\ k}(\cdot \mid \cdot)$表示由族 C_q所隐含的条件copula密度。

相应的交叉验证预测误差定义为

$$\mathrm{CVPE}(C_q) = \sum_{i=1}^{N} \sum_{k=1}^{n} \Big\{U_{ki} - \widehat{\mathbb{E}}_q^{(-i)}(U_{ki} \mid \mathbf{U}_{-k,i}, \mathbf{Y}_i)\Big\}^2 .$$

选择使 CVPE(C_q) 在 $q = 1, \ldots, Q$ 上达到最小值的 Copula 族。

该准则可解释为条件均方预测误差的经验近似。当工作 Copula 族与底层相依结构相匹配时，该准则达到最小值，从而为 Copula 族选择提供了一种无需似然且易于解释的依据。

4 主要结果

我们首先引入一些额外的记号，并说明推导主要结果所依据的概率设定。设$\mathcal{Y}_0 \subset \mathbb{R}^s$ 是一个非空紧集，且包含于协变量向量$\mathbf{Y}$的支撑内部。我们假设$\mathbf{Y}$ 关于勒贝格测度具有密度 $f_{\mathbf{Y}}$ 。本节中所有一致性陈述均指在$\mathcal{Y}_0$上成立。

在全文中，我们使用缩放局部对数似然

$$\tilde{L}_N(\boldsymbol{\gamma}; \mathbf{y}) = (Nh^s)^{-1} L_N(\boldsymbol{\gamma}; \mathbf{y}),$$

及其总体对应量

$$Q(\boldsymbol{\gamma}; \mathbf{y}) = \mathbb{E}\Big[\tilde{L}_N(\boldsymbol{\gamma}; \mathbf{y})\Big] .$$

结果　　以下内容描述局部对数似然、其得分函数及其　　s

海森矩阵作为由评估点 $\mathbf{y}$ 和局部多项式系数向量 $\boldsymbol{\gamma}$ 联合索引的随机过程。

正则性条件 4.1。 我们作出以下条件假设。

(R1) 协变量向量 $\mathbf{Y} \in \mathbb{R}^s$ 具有相对于勒贝格测度的密度 $f_{\mathbf{Y}}$ 。存在常数 $0 < c_f \le C_f < \infty$ 使得 $c_f \le f_{\mathbf{Y}}(\mathbf{y}) \le C_f$ 对所有 $\mathbf{y} \in Y_0$, 均成立，并且 $f_{\mathbf{Y}}$ 在 Y_0的一个开邻域上连续。

(R2) 核 $K : \mathbb{R}^s \to \mathbb{R}$ 有界、对称、利普希茨连续，且紧致支撑于 $[-1, 1]^s$上。它满足 $\int_{\mathbb{R}^s} K(v)\, dv = 1$，并且具有有限的 $(p+2)$阶矩 $\int_{\mathbb{R}^s} \|v\|^{p+2} |K(v)|\, dv < \infty$。

(R3) 带宽序列 $h = h_N$满足 $h_N \to 0$，当 $N \to \infty$时， $N h_N^s / \log(1/h_N) \to \infty$，且 $N h_N^{s+2(p+1)} \to 0$。

（*R4*）校准函数 $\nu :\ \mathbb{R}^s \to \mathbb{R}$ 在 Y_0, 的某个开邻域内$p+2$阶连续可微，且所有总阶数不超过 $p+2$ 的偏导数均一致有界。

（*R5*）对于包含 $\{\theta(\mathbf{y}) : \mathbf{y} \in Y_0\}$的紧集 Θ_0 中的每个 θ ，*copula* 对数密度$\ell(\theta, u) = \log c(u \mid \theta)$ 关于 θ二次连续可微，对所有 $u \in (0,\ 1)^n$均成立。存在包络函数 M_1 和 M_2 ，使得

$$\sup_{\theta \in \Theta_0} |\ell'(\theta, u)| \le M_1(u), \qquad \sup_{\theta \in \Theta_0} |\ell''(\theta, u)| \le M_2(u),$$

其中 $E[M_1(U)^2] < \infty$ 且 $E[M_2(U)^2] < \infty$。条件*Fisher*曲率

$$\sigma^2(\mathbf{y}) = -E\big[\ell''(\theta(\mathbf{y}), U) \mid \mathbf{Y} = \mathbf{y}\big]$$

关于 $\mathbf{y}$ 连续，且在 Y_0上有正的下界和有限的上界。

（*R6*）逆链接函数 ψ^{-1} 严格单调且二次连续可微，其一阶和二阶导数在 $\{\nu(\mathbf{y}) : \mathbf{y} \in Y_0\}$上有界。

(R7) 对每个 $\mathbf{y} \in Y_0$, ，总体准则函数 $Q(\boldsymbol{\gamma}; \mathbf{y})$ 在紧集 $\Gamma \subset \mathbb{R}^{d_p}$ 上存在唯一最大化点 $\boldsymbol{\gamma}^*(\mathbf{y})$ 。映射 $\mathbf{y} \mapsto \boldsymbol{\gamma}^*(\mathbf{y})$ 连续，且总体海森矩阵

$$H(\boldsymbol{\gamma}; \mathbf{y}) = E\Big[\nabla^2_{\boldsymbol{\gamma}} \ell\Big(\psi^{-1}(\phi_p(\mathbf{Y} - \mathbf{y})^\top \boldsymbol{\gamma}), U\Big)\, K\Big(\tfrac{\mathbf{Y} - \mathbf{y}}{h}\Big)\Big]$$

在 $\Gamma \times Y_0$ 上一致负定。

（*R8*）存在常数 $C_\ell^{(0)}$、$C_\ell^{(1)}$、$C_\ell^{(2)} < \infty$，使得

$$\sup_{\theta\in\Theta_0}\sup_{u\in(0,1)^n} \left|\ell^{(\kappa)}(\theta,u)\right| \le C_\ell^{(\kappa)}, \qquad \kappa = 0,1,2.$$

引理 4.1. 假设 （R1）、（R2）、（R4）、（R5）、（R6） 和 （R8） 成立，且 $Y_0 \subset \mathbb{R}^s$ 和 $\Gamma \subset \mathbb{R}^{d_p}$ 为紧致集。对于每个 $(\mathbf{y}\ \boldsymbol{\gamma}) \in Y_0 \times \Gamma$，定义

$$f_{\mathbf{y},\boldsymbol{\gamma}}(U,\mathbf{Y}) = h^{-s}\,\ell\Big(\psi^{-1}(\phi_p(\mathbf{Y}-\mathbf{y})^\top\boldsymbol{\gamma}),\,U\Big)\,K\left(\frac{\mathbf{Y}-\mathbf{y}}{h}\right),$$

并令 $F_h = \{f_{\mathbf{y},\,\boldsymbol{\gamma}}:\ \mathbf{y}\ \in Y_0,\ \ \boldsymbol{\gamma}\in\Gamma\}$。那么，当 $h\to 0$ 时，以下结论成立。

（*i*）存在 F_h 的包络函数 F_h 以及常数 $0 < c_1 \le c_2 < \infty$，且它们与 h 无关，使得

$$c_1\,h^{-s} \ \le\ \|F_h\|_{L_2(P)}\ \le\ c_2\,h^{-s}.$$

（*ii*）存在一个与 h 无关的常数 $C > 0$，使得对所有 $0 < \varepsilon \le 1$，

$$\log N\Big(\varepsilon\|F_h\|_{L_2(P)},\ \mathcal{F}_h,\ L_2(P)\Big) \le C\{1+\log(1/h)+\log(1/\varepsilon)\}\,.$$

（*iii*）令 $\mathbb{P}_N$ 表示基于 $\{(U_i,\mathbf{Y}_i)\}_{i=1}^N$ 的经验测度。那么

$$\sup_{f\in\mathcal{F}_h}\left|(\mathbb{P}_N - P)f\right| = O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

注 4.2。 引理 *4.1*中的对数因子源于熵界。在温和的附加条件下

$$\log(1/h) = O(\log N),$$

随机速率也可等价地写为

$$\sup_{f\in\mathcal{F}_h}\left|(\mathbb{P}_N - P)f\right| = O_p\left(\sqrt{\frac{\log N}{Nh^s}}\right).$$

本文其余部分均采用这一重新表述。如果初步带宽或数据驱动带宽在有限样本中超过 *1*，则可以将其截断，而不会影响函数

类别或渐近速率。

证明。 论证遵循标准的经验过程技术，并分为三个步骤。

(i) 根据 (R8)，存在 $C_\ell < \infty$ 使得 $|\ell(\psi^{-1}(\phi_p(\mathrm{Y} - \mathrm{y})^\top \boldsymbol{\gamma}) \; U)| \le C_\ell$ 对所有 $(\mathrm{y} \; \boldsymbol{\gamma}, U, \mathrm{Y})$ 均成立。根据 (R2)， K 有界且支撑于 $[-1,1]^s$，因此 $|K(v)| \le K_{\max}$ 且 $K((\mathrm{Y} - \mathrm{y})/h) \neq 0$意味着 $\|\mathrm{Y} - \mathrm{y}\|_\infty \le h$。于是，$|f_{\mathrm{y},\boldsymbol{\gamma}}(U, \mathrm{Y})| \le C_\ell K_{\max} h^{-s} \mathbf{1}\{\|\mathrm{Y} - \mathrm{y}\|_\infty \le h\}$。

对每个 $\mathbf{y} \in \mathcal{Y}_0,$，定义局部 包络函数 $F_{\mathbf{y},h}(U, \mathbf{Y}) := C_\ell K_{\max} h^{-s} \mathbf{1}\{\|\mathbf{Y} - \mathbf{y}\|_\infty \le h\}$。则 $|f_{\mathbf{y},\boldsymbol{\gamma}}(U, \mathbf{Y})| \le F_{\mathbf{y},h}(U, \mathbf{Y})$ 对所有 $\boldsymbol{\gamma} \in \Gamma$ 均成立。

此外，

$$\|F_{\mathbf{y},h}\|_{L_2(P)}^2 = C_\ell^2 K_{\max}^2 h^{-2s} \, P(\|\mathbf{Y} - \mathbf{y}\|_\infty \le h).$$

根据 (R1)， $f_{\mathbf{Y}}$ 连续，并且在$\mathcal{Y}_0$的某个开邻域上满足 $c_f \le f_{\mathbf{Y}} \le C_f$。因此，对于所有充分小的 $h > 0$，且对$\mathbf{y} \in \mathcal{Y}_0$,一致成立

$$c_f \, \lambda_s([-h,h]^s) \le P(\|\mathbf{Y} - \mathbf{y}\|_\infty \le h) \le C_f \, \lambda_s([-h,h]^s),$$

其中， λ_s 表示 $\mathbb{R}^s$上的勒贝格测度。由于 $\lambda_s([-h,h]^s) = (2h)^s$，存在常数$0 < c_1 \le c_2 < \infty$（与 h无关），使得对$\mathbf{y} \in \mathcal{Y}_0$一致有 $c_1 h^s \le P(\|\mathbf{Y} - \mathbf{y}\|_\infty \le h) \le c_2 h^s$。因此，存在常数 $0 < C_1 \le C_2 < \infty$ （与 h无关），使得

$$C_1 h^{-s} \le \|F_{\mathbf{y},h}\|_{L_2(P)}^2 \le C_2 h^{-s} \quad \text{uniformly for } \mathbf{y} \in \mathcal{Y}_0.$$

两边开平方得到

$$\sqrt{C_1} \, h^{-s/2} \le \|F_{\mathbf{y},h}\|_{L_2(P)} \le \sqrt{C_2} \, h^{-s/2} \quad \text{uniformly for } \mathbf{y} \in \mathcal{Y}_0,$$

这就证明了 (i) 部分。

(ii) 定义归一化类

$$\mathcal{G}_h := \left\{ g_{\mathbf{y},\boldsymbol{\gamma}} := \frac{f_{\mathbf{y},\boldsymbol{\gamma}}}{\|F_{\mathbf{y},h}\|_{L_2(P)}} : \; \mathbf{y} \in \mathcal{Y}_0, \; \boldsymbol{\gamma} \in \Gamma \right\}.$$

在事件 $\|(\mathbf{Y} - \mathbf{y})/h\|_\infty \le 1$ 上，向量 $\phi_p(\mathbf{Y} - \mathbf{y})$ 一致有界，因此 $\psi^{-1}(\phi_p(\mathbf{Y} - \mathbf{y})^\top \boldsymbol{\gamma})$ 的取值落在 Θ_0 的一个紧致子集内。对 $\ell(\psi^{-1}(\phi_p(\mathbf{Y} -$

$\mathbf{y})^\top\boldsymbol{\gamma})$ $U)$ 关于 $(\mathbf{y}\ \boldsymbol{\gamma})$ 作一阶泰勒展开，结合 (R4)、(R5) 和 （R6）中的导数界，存在 $L_0 > 0$ 使得

$$|f_{\mathbf{y},\boldsymbol{\gamma}}(U,\mathbf{Y}) - f_{\mathbf{y}',\boldsymbol{\gamma}'}(U,\mathbf{Y})| \le L_0 h^{-s}(\|\mathbf{Y}-\mathbf{y}'\| + \|\boldsymbol{\gamma}-\boldsymbol{\gamma}'\|)\,\mathbf{1}\{\|\mathbf{Y}-\mathbf{y}\|_\infty \le h \text{ or } \|\mathbf{Y}-\mathbf{y}'\|_\infty \le h\}.$$

特别地，只要$\|\mathbf{Y}\ -\mathbf{y}'\|_\infty \le h$，右侧的指示函数就有如下上界：$\mathbf{1}\{\|\mathbf{Y}\ -\mathbf{y}\|_\infty \le 2h\}$。仿照(i)部分使用(R1)可得 $P(\|\mathbf{Y}\ -\mathbf{y}\|_\infty \le 2h)\ \le Ch^s$，且该结论对$\mathbf{y}$一致成立，因此存在一个与 h无关的 $C_1 > 0$，使得

$$\|f_{\mathbf{y},\boldsymbol{\gamma}} - f_{\mathbf{y}',\boldsymbol{\gamma}'}\|_{L_2(P)} \le C_1 h^{-s/2}(\|\mathbf{Y}-\mathbf{y}'\| + \|\boldsymbol{\gamma}-\boldsymbol{\gamma}'\|).$$

由(i)部分可知，$\|F_{\mathbf{y},\ h}\|_{L_2(P)}\ \ge \sqrt{C_1}\,h^{-s/2}$ 对$\mathbf{y}$一致成立，因此

$$\|g_{\mathbf{y},\boldsymbol{\gamma}} - g_{\mathbf{y}',\boldsymbol{\gamma}'}\|_{L_2(P)} = \frac{\|f_{\mathbf{y},\boldsymbol{\gamma}} - f_{\mathbf{y}',\boldsymbol{\gamma}'}\|_{L_2(P)}}{\|F_{\mathbf{y},h}\|_{L_2(P)}} \le C_2(\|\mathbf{Y}-\mathbf{y}'\| + \|\boldsymbol{\gamma}-\boldsymbol{\gamma}'\|),$$

对于一个与 h无关的常数 $C_2 > 0$。因此，$\mathcal{G}_h$ 是一个有限维 Lipschitz 类，其 Lipschitz 常数与 h无关。

由于$\mathcal{Y}_0 \subset \mathbb{R}^s$ 和 $\Gamma \subset \mathbb{R}^{d_p}$ 均为紧致集，欧几里得空间有界子集的覆盖数标准界表明，$\mathcal{Y}_0$ 可由至多 $C_Y h^{-s}$ 个半径为 h 的上确界范数球覆盖，而 Γ 可由至多 $C_\Gamma \varepsilon^{-d_p}$ 个半径为 ε 的欧几里得球覆盖（参见van der Vaart 和 Wellner(1996年）第 2.1–2.2 章）。合并这些覆盖可得

$$N(\varepsilon, \mathcal{G}_h, L_2(P)) \le Ch^{-s}\varepsilon^{-d_p},$$

其中，常数 $C > 0$ 与 h无关。

最后，由于$\|F_{\mathbf{y},h}\|_{L_2(P)}$一致地关于$\mathbf{y}$被 $h^{-s/2}$ 的常数倍从上方和下方界定，将尺度还原至 $\mathcal{F}_h$ y，得到

$$\log N\Big(\varepsilon\,\|F_{\mathbf{y},h}\|_{L_2(P)},\ \mathcal{F}_h,\ L_2(P)\Big) \le C[1 + \log(1/h) + \log(1/\varepsilon)],$$

这就证明了 (ii) 部分。

(iii) 令 $\mathbb{G}_N = \sqrt{N}(\mathbb{P}_N - P)$。应用 van der Vaart 和 Wellner (1996年) 中的定理 2.14.1，得到

关于 P 可测的类 $\mathcal{F}_h$，其包络函数为 $F_{\mathbf{y},\,h}$ 由此可得一个常数 $C'>0$，使得

$$E\Big[\sup_{f\in\mathcal{F}_h}|\mathbb{G}_N f|\Big]\le C' J(\theta_N,\mathcal{F}_h)\sup_{\mathbf{y}\in\mathcal{Y}_0}\|F_{\mathbf{y},h}\|_{L_2(P)},\qquad \theta_N=\frac{\sup_{f\in\mathcal{F}_h}\|f\|_{L_2(P)}}{\sup_{\mathbf{y}\in\mathcal{Y}_0}\|F_{\mathbf{y},h}\|_{L_2(P)}}.$$

根据第 (i) 部分，$\sup_{\mathbf{y}\in\mathcal{Y}_0}\|F_{\mathbf{y},\,h}\|_{L_2(P)}\le Ch^{-s/2}$；根据与第 (i) 部分相同的局部化论证，$\sup_{f\in\mathcal{F}_h}\|f\|_{L_2(P)}\le Ch^{-s/2}$，因此对所有充分小的 h，$\theta_N\le C$。利用第 (ii) 部分的熵界，熵积分满足 $J(\theta_N,\,\mathcal{F}_h)\le C\sqrt{\log(1/h)}$。因此，

$$E\Big[\sup_{f\in\mathcal{F}_h}|\mathbb{G}_N f|\Big]\le C''h^{-s/2}\sqrt{\log(1/h)}.$$

由此可得

$$\sup_{f\in\mathcal{F}_h}|(\mathbb{P}_N-P)f|=\frac{1}{\sqrt{N}}\sup_{f\in\mathcal{F}_h}|\mathbb{G}_N f|=O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right),$$

这就证明了第 (iii) 部分。应用备注4.2，我们可以将$\log(1/h)$ 替换为$\log\,N$。 □

以下推论由引理4.1得出。

推论 **4.3.** 假设(R1)–(R8) 成立，且 $\Gamma\subset\mathbb{R}^{d_p}$ 是紧致的。那么，当 $N\to\infty$,

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\boldsymbol{\gamma}\in\Gamma}|\tilde{L}_N(\boldsymbol{\gamma};\mathbf{y})-Q(\boldsymbol{\gamma};\mathbf{y})|=O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right),\tag{4.1}$$

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\boldsymbol{\gamma}\in\Gamma}\left\|\frac{1}{Nh^s}\nabla_\gamma L_N(\boldsymbol{\gamma};\mathbf{y})-E\Big[\nabla_\gamma\tilde{L}_N(\boldsymbol{\gamma};\mathbf{y})\Big]\right\|=O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right),\tag{4.2}$$

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\boldsymbol{\gamma}\in\Gamma}\left\|\frac{1}{Nh^s}\nabla^2_\gamma L_N(\boldsymbol{\gamma};\mathbf{y})-E\Big[\nabla^2_\gamma\tilde{L}_N(\boldsymbol{\gamma};\mathbf{y})\Big]\right\|=O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right).\tag{4.3}$$

注：所述速率源于核加权经验过程的一致大数定律。该论证依赖于核 K的紧支集和有界性、条件 (R5) 中的包络条件，以及带宽要求 $Nh^s/\log(1/h)\to\infty$。

证明：回顾第 2节，在$\mathbf{y}$ 处的（核加权）局部对数似然为

$$L_N(\boldsymbol{\gamma};\mathbf{y})=\sum_{i=1}^{N}\ell\Big(\psi^{-1}(\phi_p(Y_i-\mathbf{y})^\top\boldsymbol{\gamma}),\,U_i\Big)\,K\left(\frac{Y_i-\mathbf{y}}{h}\right),$$

并定义 将缩放目标函数及其期望分别记为 $\tilde{L}_N(\boldsymbol{\gamma};\,\mathbf{y})=\frac{1}{Nh^s}L_N(\boldsymbol{\gamma};\,\mathbf{y})$ 和 $Q(\boldsymbol{\gamma};\,\mathbf{y})$ =

$E[\tilde{L}_N(\boldsymbol{\gamma};\mathbf{y})]$。令 $\mathbb{P}_N f = N^{-1}\sum_{i=1}^N f(U_i,Y_i)$ 和 $Pf = E[f(U,\mathbf{Y})]$ 分别表示经验测度和总体测度。对于 $(\mathbf{y}\ \boldsymbol{\gamma}) \in \mathcal{Y}_0 \times \Gamma$，定义

$$f_{\mathbf{y},\gamma}(U,\mathbf{Y}) := h^{-s}\,\ell\Big(\psi^{-1}(\phi_p(\mathbf{Y}-\mathbf{y})^\top\boldsymbol{\gamma}),\,U\Big)\,K\Big(\frac{\mathbf{Y}-\mathbf{y}}{h}\Big).$$

于是 $L_N(\boldsymbol{\gamma};\mathbf{y}) = \mathbb{P}_N f_{\mathbf{y},\gamma}$ 且 $Q(\boldsymbol{\gamma};\mathbf{y}) = Pf_{\mathbf{y},\gamma}$，因此

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\gamma\in\Gamma}\big|\tilde{L}_N(\boldsymbol{\gamma};\mathbf{y}) - Q(\boldsymbol{\gamma};\mathbf{y})\big| = \sup_{f\in\mathcal{F}_h}\big|(\mathbb{P}_N - P)f\big|, \qquad \mathcal{F}_h = \{f_{\mathbf{y},\gamma} : \mathbf{y}\in\mathcal{Y}_0,\ \boldsymbol{\gamma}\in\Gamma\}.$$

因此，一致收敛速率（4.1）可直接由引理4.1（iii）得出。

梯度界。根据链式法则（参见第2节），

$$\nabla_\gamma L_N(\boldsymbol{\gamma};\mathbf{y}) = \sum_{i=1}^N \ell'\Big(\psi^{-1}(\phi_p(Y_i-\mathbf{y})^\top\boldsymbol{\gamma}),\,U_i\Big)\,(\psi^{-1})'(\phi_p(Y_i-\mathbf{y})^\top\boldsymbol{\gamma})\,\phi_p(Y_i-\mathbf{y})\,K\Big(\frac{Y_i-\mathbf{y}}{h}\Big),$$

因此

$$\frac{1}{Nh^s}\,\nabla_\gamma L_N(\boldsymbol{\gamma};\mathbf{y}) = \mathbb{P}_N g_{\mathbf{y},\gamma},$$

其中，$g_{\mathbf{y},\gamma}(U,\mathbf{Y})$ 的每个坐标为

$$g_{\mathbf{y},\gamma}(U,\mathbf{Y}) = h^{-s}\,\ell'\Big(\psi^{-1}(\phi_p(\mathbf{Y}-\mathbf{y})^\top\boldsymbol{\gamma}),\,U\Big)\,(\psi^{-1})'(\phi_p(\mathbf{Y}-\mathbf{y})^\top\boldsymbol{\gamma})\,\phi_p(\mathbf{Y}-\mathbf{y})\,K\Big(\frac{\mathbf{Y}-\mathbf{y}}{h}\Big). \tag{4.4}$$

令 $\mathcal{G}_h := \{g_{\mathbf{y},\gamma} : \mathbf{y}\in\mathcal{Y}_0,\ \boldsymbol{\gamma}\in\Gamma\}$.

根据（R4）–（R6）以及（R5）中的包络条件，存在 $C_0>0$，使得 $|\ell'(\theta,U)|\le C_0$ 在相关范围内成立（特别地，对于下面的 L_2 参数，$E[M_1(U)^2]<\infty$ 即已足够）。根据（R2），$K((\mathbf{Y}-\mathbf{y})/h)\neq 0$ 意味着 $\|\mathbf{Y}-\mathbf{y}\|_\infty \le h$，因此对于某个 $C_K>0$，有 $\|\mathbf{Y}-\mathbf{y}\|\le C_K h$，从而

$$|g_{\mathbf{y},\gamma}(U,\mathbf{Y})| \le C_0 h^{-s}\mathbf{1}\{\|\mathbf{Y}-\mathbf{y}\|\le C_K h\}.$$

定义（全局）包络函数

$$F_h^{(1)}(Y) := C_0 h^{-s}\mathbf{1}\{\mathrm{dist}_\infty(Y,\mathcal{Y}_0)\le C_K h\}.$$

由于$\mathcal{Y}_0$ 是紧致集且内部非空，并且根据（R1），$f_{\mathbf{Y}}$ 在$\mathcal{Y}_0$ 的某个邻域上有界且与零保持正距离，因此存在常数 $0<c_1\le c_2<\infty$，使得对所有充分

小的 h,

$$c_1 \le P(\mathrm{dist}_\infty(Y, \mathcal{Y}_0) \le C_K h) \le c_2,$$

因此

$$C_1 h^{-s} \le \|F_h^{(1)}\|_{L_2(P)} \le C_2 h^{-s}$$

对于某些常数 $0 < C_1 \le C_2 < \infty$。

供后续使用，注意，对每个固定的 $\mathbf{y} \in \mathcal{Y}_0$,

$$\|g_{\mathbf{y},\gamma}\|_{L_2(P)}^2 \le C_0^2 h^{-2s} P(\|\mathbf{Y} - \mathbf{y}\| \le C_K h),$$

并且由（R1）可知，存在常数 $0 < c \le \underline{c} < \infty$，使得 $c\,h^s \le P(\|\mathbf{Y} - \mathbf{y}\| \le C_K h) \le c\,h^s$ 关于 $\mathbf{y} \in \mathcal{Y}_0$ 一致成立。因此，存在 $C_3 > 0$ 使得

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} \sup_{\gamma\in\Gamma} \|g_{\mathbf{y},\gamma}\|_{L_2(P)} \le C_3 h^{-s/2}.$$

对 (4.4) 关于 (y γ) 作一阶泰勒展开，并结合（R4）–（R6），可得对于某些 $L_0 > 0$, $|g_{\mathbf{y}}$, $\gamma(U, \mathrm{Y}) - g_{\mathbf{y}',\gamma'}(U, \mathrm{Y})| \le L_0 h^{-s}(\|\mathrm{Y} - \mathbf{y}'\| + \|\gamma - \gamma'\|)1\{\mathrm{dist}_\infty(Y, \mathrm{Y}_0) \le 2C_K h\}$。由于当 h 足够小时， $P(\mathrm{dist}_\infty(Y, \mathrm{Y}_0) \le 2C_K h)$ 与 0 和 1 均保持正距离，这意味着 $\|g_{\mathbf{y},\gamma} - g_{\mathbf{y}',\gamma'}\|_{L_2(P)}$ $\le Ch^{-s}(\|\mathrm{Y}-\mathbf{y}'\|+\|\gamma-\gamma'\|)$，其中 $C > 0$ 为常数。将此与 $\|F(1)_h\|_{L_2(P)} \ge C_1 h^{-s}$ 结合可知，归一化类关于 (y γ) 满足 Lipschitz 条件，且其常数与 h 无关。

用网格大小为 h 的上确界范数网格覆盖 $\mathcal{Y}_0$ （至多 $C_Y h^{-s}$ 个点），并用欧几里得 δ-网覆盖 Γ （至多 $C_\Gamma \delta^{-d_p}$ 个点）。由 Lipschitz 界可得

$$N(\delta, \mathcal{G}_h, L_2(P)) \le Ch^{-s}\delta^{-d_p}.$$

现在对具有有界包络和多项式熵的经验过程应用极大不等式（例如，van der Vaart 和 Wellner （1996年）中的$\mathcal{G}_h$定理 2.14.1）。根据$\theta_N = \sup_{g\in\mathcal{G}_h} \|g\|_{L_2(P)}/\|F_{(1)h}\|_{L_2(P)} \le (C_3/C_1)h^{s/2}$ 以及上述熵界，我们得到$J(\theta_N, \mathcal{G}_h) \le C\,\theta_N\sqrt{\log(1/h)}$，因此

$$E\Big[\sup_{g\in\mathcal{G}_h} |\mathbb{G}_N g|\Big] \le C\,J(\theta_N, \mathcal{G}_h)\,\|F_h^{(1)}\|_{L_2(P)} \le C' h^{-s/2}\sqrt{\log(1/h)}.$$

因此,

$$\sup_{g\in\mathcal{G}_h}|(\mathbb{P}_N-P)g|=O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

由于 $\frac{1}{Nh^s}\nabla_\gamma L_N(\boldsymbol{\gamma};\ \mathbf{y})-E[\nabla_\gamma\tilde{L}_N(\boldsymbol{\gamma};\ \mathbf{y})]=(\mathbb{P}_N-P)g_{\mathbf{y},\ \gamma}$，取上确界即得（4.2）。

海森界。再求一次导数,

$$\nabla^2_\gamma L_N(\boldsymbol{\gamma};\mathbf{y})=\sum_{i=1}^{N}H(U_i,Y_i;\boldsymbol{\gamma},\mathbf{y})\,K\left(\frac{Y_i-\mathbf{y}}{h}\right),$$

其中 $H(U,Y;\ \boldsymbol{\gamma},\mathrm{y})$ 的每个元素都是涉及 ℓ'', ℓ', $(\psi^{-1})'$, $(\psi^{-1})''$ 和 $\phi_p(\mathrm{Y}\ -\mathrm{y})\phi_p(\mathrm{Y}\ -\mathrm{y})^\top$ 的项的线性组合。利用 (R4)–（R6）以及 (R5) 中针对 ℓ' 和 ℓ'' 的包络函数界，$(Nh^s)^{-1}\nabla^2_\gamma L_N(\boldsymbol{\gamma};\mathrm{y})=\mathbb{P}_N h_{\mathbf{y},\boldsymbol{\gamma}}$ 的每个标量元素 $h_{\mathbf{y},\boldsymbol{\gamma}}(U,\mathrm{Y})$ 均满足 $|h_{\mathbf{y},\boldsymbol{\gamma}}(U,\mathrm{Y})|\le Ch^{-s}\mathbf{1}\{\|\mathrm{Y}\ -\mathrm{y}\|\le C_Kh\}$。令 $\mathrm{H}_h:=\{h_{\mathbf{y},\boldsymbol{\gamma}}:\mathrm{y}\ \in\mathrm{Y0},\ \boldsymbol{\gamma}\in\Gamma\}$。对 G_h 使用与前述相同的覆盖数和极大不等式论证（其中 $\sup_{h\in\mathcal{H}_h}\|h\|_{L_2(P)}\le Ch^{-s/2}$），可得

$$\sup_{h\in\mathcal{H}_h}|(\mathbb{P}_N-P)h|=O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right),$$

这意味着（4.3）。

结合这三个界即可完成证明。在这三种情况下，应用注释 4.2 后，我们可以将 $\log(1/h)$ 替换为 $\log\ N$。□

引理 4.4。假设(R1)–(R7) 成立，且局部多项式阶数 p 为奇数。令$\boldsymbol{\gamma}^*(\mathbf{y})$ 表示总体准则函数 $Q(\cdot;\mathbf{y})$的唯一最大化点，并定义

$$S_N(\mathbf{y})=\frac{1}{Nh^s}\,\nabla_{\boldsymbol{\gamma}}L_N(\boldsymbol{\gamma}^*(\mathbf{y});\mathbf{y}).$$

那么，当 $N\to\infty$ 时,

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\|S_N(\mathbf{y})\|=O_p\left(h^{p+1}+\sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

在注释 *4.2*中所述附加条件下，对数项 $\log(1/h)$ 可替换为 $\log\ N$，当 N 足够大时。

证明。我们将局部得分分解为总体部分（偏差）和经验波动，然后对$\mathbf{y}$上的每一项进行一致界定。

Recall $L_N(\boldsymbol{\gamma};\mathbf{y})=\sum_{i=1}^N\ell\Big(\psi^{-1}(\phi_p(Y_i-y)^\top\boldsymbol{\gamma}),U_i\Big)\,K\left(\frac{Y_i-y}{h}\right)$, and $\tilde{L}_N(\boldsymbol{\gamma};\mathbf{y})=\frac{1}{Nh^s}L_N(\boldsymbol{\gamma};\mathbf{y})$,

$Q(\boldsymbol{\gamma};\ \mathbf{y}) = E[\tilde{L}_N(\boldsymbol{\gamma};\ \mathbf{y})]$. 将总体得分函数在 $\boldsymbol{\gamma}^*(\mathbf{y})$ 定义为

$$s_h(\mathbf{y}) := E[S_N(\mathbf{y})] = E\Big[\frac{1}{Nh^s}\,\nabla_{\boldsymbol{\gamma}} L_N(\boldsymbol{\gamma}^*(\mathbf{y});\mathbf{y})\Big] = E[\nabla_{\boldsymbol{\gamma}}\tilde{L}_N(\boldsymbol{\gamma}^*(\mathbf{y});\mathbf{y})].$$

于是我们可以写成 $S_N(\mathbf{y}) = \underbrace{(S_N(\mathbf{y}) - s_h(\mathbf{y}))}_{\text{经验波动}} + \underbrace{s_h(\mathbf{y})}_{\text{总体项}}$. 因此

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\|S_N(\mathbf{y})\| \ \le\ \sup_{\mathbf{y}\in\mathcal{Y}_0}\|S_N(\mathbf{y}) - s_h(\mathbf{y})\| + \sup_{\mathbf{y}\in\mathcal{Y}_0}\|s_h(\mathbf{y})\|.$$

我们分别对这两项进行界定。

根据定义，

$$S_N(\mathbf{y}) = \frac{1}{Nh^s}\,\nabla_{\boldsymbol{\gamma}} L_N(\boldsymbol{\gamma}^*(\mathbf{y});\mathbf{y}) = \mathbb{P}_N g_{\mathbf{y},\boldsymbol{\gamma}^*(\mathbf{y})},$$

其中，按分量而言，$g_{\mathbf{y},\boldsymbol{\gamma}}(U, Y)$ 如 (4.4) 所给出，$\mathbb{P}_N$ 是基于 $(U_i, Y_i)_{i=1}^N$ 的经验测度。令 $\mathcal{G}_h = \{g_{\mathbf{y},\boldsymbol{\gamma}} : \mathbf{y} \in \mathcal{Y}_0,\ \boldsymbol{\gamma}\in\Gamma\}$。根据引理 4.1 中使用的相同包络函数、Lipschitz 条件和熵论证（见推论 4.3 的证明），类 $\mathcal{G}_h$ 满足

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\boldsymbol{\gamma}\in\Gamma}\|\mathbb{P}_N g_{\mathbf{y},\boldsymbol{\gamma}} - P g_{\mathbf{y},\boldsymbol{\gamma}}\| = O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

特别地，由于 $\{\boldsymbol{\gamma}^*(\mathbf{y}) : \mathbf{y} \in \mathcal{Y}_0\} \subset \Gamma$，我们有

$$\begin{aligned}\sup_{\mathbf{y}\in\mathcal{Y}_0}\|S_N(\mathbf{y}) - s_h(\mathbf{y})\| &= \sup_{\mathbf{y}\in\mathcal{Y}_0}\|\mathbb{P}_N g_{\mathbf{y},\boldsymbol{\gamma}^*(\mathbf{y})} - P g_{\mathbf{y},\boldsymbol{\gamma}^*(\mathbf{y})}\| \\ &\le \sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\boldsymbol{\gamma}\in\Gamma}\|\mathbb{P}_N g_{\mathbf{y},\boldsymbol{\gamma}} - P g_{\mathbf{y},\boldsymbol{\gamma}}\| = O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right). \qquad (4.5)\end{aligned}$$

现在证明 $\sup_{\mathbf{y}\in\mathcal{Y}_0}\|s_h(\mathbf{y})\| = O(h^{p+1})$ 在 $\mathbf{y}$ 上一致成立，其中 p 为奇数。将 $s_h(\mathbf{y})$ 更明确地写为：

$$s_h(\mathbf{y}) = E[\nabla_{\boldsymbol{\gamma}}\tilde{L}_N(\boldsymbol{\gamma}^*(\mathbf{y});\mathbf{y})] = E\Big[h^{-s}\nabla_{\boldsymbol{\gamma}}\ell\Big(\psi^{-1}(\phi_p(Y-\mathbf{y})^\top\boldsymbol{\gamma}^*(\mathbf{y})),\,U\Big)\,K\Big(\frac{Y-\mathbf{y}}{h}\Big)\Big].$$

以 $\mathbf{Y}$ 为条件并使用链式法则，被积函数的每个坐标都是由（R4）–（R6）限定的线性组合，其项的形式为

$$h^{-s}E\Big[\ell'\Big(\theta^*_{\mathbf{y}}(Y), U\Big)\mid Y\Big]\,(\psi^{-1})'(\phi_p(\mathbf{Y}-\mathbf{y})^\top\boldsymbol{\gamma}^*(\mathbf{y}))\phi_p(\mathbf{Y}-\mathbf{y})K\Big(\frac{\mathbf{Y}-\mathbf{y}}{h}\Big),$$

其中 $\theta^*_{\mathbf{y}}(Y) = \psi^{-1}(\phi_p(\mathbf{Y}-\mathbf{y})^\top \boldsymbol{\gamma}^*(\mathbf{y}))$. 令 $\theta(Y) = \psi^{-1}(\nu(Y))$ 表示在 Y 处求值的真实校准，因此 $U \mid Y$具有 Copula参数 $\theta(Y)$，特别地，$E\big[\ell'(\theta(Y)\ U)\ |Y\big] = 0$，因为给定 y 的条件 Copula 为 $C(\cdot \mid \theta(\mathbf{y}))$。因此我们可以写作$E\big[\ell'(\theta^*_{\mathbf{y}}(Y)\ U)\mid Y\big] = E\big[\ell'(\theta^*_{\mathbf{y}}(Y)\ U) - \ell'(\theta(Y)\ U) \mid Y\big]$.

根据（R5）和（R6），ℓ' 在 θ 中可微，且导数一致有界；ψ^{-1} 在相关范围内的导数也一致有界。因此，

$$\big|E\big[\ell'(\theta^*_{\mathbf{y}}(Y), U) \mid Y\big]\big| \ \le\ C\,|\theta^*_{\mathbf{y}}(Y) - \theta(Y)|$$

对于某个常数 $C < \infty$。此外，根据 $\boldsymbol{\gamma}^*(\mathbf{y})$ 的定义以及（R4），$\phi_p(\cdot)$ 是阶为p的局部多项式基，而 ν 是$(p+2)$次连续可微。随后，利用标准的局部多项式近似论证（例如参见Fan and Gijbels (1996年)），可得一致展开式

$$\nu(Y) = \phi_p(\mathbf{Y}-\mathbf{y})^\top \boldsymbol{\gamma}^*(\mathbf{y}) + R_{p+1}(\mathbf{Y}, \mathbf{y}),$$

其余项满足

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\ \sup_{\|\mathbf{Y}-\mathbf{y}\|_\infty \le Ch} \big|R_{p+1}(Y,\mathbf{y})\big| \le C'\|\mathbf{Y}-\mathbf{y}\|^{p+1}$$

对于某个 C'，$C < \infty$。由于 ψ^{-1} 具有有界的一阶导数（R6），这意味着

$$|\theta^*_{\mathbf{y}}(Y) - \theta(Y)| = |\psi^{-1}(\phi_p(\mathbf{Y}-\mathbf{y})^\top \boldsymbol{\gamma}^*(\mathbf{y})) - \psi^{-1}(\nu(Y))| \ \le\ C''\|\mathbf{Y}-\mathbf{y}\|^{p+1}$$

对于 $\|\mathbf{Y}-\mathbf{y}\|_\infty \le Ch$ 以及某个 $C'' < \infty$。综合这些界，并利用 $(\psi^{-1})'$、$\phi_p(\cdot)$ 和 K 在局部区域 $\|\mathbf{Y}-\mathbf{y}\|_\infty \le Ch$上的有界性，我们得到（对每个坐标）

$$\|s_h(\mathbf{y})\| \ \le\ C_1 E\Big[h^{-s}\|\mathbf{Y}-\mathbf{y}\|^{p+1}\mathbf{1}\{\|\mathbf{Y}-\mathbf{y}\|_\infty \le Ch\}\Big].$$

利用（R1）和变量替换 $v = (\mathbf{Y}-\mathbf{y})/h$，可得

$$\begin{aligned}
\|s_h(\mathbf{y})\| &\le C_2 \int_{\|z-y\|_\infty \le Ch} h^{-s}\,\|z-y\|^{p+1}\, f_{\mathbf{Y}}(z)\,dz \\
&= C_2 \int_{\|v\|_\infty \le C} h^{-s}\,\|hv\|^{p+1}\, f_{\mathbf{Y}}(y+hv)\,h^s\,dv \\
&= C_3 h^{p+1} \int_{\|v\|_\infty \le C} \|v\|^{p+1}\,dv \ \le\ C_4 h^{p+1}.
\end{aligned}$$

其中常数 C_j 不依赖于$\mathbf{y}$ （这是由 $f_{\mathbf{Y}}$ 在$\mathcal{Y}_0$ 上的有界性和连续性，以及（R4）–（R6）中的导数的有界性和连续性得出的）。由于$\mathcal{Y}_0$ 是紧致集，这个界关于$\mathbf{y}$是一致的，即

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} \|s_h(\mathbf{y})\| \;\le\; C_4 h^{p+1}. \tag{4.6}$$

假设 p 为奇数且 K 对称，可确保积分后所有阶数不超过 p 的多项式项都消失（因为 K 的相应奇数阶矩为零），因此第一个非零贡献的阶数为 h^{p+1}，与上文相同。

结合界（4.5）和（4.6），可得

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} \|S_N(\mathbf{y})\| \;\le\; \sup_{\mathbf{y}\in\mathcal{Y}_0} \|S_N(\mathbf{y}) - s_h(\mathbf{y})\| + \sup_{\mathbf{y}\in\mathcal{Y}_0} \|s_h(\mathbf{y})\| = O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right) + O(h^{p+1}),$$

由此得证。 □

引理 4.5。 设$\hat{}\boldsymbol{\gamma}(\mathbf{y})$ 是 $L_N(\boldsymbol{\gamma};\mathbf{y})$ 在紧集 Γ上的一个局部极大值点，并令 $\boldsymbol{\gamma}^*(\mathbf{y})$表示由(R7)保证存在的总体准则函数 $Q(\cdot;\mathbf{y})$ 的唯一最大化点。假设(R1)–(R7) 成立，且多项式阶数 p 为奇数。则

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} \|\hat{\boldsymbol{\gamma}}(\mathbf{y}) - \boldsymbol{\gamma}^*(\mathbf{y})\| = O_p\left(h^{p+1} + \sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

在注 *4.2* 中的附加条件下，对数项$\log(1/h)$ 对于所有充分大的 N，可以替换为$\log\, N$。

证明。 证明基于在 $\boldsymbol{\gamma}^*(\mathbf{y})$处对经验得分作泰勒展开，同时利用得分函数的一致随机展开以及海森矩阵的一致控制。

对于每个 $\mathbf{y} \in \mathbf{Y}_0$,， $L_N(\cdot;\mathbf{y})$ 的局部极大值点 $\hat{}\boldsymbol{\gamma}(\mathbf{y})$ 满足得分方程 $\nabla_\gamma L_N(\hat{}\boldsymbol{\gamma}(\mathbf{y});\mathbf{y}) = \mathbf{0}$。回顾引理 4.4 中在 $\gamma^*(\mathbf{y})$ 处的缩放得分函数 $S_N(\mathbf{y}) = \frac{1}{Nh^s}\nabla_\gamma L_N(\gamma^*(\mathbf{y});\mathbf{y})$，并将缩放经验海森矩阵定义为 $H_N(\gamma;\mathbf{y}) := \frac{1}{Nh^s}\nabla^2_\gamma L_N(\gamma;\mathbf{y})$ $H(\gamma;\mathbf{y}) := E[H_N(\gamma;\mathbf{y})] = E[\nabla^2_\gamma \tilde{L}_N(\gamma;\mathbf{y})]$.固定 $\mathbf{y} \in \mathbf{Y}_0$，并在 $\gamma^*(\mathbf{y})$ 处对 $\nabla_\gamma L_N(\gamma;\mathbf{y})$（等价地，对其缩放版本）作中值展开：

$$0 = \frac{1}{Nh^s}\nabla_\gamma L_N(\hat{\gamma}(\mathbf{y});\mathbf{y}) = \frac{1}{Nh^s}\nabla_\gamma L_N(\gamma^*(\mathbf{y});\mathbf{y}) + \left[\int_0^1 H_N(\gamma^*(\mathbf{y}) + t(\hat{\gamma}(\mathbf{y}) - \gamma^*(\mathbf{y}));\mathbf{y})\, dt\right](\hat{\gamma}(\mathbf{y}) - \gamma^*(\mathbf{y})).$$

换言之，$0 = S_N(\mathbf{y}) + H_N^*(\mathbf{y})(\hat{}\gamma(\mathbf{y}) - \gamma^*(\mathbf{y}))$，其中，我们将连接 $\gamma^*(\mathbf{y})$ 与 $\hat{}\gamma(\mathbf{y})$ 的线段上的平均海森矩阵定义为

$$H_N^*(\mathbf{y}) := \int_0^1 H_N(\gamma^*(\mathbf{y}) + t(\hat{\gamma}(\mathbf{y}) - \gamma^*(\mathbf{y})); \mathbf{y})\, dt.$$

重新排列可得

$$\hat{\gamma}(\mathbf{y}) - \gamma^*(\mathbf{y}) = -\left[H_N^*(\mathbf{y})\right]^{-1} S_N(\mathbf{y}), \tag{4.7}$$

前提是 $H_N^*(\mathbf{y})$ 可逆。

因此，要使 $\|\hat{}\gamma(\mathbf{y}) - \gamma^*(\mathbf{y})\|$ 关于 $\mathbf{y}$ 一致有界，只需得到 $\|S_N(\mathbf{y})\|$ 的一致界，并得到 $-H_N^*(\mathbf{y})$ 的特征值一致下界。

根据引理 4.4，在满足 (R1)–(R7) 且 p为奇数时，

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} \|S_N(\mathbf{y})\| = O_p\left(h^{p+1} + \sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

为了一致地控制海森矩阵，推论4.3 给出

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\gamma\in\Gamma} \|H_N(\gamma;\mathbf{y}) - H(\gamma;\mathbf{y})\| = O_p\left(\sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

另一方面，假设 (R7) 意味着总体海森矩阵一致负定：

$$v^\top[-H(\gamma;\mathbf{y})]\, v \;\geq\; c_H\|v\|^2, \qquad \text{for all } v\in\mathbb{R}^{d_p},\ (\gamma,\mathbf{y})\in\Gamma\times\mathcal{Y}_0.$$

由于在（R3）下 $\sqrt{\log(1/h)/(Nh^s)} \to 0$，因此可得

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\gamma\in\Gamma} \|H_N(\gamma;\mathbf{y}) - H(\gamma;\mathbf{y})\| = o_p(1).$$

因此，对于任意 $\eta\in(0,\ c_H)$，存在 N_0，使得对所有 $N\geq N_0$，

$$\mathbb{P}\left(\sup_{\mathbf{y}\in\mathcal{Y}_0}\sup_{\gamma\in\Gamma} \|H_N(\gamma;\mathbf{y}) - H(\gamma;\mathbf{y})\| \leq \eta\right) \to 1.$$

在此事件下，韦伊尔不等式Bhatia （1997）（或一个初等扰动界）可推出

that

$$v^\top[-H_N(\gamma;\mathbf{y})]\,v \;\geq\; (c_H-\eta)\,\|v\|^2, \qquad \forall v,\ (\gamma,\mathbf{y})\in\Gamma\times\mathcal{Y}_0.$$

选择 $\eta = c_H/2$，我们以趋于 1 的概率得到，

$$v^\top[-H_N(\gamma;\mathbf{y})]\,v \;\geq\; \frac{c_H}{2}\,\|v\|^2, \qquad \forall v,\ (\gamma,\mathbf{y})\in\Gamma\times\mathcal{Y}_0.$$

由于 $\gamma^*(\mathbf{y})$ 和 $\hat\gamma(\mathbf{y})$ 都属于 Γ，因此线段 $\gamma^*(\mathbf{y}) + t(\hat\gamma(\mathbf{y}) - \gamma^*(\mathbf{y}))$ 上的每一点也都属于 $t\in[0,1]$（因为 Γ 是凸集；如有需要，可以将 Γ 取为原紧集的闭凸包，而不影响 (R7)）。因此，在同一个高概率事件上，

$$v^\top[-H_N^*(\mathbf{y})]\,v = \int_0^1 v^\top[-H_N(\gamma^*(\mathbf{y}) + t(\hat\gamma(\mathbf{y})-\gamma^*(\mathbf{y})),\mathbf{y})]\,v\,dt \;\geq\; \frac{c_H}{2}\,\|v\|^2, \qquad \forall v,\ \mathbf{y}\in\mathcal{Y}_0.$$

这说明，对于所有 $\mathbf{y}\in\mathcal{Y}_0$，$H_N^*(\mathbf{y})$ 都是可逆的，并且

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\left\|\left[H_N^*(\mathbf{y})\right]^{-1}\right\| \;\leq\; \frac{2}{c_H} \quad \text{with probability tending to one.}$$

回到展开式（4.7），在 $(\mathbf{y})$ 对所有 $\mathbf{y}$均可逆的事件上，我们有$\hat\gamma(\mathbf{y}) - \gamma^*(\mathbf{y}) = -[H_N^*(\mathbf{y})]^{-1}S_N(\mathbf{y})$，因此

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\|\hat\gamma(\mathbf{y})-\gamma^*(\mathbf{y})\| \;\leq\; \sup_{\mathbf{y}\in\mathcal{Y}_0}\left\|\left[H_N^*(\mathbf{y})\right]^{-1}\right\|\cdot\sup_{\mathbf{y}\in\mathcal{Y}_0}\|S_N(\mathbf{y})\| \;\leq\; \frac{2}{c_H}\sup_{\mathbf{y}\in\mathcal{Y}_0}\|S_N(\mathbf{y})\|.$$

利用引理 4.4中的界，我们最终得到

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\|\hat\gamma(\mathbf{y})-\gamma^*(\mathbf{y})\| = O_p\left(h^{p+1}+\sqrt{\frac{\log(1/h)}{Nh^s}}\right),$$

如所要证明的。 □

定理 4.6. 假设 *(R1)–(R7)* 在 p 为奇数时成立。令 $\hat\nu(\mathbf{y}) = e_0^\top\hat\gamma(\mathbf{y})$ $\nu(\mathbf{y}) = e_0^\top\gamma^*(\mathbf{y})$，并定义 $\theta(\mathbf{y}) = \psi^{-1}(\nu(\mathbf{y}))$ 和 $\hat\theta(\mathbf{y}) = \psi^{-1}(\hat\nu(\mathbf{y}))$。那么，当 $N\to\infty$ 时，

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}|\hat\nu(\mathbf{y})-\nu(\mathbf{y})| = O_p\left(h^{p+1}+\sqrt{\frac{\log(1/h)}{Nh^s}}\right), \tag{4.8}$$

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} |\hat{\theta}(\mathbf{y}) - \theta(\mathbf{y})| = O_p\left(h^{p+1} + \sqrt{\frac{\log(1/h)}{Nh^s}}\right). \tag{4.9}$$

特别地，在 （R3） 条件下，ˆ$\nu(\cdot)$ 和 ˆ$\theta(\cdot)$ 在 $\mathcal{Y}_0$ 上一致相合。应用备注 *4.2*，当 $\log(1/h)$ 足够大时，可将其替换为 $\log N$。 N。

证明。ˆ$\nu(\cdot)$ 的结果直接由引理 4.5 中 ˆ$\gamma(\cdot)$ 的一致收敛速率得出。回顾 $\nu(\mathbf{y}) = e_0^\top \gamma^*(\mathbf{y})$ 且 ˆ$\nu(\mathbf{y})$ $= e_0^\top$ˆ$\gamma(\mathbf{y})$，其中 $e_0 = (1,0,\ldots,0)^\top \in \mathbb{R}^{d_p}$，因此

$$\hat{\nu}(\mathbf{y}) - \nu(\mathbf{y}) = e_0^\top (\hat{\gamma}(\mathbf{y}) - \gamma^*(\mathbf{y})).$$

因此，对每个 $\mathbf{y}$ $\in\mathcal{Y}_0$,

$$|\hat{\nu}(\mathbf{y}) - \nu(\mathbf{y})| \le \|e_0\| \, \|\hat{\gamma}(\mathbf{y}) - \gamma^*(\mathbf{y})\| = \|\hat{\gamma}(\mathbf{y}) - \gamma^*(\mathbf{y})\|,$$

并对 $\mathbf{y}$ 取上确界，得到

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} |\hat{\nu}(\mathbf{y}) - \nu(\mathbf{y})| \le \sup_{\mathbf{y}\in\mathcal{Y}_0} \|\hat{\gamma}(\mathbf{y}) - \gamma^*(\mathbf{y})\|.$$

因此，由引理 4.5,

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} |\hat{\nu}(\mathbf{y}) - \nu(\mathbf{y})| = O_p\left(h^{p+1} + \sqrt{\frac{\log(1/h)}{Nh^s}}\right),$$

这就证明了（4.8）。

现在转而讨论 ˆ$\theta(\cdot)$。根据定义,

$$\theta(\mathbf{y}) = \psi^{-1}(\nu(\mathbf{y})), \qquad \hat{\theta}(\mathbf{y}) = \psi^{-1}(\hat{\nu}(\mathbf{y})).$$

对于每个固定的 $y \in\mathcal{Y}_0$, ，由中值定理可得

$$\hat{\theta}(\mathbf{y}) - \theta(\mathbf{y}) = (\psi^{-1})'(\xi_N(\mathbf{y})) \, (\hat{\nu}(\mathbf{y}) - \nu(\mathbf{y})),$$

存在一个介于 ˆ $\xi_N(\mathbf{y})$ 和 $\nu(\mathbf{y})$ 之间的中间点 $\xi_N(\mathbf{y})$。

由 (R 6)，ψ^{-1} 连续可微，且其导数 $(\psi^{-1})\prime$ 一致有界

n

紧致　　紧致集 $\{\nu(\mathbf{y})\colon \mathbf{y}\in\mathcal{Y}_0\}$。此外，根据（4.8）以及 $\mathcal{Y}_0$ 是紧致的这一事实，我们有　　e

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\left|\hat{\nu}(\mathbf{y})-\nu(\mathbf{y})\right| \xrightarrow{p} 0,$$

因此，对于所有充分大的 N，随机值 $\hat{\nu}(\mathbf{y})$ 以趋于 1 的概率落在 $\{\nu(\mathbf{y}):\mathbf{y}\in\mathcal{Y}_0\}$ 的某个固定紧邻域内。由 $(\psi^{-1})'$ 的连续性，存在一个常数$C_\psi<\infty$，使得

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\left|(\psi^{-1})'(\xi_N(\mathbf{y}))\right|\le C_\psi \quad \text{with probability tending to one.}$$

因此，

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\left|\hat{\theta}(\mathbf{y})-\theta(\mathbf{y})\right|\le C_\psi \sup_{\mathbf{y}\in\mathcal{Y}_0}\left|\hat{\nu}(\mathbf{y})-\nu(\mathbf{y})\right| = O_p\left(h^{p+1}+\sqrt{\frac{\log(1/h)}{Nh^s}}\right),$$

由此得证（4.9）。

最后，由（R3）可知 $h_N\to 0$ 且 $Nh_N^s/\log(1/h_N)\to\infty$，因此 h^{p+1} 和 $\sqrt{\log(1/h)/(Nh^s)}$ 均收敛于零。因此，（4.8）和（4.9）右端均收敛于零，且 $\hat{\nu}(\cdot)$ 和 $\hat{\theta}(\cdot)$ 在 $\mathcal{Y}_0$ 上一致相合。

□

推论 *4.7.* 令 $\tau(\theta)$ 表示与所选 *Copula* 族相关联的 *Kendall's* τ，并定义 $\tau(\mathbf{y})=\tau(\theta(\mathbf{y}))$ 和 $\hat{\tau}(\mathbf{y})=\tau(\hat{\theta}(\mathbf{y}))$。在定理 *4.6* 的假设下，进一步假设 $\tau(\cdot)$ 在 Θ_0 上连续可微。那么，当 $N\to\infty$ 时，

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\left|\hat{\tau}(\mathbf{y})-\tau(\mathbf{y})\right| = O_p\left(h^{p+1}+\sqrt{\frac{\log(1/h)}{Nh^s}}\right),$$

特别是$\hat{\tau}(\cdot)$ 在$\mathcal{Y}_0$上一致相合。应用注*4.2*，对于充分大的$\log(1/h)$，我们可以将其替换为$\log N$。
N

证明。根据定理4.6，在（R1）–（R7）成立且 p 为奇数时，有

$$\sup_{\mathbf{y}\in\mathcal{Y}_0}\left|\hat{\theta}(\mathbf{y})-\theta(\mathbf{y})\right| = O_p\left(h^{p+1}+\sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

对于每个固定的$\mathbf{y}\in\mathcal{Y}_0$,，将中值定理应用于一维映射 $\tau(\cdot)$ 可得

$$\hat{\tau}(\mathbf{y})-\tau(\mathbf{y})=\tau(\hat{\theta}(\mathbf{y}))-\tau(\theta(\mathbf{y}))=\tau'(\xi_N(\mathbf{y}))\,(\hat{\theta}(\mathbf{y})-\theta(\mathbf{y})),$$

对于某个介于 $\hat{\theta}(\mathbf{y})$ 与 $\theta(\mathbf{y})$ 之间的中间点 $\xi_N(\mathbf{y})$。

根据假设，$\tau(\cdot)$ 在 Θ_0上连续可微。此外，根据定理 4.6 以及 $\mathcal{Y}_0$,的紧致性，我们有

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} |\hat{\theta}(\mathbf{y}) - \theta(\mathbf{y})| \xrightarrow{p} 0,$$

因此，对于所有充分大的 N，随机值 $\hat{\theta}(\mathbf{y})$ 以趋于 1 的概率落在 $\{\theta(\mathbf{y}) : \mathbf{y} \in \mathcal{Y}_0\} \subset \Theta_0$的任意固定紧邻域内。由 τ' 在 Θ_0上的连续性，存在一个有限常数 C_τ，使得

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} |\tau'(\xi_N(\mathbf{y}))| \le C_\tau \quad \text{with probability tending to one.}$$

因此,

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} |\hat{\tau}(\mathbf{y}) - \tau(\mathbf{y})| \le \sup_{\mathbf{y}\in\mathcal{Y}_0} |\tau'(\xi_N(\mathbf{y}))| \sup_{\mathbf{y}\in\mathcal{Y}_0} |\hat{\theta}(\mathbf{y}) - \theta(\mathbf{y})| \le C_\tau \sup_{\mathbf{y}\in\mathcal{Y}_0} |\hat{\theta}(\mathbf{y}) - \theta(\mathbf{y})|,$$

从而

$$\sup_{\mathbf{y}\in\mathcal{Y}_0} |\hat{\tau}(\mathbf{y}) - \tau(\mathbf{y})| = O_p\left(h^{p+1} + \sqrt{\frac{\log(1/h)}{Nh^s}}\right).$$

在带宽条件（R3）下，h^{p+1} 和 $\sqrt{\log(1/h)/(Nh^s)}$均趋于零，因此右侧趋于零，且 $\hat{\tau}(\cdot)$ 在 $\mathcal{Y}_0$上一致相合。 □

5 数值分析

我们研究所提出的局部似然估计量的有限样本表现，并评估第 4 节中推导的一致收敛速率在经验上的相关性。特别强调评估实践中涉及$\log(1/h)$ 的随机项是否能够被$\log N$ 准确近似。

数据生成机制。协变量 $\mathbf{Y}$ 服从截断正态分布 TN(0，4; $[-2, 2]$)。给定 $Y = y(s = 1)$ 后，数对 (U_1, U_2) 服从参数为 $\theta(y) = \psi^{-1}(\nu(y))$ 的 Frank Copula，其中逆链接函数为 $\psi^{-1}(\nu) = 1 + \frac{4}{1+e^{-\nu}}$，因此 $\theta(y) \in (1, 5)$。采用 $p = 1$ 和 Epanechnikov核进行局部线性似然平滑估计。使用 BFGS算法最大化局部对数似然。通过最大化最小二乘交叉验证来选择带宽 h

（3.1），且不施加上界。

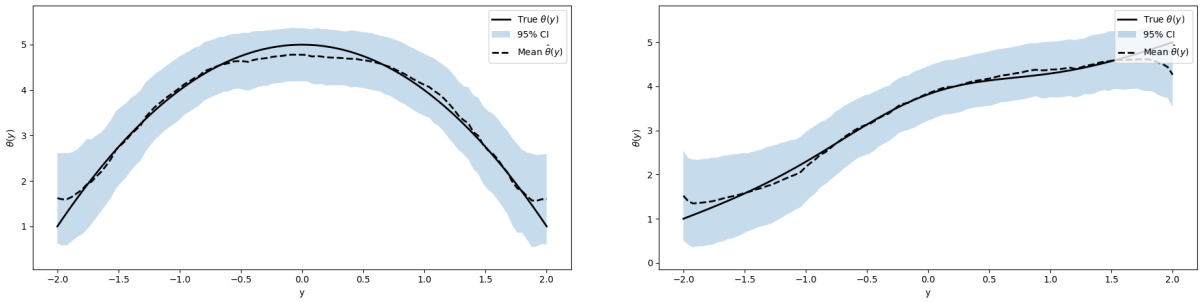
模型。考虑了两个校准函数：

$$\theta_{\mathrm{M1}}(y) = 5 - y^2, \tag{5.1}$$

$$\theta_{\mathrm{M2}}(y) = 3 + y + \frac{5}{6}\left(e^{-y^2} - e^{-4}\right). \tag{5.2}$$

模型 M1 呈现出显著的曲率和较大的空间变化，而模型 M2 则代表一种更平滑的非线性趋势，局部变化也较为温和。这两种设计使我们能够考察一致收敛速率近似在不同函数形状下的稳健性。

函数估计示例。图 1 比较了代表性样本量下的真实函数 $\theta(\cdot)$与估计量 $\widehat{\theta}(\cdot)$的蒙特卡洛均值。图中还展示了逐点 95% 蒙特卡洛置信带。结果基于样本量为 $N = 500$ 且进行了 100 次蒙特卡洛重复试验的样本。在两个模型中，估计量都紧密跟随 $\theta(\cdot)$的整体形状。正如预期，在协变量域的边界附近，变异性会增大，因为有效局部样本量较小，核平滑也因此变得不太稳定。计算下文报告的一致误差时，不包括这些边界区域。关于式 (2.3) 中 $\theta(\cdot)$和 $\widehat{\gamma}(\cdot)$的偏差、方差及渐近正态性的更多细节见附录。



(a) 模型 M1（Frank Copula）。　　(b) 模型 M2（Frank Copula）。

图 1：真实参数曲线 $\theta(y)$（实线）、$\widehat{\theta}(y)$的蒙特卡洛均值（虚线），以及跨重复试验的逐点95% 蒙特卡洛置信带（阴影部分）；结果展示了一个有代表性的样本量。

性能指标和速率代理量。对于每次蒙特卡洛重复实验，性能通过一致估计误差 $\sup_{y\in\mathcal{Y}_0}|\widehat{\theta}(y) - \theta(y)|$，$\mathcal{Y}_0 = [-2+\delta, 2-\delta]$，$\delta = 0.2$，进行总结，该误差可降低边界效应。为将经验性能与理论结果进行比较，我们还使用通过交叉验证选定的相同带宽 h 计算速率代理量 $h^{p+1} + \sqrt{\log N / Nh^s}$，$p = s = 1$，

。如第 4 节所述，当 $\log N$ 替代 $\log(1/h)$ 时，该量刻画了主要的一致随机波动；参见注释 4.2。

针对样本量 $N \in \{100,200,500,750,1000,1500,2000\}$进行了模拟，每种配置重复试验 $R = 100$ 次。为简洁起见并便于直接比较，将两个模型的结果一并报告。

主要发现。 表 1 报告了一致估计误差（上确界）及相应速率代理量（速率）的经验均值、标准差和中位数。对于两个模型，随着样本量增加，一致误差逐渐减小，速率代理量也呈现出高度一致的下降趋势。经验误差与理论速率在中等和大样本量下尤其吻合，这支持了用 $\log(1/h)$ 替换为 $\log\ N$的实际做法，即使交叉验证有时会为较小样本选择相对较大的带宽。

模型	N	上确界$_{\text{mean}}$	上确界$_{\text{sd}}$	上确界$_{\text{median}}$	速率$_{\text{mean}}$	速率$_{\text{sd}}$	速率$_{\text{median}}$
M1	100	2.4138	0.7896	2.5546	1.4981	0.8618	1.3600
	200	1.8559	0.8405	1.9266	1.3065	0.6611	1.5206
	500	0.9964	0.4890	0.9022	0.9802	0.4305	1.0788
	750	0.9342	0.5532	0.7647	0.8096	0.3839	0.8022
	1000	0.7907	0.4390	0.6666	0.7421	0.3385	0.8291
	1500	0.6189	0.3373	0.5364	0.6483	0.2862	0.7090
	2000	0.5752	0.2485	0.5313	0.5990	0.2332	0.6337
M2	100	2.0354	1.0721	1.9338	1.7376	0.8866	2.0050
	200	1.3672	0.8389	1.1071	1.3800	0.6337	1.5371
	500	0.8417	0.5490	0.7144	1.1297	0.4066	1.4144
	750	0.7060	0.5676	0.5560	0.9808	0.3281	1.2102
	1000	0.6776	0.5175	0.5370	0.8176	0.3185	1.0783
	1500	0.5601	0.3242	0.4595	0.7372	0.2664	0.9269
	2000	0.4581	0.1902	0.4285	0.6803	0.2091	0.8277

表1：一致估计误差 上确界$_{\mathbf{y}\in\mathcal{Y}_0}|\widehat{\theta}(\mathbf{y}) - \theta(\mathbf{y})|$ 以及理论速率代理量 $h^2 + \sqrt{\text{对数}\ N/(Nh)}$ 在 $R = 100$ 次重复试验中的表现。

为量化尺度一致性，表2 报告了蒙特卡洛均值之比。比值接近 1 表明代理量捕捉到了正确的阶和尺度。对于两个模型，随着 N，比值趋于稳定，这与理论代理量描述主导一致波动的结论相符。具体而言，即使在较小样本中带宽偶尔可能超过 1（见下文），比值的表现仍与理论相符。

带宽表现（交叉验证）。表3 总结了选定带宽的分布。对于较小和中等的 N，交叉验证可能倾向于全局平滑，并且

模型	N	$\sup_{\text{mean}}/$速率$_{\text{mean}}$	速率$_{\text{mean}}/\sup_{\text{mean}}$
M1	100	1.6112	0.6207
	200	1.4205	0.7040
	500	1.0165	0.9837
	750	1.1539	0.8666
	1000	1.0655	0.9386
	1500	0.9546	1.0475
	2000	0.9602	1.0414
M2	100	1.1724	0.8533
	200	0.9901	1.0100
	500	0.7456	1.3410
	750	0.7204	1.3897
	1000	0.8290	1.2067
	1500	0.7606	1.3162
	2000	0.6748	1.4851

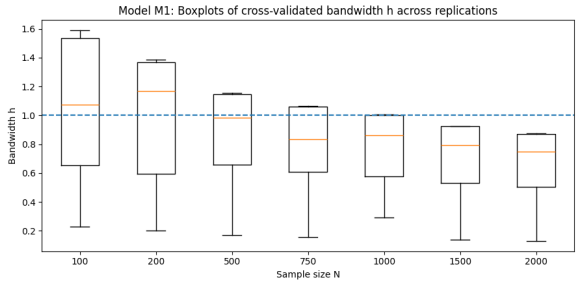
表2： 一致估计误差的蒙特卡洛均值与理论速率之比。

偶尔会得到 $h > 1$。随着 N 增大，分布向下移动并趋于集中，这与 $h \to 0$ 依概率成立一致。重要的是，出现一些 $h > 1$ 值并不与论文中使用的理论随机项相矛盾：根据引理4.1 和注4.2中的熵控制，一致界以$\log N$ 而非$\log(1/h)$表示，因此该模拟设计与所评估的理论直接相关。

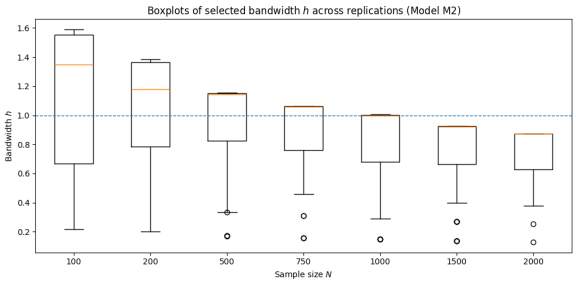
模型	N	h_{mean}	h_{sd}	h_{median}	$h_{\min}$	$h_{\max}$
M1	100	1.0145	0.4780	1.0737	0.2272	1.5875
	200	0.9791	0.4069	1.1705	0.2029	1.3856
	500	0.8813	0.2809	0.9830	0.1709	1.1534
	750	0.7824	0.2883	0.8331	0.1576	1.0637
	1000	0.7580	0.2586	0.8599	0.2904	1.0047
	1500	0.7100	0.2381	0.7942	0.1388	0.9264
	2000	0.6939	0.2007	0.7500	0.1312	0.8744
M2	100	1.1382	0.4694	1.3499	0.2181	1.5871
	200	1.0312	0.3793	1.1785	0.2030	1.3863
	500	0.9644	0.2802	1.1458	0.1719	1.1530
	750	0.9043	0.2431	1.0588	0.1589	1.0644
	1000	0.8090	0.2529	0.9984	0.1507	1.0053
	1500	0.7702	0.2400	0.9243	0.1396	0.9267
	2000	0.7601	0.1747	0.8731	0.1311	0.8752

表3： h 在 $R = 100$ 次重复试验中的交叉验证带宽汇总。

图 2 并排展示了模型 M1 和 M2 对应的箱线图。随着 N 变化，带宽分布向下移动且更加集中，这一点在两种设计中都清晰可见。



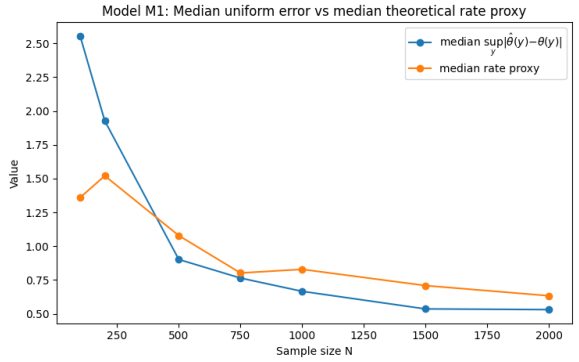
(a) 模型 M1。



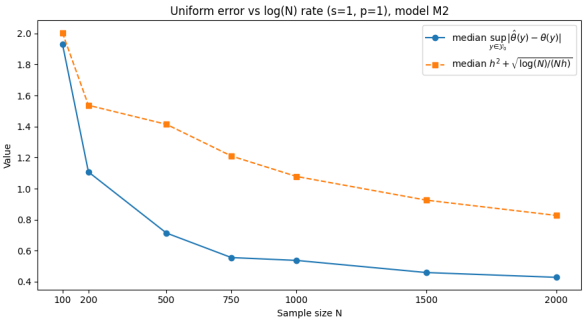
(b) 模型 M2。

图 2： 各样本量下， $R = 100$ 次重复试验中交叉验证带宽 h 的箱线图。虚线水平线标示 $h = 1$。

中位数趋势（一致误差与速率）。 图 3 并排比较了两个模型中一致误差的中位数与速率的中位数随 N变化的情况。在两种情形下，两条曲线都同步变化，这是关键的经验检验：当使用 $\log N$ 时，速率代理量不仅捕捉了一致误差的变化阶数，也捕捉了其实际衰减模式。



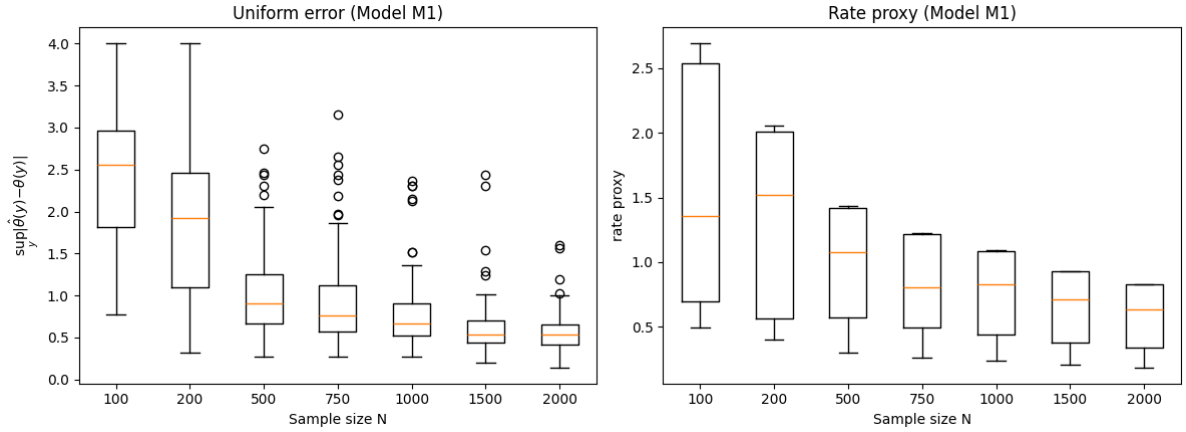
(a) 模型 M1。



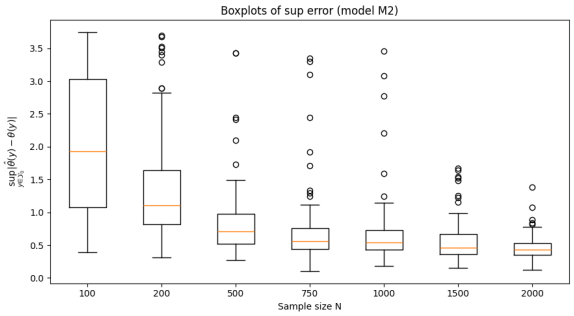
(b) 模型 M2。

图 3： 一致估计误差的中位数和理论速率的中位数 $h^2 + \sqrt{\log N/(Nh)}$ 随 N变化的函数。

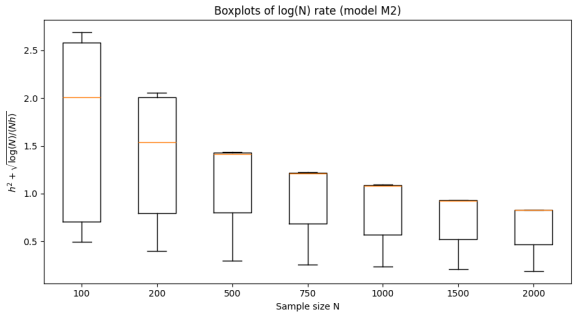
重复试验间的分布比较。最后，图4 报告了重复试验中一致估计误差和理论速率的箱线图，并以配对面板（误差与速率）的形式排列，先展示模型 M1，再展示模型 M2。在两个模型中，随着N 增大，离散程度均有所收缩。速率分布在收缩趋势及其尺度变化方面均反映了误差分布，从而在现实的数据驱动带宽下，为一致速率近似提供了严格的验证。



(a) 模型 M1：一致误差和速率。



(b) 模型 M2：一致误差。



(c) 模型 M2：速率。

图 4： 各样本量下 $h^2 + \sqrt{\log\, N/(Nh)}$一致估计误差和理论速率的箱线图，共进行 $R = 100$ 次重复试验；配对面板分别对应模型 *M1*（上排）和模型 *M2*（下排）。

6 讨论

本文针对多元情形下协变量依赖的 Copula 参数的局部似然估计量，建立了一致性结果。不同于现有文献中相当一部分聚焦于固定协变量取值处的逐点渐近行为，本文的分析对协变量空间的紧致区域提供了一致控制。这一区别具有实质意义，而不仅仅是技术层面的区别。一致收敛对于确保局部优化程序的可靠性、论证在所有协变量取值上同时成立的推断，以及支持依赖整体而非逐点性能的带宽选择策略，都是必不可少的。

从方法论角度来看，主要贡献在于推导出局部对数似然及其前两阶导数的显式一致收敛速率。这些结果表明，局部最大似然估计量一致收敛于其总体目标，并且估计得到的校准函数和 Copula 参数函数继承相同的速率。对海森矩阵的一致控制

发挥着核心作用，因为它以高概率保证了局部极大值点的存在性与稳定性，并确保牛顿–拉夫森算法或拟牛顿算法在整个协变量域内表现良好。这为局部似然方法在条件 Copula 建模中的广泛实证应用提供了理论支持。

理论论证依赖于经验过程技术，用于研究由协变量位置和局部多项式系数联合索引的核加权函数类。一项主要困难来自核局部化：它会产生收缩邻域，并且随着带宽减小，包络函数会增大。结合多项式熵界与协变量空间的几何性质，可以在相对温和的光滑性和矩条件假设下应对这些挑战。所得收敛速率明确取决于协变量的维数，体现了多元平滑中常见的维数灾难。

本文所得结果扩展了关于条件 Copula 模型局部似然估计的早期研究。尽管逐点偏差与方差展开有助于了解局部行为，但要确立依赖性度量（如 Kendall's τ）的全局一致性，并验证基于上确界型误差界的推断程序，一致收敛至关重要。因此，本文建立的一致性理论为在协变量依赖的 copula 模型中构造一致置信带、开展同时假设检验以及开发其他全局推断工具提供了自然基础。

未来研究仍有若干方向有待探索。尽管本文结果为条件 Copula 模型中的数据驱动带宽选择提供了理论依据，但针对交叉验证带宽的完整渐近分析尚未建立。此外，当前框架假设观测值相互独立。将一致性结果推广到依赖情形，包括时间序列和空间数据，将大幅拓宽其适用范围。最后，本文聚焦于单参数 copula 族，但许多应用涉及多参数 copula 或藤 copula 构造，而要实现一致控制，可能需要额外的结构条件。

本文所使用的经验过程技术并不限于 Copula 模型，也可能适用于更广泛的一类带协变量索引参数的核加权 M-估计问题。因此，本文提出的方法和结果或可应用于条件 Copula 模型之外的领域，尤其是结合局部平滑与全局约束的半参数框架。

A 附录

A.1 多元局部似然辅助结果

本附录汇集了多元局部似然估计量的逐点渐近结果，以补充正文中建立的一致收敛理论。此处总结的结果将标准的局部多项式似然论证（Fan and GijbelsFanand Gijbels (1996年）推广至带链接变换参数的条件 Copula 情形，并给出明确的偏差、方差和渐近正态性结论（参见 Acar et al. Acar et al.(2011b）、Muia et al. Muia et al. (2025c）以及 Muia et al. (2025d））。这些结果还阐明了第4节所建立的一致速率中平滑项和随机项的来源。

全文中，令 $\mathbf{Y} \in \mathbb{R}^s$ 为密度为 $f_{\mathbf{Y}}$的协变量向量，并令 $\nu(\mathbf{y})$ 为实值校准函数。Copula参数定义为 $\theta(\mathbf{y}) = \psi^{-1}\{\nu(\mathbf{y})\}$，其中 ψ^{-1} 是严格单调的逆链接函数。

令 $\alpha \in \mathbb{N}^s$ 为多重指标，其中 $|\alpha| = \sum_{j=1}^s \alpha_j$，$\alpha! = \prod_{j=1}^s \alpha_j!$，$\mathbf{v}^\alpha = \prod_{j=1}^s v_j^{\alpha_j}$, $D^\alpha = \underbrace{\partial^{|\alpha|}}_{\partial y_1^{\alpha_1} \cdots \partial y_s^{\alpha_s}}$ · 我们在局部用 $\nu(\mathbf{Y}_i) \approx \boldsymbol{\phi}_p(\mathbf{Y}_i - \mathbf{y})^\top \boldsymbol{\gamma}$，近似 $\nu(\mathbf{Y}_i)$，并通过 $D^\alpha \nu(\mathbf{y}) = \alpha! \, \widehat{\gamma}_\alpha$，$|\alpha| \le p$，估计导数$\widehat{\ }$，其中 $\widehat{\gamma}_\alpha$ 表示与 α 对应的系数，而 $\mathbf{e}_\alpha$ 在 $\boldsymbol{\gamma}$中选取该坐标。

我们考虑使用各向同性带宽矩阵$H = h\mathbf{I}_s$进行 p 阶局部多项式似然拟合。由于 H 是对角矩阵且对角线元素相同，因此沿协变量的各坐标轴进行均匀平滑。

定义核矩矩阵 $S = \int \boldsymbol{\phi}_p(\mathbf{v})\boldsymbol{\phi}_p(\mathbf{v})^\top K(\mathbf{v}) \, d\mathbf{v}$ $S^* = \int \boldsymbol{\phi}_p(\mathbf{v})\boldsymbol{\phi}_p(\mathbf{v})^\top K^2(\mathbf{v}) \, d\mathbf{v}$，以及核矩向量$\mathbf{c}_p = \int \boldsymbol{\phi}_p(\mathbf{v})\mathbf{r}_{p+1}(\mathbf{v})K(\mathbf{v}) \, d\mathbf{v}$ $\tilde{\ }\mathbf{c}_p = \int \boldsymbol{\phi}_p(\mathbf{v})\mathbf{r}_{p+2}(\mathbf{v})K(\mathbf{v}) \, d\mathbf{v}$，其中$\mathbf{r}_{p+1}$ 和$\mathbf{r}_{p+2}$ 分别由总次数为 $p+1$ 和 $p+2$的单项式组成。

令 $\widehat{\boldsymbol{\gamma}}(\mathbf{y})$ 表示在$\mathbf{y}$ 处给出的核加权对数似然的局部极大值点（2.3）。定义条件Fisher曲率 $\sigma^2(\mathbf{y}) = -\mathbb{E}[\ell''(\psi^{-1}(\nu(\mathbf{y})), \mathbf{U}) \mid \mathbf{Y} = \mathbf{y}]$，并假设$\sigma^2(\mathbf{y}) > 0$ 在所关注的邻域内成立。

A.2 渐近偏差与方差

定理 A.1。设 $f_{\mathbf{Y}}(\mathbf{y}) > 0$，并假设 $f_{\mathbf{Y}}$，ν 的所有不超过 $p+2$ 阶的偏导数、$\sigma^2(\cdot)$，以及 $(\psi^{-1})'(\cdot)$ 在 $\mathbf{y}$ 的某个邻域内连续。此外，假设 $\ell'(\theta,\mathbf{U})$ 和 $\ell''(\theta,\mathbf{U})$ 存在，在 θ 上连续，且由可积包络函数控制。令 $h \to 0$ 和 $Nh^s \to \infty$，当 $N \to \infty$ 时。

(i) 若 $p-|\alpha|$ 为奇数，则

$$\operatorname{Bias}\{\widehat{D^{\alpha}\nu}(\mathbf{y}) \mid \mathbb{Y}\} = \alpha!\, h^{p+1-|\alpha|}\mathbf{e}_{\alpha}^{\top} S^{-1}\left(\sum_{|\beta|=p+1} \frac{D^{\beta}\nu(\mathbf{y})}{\beta!}\,\mathbf{c}_{p,\beta}\right) + o_P(h^{p+1-|\alpha|}).$$

(ii) 对任意 $|\alpha| \le p$,

$$\operatorname{Var}\{\widehat{D^{\alpha}\nu}(\mathbf{y}) \mid \mathbb{Y}\} = \frac{(\alpha!)^2}{Nh^{s+2|\alpha|}\,\sigma^2(\mathbf{y})\{(\psi^{-1})'(\nu(\mathbf{y}))\}^2 f_{\mathbf{Y}}(\mathbf{y})}\mathbf{e}_{\alpha}^{\top} S^{-1}S^{*}S^{-1}\mathbf{e}_{\alpha} + o_P\Big(\tfrac{1}{Nh^{s+2|\alpha|}}\Big).$$

此处 S 和 S^{*} 分别为由核函数 K 和基函数 ϕ_p的选择所确定的核矩矩阵和核向量。证明遵循标准的局部似然论证，见 Muia et al. Muia et al. (2025d)。

A.3 校准函数的点估计

推论 A.2. 设 p 为奇数，并定义 $\widehat{\nu}(\mathbf{y}) = \mathbf{e}_0^{\top}\widehat{\boldsymbol{\gamma}}(\mathbf{y})$。则

$$\operatorname{Bias}\{\widehat{\nu}(\mathbf{y}) \mid \mathbb{Y}\} = h^{p+1}\,\mathbf{e}_0^{\top} S^{-1}\left(\sum_{|\beta|=p+1} \frac{D^{\beta}\nu(\mathbf{y})}{\beta!}\,\mathbf{c}_{p,\beta}\right) + o_P(h^{p+1}),$$

$$\operatorname{Var}\{\widehat{\nu}(\mathbf{y}) \mid \mathbb{Y}\} = \frac{1}{Nh^{s}\,\sigma^2(\mathbf{y})\{(\psi^{-1})'(\nu(\mathbf{y}))\}^2 f_{\mathbf{Y}}(\mathbf{y})}\mathbf{e}_0^{\top} S^{-1}S^{*}S^{-1}\mathbf{e}_0 + o_P\Big(\tfrac{1}{Nh^{s}}\Big).$$

A.4 Copula参数估计

推论 A.3. 设 $\widehat{\theta}(\mathbf{y}) = \psi^{-1}\{\widehat{\nu}(\mathbf{y})\}$，并假设 $\psi'(\theta(\mathbf{y})) \neq 0$。则

$$\operatorname{Bias}\{\widehat{\theta}(\mathbf{y}) \mid \mathbb{Y}\} = \frac{1}{\psi'(\theta(\mathbf{y}))}\,h^{p+1}\,\mathbf{e}_0^{\top} S^{-1}\left(\sum_{|\beta|=p+1} \frac{D^{\beta}\nu(\mathbf{y})}{\beta!}\,\mathbf{c}_{p,\beta}\right) + o_P(h^{p+1}),$$

$$\operatorname{Var}\{\widehat{\theta}(\mathbf{y}) \mid \mathbb{Y}\} = \frac{1}{Nh^{s}\,\sigma^2(\mathbf{y}) f_{\mathbf{Y}}(\mathbf{y})}\mathbf{e}_0^{\top} S^{-1}S^{*}S^{-1}\mathbf{e}_0 + o_P\Big(\tfrac{1}{Nh^{s}}\Big).$$

A.5 渐近正态性

定理 A.4 （渐近正态性）。 在定理 *A.1*的假设下，

$$\sqrt{Nh^{s+2|\alpha|}}\left(\widehat{D^{\alpha}\nu}(\mathbf{y}) - \operatorname{Bias}\{\widehat{D^{\alpha}\nu}(\mathbf{y}) \mid \mathbb{Y}\}\right) \xrightarrow{d} \mathcal{N}\left(0,\ \frac{(\alpha!)^2\mathbf{e}_{\alpha}^{\top} S^{-1}S^{*}S^{-1}\mathbf{e}_{\alpha}}{\sigma^2(\mathbf{y})\{(\psi^{-1})'(\nu(\mathbf{y}))\}^2 f_{\mathbf{Y}}(\mathbf{y})}\right).$$

推论 A.5 （ $\widehat{\nu}(\mathbf{y})$ 的渐近正态性）。 设 p 为奇数。则

$$\sqrt{Nh^{s}}\,(\widehat{\nu}(\mathbf{y})-\mathrm{Bias}\{\widehat{\nu}(\mathbf{y})\mid\mathbb{Y}\})\xrightarrow{d}\mathcal{N}\left(0,\,\frac{\mathbf{e}_0^{\top}S^{-1}S^{*}S^{-1}\mathbf{e}_0}{\sigma^2(\mathbf{y})\{(\psi^{-1})'(\nu(\mathbf{y}))\}^2 f_{\mathbf{Y}}(\mathbf{y})}\right).$$

推论 A.6 （ $\widehat{\theta(\mathbf{y})}$的渐近正态性）。

$$\sqrt{Nh^{s}}\left(\widehat{\theta}(\mathbf{y})-\theta(\mathbf{y})-\mathrm{Bias}\{\widehat{\theta}(\mathbf{y})\mid\mathbb{Y}\}\right)\xrightarrow{d}\mathcal{N}\left(0,\,\frac{\mathbf{e}_0^{\top}S^{-1}S^{*}S^{-1}\mathbf{e}_0}{\sigma^2(\mathbf{y})f_{\mathbf{Y}}(\mathbf{y})}\right).$$

A.6 均方误差与最优带宽

$\widehat{D^{\alpha}\nu(\mathbf{y})}$的条件均方误差满足 $\mathrm{MSE}\{\widehat{D^{\alpha}\nu(\mathbf{y})}\}\approx h^{2(p+1-|\alpha|)}+\frac{1}{Nh^{s+2|\alpha|}}$。平衡这些项可得最优带宽速率 $h_{\mathrm{opt}}\propto N^{-1/(2p+2+s)}$，它取决于多项式次数 p和协变量维数 s，但与导数阶数 $|\alpha|$无关。

A.7 与一致理论的联系

上述逐点展开和渐近正态性结果揭示了一致收敛速率背后的局部偏差—方差结构。结合核加权似然类的熵界，这些结果为如下形式的一致误差速率提供了依据： $h^{p+1}+\sqrt{\frac{\log N}{Nh^{s}}}$，即使在采用数据驱动的带宽选择时也是如此。

资助 作者声明本文未获得任何资助。

作者贡献 作者独自负责本研究的所有方面，包括研究构思与设计、数据生成与分析、理论研究、稿件撰写及所有修订。

利益冲突 作者声明不存在任何利益冲突 .

参考文献

F. Abegaz、I. Gijbels 和 N. Veraverbeke，条件 *copula* 的半参数估计，Journal of Multivariate Analysis，110（2012），43–73。

E. F. Acar、R. V. Craiu 和 F. Yao，条件 *copula* 中的依赖性校准，Journal of the American Statistical Association，106(494)(2011)，708–718。

E. F. Acar、R. V. Craiu 和 F. Yao，条件 *copula* 中的依赖性校准：一种非参数方法，Biometrics，67(2)（2011），445–453。R. Bhatia，矩阵分析，施普林格，纽约，1997年。F. Durante 和 C. Sempi，*Copula* 理论原理，CRC出版社，博卡拉顿，佛罗里达州，2016年。J. Fan 和 I. Gijbels，可变带宽与局部线性回归平滑器，统计年鉴，20（1992），2008–2036。J. Fan 和 I. Gijbels，局部多项式建模及其应用，Chapman & Hall，伦敦，1996年。R. A. Fisher，研究人员统计方法，Oliver and Boyd，爱丁堡，1928年。C. Genest、K. Ghoudi 和 L.-P. Rivest，多元分布族中依赖参数的半参数估计程序，Biometrika，82 （ 1995 年），543–552。I. Gijbels 和 J. Mielniczuk，估计 *copula* 函数的密度，统计学通讯——理论与方法，19（1990），445–464。I. Gijbels、N. Veraverbeke 和 M. Omelka，条件 *Copula*、关联度量及其应用，计算统计学与数据分析，55(5)（2011），1919–1932。C. M. Hafner 和 O. Reznikova，半参数动态 *copula* 模型的有效估计，计算统计学与数据分析，54（2010），2609–2627。H. Joe，多元模型与多元依赖概念，CRC出版社，博卡拉顿，佛罗里达州，1997年。G. Kim、M. Silvapulle 和 P. Silvapulle，用于估计 *copula* 的半参数方法与参数方法比较，计算统计学与数据分析，51（2007），2836–2850。A. J. McNeil 和 J. Nešlehová，多元阿基米德 *copula*、 d-单调函数与ℓ_1-范数对称分布，统计年鉴，37（2009），3059–3097。M. Muia、O. Atutey 和 M. Hasan，有界 *copula* 密度函数的核平滑，欧洲统计学期刊，5（2025），7–7。M. N. Muia、O. Atutey 和 M. Hasan，有界 *copula* 密度的核平滑，arXiv 预印本 arXiv:2502.05470，2025年。

M. N. Muia、M. Hasan 和 M.-A. Hamadou，协变量依赖的 *Copula* 参数的局部多项式估计，投稿稿件，2025年。M. N. Muia、M. Hasan 和 M.-A. Hamadou，协变量依赖的 *copula* 的多元局部似然框架，投稿稿件，2025年。R. B. Nelsen，多元关联的非参数度量，载于：固定边缘分布及相关主题，IMS讲义—专著系列，第28卷（1996年），223–232。R. B. Nelsen，*Copula*导论，施普林格，纽约，2006年。M. Omelka、I. Gijbels 和 N. Veraverbeke，改进的 *copula* 核估计：弱收敛与拟合优度检验，《Bernoulli》，15（2009年），1133–1174。A. J. Patton，非对称汇率依赖建模，《国际经济评论》，47(2)（2006年），527–556。M. Rosenblatt，关于多元变换的评注，《数理统计年刊》，23(3)（1952年），470–472。F. Schmid、R. Schmidt、T. Blumentritt、S. Gaißer 和 M. Ruppert，*Copula*理论及其应用，施普林格，柏林，2010年。M. Sklar，多维分布函数及其边缘分布，巴黎大学统计研究所出版物，8（1959年），229–231。P. K. Trivedi 和 D. M. Zimmer，*Copula*建模：面向实践者的导论，《计量经济学基础与趋势》，1(1)（2007年），1–111。H. Tsukahara，*Copula* 模型中的半参数估计，《加拿大统计学杂志》，33（2005年），357–375。A. W. van der Vaart 和 J. A. Wellner，弱收敛与经验过程，施普林格，纽约，1996年。N. Veraverbeke、M. Omelka 和 I. Gijbels，条件 *Copula* 与关联度量的估计，《斯堪的纳维亚统计学杂志》，38（2011年），766–780。M. P. Wand 和 M. C. Jones，核平滑，Chapman & Hall，纽约，1995年。