

误差分布未知下时空模型的自适应非参数估计*

汪红霞¹ 罗学洪¹ 林金官² 唐 星¹

摘要 极大似然估计作为参数估计中较为有效的一种估计方法,在误差分布未知下无法进行,另一方面,时空数据经常含有奇异点或来自重尾分布,此时基于最小二乘的估计方法效果欠佳.考虑时空异质性和相关性,针对误差分布未知的时空模型,本文提出基于核密度估计的自适应非参数估计方法.在较弱的条件下证明了该估计量和已知误差分布下的局部极大似然估计量是渐近等效,比基于最小二乘的局部多项式估计量有效.模拟和实证都验证了该方法对于有限样本的有效性,尤其奇异点的存在,该方法在边界的拟合效果显著优于基于最小二乘的方法.

关键词 时空模型,核密度估计,局部多项式方法,局部极大似然方法

MR (2000) 主题分类 62M30, 62G08

中图法分类 O212.7

文献标志码 A

文章编号 1000-8314(2021)02-0125-22

1 引 言

近年来,时空数据统计分析广泛应用于生物自然科学、经济社会科学等诸多领域中(参见文 [1–2]),在学术界与业界都扮演着重要的角色.时空数据的估计方法也相继产生,如文 [3] 提出局部线性方法以刻画空间数据的异质性,在此基础上,文 [4] 将局部线性方法的应用范围进行拓展,实现了对时空数据的异质性估计.目前对时空数据估计常采用局部多项式估计的方法,如文 [5],该方法不拘泥于线性估计,提高了实际中时空数据的估计精确度.同时,文 [6] 提出了一个对数据进行线性缩放的时空有限混合模型,以缩短估计计算时间.然而这些方法都建立在误差服从高斯正态分布假设之上,在样本数据量不足以支撑中心极限定理成立的情况下,已有文献中估计方法的估计效果欠佳.

针对误差分布未知的一般空间非参数回归问题,可通过自适应方法解决.自适应估计最早由文 [7] 提出,用于处理存在无穷维数的厌恶参数的参数估计和假设检验问题.随后,自适应估计思想逐渐被推广至线性回归和其他模型中(参见文 [8–10]),通过提供一个可替代传统高斯极大似然估计量,即通过用对数似然得分函数的一个非参数估计以代替高斯

本文 2020 年 9 月 7 日收到, 2020 年 11 月 24 日收到修改稿.

¹南京审计大学统计与数学学院, 南京 211815.

E-mail: hxwang@nau.edu.cn; xueh_luo@163.com; xtang_tj@163.com

²通信作者. 南京审计大学统计与数学学院, 南京 211815. E-mail: jglin@nau.edu.cn

*本文受到国家自然科学基金 (No. 11831008, No. 11971235), 国家社会科学基金 (No. 17CTJ016) 和江苏省研究生科研与实践创新计划项目 (No. KYCX19_1526, No. KYCX20_1677) 的资助.

密度函数,进而解决高斯假设不合理的估计问题.文 [11] 在具有局部渐近混合正态性的非平稳模型中考虑自适应估计方法,发现无需对误差项进行具体的分布假设,也能够实现高斯假设下的类似估计结果.文 [12] 将该思想在非参数回归模型中进行尝试,以此提出了一个适用于误差分布未知情形下的非参数回归估计方法,并指出该方法具有自适应性,且渐近方差与已知误差真实分布的极大似然估计一致,文 [13] 在对经过过滤处理的数据进行非参数回归时,再次验证了该方法的渐近方差与极大似然估计一致.以上研究都是针对空间数据或时间序列数据,并未对兼具空间与时间的时空数据进行研究.

基于自适应估计方法有很好的性质,本文考虑如下非参数时空模型:

$$Y_{i,t} = m(X_{i,t}) + R_{i,t}, \quad (1.1)$$

其中 $Y_{i,t} \in \mathbb{R}$, $X_{i,t} = (X_{i,t}(1), \dots, X_{i,t}(d)) \in \mathbb{R}^d$, $m(\cdot)$ 是未知的光滑函数, $R_{i,t}$ 是随机误差,满足 $E(R_{i,t}) = 0$, $\text{Var}(R_{i,t}) = \sigma^2$, $\mathbf{i} \in \Lambda_n = \{\mathbf{1}, \dots, \mathbf{n}_1\} \times \dots \times \{\mathbf{1}, \dots, \mathbf{n}_N\}$, $t \in T_n = \{1, \dots, n_{N+1}\}$, $\mathbf{n} = (\mathbf{n}_1, \dots, \mathbf{n}_{N+1})$, $n^* = n_1 \times \dots \times n_{N+1}$. 上述文献对时空数据估计的良好性质是在假定 $R_{i,t}$ 为高斯分布情形下得到的,然而在许多研究领域中,常出现误差项不服从该假定分布的情形,如经济学、生物学和医学化学.因此,当非参数时空模型的误差项不满足高斯分布时,需要一种拟合效率较好的估计方法能够适用该情形.目前自适应方法未应用到非参数时空模型,故本文主要做出以下改进: (1) 将自适应估计方法推广到非参数时空模型 (1.1) 式,期望同时达到稳健性和有效性.理论、模拟和实例结果显示,自适应非参数估计的渐近方差与极大似然估计方法一致.另外,在误差项呈原点中心对称的非正态分布时,自适应估计与基于最小二乘的局部多项式估计结果相差较小,但自适应估计方法效果仍优于局部多项式;在误差项服从偏态分布或存在异常点的情形时,自适应方法的估计效率显著高于常用的局部多项式估计; (2) 模型更贴近实际数据,同时考虑异质性和相依性,本文考虑非参模型且假设误差服从 α 混合相依; (3) 在较弱的条件下得到估计量的大样本性质,为后续统计推断提供理论依据.本文没有假设 $R_{i,t}$ 关于时间和空间是严平稳的,只是假设在固定空间点 \mathbf{i}_0 , $\{R_{i_0,t}, t \in T_n\}$ 是严平稳的,在固定时间点 t_0 , $\{R_{i,t_0}, \mathbf{i} \in \Lambda_n\}$ 是严平稳的.

本文结构如下: 第二节给出误差分布已知的局部极大似然估计、误差分布未知的非参数估计. 第三节是估计量的理论性质. 第四节是模拟和实例. 第五节是理论证明. 第六节对本文进行简要的总结.

2 理论模型及估计量

由最小二乘局部多项式估计方法 (见 [14]), 若 \mathbf{x} 是 $X_{i,t}$ 邻域的点, 极小化如下最小二乘损失函数:

$$\frac{1}{n^* h_1^d} \sum_{i \in \Lambda_n} \sum_{t \in T_n} (Y_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))^2 \mathcal{K}((\mathbf{X}_{i,t} - \mathbf{x})/\mathbf{h}_1), \quad (2.1)$$

其中 $\mathcal{K}(\cdot)$ 为 d 维核函数, 假定 \mathbf{h}_1 是 d 维自变量的共同带宽, $\mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x})$ 是 $\mathbf{X}_{i,t} - \mathbf{x}$ 的 p 阶多项式, β 为系数向量, 核函数的选取不那么重要, 而带宽的选取将在后面的章节介绍. 若误差 $R_{i,t}$ 服从 Gauss 分布, 则 (2.1) 式同时表现为加权似然损失函数; 若误差项的高斯分布假设不成立, 该估计量的渐近结果仍具有相同性质, 但是较分布已知下的估计量而言, 均方误差显示该估计量缺乏有效性. 极小化 (2.1) 式得最小二乘局部多项式估计量为 $\tilde{\beta}(\mathbf{x})$.

2.1 误差分布已知的局部极大似然估计

在误差的密度函数 f 已知的情况下, 由文 [15–16] 可知, 将 (2.1) 式中的 $(Y_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))^2$ 替换为 $\log f(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))$ 以构造似然估计量, 令 $\bar{\beta}$ 为最优解, 其中第一个分量记为 $\bar{\beta}_0 = \bar{m}(\mathbf{x})$. 由 Cramér-Rao 不等式可得, 局部似然估计量比局部最小二乘有更低的方差上界; 同时, 二者具有同样的偏差, 因此局部似然估计量有更小的 MSE. 若误差项 $R_{i,t}$ 服从正态分布, 则系数的最小二乘与极大似然估计量均为无偏估计量, 且最小二乘局部多项式估计是极大似然估计的特殊情况.

为了构造似然函数, 引用文 [17–18] 的注释, 满足 $\sum_{i=1}^d \alpha_i = \ell$ 条件的 d 维元组 $\alpha = (\alpha_1, \dots, \alpha_d)$ 有 $N_\ell = C_{\ell+d-1} \ell!$ 个, 将该 N_ℓ 个 d 维元组依照先 ℓ 后 α 的字典序排为一列, 同时用 ϕ_ℓ^{-1} 作为标记, 以与原元组一一映射, 例如 $\phi_1(2) = \{(0, 1), (1, 0)\}$, $\phi_1(d) = \{(0, 0, \dots, 1), \dots, (1, 0, \dots, 0)\}$. 一般而言, 估计量 $\bar{m}(\mathbf{x})$ 没有显示表达式, 因此计算 $\bar{\beta}$ 值须采用迭代算法, 如修正 Newton-Raphson 迭代. 假设 $\tilde{\beta}$ 为初值, 则单步 Newton-Raphson 估计量为 $\bar{\beta}_{NR}(\mathbf{x})$,

$$\bar{\beta}_{NR}(\mathbf{x}) = \tilde{\beta}(\mathbf{x}) + \mathcal{I}_{\mathbf{n}^*}(\tilde{\beta}(\mathbf{x}); \mathbf{f})^{-1} \mathbf{S}_{\mathbf{n}^*}(\tilde{\beta}(\mathbf{x}); \mathbf{f}), \quad (2.2)$$

$\mathbf{S}_{\mathbf{n}^*}(\tilde{\beta}(\mathbf{x}); \mathbf{f})$ 是光滑得分函数, 其中

$$\mathbf{S}_{\mathbf{n}^*}(\beta; \mathbf{f}) = -\frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{\mathbf{h}}\right) \frac{\mathbf{f}'(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))}{\mathbf{f}(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))} \tilde{\mathbf{X}}_{i,t}, \quad (2.3)$$

$$\mathcal{I}_{\mathbf{n}^*}(\beta; \mathbf{f}) = \frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{\mathbf{h}}\right) \left[\frac{\mathbf{f}'(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))}{\mathbf{f}(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))} \right]^2 \tilde{\mathbf{X}}_{i,t} \tilde{\mathbf{X}}_{i,t}^T. \quad (2.4)$$

$\tilde{\mathbf{X}}_{i,t} = [\mathbf{X}_{i,t,0}^T(\mathbf{x}), \dots, \mathbf{X}_{i,t,p}^T(\mathbf{x})]^T$, $\mathbf{X}_{i,t,\ell}(x) (0 \leq \ell \leq p)$ 是 N_ℓ 维向量, 且向量的第 r 个元素 $[\mathbf{X}_{i,t,\ell}]_r = (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_\ell(r)}$, $\phi_\ell(r)$ 是映射分量之和等于 ℓ 的第 r 个元组的标记.

注 2.1 Newton-Raphson 迭代, $\mathcal{I}_{\mathbf{n}^*}(\beta; \mathbf{f})$ 中应该含有 $\left[\frac{f'(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))}{f(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))} \right]^2 - \frac{f''(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))}{f(\mathbf{Y}_{i,t} - \mathcal{P}_p(\beta, \mathbf{X}_{i,t} - \mathbf{x}))}$, 我们只保留前面一项是基于以下三个原因: (1) 信息矩阵 $\mathcal{I}_{\mathbf{n}^*}(\tilde{\beta}(\mathbf{x}); \mathbf{f})$ 是 $\mathcal{I}_\beta(\mathbf{x}) = f_X(\mathbf{x})I(f)$ 的一个相合估计, $f_X(\mathbf{x})$ 是联合密度函数, $I(f)$ 是关于误差密度函数的费希尔信息量; (2) 我们常见的对称分布, 如零均值正态分布、 t 分布、Cauchy 分布等的 $\int_{\mathbb{R}} f''(z)dz = 0$; (3) 由理论和模拟结果发现, 只保留前面一项效果更好.

在初值 $\tilde{\beta}$ 给定下, 迭代至 Newton-Raphson 估计量 $\bar{\beta}_{NR}$ 收敛为止. 然而, 由经验分布可知, 随着样本量的逐渐增大, 误差倾向于服从 t 分布直至正态分布, 在迭代过程中, 对于误差较大的样本点, 其 $f(\varepsilon)$ 值接近于 0, 导致迭代无法收敛. 另外, 现实中误差分布的密度函数通常是未知的, 意味着局部似然估计是不可能实现的.

2.2 误差分布未知的非参数估计

在这节里, 本文将推广文 [12] 介绍的自适应非参数估计量, 以解决 f 未知的难题. 对误差的密度函数进行估计 (如非参数核密度估计) 得到 \hat{f} . 同样地, 迭代过程可能存在 $\hat{f}(\varepsilon)$ 接近于 0 的样本点以至于迭代结果无法收敛, 故考虑文 [8, 19] 的函数调节思想, 常用的调节方法是在 $\mathcal{I}_{\mathbf{n}^*}(\tilde{\beta}(\mathbf{x}); \mathbf{f})$ 、 $\mathbf{S}_{\mathbf{n}^*}(\tilde{\beta}(\mathbf{x}); \mathbf{f})$ 中加入调节函数. 本文引用文 [20] 中的光滑调节函数, $G_{\mathbf{n}^*}(x)$ 的具体形式如下:

$$G_{\mathbf{n}^*}(x) = \begin{cases} 0, & x < c_0 n^{*-\varepsilon}, \\ \frac{\exp\{(2c_0 n^{*-\varepsilon} - x)^{-1}\}}{\exp\{(2c_0 n^{*-\varepsilon} - x)^{-1}\} + \exp\{(x - c_0 n^{*-\varepsilon})^{-1}\}}, & c_0 n^{*-\varepsilon} \leq x \leq 2c_0 n^{*-\varepsilon}, \\ 1, & x > 2c_0 n^{*-\varepsilon}, \end{cases}$$

其中 $c_0 > 0, \varepsilon > 0$ 为调节参数 (一般比较小), 可证明函数 $G_{\mathbf{n}^*}(x)$ 在 \mathbb{R} 上 p 次连续可微. 因此, 在误差分布未知下, 时空模型的非参数估计算法如下.

第一步 给定 \mathbf{x} , 极小化 (2.1) 式, 进行局部多项式估计, 得到参数向量 β 的一个初步相合估计量 $\tilde{\beta} = \tilde{\beta}_{h_1}$. 对所有的 $\mathbf{i} \in \Lambda_{\mathbf{n}}, \mathbf{t} \in \mathbf{T}_{\mathbf{n}}$, 计算残差 $\tilde{R}_{i,t} = Y_{i,t} - \tilde{m}(X_{i,t})$.

第二步 用核密度估计方法计算误差 $R_{i,t}$ 的概率密度函数 $\tilde{f}(z)$, 并计算其导数 $\tilde{f}'(z)$.

$$\tilde{f}(z) = \tilde{f}_{n^*, h_0}(z) = \frac{1}{n^* h_0} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} K\left(\frac{\tilde{R}_{i,t} - z}{h_0}\right),$$

$$\tilde{f}'(z) = \tilde{f}'_{n^*, h_0}(z) = -\frac{1}{n^* h_0^2} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} K'\left(\frac{\tilde{R}_{i,t} - z}{h_0}\right).$$

第三步 将误差密度函数 \tilde{f} 及其导数 \tilde{f}' 代入得分函数与信息矩阵, 有

$$\begin{aligned}\tilde{\mathbf{S}}_{\mathbf{n}^*}(\beta; \tilde{f}) &= -\frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in T_{\mathbf{n}}} \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{\mathbf{h}}\right) \frac{\tilde{f}'(\mathbf{Y}_{\mathbf{i},\mathbf{t}} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}))}{\tilde{f}(\mathbf{Y}_{\mathbf{i},\mathbf{t}} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}))} \\ &\quad \cdot G_{\mathbf{n}^*}(\tilde{\mathbf{f}}(\mathbf{Y}_{\mathbf{i},\mathbf{t}} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}))) \tilde{\mathbf{X}}_{\mathbf{i},\mathbf{t}}, \\ \tilde{\mathcal{I}}_{\mathbf{n}^* \mathbf{h}^d}(\beta; \tilde{\mathbf{f}}) &= \frac{1}{n^*} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in T_{\mathbf{n}}} \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{\mathbf{h}}\right) \left[\frac{\tilde{f}'(\mathbf{Y}_{\mathbf{i},\mathbf{t}} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}))}{\tilde{f}(\mathbf{Y}_{\mathbf{i},\mathbf{t}} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}))} \right]^2 \\ &\quad \cdot G_{\mathbf{n}^*}(\tilde{f}(\mathbf{Y}_{\mathbf{i},\mathbf{t}} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}))) \tilde{\mathbf{X}}_{\mathbf{i},\mathbf{t}} \tilde{\mathbf{X}}_{\mathbf{i},\mathbf{t}}^{\mathbf{T}}.\end{aligned}$$

计算单步 Newton-Raphson 估计量,

$$\hat{\beta}(\mathbf{x}) = \tilde{\beta}(\mathbf{x}) + \tilde{\mathcal{I}}_{\mathbf{n}^*}^{-1}(\tilde{\beta}(\mathbf{x}); \tilde{\mathbf{f}}) \tilde{\mathbf{S}}_{\mathbf{n}^*}(\tilde{\beta}(\mathbf{x}); \tilde{\mathbf{f}}). \quad (2.5)$$

第四步 返回第 1 步, 直至收敛.

在获得初始估计值 $\tilde{\beta}(\mathbf{x})$ 的第一步与计算 $\mathbf{S}_{\mathbf{n}^*}(\beta; \tilde{\mathbf{f}})$, $\mathcal{I}_{\mathbf{n}^*}(\beta; \tilde{\mathbf{f}})$ 的第三步中, 所采用的是多元核函数 $\mathcal{K}(\cdot)$; 而在误差信息估计的第二步中, 核函数选取为一元核函数 $K(\cdot)$, 且 $\mathcal{K} = K \times \cdots \times K$. 带宽 h_0, h, h_1 的具体选择方法将在模拟部分给出.

实际数据误差更倾向于非正态分布, 甚至和正态分布相差较远, 在这种情况下局部似然估计量比最小二乘估计量有更小的方差, 所以更具有有效性. 误差分布未知的非参数估计量 $\hat{\beta}$ 和误差分布已知下的估计量 $\bar{\beta}$ 具有完全一样的偏差. 另外, 该估计量在相同方差的条件下还具备自适应性, 因此大样本下的自适应估计达到了和局部似然估计相同的 MSE. 在实际的模型问题中, 将误差密度函数视为未知的是更自然的一种处理方式.

3 理论性质

假设对于固定时刻 t_0 , $\{(X_{\mathbf{i},t_0}, Y_{\mathbf{i},t_0})\}, \mathbf{i} \in \Lambda_{\mathbf{n}}\}$ 满足如下混合条件: 存在一函数 $\varphi_1(u) \searrow 0$, 当 $u \rightarrow \infty$, $\varphi_1(0) = 1$, 并且当 $E, E' \subset \Lambda_n$ 时, 有

$$\begin{aligned}\alpha(\beta(\mathbf{E}), \beta(\mathbf{E}')) &= \sup\{P(AB) - P(A)P(B) | A \in \beta(\mathbf{E}), \mathbf{B} \in \beta(\mathbf{E}')\} \\ &\leq \psi_1(\text{Card}(E), \text{Card}(E')) \varphi_1(d(E, E')), \quad (3.1)\end{aligned}$$

其中 $\beta(\mathbf{E})$ 表示由 $\{R_{\mathbf{i},t'}\}_{\mathbf{i} \in \mathbf{E}}$ 产生的 Borel σ -域, $\text{Card}(E)$ 表示集合 E 中元素的个数, $d(E, E')$ 为 Euclid 距离, $\psi_1: \mathbb{N}^2 \rightarrow \mathbb{R}^+$ 为一对称非减正函数.

类似地, 我们假设对任一固定位置 \mathbf{i}_0 , $\{(X_{\mathbf{i},t_0}, Y_{\mathbf{i},t_0})\}, t \in T_n\}$ 满足如下混合条件: 存在函数 $\varphi_2(u) \searrow 0$, 当 $t \rightarrow \infty$, $\varphi_2(0) = 1$, 并且当 $G, G' \subset T_n$ 时, 有

$$\begin{aligned}\alpha(\beta(\mathbf{G}), \beta(\mathbf{G}')) &= \sup\{P(AB) - P(A)P(B) | A \in \beta(\mathbf{G}), \mathbf{B} \in \beta(\mathbf{G}')\} \\ &\leq \psi_2(\text{Card}(\mathbf{G}), \text{Card}(\mathbf{G}')) \varphi_2(d(\mathbf{G}', \mathbf{G}')), \quad (3.2)\end{aligned}$$

其中 $\beta(\mathbf{G})$ 表示由 $\{R_{i_0, t}\}_{t \in G}$ 产生的 Borel σ -域, $\psi_2: \mathbb{N}^2 \rightarrow \mathbb{R}^+$ 为一对称非减函数.

下面的假设对于推导估计量的渐近性质是必要的, 从现在起, c 表示常数, 值不重要, 在不同行值可能不一样.

假设 (A1) 随机域 $\{\mathbf{X}_{i, t}\}$ 是严平稳的. 对于 $\Lambda_n \times \mathbf{T}_n$ 中不同的点 (i, t) 和 (j, τ) , 向量 $\mathbf{X}_{i, t}$ 和 $X_{j, \tau}$ 有联合概率密度函数记 $f_{(i, t)(j, \tau)}$. 记 f_X 为 $\mathbf{X}_{i, t}$ 的边缘概率密度函数, 在 \mathbf{x}_0 点处有连续的一阶导数且 $f_X(\mathbf{x}_0) > 0$. $f_X(\mathbf{x})$ 的 k 阶导函数 $(D^k f_X)(\mathbf{x})$ 有界且在 \mathbb{R}^d 上一直连续, 即存在 $c < \infty$, 使得 $|(D^k f_X)(\mathbf{x}_1) - (D^k f_X)(\mathbf{x}_2)| \leq c \|\mathbf{x}_1 - \mathbf{x}_2\|$, $|\mathbf{k}| = \mathbf{p} + 1$, 其中 $\mathbf{k} = (\mathbf{k}_1, \dots, \mathbf{k}_d)^T$, $|\mathbf{k}| = \mathbf{k}_1 + \dots + \mathbf{k}_d$.

(A2) $(D^k m)(\mathbf{x})$ 有界且在 \mathbb{R}^d 上连续, 即存在 $c < \infty$, 使得

$$|(D^k m)(\mathbf{x}_1) - (D^k m)(\mathbf{x}_2)| \leq c \|\mathbf{x}_1 - \mathbf{x}_2\|, \quad |\mathbf{k}| = \mathbf{p} + 1.$$

(A3) 对于固定位置 \mathbf{i}_0 , $\{R_{i_0, t}, t \in T_n\}$ 是严平稳过程, 对于固定时间点 t_0 , $\{R_{i, t_0}, \mathbf{i} \in \Lambda_n\}$ 是严平稳过程. f 记为 $R_{i, t}$ 的概率密度函数, 关于 0 对称, 满足 $ER_{i, t}^4 < \infty$, $f''(z)$ 是 Lipschitz 连续, 即存在 $c < \infty$, 使得任意 $z, z' \in \mathbb{R}$, 有 $|f''(z) - f''(z')| < c|z - z'|$.

(A4) $K(\cdot)$ 是对称的概率密度函数, 满足 $\int_{\mathbb{R}} K'^2(u) du < \infty$, $\int_{\mathbb{R}} u^2 K(u) du < \infty$. 本文采用乘积核 $\mathcal{K}(\mathbf{x}) = \mathbf{K}(\mathbf{x}_1) \times \dots \times \mathbf{K}(\mathbf{x}_d)$, 满足

$$\int_{\mathbb{R}^d} \mathbf{u}^{\phi_j(\mathbf{k})} \mathcal{K}(\mathbf{u}) d\mathbf{u} < \infty, \\ \int_{\mathbb{R}^d} \mathbf{u}^{\phi_j(\mathbf{k})} \mathcal{K}^2(\mathbf{u}) d\mathbf{u} < \infty, \quad j = 1, \dots, p, \quad k = 1, \dots, 2(p+1).$$

(A5) $E[(\log f(R_{i, t}))^2] < \infty$, $E[|(\log f(R_{i, t}))''|] < \infty$, $E[|(\log f(R_{i, t}))'''|] < \infty$.

(A6) $n^* \rightarrow \infty$ 时, $h_0, h, h_1 \rightarrow 0$, 满足 $n^* h_0^4 \rightarrow \infty$, $n^* h^{2d} \rightarrow \infty$, $n^* h_1^d \rightarrow \infty$, $h/h_1 \rightarrow 0$, $(n^* h^d)^{\frac{1}{2}} h_0^2 \rightarrow 0$.

(A7) 令 $c_n = (n^* h_0)^{\frac{1}{4N}}$, $t_n = (n^* h_0)^{\frac{1}{4}}$, 有

$$t_n \sum_{\|i\| > c_n} i^{N-1} \varphi_1(\|i\|) < \infty, \quad c_n^N \sum_{|t| > t_n} \varphi_2(|t|) < \infty.$$

(A8) $\forall (i, t), (j, \tau) \in \Lambda_n \times \mathbf{T}_n$,

$$|\text{Cov}(R_{i, t} I(|R_{i, t}| \leq \gamma_n), R_{j, \tau} I(|R_{j, \tau}| \leq \gamma_n))| \\ \leq |\text{Cov}(R_{i, t} I(|R_{i, t}| \leq \gamma_n), R_{j, t} I(|R_{j, t}| \leq \gamma_n))|.$$

注 3.1 (A1) 和文 [21] 中 (A1)、文 [12] 类似. (A3) 和文 [21] 中的 (A4)、文 [12] 中的 (A1) 类似. (A2), (A4), (A5) 分别和文 [12] 中的 (A5), (A3), (A2) 类似, 其中 (A4) 中关于核函数的假设对支集是有限区间的核函数或者正态分布核函数都是满足的. (A7), (A8) 和

文 [21] 中的 (A7), (A8) 类似. (A6) 和文 [21] 中的 (A6) 类似, 根据理论最优带宽的取法, (A6) 是满足的, 其中若取 $h_1 = h \cdot \log n^*$, $h_0 = h^d \cdot \log n^*$, 则 h_0, h, h_1 间关系满足.

文中 \mathcal{P} 都表示概率意义下, $\xrightarrow{\mathcal{D}}$ 表示依分布收敛. 下面给出主要结果, 证明将在第五节给出.

令 $H = \text{diag}\{H_0, H_1, \dots, H_p\}$ 是 $N_0 + N_1 + \dots + N_p$ 维对角阵, 其中 H_ℓ 是 N_ℓ 维对角阵, 对角线上第 r 个元素为 $h^{\phi_\ell(r)}$. 令

$$\mathbf{U} = \begin{pmatrix} U_{0,0} & U_{0,1} & \cdots & U_{0,p} \\ U_{1,0} & U_{1,1} & \cdots & U_{1,p} \\ \vdots & \vdots & & \vdots \\ U_{p,0} & U_{p,1} & \cdots & U_{p,p} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{0,0} & \Sigma_{0,1} & \cdots & \Sigma_{0,p} \\ \Sigma_{1,0} & \Sigma_{1,1} & \cdots & \Sigma_{1,p} \\ \vdots & \vdots & & \vdots \\ \Sigma_{p,0} & \Sigma_{p,1} & \cdots & \Sigma_{p,p} \end{pmatrix},$$

其中 $U_{i,j}, \Sigma_{i,j}$ 都是 $N_i \times N_j$ 维矩阵, 它们的第 l 行, 第 k 列元素分别为 $\mu_{\phi_i(l)+\phi_j(k)}, \nu_{\phi_i(l)+\phi_j(k)}$,

$$\mu_i = \int_{\mathbb{R}^d} \mathbf{u}^i \mathcal{K}(\mathbf{u}) d\mathbf{u}, \quad \nu_i = \int_{\mathbb{R}^d} \mathbf{u}^i \mathcal{K}^2(\mathbf{u}) d\mathbf{u}.$$

定理 3.1 假设 (A1)–(A8) 成立, 有

$$\sqrt{n^* h^d} \{H[\widehat{\beta}(\mathbf{x}) - \beta(\mathbf{x})] - \mathbf{a} h^{p+1} \mathbf{U}^{-1} \bar{\mathbf{U}} \mathbf{m}^{(p+1)}(\mathbf{x})\} \xrightarrow{\mathcal{D}} \mathbf{N}(\mathbf{0}, \mathbf{I}_\beta^{-1}(\mathbf{x}) \mathbf{U}^{-1} \Sigma \mathbf{U}^{-1}),$$

其中

$$\mathbf{a} = \frac{\int_{\mathbb{R}} \frac{f''(z)}{f(z)} dz - \int_{\mathbb{R}} f''(z) dz}{\int_{\mathbb{R}} \frac{f''(z)}{f(z)} dz}, \quad \bar{\mathbf{U}} = (U_{0,p+1}, \dots, U_{p,p+1})^T.$$

故更具体的, 有

$$\sqrt{n^* h^d} \{[\widehat{\mathbf{m}}(\mathbf{x}) - \mathbf{m}(\mathbf{x})] - \mathbf{a} h^{p+1} [\mathbf{U}^{-1} \bar{\mathbf{U}} \mathbf{m}^{(p+1)}(\mathbf{x})]_{\mathbf{0}}\} \xrightarrow{\mathcal{D}} \mathbf{N}(\mathbf{0}, \mathbf{I}_\beta^{-1}(\mathbf{x}) [\mathbf{U}^{-1} \Sigma \mathbf{U}^{-1}]_{\mathbf{0},\mathbf{0}}),$$

其中 $[\cdot]_{\mathbf{0}}$ 表示向量的第一个元素, $[\cdot]_{\mathbf{0},\mathbf{0}}$ 表示矩阵的第一行, 第一列元素.

注 3.2 最小二乘的方差为 $\sigma^2 f_X^{-1}(\mathbf{x}) [\mathbf{U}^{-1} \Sigma \mathbf{U}^{-1}]_{\mathbf{0},\mathbf{0}}$, 由 C-R 不等式, 有 $\sigma^2 \geq I^{-1}(f)$, 所以最小二乘的方差更大. 我们常见的对称分布, 如零均值正态分布、 t 分布、Cauchy 分布等的 $\int_{\mathbb{R}} f''(z) dz = 0$, 则 $\mathbf{a} = 1$. 此时和基于最小二乘的局部多项式方法比较, 偏差一样 (见引理 5.3). 计算了非对称的卡方分布, 当自由度 > 4 时, 同样得到 $\int_{\mathbb{R}} f''(z) dz = 0$. 参数为 λ 的指数分布 $\int_{\mathbb{R}} f''(z) dz = \lambda^2 > 0$, 而 $\int_{\mathbb{R}} \frac{f''(z)}{f(z)} dz = \lambda^2$, 此时偏差为 0, 比最小二乘的偏差小, 我们的方法更有效. 若 $\int_{\mathbb{R}} f''(z) dz = \lambda^2 < 0$, 则我们的方法偏差大, 方差小, 均方误差得具体问题具体分析.

4 模拟和实例

4.1 数值模拟

本节将通过六个有限样本的蒙特卡罗模拟, 来展示时空模型中误差分布未知下非参数估计效果. 为简单起见, 令 $d = 1, N = 2$, 故考虑如下两个模型:

$$\text{模型 1: } Y_{i,j,t} = x^3 + R_{i,j,t};$$

$$\text{模型 2: } Y_{i,j,t} = 2 \sin(2\pi X_{i,j,t}) + R_{i,j,t},$$

其中 $i = 1, \dots, n_1, j = 1, \dots, n_2, t = 1, \dots, n_3, X_{i,j,t}$ 是均值为 0 的二阶平稳高斯过程. 我们采用谱方法 (见 [22]) 产生时空过程 $X_{i,j,t}$,

$$X_{i,j,t} = \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{k=1}^M \cos(w(1, k) * i + w(2, k) * j + q(k) * t + r(k)),$$

其中 $w(i, k), i = 1, 2, q(k), k = 1, \dots, M$ 独立同分布, 来自标准正态分布, 且独立于 $r(k), k = 1, \dots, M$, 而 $r(k), k = 1, \dots, M$ 独立同分布, 来自 $[-\pi, \pi]$ 上的均匀分布. 当 $n \rightarrow \infty, X_{i,j,t}$ 收敛于高斯遍历过程 (见 [23]). 我们主要考虑以下三种情形的随机误差分布:

情形 1 $R_{i,j,t} \sim t(4)$;

情形 2 $R_{i,j,t} \sim 0.4N(-0.6, 0.4^2) + 0.6N(1.2, 1)$;

情形 3 $R_{i,j,t} \sim 0.9N(0, 1) + 0.1N(0, 10^2)$.

情形 1 是自由度为 4 的 t 分布; 情形 2 是右偏的混合高斯过程; 情形 3 是含有奇异点的误差正态分布, 其中 10% 的数据可视为奇异点.

众所周知, 带宽的选取对非参数估计有极大的影响. 目前广泛用于选取带宽的方法有两种, 分别是大拇指准则方法与交叉验证 (CV) 法. 本文采用交叉验证法选取最优带宽 h , 具体地, $h = \arg \min CV$, 其中

$$CV = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{t=1}^{n_3} [m(X_{i,j,t}) - m_{\neq(i,j,t)}(h)]^2.$$

对于误差分布的三种情形, 令 $M = 1000, c_0 = \frac{1}{20}, \varepsilon = 0.01$, 重复模拟 100 次, 样本量分别为 $8 \times 8 \times 8, 10 \times 10 \times 10, 15 \times 15 \times 15$. 选取 1 维核函数 $K(\cdot)$ 为 Gauss 核, 即

$$K(u) = \frac{1}{\sqrt{2\pi} \exp\left(-\frac{u^2}{2}\right)}.$$

由于存在接近于 0 的得分函数分母 $f(z)$, 引致极大似然估计迭代过程出现不收敛, 故在模拟运算中加入调节函数 $G_{n^*}(x)$.

图 1 是样本量为 $10 \times 10 \times 10$ 的三阶局部多项式估计与非参数自适应估计的结果. 从不同误差分布的拟合图中可直观地看出, 非参数自适应方法拟合效果优于三阶局部多项式, 且在边界点处表现得更为显著.

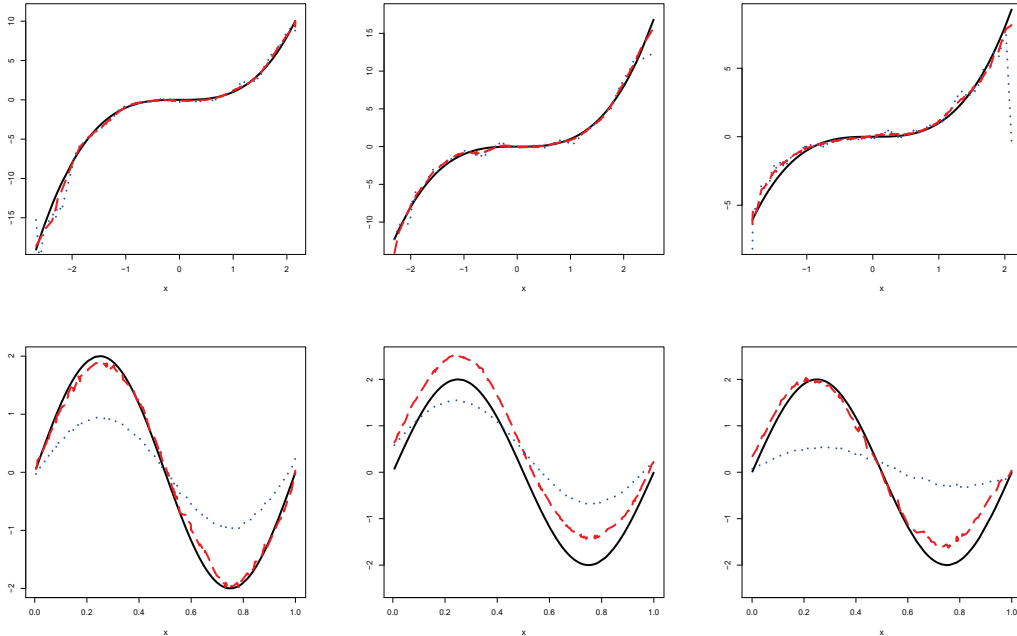


图 1 左上图与左下图分别为情形 1 下的 $m(x) = x^3$, $m(x) = 2 \sin(2\pi x)$: $R_{i,j,t} \sim t(4)$, 中上图与中下图分别为情形 2 下的 $m(x) = x^3$, $m(x) = 2 \sin(2\pi x)$: $R_{i,j,t} \sim 0.4N(-0.6, 0.4^2) + 0.6N(1.2, 1)$, 右上图与右下图分别为情形 3 下的 $m(x) = x^3$, $m(x) = 2 \sin(2\pi x)$: $R_{i,j,t} \sim 0.9N(0, 1) + 0.1N(0, 10^2)$. 黑色实线为真实曲线, 蓝色点线是三阶最小二乘局部多项式方法拟合曲线, 红色虚线是非参数自适应方法拟合曲线

为了进一步比较估计效果, 本文采用均方误差 (MSE) 与可决系数 (R^2) 来衡量各估计量的效率, 其分别定义如下:

$$\text{MSE} = \frac{1}{n^*} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{t=1}^{n_3} [m(X_{i,j,t}) - \hat{m}(X_{i,j,t})]^2, \quad R^2 = 1 - \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{t=1}^{n_3} [m(X_{i,j,t}) - \hat{m}(X_{i,j,t})]^2}{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{t=1}^{n_3} [m(X_{i,j,t}) - \bar{m}]^2},$$

其中 \bar{m} 是 $\{m(X_{i,j,t}), i = 1, \dots, n_1, j = 1, \dots, n_2, t = 1, \dots, n_3\}$ 的均值. 由 CV 方法得到的 h 也不是整体绝对最优, 因此在细化比较估计效果时, 同时计算大拇指准则最优带宽 h' , 在 h 与 h' 组成的区间内等间距地取 8 个带宽值, 以分别进行拟合, 得到不同带宽值下的 MSE 与 R^2 , 再进行不同方法估计效果的比较. 图 2 是样本量为 $10 \times 10 \times 10$ 在不同带宽值下重复模拟 100 次的平均拟合效率比较图. 从图中可直观地看到, 非参数自适应估计较局部多项式而言, 更充分地利用了样本的信息, 且拟合值更接近真实值. 对于不同的模型, 所利用样本信息程度也不同, 在误差服从 $t(4)$ 分布的 x^3 模型中, 两种估计方法效率相差无几, 但在另外两种误差分布的 x^3 模型中, 非参数自适应方法估计的 MSE 明显低于局部多项式估计方法. 另一方面, 由于时空数据中奇异点的存在, 非参数自适应估计方法在两侧的趋势拟合显著优于传统方法. 对于三角函数模型, 两种估计方法对样本信息的利用程度都较低. 另外, 我们发现局部多项式估计的 MSE 值对带宽显得较为敏感, 由于篇幅的

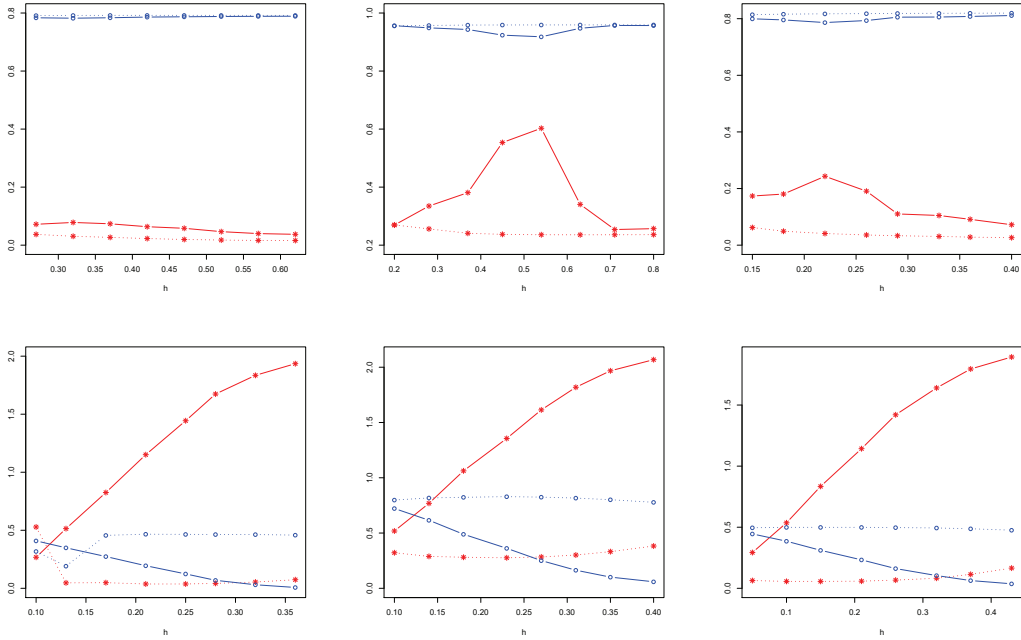


图 2 左上图与左下图分别为情形 1 下的 $m(x) = x^3$, $m(x) = 2 \sin(2\pi x)$, 中上图与中下图分别为情形 2 下的 $m(x) = x^3$, $m(x) = 2 \sin(2\pi x)$, 右上图与右下图分别为情形 3 下的 $m(x) = x^3$, $m(x) = 2 \sin(2\pi x)$. 蓝色圆点实线为三阶局部多项式估计的 R^2 , 蓝色圆点虚线为非参数自适应估计的 R^2 , 红色星号实线为三阶局部多项式估计的 MSE, 红色星号虚线为非参数自适应估计的 MSE

限制, 此处我们仅列出最优带宽下不同样本量三阶局部多项式估计 (Lpe) 与非参数自适应估计 (Adaptive) 的具体 R^2 , MSE 值, 如表 1 和表 2 所示, 读者若对等间距带宽下的拟合结果感兴趣, 可随时联系作者获取. 综上可看出, 非参数自适应估计对误差分布未知的时空数据拟合效果明显优于目前常用的最小二乘局部多项式方法.

表 1 100 次模拟 MSE 的均值

	$8 \times 8 \times 8$		$10 \times 10 \times 10$		$15 \times 15 \times 15$	
	Lpe	Adaptive	Lpe	Adaptive	Lpe	Adaptive
模型 1+ 情形 1	0.1269	0.0935	0.0782	0.0306	0.0271	0.0265
模型 1+ 情形 2	0.2313	0.2171	0.3810	0.2409	0.2253	0.2181
模型 1+ 情形 3	0.0813	0.0342	0.1103	0.0333	0.0492	0.0281
模型 2+ 情形 1	1.8269	0.1285	1.9351	0.0740	1.8464	0.0682
模型 2+ 情形 2	1.9130	0.3193	1.8186	0.3011	1.9185	0.2990
模型 2+ 情形 3	1.2052	0.1152	1.1427	0.0586	1.1682	0.0480

表 2 100 次模拟 R^2 的均值

	$8 \times 8 \times 8$		$10 \times 10 \times 10$		$15 \times 15 \times 15$	
	Lpe	Adaptive	Lpe	Adaptive	Lpe	Adaptive
模型 1+ 情形 1	0.6887	0.6918	0.7821	0.7917	0.7566	0.7580
模型 1+ 情形 2	0.9252	0.9260	0.9432	0.9585	0.9554	0.9560
模型 1+ 情形 3	0.6977	0.7104	0.8054	0.8186	0.8386	0.8395
模型 2+ 情形 1	0.0096	0.3920	0.0074	0.4585	0.0263	0.5244
模型 2+ 情形 2	0.1493	0.8261	0.1638	0.8164	0.1342	0.8347
模型 2+ 情形 3	0.1995	0.4773	0.2383	0.4990	0.2295	0.5365

4.2 实证分析

本节考虑南京市空气污染数据, 该数据来源于中国气象数据网和南京环境监测中心, 数据包含了南京市 2018 年 9 月 19 日至 2019 年 10 月 20 日的空气质量指数 (AQI)、 $PM_{2.5}$ 浓度 ($\mu g/m^3$)、CO 浓度 (mg/m^3)、 SO_2 浓度 ($\mu g/m^3$)、 NO_2 浓度 ($\mu g/m^3$)、臭氧 8 小时 O_3-8h ($\mu g/m^3$)、温度 ($^{\circ}C$)、大气相对湿度 (%)、风速 (m/s) 这 9 个变量, 其中 AQI 为响应变量 (Y), 其余变量按顺序记为协变量 X_1, X_2, \dots, X_8 . 具体考虑各污染物与环境因素对空气质量的影响关系.

随着协变量维数的增加, 非参数估计难度迅速增加, 为降低维数带来的风险, 对协变量标准化后进行主成分分析, 取累积方差贡献率达到 88.75% 的前四个主成分, 分别表示为:

$$Z1 = 0.479X_1^* + 0.468X_2^* + 0.338X_3^* + 0.483X_4^* - 0.148X_5^* - 0.297X_6^* - 0.297X_8^*,$$

$$Z2 = -0.43X_3^* - 0.603X_5^* - 0.399X_6^* + 0.519X_7^*,$$

$$Z3 = 0.292X_2^* - 0.191X_3^* + 0.221X_5^* + 0.49X_6^* + 0.567X_7^* - 0.514X_8^*,$$

$$Z4 = 0.411X_1^* + 0.373X_2^* + 0.187X_5^* + 0.227X_6^* + 0.186X_7^* + 0.751X_8^*,$$

其中 X^* 为标准化的解释变量. 利用 $Z1, Z2, Z3, Z4$ 四个主成分进行分析, 由于维数与 Taylor 展开的复杂性, 局部多项式估计与非参数自适应估计都将只展开至一阶. 核函数与带宽的选取方法与模拟部分相同, 且估计效率的衡量仍采用 MSE 与 R^2 , 结果如表 3 所示. 由本例可知, 对城市污染数据, 自适应非参数估计效果优于常用的局部多项式估计方法.

表 3 两种方法 MSE 和 R^2 的值

	MSE	R^2
局部多项式方法	207.2645	0.8260
自适应方法	194.9940	0.8363

5 定理的证明

在这部分, 我们将给出定理的证明, 证明依赖于一些基本的结论, 引理 5.1 和引理 5.2 类似文 [24–25] 的结论, 引理 5.3 类似于文 [3, 21] 的结论, 我们就不再给出证明.

引理 5.1 (a) 假定 (3.1) 式成立, 令 $\mathcal{L}_r(\mathcal{F})$ 为所有满足条件 $\|X\|_r = (\mathbb{E}|X|^r)^{\frac{1}{r}} < \infty$ 的 \mathcal{F} -可测随机变量 X 组成的集合. 令 $X \in \mathcal{L}_r(\mathcal{B}(E))$, $Y \in \mathcal{L}_s(\mathcal{B}(E'))$, 则对任意满足 $r_1^{-1} + r_2^{-1} + r_3^{-1} = 1$ 的 $1 \leq r_1, r_2, r_3 < \infty$, 有

$$|\mathbb{E}XY - \mathbb{E}X\mathbb{E}Y| \leq C\|X\|_{r_1}\|Y\|_{r_2}\{\psi_1(\text{Card}E, \text{Card}E')\varphi_1(d(E, E'))\}^{\frac{1}{r_3}}. \quad (5.1)$$

(b) 进一步, 若 $|X|$ 和 $|Y|$ 几乎处处有界, 则 (5.1) 的右边可以由 $C\psi_1(\text{Card}E, \text{Card}E') \cdot \varphi_1(d(E, E'))$ 代替.

引理 5.2 (a) 假定 (3.2) 式成立, 令 $\mathcal{L}_r(\mathcal{F})$ 为所有满足条件 $\|X\|_r = (\mathbb{E}|X|^r)^{\frac{1}{r}} < \infty$ 的 \mathcal{F} -可测随机变量 X 组成的集合. 令 $X \in \mathcal{L}_r(\mathcal{B}(G))$, $Y \in \mathcal{L}_s(\mathcal{B}(G'))$, 则对任意满足 $r_1^{-1} + r_2^{-1} + r_3^{-1} = 1$ 的 $1 \leq r_1, r_2, r_3 < \infty$, 有

$$|\mathbb{E}XY - \mathbb{E}X\mathbb{E}Y| \leq C\|X\|_{r_1}\|Y\|_{r_2}\{\psi_2(\text{Card}G, \text{Card}G')\varphi_2(d(G, G'))\}^{\frac{1}{r_3}}. \quad (5.2)$$

(b) 进一步, 如果 $|X|$ 和 $|Y|$ 几乎处处有界, 那么 (5.2) 式的右边可以由 $C\psi_2(\text{Card}G, \text{Card}G')\varphi_2(d(G, G'))$ 代替.

引理 5.3 假设 (A1)–(A8) 成立, 进一步加上条件: 存在正整数序列 $q = o(n^*h_1^d)^{\frac{1}{2d}} \rightarrow \infty$, 使得 $n^* \sum_{i=1}^{\infty} i^d \varphi_1(iq) < \infty$, $n^* \sum_{i=1}^{\infty} \varphi_2(iq) < \infty$, 有

$$(n^*h_1^d)^{\frac{1}{2}} \left[\begin{pmatrix} \tilde{m}(\mathbf{x}) \\ h_1 \tilde{m}'(\mathbf{x}) \end{pmatrix} - \begin{pmatrix} m(\mathbf{x}) \\ h_1 m'(\mathbf{x}) \end{pmatrix} - h^{p+1} \Psi^{-1} \begin{pmatrix} \mathbf{a}_0(\mathbf{x}) \\ \mathbf{a}_1(\mathbf{x}) \end{pmatrix} \right] \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \sigma^2 \Psi^{-1} \Gamma (\Psi^{-1})^T),$$

其中

$$\begin{aligned} \mathbf{a}_0(\mathbf{x}) &= \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \mathbf{m}_{ij}(\mathbf{x}) \int_{\mathbb{R}^d} \mathbf{u}_i \mathbf{u}_j \mathcal{K}(\mathbf{u}) \mathrm{d}\mathbf{u}, \\ \mathbf{a}_1(\mathbf{x}) &= \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \mathbf{m}_{ij}(\mathbf{x}) \int_{\mathbb{R}^d} \mathbf{u}_i \mathbf{u}_j \mathbf{u} \mathcal{K}(\mathbf{u}) \mathrm{d}\mathbf{u}, \\ \mathbf{m}_{ij}(\mathbf{x}) &= \frac{\partial^2 \mathbf{m}(\mathbf{x})}{\partial \mathbf{x}_i \partial \mathbf{x}_j}, \quad i, j = 1, \dots, d, \quad \mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_d)^T. \\ \Psi &= \begin{pmatrix} f_X(x) \int_{\mathbb{R}^d} \mathcal{K}(\mathbf{u}) \mathrm{d}\mathbf{u} & f_X(x) \int_{\mathbb{R}^d} \mathbf{u}^T \mathcal{K}(\mathbf{u}) \mathrm{d}\mathbf{u} \\ f_X(x) \int_{\mathbb{R}^d} \mathbf{u} \mathcal{K}(\mathbf{u}) \mathrm{d}\mathbf{u} & f_X(x) \int_{\mathbb{R}^d} \mathbf{u} \mathbf{u}^T \mathcal{K}(\mathbf{u}) \mathrm{d}\mathbf{u} \end{pmatrix}, \\ \Gamma &= \begin{pmatrix} f_X(x) \int_{\mathbb{R}^d} K^2(\mathbf{u}) \mathrm{d}\mathbf{u} & f_X(x) \int_{\mathbb{R}^d} \mathbf{u}^T K^2(\mathbf{u}) \mathrm{d}\mathbf{u} \\ f_X(x) \int_{\mathbb{R}^d} \mathbf{u} K^2(\mathbf{u}) \mathrm{d}\mathbf{u} & f_X(x) \int_{\mathbb{R}^d} \mathbf{u} \mathbf{u}^T K^2(\mathbf{u}) \mathrm{d}\mathbf{u} \end{pmatrix}. \end{aligned}$$

引理 5.4 假设 (A1)–(A8) 成立, 有对任意 $z \in \mathbb{R}$,

$$\tilde{f}(z) - f(z) = O_{\mathcal{P}}((n^*h_1^d)^{-\frac{1}{2}} + h_0^2).$$

证 由引理 5.3 和 Taylor 展式, 有

$$\begin{aligned} \tilde{f}(z) - f(z) &= \frac{1}{n^*h_0} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K\left(\frac{\tilde{R}_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) - \frac{1}{n^*h_0} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) \\ &\quad + \frac{1}{n^*h_0} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) - f(z) = A_1 + A_2. \end{aligned} \quad (5.3)$$

$$\begin{aligned} A_1 &= \frac{1}{n^*h_0} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K\left(\frac{\tilde{R}_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) - \frac{1}{n^*h_0} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) \\ &= \frac{1}{n^*h_0^2} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K'\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) [m(\mathbf{X}_{\mathbf{i},\mathbf{t}}) - \tilde{m}(\mathbf{X}_{\mathbf{i},\mathbf{t}}) + o_{\mathcal{P}}((n^*h_1^d)^{-\frac{1}{2}})] \\ &= O_{\mathcal{P}}((n^*h_1^d)^{-\frac{1}{2}}) \cdot \frac{1}{n^*h_0^2} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K'\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right). \end{aligned} \quad (5.4)$$

由假设条件 (A4), (A6) 和 (A7), 有

$$\mathbb{E}\left[\frac{1}{n^*h_0} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K'\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right)\right] = f'(z) \int_{\mathbb{R}} u K'(u) du.$$

令 $\mathbf{c}_n = (n^*h_0)^{\frac{1}{4N}}$, $\mathbf{t}_n = (n^*h_0)^{\frac{1}{4}}$, $\mathbf{S}_n = \{(\mathbf{i}, \mathbf{t}) \neq (\mathbf{j}, \tau) : \|\mathbf{i} - \mathbf{j}\| \leq \mathbf{c}_n, |\mathbf{t} - \tau| \leq \mathbf{t}_n\}$, 则 $\text{Card}(\mathbf{S}_n) \leq 2^{N+1} n^* c_n^N t_n$, 类似地,

$$\begin{aligned} &\text{Var}\left[\frac{1}{n^*h_0^2} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K'\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right)\right] \\ &= \frac{1}{(n^*h_0)^2} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} \mathbb{E}\left[K'\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right)\right]^2 + \frac{1}{(n^*h_0^2)^2} \sum_{((\mathbf{i},\mathbf{t}),(\mathbf{j},\tau)) \in \mathbf{S}_n} \mathbb{E}K'\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) K'\left(\frac{R_{\mathbf{j},\tau} - z}{h_0}\right) \\ &\quad + \frac{1}{(n^*h_0^2)^2} \sum_{((\mathbf{i},\mathbf{t}),(\mathbf{j},\tau)) \notin \mathbf{S}_n} \mathbb{E}K'\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) K'\left(\frac{R_{\mathbf{j},\tau} - z}{h_0}\right) \\ &= O\left[\frac{1}{n^*h_0^3} f(z) \int_{\mathbb{R}} K'^2(u) du + \frac{2^{N+1} c_n^N t_n}{n^*h_0} + \frac{2^N c_n^N}{n^*h_0^4} \sum_{|t| > t_n} \varphi_2(|t|) + \frac{C t_n}{n^*h_0^4} \sum_{\|\mathbf{i}\| > \mathbf{c}_n} \varphi_1(\|\mathbf{i}\|)\right] \\ &= O\left(\frac{1}{(n^*h_0)^{\frac{1}{2}}} + \frac{1}{n^*h_0^4}\right) \rightarrow 0. \end{aligned}$$

故

$$\frac{1}{n^*h_0^2} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K'\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right) \xrightarrow{\mathcal{P}} f'(z) \int_{\mathbb{R}} u K'(u) du. \quad (5.5)$$

由 (5.4) 式和 (5.5) 式, 有

$$A_1 = O_{\mathcal{P}}((n^*h_1^d)^{-\frac{1}{2}}). \quad (5.6)$$

类似地, 可得

$$\begin{aligned} &\mathbb{E}\left[\frac{1}{n^*h_0} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K\left(\frac{R_{\mathbf{i},\mathbf{t}} - z}{h_0}\right)\right] \\ &= f(z) + \frac{1}{2} h_0^2 f''(z) \int_{\mathbb{R}} u^2 K(u) du + o(h_0^2), \end{aligned}$$

$$\begin{aligned}
& \text{Var} \left[\frac{1}{n^* h_0} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} K \left(\frac{R_{\mathbf{i}, \mathbf{t}} - z}{h_0} \right) \right] \\
&= O \left[\frac{1}{n^* h_0} f(z) \int_{\mathbb{R}} K^2(u) du + \frac{2^{N+1} c_n^N t_n}{n^*} + \frac{2^N c_n^N}{n^* h_0^2} \sum_{|t| > t_n} \varphi_2(|t|) + \frac{t_n}{n^* h_0^2} \sum_{\|\mathbf{i}\| > c_n} \varphi_1(\|\mathbf{i}\|) \right] \\
&= O(n^{*-\frac{1}{2}} h_0^{\frac{1}{2}} + n^{*-1} h_0^{-2}) \rightarrow 0.
\end{aligned}$$

$$A_2 = O_{\mathcal{P}}(h_0^2). \quad (5.7)$$

由 (5.3) 式, (5.6) 式和 (5.7) 式, 有

$$\tilde{f}(z) - f(z) = O_{\mathcal{P}}((n^* h_1^d)^{-\frac{1}{2}} + h_0^2).$$

为方便起见, 令 $\delta_{\mathbf{i}, \mathbf{t}} = m(\mathbf{X}_{\mathbf{i}, \mathbf{t}}) - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x})$, $\zeta_{\mathbf{i}, \mathbf{t}} = \mathcal{P}_{\mathbf{p}}(\tilde{\beta}, \mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}) - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x})$, $\bar{R}_{\mathbf{i}, \mathbf{t}} = \mathbf{Y}_{\mathbf{i}, \mathbf{t}} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}) = R_{\mathbf{i}, \mathbf{t}} + \delta_{\mathbf{i}, \mathbf{t}}$, $\check{R}_{\mathbf{i}, \mathbf{t}} = \mathbf{Y}_{\mathbf{i}, \mathbf{t}} - \mathcal{P}_{\mathbf{p}}(\tilde{\beta}, \mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}) = \bar{R}_{\mathbf{i}, \mathbf{t}} - \zeta_{\mathbf{i}, \mathbf{t}} = R_{\mathbf{i}, \mathbf{t}} + \delta_{\mathbf{i}, \mathbf{t}} - \zeta_{\mathbf{i}, \mathbf{t}}$.

引理 5.5 假设 (A1)–(A8) 成立, 有

$$H^{-1} \tilde{I}_{n^*}(\tilde{\beta}, \tilde{f}) H^{-1} \xrightarrow{\mathcal{P}} I_{\beta}(x) \mathbf{U}.$$

证 因

$$\begin{aligned}
[\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}})]^2 &= [f'(\bar{R}_{\mathbf{i}, \mathbf{t}})]^2 + [\tilde{f}'(\check{R}_{\mathbf{i}, \mathbf{t}}) - f'(\bar{R}_{\mathbf{i}, \mathbf{t}})]^2 + 2f'(\bar{R}_{\mathbf{i}, \mathbf{t}})[\tilde{f}'(\check{R}_{\mathbf{i}, \mathbf{t}}) - f'(\bar{R}_{\mathbf{i}, \mathbf{t}})], \\
\frac{1}{\tilde{f}^2(\check{R}_{\mathbf{i}, \mathbf{t}})} &= \frac{1}{f^2(\bar{R}_{\mathbf{i}, \mathbf{t}})} - \frac{[\tilde{f}'(\check{R}_{\mathbf{i}, \mathbf{t}}) - f'(\bar{R}_{\mathbf{i}, \mathbf{t}})][\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}}) + f(\bar{R}_{\mathbf{i}, \mathbf{t}})]}{f^2(\bar{R}_{\mathbf{i}, \mathbf{t}}) \tilde{f}^2(\check{R}_{\mathbf{i}, \mathbf{t}})},
\end{aligned}$$

故

$$\begin{aligned}
& H^{-1} \tilde{I}_{n^*}(\tilde{\beta}, \tilde{f}) H^{-1} \\
&= \frac{1}{n^* h^d} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} \kappa \left(\frac{\mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}}{h} \right) \left[\frac{f'(\bar{R}_{\mathbf{i}, \mathbf{t}})}{f(\bar{R}_{\mathbf{i}, \mathbf{t}})} \right]^2 H^{-1} \tilde{X}_{\mathbf{i}, \mathbf{t}} \tilde{X}_{\mathbf{i}, \mathbf{t}}^T H^{-1} \\
&\quad - \frac{1}{n^* h^d} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} \kappa \left(\frac{\mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}}{h} \right) \left[\frac{f'(\bar{R}_{\mathbf{i}, \mathbf{t}})}{f(\bar{R}_{\mathbf{i}, \mathbf{t}})} \right]^2 [1 - G_{n^*}(\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}}))] H^{-1} \tilde{X}_{\mathbf{i}, \mathbf{t}} \tilde{X}_{\mathbf{i}, \mathbf{t}}^T H^{-1} \\
&\quad + \frac{1}{n^* h^d} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} \kappa \left(\frac{\mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}}{h} \right) \left[\frac{\tilde{f}'(\check{R}_{\mathbf{i}, \mathbf{t}}) - f'(\bar{R}_{\mathbf{i}, \mathbf{t}})}{f(\bar{R}_{\mathbf{i}, \mathbf{t}})} \right]^2 G_{n^*}(\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}})) H^{-1} \tilde{X}_{\mathbf{i}, \mathbf{t}} \tilde{X}_{\mathbf{i}, \mathbf{t}}^T H^{-1} \\
&\quad + \frac{2}{n^* h^d} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} \kappa \left(\frac{\mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}}{h} \right) \frac{f'(\bar{R}_{\mathbf{i}, \mathbf{t}})[\tilde{f}'(\check{R}_{\mathbf{i}, \mathbf{t}}) - f'(\bar{R}_{\mathbf{i}, \mathbf{t}})]}{f^2(\bar{R}_{\mathbf{i}, \mathbf{t}})} G_{n^*}(\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}})) H^{-1} \tilde{X}_{\mathbf{i}, \mathbf{t}} \tilde{X}_{\mathbf{i}, \mathbf{t}}^T H^{-1} \\
&\quad - \frac{1}{n^* h^d} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} \kappa \left(\frac{\mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}}{h} \right) \frac{[f'(\bar{R}_{\mathbf{i}, \mathbf{t}})]^2 [\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}}) + f(\bar{R}_{\mathbf{i}, \mathbf{t}})][\tilde{f}'(\check{R}_{\mathbf{i}, \mathbf{t}}) - f'(\bar{R}_{\mathbf{i}, \mathbf{t}})]}{f^2(\bar{R}_{\mathbf{i}, \mathbf{t}}) \tilde{f}^2(\check{R}_{\mathbf{i}, \mathbf{t}})} \\
&\quad \cdot G_{n^*}(\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}})) H^{-1} \tilde{X}_{\mathbf{i}, \mathbf{t}} \tilde{X}_{\mathbf{i}, \mathbf{t}}^T H^{-1} \\
&\quad - \frac{1}{n^* h^d} \sum_{\mathbf{i} \in \Lambda_n} \sum_{\mathbf{t} \in \mathbf{T}_n} \kappa \left(\frac{\mathbf{X}_{\mathbf{i}, \mathbf{t}} - \mathbf{x}}{h} \right) \frac{[\tilde{f}'(\check{R}_{\mathbf{i}, \mathbf{t}}) - f'(\bar{R}_{\mathbf{i}, \mathbf{t}})]^2 [\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}}) + f(\bar{R}_{\mathbf{i}, \mathbf{t}})][\tilde{f}'(\check{R}_{\mathbf{i}, \mathbf{t}}) - f'(\bar{R}_{\mathbf{i}, \mathbf{t}})]}{f^2(\bar{R}_{\mathbf{i}, \mathbf{t}}) \tilde{f}^2(\check{R}_{\mathbf{i}, \mathbf{t}})} \\
&\quad \cdot G_{n^*}(\tilde{f}(\check{R}_{\mathbf{i}, \mathbf{t}})) H^{-1} \tilde{X}_{\mathbf{i}, \mathbf{t}} \tilde{X}_{\mathbf{i}, \mathbf{t}}^T H^{-1}
\end{aligned}$$

$$\begin{aligned}
& - \frac{2}{\mathbf{n}^* \mathbf{h}^d} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right) \frac{[\tilde{f}'(\tilde{R}_{\mathbf{i},\mathbf{t}}) - f'(\bar{R}_{\mathbf{i},\mathbf{t}})][\tilde{f}(\tilde{R}_{\mathbf{i},\mathbf{t}}) + f(\bar{R}_{\mathbf{i},\mathbf{t}})][\tilde{f}(\tilde{R}_{\mathbf{i},\mathbf{t}}) - f(\bar{R}_{\mathbf{i},\mathbf{t}})]}{f^2(\bar{R}_{\mathbf{i},\mathbf{t}})\tilde{f}^2(\tilde{R}_{\mathbf{i},\mathbf{t}})} \\
& \cdot G_{\mathbf{n}^*}(\tilde{f}(\tilde{R}_{\mathbf{i},\mathbf{t}}))H^{-1}\tilde{X}_{\mathbf{i},\mathbf{t}}\tilde{X}_{\mathbf{i},\mathbf{t}}^T H^{-1} \\
= & B_1 + B_2 + B_3 + B_4 - B_5 - B_6 - B_7. \tag{5.8}
\end{aligned}$$

下证 $B_1 \xrightarrow{\mathcal{P}} I_{\beta}(\mathbf{x})\mathbf{U}$, $B_i \xrightarrow{\mathcal{P}} 0, i = 2, \dots, 7$, 则引理得证. $B_1 - B_7$ 都含有

$$\mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right)H^{-1}\tilde{X}_{\mathbf{i},\mathbf{t}}\tilde{X}_{\mathbf{i},\mathbf{t}}^T H^{-1} = \begin{pmatrix} M_{0,0} & M_{0,1} & \cdots & M_{0,p} \\ M_{1,0} & M_{1,1} & \cdots & M_{1,p} \\ \vdots & \vdots & & \vdots \\ M_{p,0} & M_{p,1} & \cdots & M_{p,p} \end{pmatrix},$$

其中 M_{kl} 是 $N_k \times N_l$ 维矩阵, 它的 r 行 γ 列元素为 $\mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right)\frac{1}{h^k}(\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x})^{\phi_{\mathbf{k}}(\mathbf{r})} \cdot \frac{1}{h^l}(\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x})^{\phi_{\mathbf{l}}(\gamma)}$.

$$\begin{aligned}
B_1 &= \frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right) \left[\frac{f'(R_{\mathbf{i},\mathbf{t}})}{f(R_{\mathbf{i},\mathbf{t}})}\right]^2 H^{-1}\tilde{X}_{\mathbf{i},\mathbf{t}}\tilde{X}_{\mathbf{i},\mathbf{t}}^T H^{-1} \\
& + \frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right) \left\{ \left[\frac{f'(R_{\mathbf{i},\mathbf{t}} + \delta_{\mathbf{i},\mathbf{t}})}{f(R_{\mathbf{i},\mathbf{t}} + \delta_{\mathbf{i},\mathbf{t}})}\right]^2 - \left[\frac{f'(R_{\mathbf{i},\mathbf{t}})}{f(R_{\mathbf{i},\mathbf{t}})}\right]^2 \right\} H^{-1}\tilde{X}_{\mathbf{i},\mathbf{t}}\tilde{X}_{\mathbf{i},\mathbf{t}}^T H^{-1} \\
= & B_{11} + B_{12}. \tag{5.9}
\end{aligned}$$

为了简单起见, 令 $\xi_{\mathbf{i},\mathbf{t}} = \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right) \left[\frac{f'(R_{\mathbf{i},\mathbf{t}})}{f(R_{\mathbf{i},\mathbf{t}})}\right]^2 \frac{1}{h^k}(\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x})^{\phi_{\mathbf{k}}(\mathbf{r})} \cdot \frac{1}{h^l}(\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x})^{\phi_{\mathbf{l}}(\gamma)}$, 类似于引理 5.4, 有

$$\begin{aligned}
& E\left\{ \frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right) \left[\frac{f'(R_{\mathbf{i},\mathbf{t}})}{f(R_{\mathbf{i},\mathbf{t}})}\right]^2 \left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h^k}\right)^{\phi_{\mathbf{k}}(\mathbf{r})} \cdot \left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h^l}\right)^{\phi_{\mathbf{l}}(\gamma)} \right\} \\
= & I_{\beta}(\mathbf{x})\mu_{\phi_{\mathbf{k}}(\mathbf{r}) + \phi_{\mathbf{l}}(\gamma)}, \\
& \text{Var}\left[\frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} \xi_{\mathbf{i},\mathbf{t}}\right] \\
= & \frac{1}{(n^* h^d)^2} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} E(\xi_{\mathbf{i},\mathbf{t}} - E\xi_{\mathbf{i},\mathbf{t}})^2 + \frac{1}{(n^* h^d)^2} \sum_{(\mathbf{i},\mathbf{t}) \neq (\mathbf{j},\tau)} E(\xi_{\mathbf{i},\mathbf{t}} - E\xi_{\mathbf{i},\mathbf{t}})(\xi_{\mathbf{j},\tau} - E\xi_{\mathbf{j},\tau}) \\
= & O\left[\frac{1}{\mathbf{n}^* \mathbf{h}^d} + \frac{(n^* h_0)^{\frac{1}{2}}}{n^*} + \frac{1}{n^* h^{2d}}\right] \rightarrow 0.
\end{aligned}$$

故有

$$B_{11} \xrightarrow{\mathcal{P}} E(B_{11}) = I_{\beta}(\mathbf{x})\mathbf{U}. \tag{5.10}$$

由 Taylor 展式, 有

$$\begin{aligned}
B_{12} &= \frac{2}{\mathbf{n}^* \mathbf{h}^d} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right) \left\{ \frac{f(R_{\mathbf{i},\mathbf{t}})f'(R_{\mathbf{i},\mathbf{t}})f''(R_{\mathbf{i},\mathbf{t}}) - [f'(R_{\mathbf{i},\mathbf{t}})]^3}{f^3(R_{\mathbf{i},\mathbf{t}})} \delta_{\mathbf{i},\mathbf{t}} + O_{\mathcal{P}}(h^{p+2}) \right\} \\
& \cdot H^{-1}\tilde{X}_{\mathbf{i},\mathbf{t}}\tilde{X}_{\mathbf{i},\mathbf{t}}^T H^{-1}.
\end{aligned}$$

由 $\delta_{i,t}$ 的定义知 $\delta_{i,t} = O_{\mathcal{P}}(h^p)$, 类似地可以证明

$$\begin{aligned} & \mathbb{E}\left\{\frac{2}{\mathbf{n}^*\mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \frac{f(R_{i,t})f'(R_{i,t})f''(R_{i,t}) - [f'(R_{i,t})]^3}{f^3(R_{i,t})} \delta_{i,t} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{1}{h^k} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_{\mathbf{k}}(\mathbf{r})} \right. \\ & \quad \left. \cdot \frac{1}{h^l} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_1(\gamma)}\right\} \\ &= O_{\mathcal{P}}(h^{p+1}) \rightarrow 0, \\ & \text{Var}\left\{\frac{2}{\mathbf{n}^*\mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \frac{f(R_{i,t})f'(R_{i,t})f''(R_{i,t}) - [f'(R_{i,t})]^3}{f^3(R_{i,t})} \delta_{i,t} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \right. \\ & \quad \left. \cdot \frac{1}{h^k} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_{\mathbf{k}}(\mathbf{r})} \frac{1}{h^l} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_1(\gamma)}\right\} \\ &= O\left(\frac{h^{2(p+1)}}{\mathbf{n}^*\mathbf{h}^d} + \frac{h^{2(p+1)}(n^*h_0)^{\frac{1}{2}}}{n^*} + \frac{h^{2(p+1)}}{n^*h^{2d}}\right) \rightarrow 0. \end{aligned}$$

故有

$$\mathbf{B}_{12} \xrightarrow{\mathcal{P}} \mathbb{E}(\mathbf{B}_{12}) = 0. \quad (5.11)$$

由 (5.9) 式和 (5.11) 式知,

$$\mathbf{B}_1 \xrightarrow{\mathcal{P}} I_{\beta}(x)\mathbf{U}. \quad (5.12)$$

$\mathbf{B}_3 - \mathbf{B}_7$ 项要么含有 $\tilde{f}(\check{R}_{i,t}) - f(\bar{R}_{i,t})$, 要么含有 $\tilde{f}'(\check{R}_{i,t}) - f'(\bar{R}_{i,t})$. 易得 $\zeta_{i,t} = O_{\mathcal{P}}(h^{p+1} + n^{*-\frac{1}{2}}h_1^{-\frac{d}{2}})$, 由引理 5.4 知,

$$\begin{aligned} \tilde{f}(\check{R}_{i,t}) - f(\bar{R}_{i,t}) &= \tilde{f}(\check{R}_{i,t}) - \tilde{f}(\bar{R}_{i,t}) + \tilde{f}(\bar{R}_{i,t}) - f(\bar{R}_{i,t}) \\ &= \frac{1}{n^*h_0^2} \sum_{j \in \Lambda_n} \sum_{\tau \in \mathbf{T}_n} \zeta_{i,t} K' \left(\frac{\check{R}_{j,\tau} - \bar{R}_{i,t}}{h_0} \right) + O_{\mathcal{P}}(n^{*-\frac{1}{2}}h_1^{-\frac{d}{2}} + h_0^2) \\ &= O_{\mathcal{P}}(h^{p+1} + n^{*-\frac{1}{2}}h_1^{-\frac{d}{2}} + h_0^2). \end{aligned} \quad (5.13)$$

对任意 $z \in \mathbb{R}$, 有

$$\tilde{f}'(z) - f'(z) = -\frac{1}{n^*h_0^2} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} K' \left(\frac{\check{R}_{i,t} - z}{h_0} \right) - f'(z).$$

由分部积分法和 Taylor 展式, 有

$$\begin{aligned} & \mathbb{E}\left[-\frac{1}{n^*h_0^2} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} K' \left(\frac{\check{R}_{i,t} - z}{h_0} \right)\right] \\ &= -\frac{1}{h_0^2} \int_{\mathbb{R}} K' \left(\frac{y}{h_0} \right) f(y) dy \\ &= \frac{1}{h_0} \int_{\mathbb{R}} K \left(\frac{y}{h_0} \right) f'(y+z) dy = f'(z) + f'''(z)h_0^2 \int_{\mathbb{R}} u^2 K(u) du + o(h_0^2), \\ & \text{Var}\left[-\frac{1}{n^*h_0^2} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} K' \left(\frac{\check{R}_{i,t} - z}{h_0} \right)\right] = O\left(\frac{1}{n^*h_0^3} + \frac{(n^*h_0)^{\frac{1}{2}}}{n^{*2}h_0^4} + \frac{1}{n^*h_0^4}\right) \rightarrow 0. \end{aligned}$$

故类似可得

$$\tilde{f}'(z) - f'(z) = O_{\mathcal{P}}(h_0^2), \quad (5.14)$$

$$\begin{aligned}
\tilde{f}'(\check{R}_{i,t}) - f'(\bar{R}_{i,t}) &= \tilde{f}'(\check{R}_{i,t}) - \tilde{f}'(\bar{R}_{i,t}) + \tilde{f}'(\bar{R}_{i,t}) - f'(\bar{R}_{i,t}) \\
&= \frac{1}{n^*h_0^3} \sum_{j \in \Lambda_n} \sum_{\tau \in T_n} \zeta_{i,t} K'' \left(\frac{\tilde{R}_{j,\tau} - \bar{R}_{i,t}}{h_0} \right) + O_{\mathcal{P}}(h_0^2) \\
&= O_{\mathcal{P}}(h^{p+1} + n^{*-\frac{1}{2}}h_1^{-\frac{d}{2}} + h_0^2). \tag{5.15}
\end{aligned}$$

由 (5.13) 式和 (5.15) 式, 类似证 $B_1 \xrightarrow{\mathcal{P}} I_{\beta}(x)\mathbf{U}$, 我们可以证 B_3, B_4, B_5, B_6, B_7 都以概率收敛于 0. 由引理 5.3 和引理 5.4, 有

$$G_{\mathbf{n}^*}(\tilde{f}(\check{R}_{i,t})) = G_{\mathbf{n}^*}(f(R_{i,t})) + o_{\mathcal{P}}(1).$$

而根据 $G_{\mathbf{n}^*}(\cdot)$ 的表达式, 有

$$\begin{aligned}
E[1 - G_{\mathbf{n}^*}(f(R_{i,t}))] &< \mathcal{P}(f(R_{i,t}) < 2c_0n^{-\varepsilon}) = \int_{f(z) < 2c_0n^{-\varepsilon}} f(z)dz = o(1), \\
\text{Var}[1 - G_{\mathbf{n}^*}(f(R_{i,t}))] &< \mathcal{P}(f(R_{i,t}) < 2c_0n^{-\varepsilon}) = \int_{f(z) < 2c_0n^{-\varepsilon}} f(z)dz = o(1).
\end{aligned}$$

故

$$1 - G_{\mathbf{n}^*}(\tilde{f}(\check{R}_{i,t})) = o_{\mathcal{P}}(1). \tag{5.16}$$

类似可证 $B_2 \xrightarrow{\mathcal{P}} 0$. 由 (5.8) 式, 引理 5.5 得证.

引理 5.6 假设 (A1)–(A8) 成立, 有

$$H^{-1}[\tilde{I}_{\mathbf{n}^*}(\tilde{\beta}, \tilde{\mathbf{f}}) - \tilde{I}_{\mathbf{n}^*}(\beta, \tilde{\mathbf{f}})]H^{-1} \xrightarrow{\mathcal{P}} \mathbf{0}.$$

证 方便起见, 令 $\psi(z) = \left[\frac{\tilde{f}'(z)}{\tilde{f}(z)} \right]^2$, 则由 Taylor 展式, 有

$$\begin{aligned}
&H^{-1}[\tilde{I}_{\mathbf{n}^*}(\tilde{\beta}, \tilde{\mathbf{f}}) - \tilde{I}_{\mathbf{n}^*}(\beta, \tilde{\mathbf{f}})]H^{-1} \\
&= \frac{1}{\mathbf{n}^*h^d} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K} \left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h} \right) \left\{ \left[\frac{\tilde{f}'(\check{R}_{i,t})}{\tilde{f}(\check{R}_{i,t})} \right]^2 - \left[\frac{\tilde{f}'(\bar{R}_{i,t})}{\tilde{f}(\bar{R}_{i,t})} \right]^2 \right\} G_{\mathbf{n}^*}(\tilde{f}(\check{R}_{i,t})) H^{-1} \tilde{X}_{i,t} \tilde{X}_{i,t}^T H^{-1} \\
&= \frac{-2}{\mathbf{n}^*h^d} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \zeta_{i,t} \mathcal{K} \left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h} \right) \psi'(\bar{R}_{i,t}) G_{\mathbf{n}^*}(\tilde{f}(\check{R}_{i,t})) H^{-1} \tilde{X}_{i,t} \tilde{X}_{i,t}^T H^{-1} \\
&\quad + o_{\mathcal{P}} \left(h^{p+1} + \frac{1}{(n^*h_1^d)^{\frac{1}{2}}} \right).
\end{aligned}$$

类似于引理 5.5, 易证

$$\begin{aligned}
&E \left[\frac{1}{\mathbf{n}^*h^d} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K} \left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h} \right) \psi'(\bar{R}_{i,t}) G_{\mathbf{n}^*}(\tilde{f}(\check{R}_{i,t})) \frac{1}{h^k} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_{\mathbf{k}}(\mathbf{r})} \cdot \frac{1}{h^l} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_1(\gamma)} \right] \\
&= \mu_{\phi_{\mathbf{k}}(\mathbf{r}) + \phi_1(\gamma)} \int_{\mathbb{R}} \frac{f'(z)f''(z)f'''(z) - [f'(z)]^3}{f^2(z)} dz = O_{\mathcal{P}}(1), \\
&\text{Var} \left[\frac{1}{\mathbf{n}^*h^d} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K} \left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h} \right) \psi'(\bar{R}_{i,t}) G_{\mathbf{n}^*}(\tilde{f}(\check{R}_{i,t})) \left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h^k} \right)^{\phi_{\mathbf{k}}(\mathbf{r})} \cdot \left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h^l} \right)^{\phi_1(\gamma)} \right] \\
&= O \left(\frac{1}{\mathbf{n}^*h^d} + \frac{(n^*h_0)^{\frac{1}{2}}}{n^*} + \frac{1}{n^*h^{2d}} \right) \rightarrow 0.
\end{aligned}$$

因 $\zeta_{i,t} = O_{\mathcal{P}} \left(h^{p+1} + \frac{1}{(n^*h_1^d)^{\frac{1}{2}}} \right)$, 故 $H^{-1}[\tilde{I}_{\mathbf{n}^*}(\tilde{\beta}, \tilde{\mathbf{f}}) - \tilde{I}_{\mathbf{n}^*}(\beta, \tilde{\mathbf{f}})]H^{-1} \xrightarrow{\mathcal{P}} \mathbf{0}$, 得证.

令

$$Q(\beta, \tilde{f}) = \frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \kappa\left(\frac{X_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}''(\bar{R}_{i,t})}{\tilde{f}(\bar{R}_{i,t})} G_{\mathbf{n}^*}(\tilde{f}(\bar{R}_{i,t})) \tilde{X}_{i,t} \tilde{X}_{i,t}^T \\ + \frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \kappa\left(\frac{X_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}'^2(\bar{R}_{i,t})}{\tilde{f}(\bar{R}_{i,t})} G'_{\mathbf{n}^*}(\tilde{f}(\bar{R}_{i,t})) \tilde{X}_{i,t} \tilde{X}_{i,t}^T.$$

引理 5.7 假设 (A1)–(A8) 成立, 有

$$\sqrt{\mathbf{n}^* \mathbf{h}^d} [H^{-1} \tilde{I}_{\mathbf{n}^*}(\tilde{\beta}, \tilde{\mathbf{f}}) H^{-1}]^{-1} H^{-1} Q(\beta, \tilde{\mathbf{f}}) (\tilde{\beta} - \beta) \xrightarrow{\mathcal{P}} \mathbf{0}.$$

证 由假设条件 (A1) 和 (A4), 类似引理 5.4 知

$$\mathbb{E} \left[\frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \kappa\left(\frac{X_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}''(\bar{R}_{i,t})}{\tilde{f}(\bar{R}_{i,t})} G_{\mathbf{n}^*}(\tilde{f}(\bar{R}_{i,t})) \frac{1}{h^k} (X_{i,t} - \mathbf{x})^{\phi_k(r)} \cdot \frac{1}{h^l} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_l(\gamma)} \right. \\ \left. + \frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \kappa\left(\frac{X_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}'^2(\bar{R}_{i,t})}{\tilde{f}(\bar{R}_{i,t})} G'_{\mathbf{n}^*}(\tilde{f}(\bar{R}_{i,t})) \frac{1}{h^k} (X_{i,t} - \mathbf{x})^{\phi_k(r)} \cdot \frac{1}{h^l} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_l(\gamma)} \right] \\ = \mu_{\phi_k(r) + \phi_l(\gamma)} \cdot \int_{\mathbb{R}} f''(z) dz,$$

且

$$\text{Var} \left[\frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \kappa\left(\frac{X_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}''(\bar{R}_{i,t})}{\tilde{f}(\bar{R}_{i,t})} G_{\mathbf{n}^*}(\tilde{f}(\bar{R}_{i,t})) \frac{1}{h^k} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_k(r)} \cdot \frac{1}{h^l} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_l(\gamma)} \right. \\ \left. + \frac{1}{\mathbf{n}^* \mathbf{h}^d} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \kappa\left(\frac{X_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}'^2(\bar{R}_{i,t})}{\tilde{f}(\bar{R}_{i,t})} G'_{\mathbf{n}^*}(\tilde{f}(\bar{R}_{i,t})) \frac{1}{h^k} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_k(r)} \cdot \frac{1}{h^l} (\mathbf{X}_{i,t} - \mathbf{x})^{\phi_l(\gamma)} \right] \\ = O\left(\frac{1}{\mathbf{n}^* \mathbf{h}^d} + \frac{(n^* h_0)^{\frac{1}{2}}}{n^*} + \frac{1}{n^* h^{2d}}\right) \rightarrow 0.$$

故有

$$H^{-1} Q(\beta, \tilde{f}) H^{-1} \xrightarrow{\mathcal{P}} \mathbf{O}_{\mathcal{P}}(\mathbf{1}), \quad (5.17)$$

其中 $\mathbf{1}$ 表示 $N_0 + \dots + N_p$ 维向量, 分量全是 1. 由引理 5.3 知, $(n^* h_1^d)^{\frac{1}{2}} H(\tilde{\beta} - \beta) = \mathbf{O}_{\mathcal{P}}(\mathbf{1})$, 由假设条件, 有 $(n^* h^d)^{\frac{1}{2}} H(\tilde{\beta} - \beta) \xrightarrow{\mathcal{P}} \mathbf{0}$. 再由引理 5.5 和 (5.17) 式, 引理 5.7 得证.

引理 5.8 假设成立, 有

$$\sqrt{\mathbf{n}^* \mathbf{h}^d} [H^{-1} \tilde{S}_{\mathbf{n}^*}(\beta, \tilde{\mathbf{f}}) - \mathbf{a} \mathbf{h}^{\mathbf{p}+1} \mathbf{U}^{-1} \bar{\mathbf{U}} \mathbf{m}^{(\mathbf{p}+1)}(\mathbf{x}) \mathbf{t}] \xrightarrow{\mathcal{P}} \mathbf{N}(\mathbf{0}, \mathbf{I}_{\beta}^{-1}(\mathbf{x}) \mathbf{U}^{-1} \boldsymbol{\Sigma} \mathbf{U}^{-1}) \\ \xrightarrow{\mathcal{P}} N(0, I_{\beta}(\mathbf{x}) \boldsymbol{\Sigma}).$$

证 由 $\tilde{S}_{\mathbf{n}^*}(\beta, \tilde{f})$ 的定义, 将 $\tilde{f}'(\bar{R}_{i,t}) / \tilde{f}(\bar{R}_{i,t})$ 分解如下:

$$\tilde{f}'(\bar{R}_{i,t}) = f'(\bar{R}_{i,t}) + \tilde{f}'(\bar{R}_{i,t}) - f'(\bar{R}_{i,t}), \\ \frac{1}{\tilde{f}(\bar{R}_{i,t})} = \frac{[\tilde{f}(\bar{R}_{i,t}) - f(\bar{R}_{i,t})]^2}{\tilde{f}(\bar{R}_{i,t}) f^2(\bar{R}_{i,t})} + \frac{1}{f(\bar{R}_{i,t})} - \frac{\tilde{f}(\bar{R}_{i,t}) - f(\bar{R}_{i,t})}{f^2(\bar{R}_{i,t})}, \\ \sqrt{\mathbf{n}^* \mathbf{h}^d} H^{-1} \tilde{S}_{\mathbf{n}^*}(\beta, \tilde{f}) = -\frac{1}{\sqrt{\mathbf{n}^* \mathbf{h}^d}} \sum_{i \in \Lambda_n} \sum_{t \in \mathbf{T}_n} \kappa\left(\frac{X_{i,t} - \mathbf{x}}{h}\right) \frac{f'(\bar{R}_{i,t})}{f(\bar{R}_{i,t})} \mathbf{H}^{-1} \tilde{X}_{i,t}$$

$$\begin{aligned}
& + \frac{1}{\sqrt{n^*h^d}} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{f'(\bar{R}_{i,t})}{f(\bar{R}_{i,t})} [1 - G_{n^*}(\tilde{f}(\bar{R}_{i,t}))] H^{-1} \tilde{X}_{i,t} \\
& - \frac{1}{\sqrt{n^*h^d}} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}'(\bar{R}_{i,t}) - f'(\bar{R}_{i,t})}{f(\bar{R}_{i,t})} \\
& \cdot G_{n^*}(\tilde{f}(\bar{R}_{i,t})) H^{-1} \tilde{X}_{i,t} \\
& + \frac{1}{\sqrt{n^*h^d}} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}'(\bar{R}_{i,t})[\tilde{f}(\bar{R}_{i,t}) - f(\bar{R}_{i,t})]}{f^2(\bar{R}_{i,t})} \\
& \cdot G_{n^*}(\tilde{f}(\bar{R}_{i,t})) H^{-1} \tilde{X}_{i,t} \\
& - \frac{1}{\sqrt{n^*h^d}} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{\tilde{f}'(\bar{R}_{i,t})[\tilde{f}(\bar{R}_{i,t}) - f(\bar{R}_{i,t})]^2}{f(\bar{R}_{i,t})f^2(\bar{R}_{i,t})} \\
& \cdot G_{n^*}(\tilde{f}(\bar{R}_{i,t})) H^{-1} \tilde{X}_{i,t} \\
& = D_1 + D_2 + D_3 + D_4 + D_5. \tag{5.18}
\end{aligned}$$

类似于引理 5.4 的证明, 由 (5.16) 式, 我们易得 $D_2 \xrightarrow{\mathcal{P}} \mathbf{0}$, 由 (5.14) 式和假设条件 (A6), 有 $D_3 \xrightarrow{\mathcal{P}} \mathbf{0}$. 由引理 5.4 和假设条件 (A6), 有 $D_4 \xrightarrow{\mathcal{P}} \mathbf{0}$, $D_5 \xrightarrow{\mathcal{P}} \mathbf{0}$. 所以下面只需要 $D_1 - \sqrt{n^*h^d}h^{p+1}\zeta(\mathbf{x}) \xrightarrow{\mathcal{P}} \mathbf{N}(\mathbf{0}, \mathbf{I}_\beta(\mathbf{x})\boldsymbol{\Sigma})$, 则引理 5.8 成立. 类似分解 $f'(\bar{R}_{i,t})$, $\frac{1}{f(\bar{R}_{i,t})}$ 为

$$f'(\bar{R}_{i,t}) = [f'(\bar{R}_{i,t}) - f'(R_{i,t})] + f'(R_{i,t}), \quad -\frac{1}{f(\bar{R}_{i,t})} = \frac{f(\bar{R}_{i,t}) - f(R_{i,t})}{f(\bar{R}_{i,t})f(R_{i,t})} - \frac{1}{f(R_{i,t})},$$

则

$$\begin{aligned}
D_1 & = -\frac{1}{\sqrt{n^*h^d}} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{f'(R_{i,t})}{f(R_{i,t})} H^{-1} \tilde{X}_{i,t} \\
& - \frac{1}{\sqrt{n^*h^d}} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{f'(R_{i,t} + \delta_{i,t}) - f'(R_{i,t})}{f(R_{i,t})} H^{-1} \tilde{X}_{i,t} \\
& + \frac{1}{\sqrt{n^*h^d}} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{f'(R_{i,t})[f(R_{i,t} + \delta_{i,t}) - f(R_{i,t})]}{f(\bar{R}_{i,t})f(R_{i,t})} H^{-1} \tilde{X}_{i,t} \\
& + \frac{1}{\sqrt{n^*h^d}} \sum_{i \in \Lambda_n} \sum_{t \in T_n} \mathcal{K}\left(\frac{\mathbf{X}_{i,t} - \mathbf{x}}{h}\right) \frac{[f'(R_{i,t} + \delta_{i,t}) - f'(R_{i,t})][f(R_{i,t} + \delta_{i,t}) - f(R_{i,t})]}{f(\bar{R}_{i,t})f(R_{i,t})} \\
& \cdot H^{-1} \tilde{X}_{i,t} \\
& = D_{11} + D_{12} + D_{13} + D_{14}. \tag{5.19}
\end{aligned}$$

类似于引理 5.4, 我们可以证明 $D_{14} = o_{\mathcal{P}}(\sqrt{n^*h^d}h^{p+1})$,

$$\begin{aligned}
D_{12} & \simeq -\sqrt{n^*h^d}h^{p+1}f_X(\mathbf{x}) \int_{\mathbb{R}} \mathbf{f}''(\mathbf{z})d\mathbf{z} \bar{\mathbf{U}}\mathbf{m}^{(p+1)}(\mathbf{x}), \\
D_{13} & \simeq \sqrt{n^*h^d}h^{p+1}f_X(\mathbf{x}) \int_{\mathbb{R}} \frac{\mathbf{f}'^2(\mathbf{z})}{\mathbf{f}(\mathbf{z})}d\mathbf{z} \bar{\mathbf{U}}\mathbf{m}^{(p+1)}(\mathbf{x}).
\end{aligned}$$

下面只需证 $D_{11} \xrightarrow{\mathcal{L}} \mathbf{N}(0, \mathcal{I}_\beta(\mathbf{x})\boldsymbol{\Sigma})$, 由假设条件, 有

$$E D_{11} = -\sqrt{n^*h^d}f_X(\mathbf{x}) \int_{\mathbb{R}} \mathbf{f}'(\mathbf{z})d\mathbf{z} (\mathbf{U}_{0,0}, \dots, \mathbf{U}_{p,0})^T = \mathbf{0},$$

$$\text{Var}(\mathbf{D}_{11}) = \mathcal{I}_\beta(\mathbf{x})\boldsymbol{\Sigma}.$$

利用 Cramér-Wold 工具, 只需证对任意 $C \in \mathbb{R}^{N_0+\dots+N_p}$ ($C \neq \mathbf{0}$), 有

$$C^T \mathbf{D}_{11} \xrightarrow{\mathcal{L}} N(0, \mathcal{I}_\beta(\mathbf{x})C^T \boldsymbol{\Sigma} C).$$

为了简单起见, 记

$$C^T \mathbf{D}_{11} = -\frac{1}{\sqrt{\mathbf{n}^* \mathbf{h}^d}} \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} \mathcal{K}\left(\frac{\mathbf{X}_{\mathbf{i},\mathbf{t}} - \mathbf{x}}{h}\right) \frac{f'(R_{\mathbf{i},\mathbf{t}})}{f(R_{\mathbf{i},\mathbf{t}})} C^T H^{-1} \tilde{X}_{\mathbf{i},\mathbf{t}} = \sum_{\mathbf{i} \in \Lambda_{\mathbf{n}}} \sum_{\mathbf{t} \in \mathbf{T}_{\mathbf{n}}} \Delta_{\mathbf{i},\mathbf{t}}.$$

利用时空分块思想 (参见文 [4]), 大块中所有的边等于 $p_k, k = 1, \dots, N + 1$, 而小块中至少有一边等于 q . 不失一般性, 假设存在正整数 l_1, \dots, l_{N+1} , 使得 $n_k = l_k(p_k + q)$ 且满足 $\lim_{n \rightarrow \infty} \frac{p_k}{q} = \infty, k = 1, \dots, N + 1$.

记

$$\begin{aligned} U(1, \mathbf{n}, \mathbf{j}) &= \sum_{\substack{i_k = j_k(p_k + q) + 1 \\ k = 1, \dots, N}}^{j_k(p_k + q) + p_k} \sum_{t = j_{N+1}(p_{N+1} + q) + 1}^{j_{N+1}(p_{N+1} + q) + p_{N+1}} \Delta_{\mathbf{i}, \mathbf{t}}, \\ U(2, \mathbf{n}, \mathbf{j}) &= \sum_{\substack{i_k = j_k(p_k + q) + 1 \\ k = 1, \dots, N}}^{j_k(p_k + q) + p_k} \sum_{t = j_{N+1}(p_{N+1} + q) + p_{N+1} + 1}^{(j_{N+1} + 1)(p_{N+1} + q)} \Delta_{\mathbf{i}, \mathbf{t}}, \\ U(3, \mathbf{n}, \mathbf{j}) &= \sum_{\substack{i_k = j_k(p_k + q) + 1 \\ k = 1, \dots, N - 1}}^{j_k(p_k + q) + p_k} \sum_{i_N = j_N(p_N + q) + p_N + 1}^{(j_N + 1)(p_N + q)} \sum_{t = j_{N+1}(p_{N+1} + q) + 1}^{j_{N+1}(p_{N+1} + q) + p_{N+1}} \Delta_{\mathbf{i}, \mathbf{t}}, \\ U(4, \mathbf{n}, \mathbf{j}) &= \sum_{\substack{i_k = j_k(p_k + q) + 1 \\ k = 1, \dots, N - 1}}^{j_k(p_k + q) + p_k} \sum_{i_N = j_N(p_N + q) + p_N + 1}^{(j_N + 1)(p_N + q)} \sum_{t = j_{N+1}(p_{N+1} + q) + p_{N+1} + 1}^{(j_{N+1} + 1)(p_{N+1} + q)} \Delta_{\mathbf{i}, \mathbf{t}}, \end{aligned}$$

等等. 而

$$U(2^N, \mathbf{n}, \mathbf{j}) = \sum_{\substack{i_k = j_k(p_k + q) + p_k + 1 \\ k = 1, \dots, N}}^{(\mathbf{j} + \mathbf{1})(\mathbf{p} + \mathbf{q})} \sum_{t = \mathbf{j}_{N+1}(\mathbf{p}_{N+1} + \mathbf{q}) + \mathbf{1}^{\mathbf{j}_{N+1}(\mathbf{p}_{N+1} + \mathbf{q}) + \mathbf{p}_{N+1}}} \Delta_{\mathbf{i}, \mathbf{t}}.$$

最终

$$U(2^{N+1}, \mathbf{n}, \mathbf{j}) = \sum_{\substack{i_k = j_k(p_k + q) + p_k + 1 \\ k = 1, \dots, N}}^{(\mathbf{j} + \mathbf{1})(\mathbf{p} + \mathbf{q})} \sum_{t = \mathbf{j}_{N+1}(\mathbf{p}_{N+1} + \mathbf{q}) + \mathbf{p}_{N+1} + \mathbf{1}}^{(\mathbf{j}_{N+1} + \mathbf{1})(\mathbf{p}_{N+1} + \mathbf{q})} \Delta_{\mathbf{i}, \mathbf{t}}.$$

令 $\mathcal{R} = \{0, \dots, l_1\} \times \dots \times \{0, \dots, l_{N+1}\}$, 对每个正整数 $i = 1, \dots, 2^{N+1}$, 定义

$$T(\mathbf{n}, \mathbf{i}) = \sum_{\mathbf{j} \in \mathcal{R}} \mathbf{U}(\mathbf{i}, \mathbf{n}, \mathbf{j}),$$

则有

$$C^T \mathbf{D}_{11} = \sum_{i=1}^{2^{N+1}} T(\mathbf{n}, \mathbf{i}).$$

我们注意到 $T(n, 1)$ 是大块随机变量 $\Delta_{i,t}$ 的和, 而所有 $T(n, i)$, $2 \leq i \leq 2^{N+1}$ 是小块随机变量 $\Delta_{i,t}$ 的和. 若对正整数 l_1, \dots, l_{N+1} 不是正好 $n_k = l_k(p_k + q)$, $k = 1, \dots, N + 1$, 那么可另添加一项 $T(n, 2^{d+1} + 1)$. 在这一项中, 包含所有的既不在大块中也不在小块中的随机变量 $\Delta_{i,t}$. 添加这一项后, 证明无需做太大的改变. 类似于文 [4] 中的证明, 当 $\mathbf{n} \rightarrow +\infty$ 时, 有

$$\begin{aligned} \mathbf{Q}_1 &= \left| \mathbb{E} \exp[iuT(\mathbf{n}, \mathbf{1})] - \prod_{\substack{j_k=0 \\ k=1, \dots, N+1}}^{l_k-1} \mathbb{E} \exp[iuU(\mathbf{1}, \mathbf{n}, \mathbf{j})] \right| \rightarrow 0, \\ \mathbf{Q}_2 &= \left(\sum_{i=2}^{2^{N+1}} T(\mathbf{n}, \mathbf{1}) \right)^2 \rightarrow 0, \\ \mathbf{Q}_3 &= \sum_{\substack{j_k=0 \\ k=1, \dots, N+1}}^{l_k-1} \mathbb{E}[U(\mathbf{1}, \mathbf{n}, \mathbf{j})]^2 \rightarrow \mathcal{I}_\beta(\mathbf{x})\mathbf{C}^T \Sigma \mathbf{C}, \\ \mathbf{Q}_4 &= \sum_{\substack{j_k=0 \\ k=1, \dots, N+1}}^{l_k-1} \mathbb{E}[(U(\mathbf{1}, \mathbf{n}, \mathbf{j}))^2 \mathbf{I}\{|\mathbf{U}(\mathbf{1}, \mathbf{n}, \mathbf{j})| > \varepsilon \mathcal{I}_\beta(\mathbf{x})\mathbf{C}^T \Sigma \mathbf{C}\}] \rightarrow 0. \end{aligned}$$

故引理 5.8 得证.

定理 3.1 的证明 为简单起见, 令

$$\tilde{s}_{i,t}(\tilde{\beta}) = \frac{\tilde{f}'(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\tilde{\beta}, \mathbf{X}_{i,t} - \mathbf{x}))}{\tilde{f}(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\tilde{\beta}, \mathbf{X}_{i,t} - \mathbf{x}))} G_{\mathbf{n}^*}(\tilde{f}(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\tilde{\beta}, \mathbf{X}_{i,t} - \mathbf{x}))).$$

$\tilde{s}_{i,t}(\tilde{\beta})$ 在 β 处 Taylor 展开为

$$\begin{aligned} \tilde{s}_{i,t}(\tilde{\beta}) &= \tilde{s}_{i,t}(\beta) + \frac{1}{2}(\tilde{\beta} - \beta)^T \left[\frac{\partial^2 \tilde{s}_{i,t}(\beta^*)}{\partial \beta \partial \beta^T} \right] (\tilde{\beta} - \beta) \\ &\quad - \frac{\tilde{f}''(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))}{\tilde{f}'(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))} G_{\mathbf{n}^*}(\tilde{f}(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))) \tilde{\mathbf{X}}_{i,t}(\tilde{\beta} - \beta) \\ &\quad + \frac{\tilde{f}''^2(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))}{\tilde{f}^2(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))} G_{\mathbf{n}^*}(\tilde{f}(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))) \tilde{\mathbf{X}}_{i,t}(\tilde{\beta} - \beta) \\ &\quad - \frac{\tilde{f}'(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))}{\tilde{f}(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))} G_{\mathbf{n}^*}'(\tilde{f}(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x}))) \\ &\quad \cdot \tilde{\mathbf{f}}'(\mathbf{Y}_{i,t} - \mathcal{P}_{\mathbf{p}}(\beta, \mathbf{X}_{i,t} - \mathbf{x})) \tilde{\mathbf{X}}_{i,t}(\tilde{\beta} - \beta). \end{aligned}$$

由 (2.5) 式和上式, 我们有

$$\begin{aligned} &\sqrt{\mathbf{n}^* \mathbf{h}^d} \{H[\hat{\beta}(\mathbf{x}) - \beta(\mathbf{x})]\} \\ &= \sqrt{\mathbf{n}^* \mathbf{h}^d} H \tilde{\mathcal{L}}_{\mathbf{n}^*}^{-1}(\tilde{\beta}, \tilde{\mathbf{f}}) \tilde{\mathbf{S}}_{\mathbf{n}^*}(\tilde{\beta}, \tilde{\mathbf{f}}) + \sqrt{\mathbf{n}^* \mathbf{h}^d} H[\hat{\beta}(\mathbf{x}) - \beta(\mathbf{x})] \\ &= \sqrt{\mathbf{n}^* \mathbf{h}^d} [H^{-1} \tilde{\mathcal{L}}_{\mathbf{n}^*}(\tilde{\beta}; \tilde{\mathbf{f}}) H^{-1}]^{-1} H^{-1} \tilde{\mathbf{S}}_{\mathbf{n}^*}(\tilde{\beta}, \tilde{\mathbf{f}}) + \sqrt{\mathbf{n}^* \mathbf{h}^d} [H^{-1} \tilde{\mathcal{L}}_{\mathbf{n}^*}(\tilde{\beta}; \tilde{\mathbf{f}}) H^{-1}]^{-1} \\ &\quad \cdot H^{-1} \mathbf{Q}(\beta, \tilde{f})(\tilde{\beta} - \beta) + \sqrt{\mathbf{n}^* \mathbf{h}^d} [H^{-1} \tilde{\mathcal{L}}_{\mathbf{n}^*}(\tilde{\beta}; \tilde{\mathbf{f}}) H^{-1}]^{-1} H^{-1} [\tilde{\mathcal{L}}_{\mathbf{n}^*}(\tilde{\beta}; \tilde{\mathbf{f}}) - \tilde{\mathcal{L}}_{\mathbf{n}^*}(\beta; \tilde{\mathbf{f}})] (\tilde{\beta} - \beta). \end{aligned}$$

由引理 5.5- 引理 5.8, 定理 3.1 得证.

6 结 语

本文基于误差分布未知的时空模型, 结合一般空间的非参数自适应估计方法, 给出了用以解决误差分布未知下时空模型稳健问题的估计方法, 并建立了估计量的渐近性质, 其渐近方差与已知误差具体分布下的极大似然估计方差一致. 数值模拟的结果显示, 在误差分布未知的条件下, 本文给出的估计方法较常用的局部多项式估计有更好的估计表现, 且小样本下表现更为突出. 最后通过一个实际例子验证了非参数自适应方法的有效性和可行性. 事实上, 本文给出的估计方法可推广到含有缺失数据的时空模型中; 此外, 本文只考虑了同期一般误差项时空数据估计的情形, 在未来的研究中可将类似的方法考虑到包含自回归误差项的时空模型中.

致谢 感谢编辑和审稿人对本文提出的宝贵意见.

参 考 文 献

- [1] Li Y H, Wang N, Hong M, et al. Nonparametric estimation of correlation functions in longitudinal and spatial data, with application to colon carcinogenesis experiments [J]. *The Annals of Statistics*, 2007, 35(4):1608–1643.
- [2] Heaton M J, Berrett C, Pugh S, et al. Modeling bronchiolitis incidence proportions in the presence of spatio-temporal uncertainty [J]. *Journal of the American Statistical Association*, 2019, 115(529):66–78.
- [3] Hallin M, Lu Z D, Tran L T. Local linear spatial regression [J]. *The Annals of Statistics*, 2004, 32(6):2469–2500.
- [4] Wang H X, Wang J D. Estimation of the trend function for spatio-temporal models [J]. *Journal of Nonparametric Statistics*, 2009, 21(5):567–588.
- [5] Yue H, Rilett L R, Revesz P Z. Spatio-Temporal traffic video data archiving and retrieval system [J]. *GeoInformatica*, 2016, 20(3):59–94.
- [6] Zhou Z, Matteson D S, Woodard D B, et al. A spatio-temporal point process model for ambulance demand [J]. *Journal of the American Statistical Association*, 2015, 110(509):6–15.
- [7] Stein C. Efficient nonparametric testing and estimation [C]. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Berkeley: University of California Press, 1956:187–195.
- [8] Bickel P J. On adaptive estimation [J]. *The Annals of Statistics*, 1982, 10(3):647–671.
- [9] Ling S, McAleer M. On adaptive estimation in nonstationary ARMA models with GARCH errors [J]. *Annals of Statistics*, 2003, 31(2):642–674.

- [10] Xu A P, Zhang Q H. Nonlinear system fault diagnosis based on adaptive estimation [J]. *Automatica*, 2004, 40(7):1181–1193.
- [11] Jeganathan P. Some aspects of asymptotic theory with applications to time series models [J]. *Econometric Theory*, 1995, 11(5):818–887.
- [12] Linton O, Xiao Z J. A nonparametric regression estimator that adapts to error distribution of unknown form [J]. *Econometric Theory*, 2007, 23(3):371–413.
- [13] Linton O, Mammen E, Nielsen J P, et al. Nonparametric regression with filtered data [J]. *Bernoulli*, 2011, 17(1):60–87.
- [14] Fan J Q, Gijbels I. Local polynomial modelling and its applications [M]. London: Chapman and Hall, 1996.
- [15] Fan J Q, Farnen M, Gijbels I. Local maximum likelihood estimation and inference [J]. *Journal of the Royal Statistical Society*, 1998, 60(3):591–608.
- [16] Fan J Q, Chen J W. One-Step local quasi-likelihood estimation [J]. *Journal of the Royal Statistical Society Series B*, 1999, 61(4):927–943.
- [17] Masry E. Multivariate local polynomial regression for time series: uniform strong consistency and rates [J]. *Journal of Time Series Analysis*, 1996, 17(6):571–599.
- [18] Masry E. Multivariate regression estimation: local polynomial fitting for time series [J]. *Stochastic Processes & Their Applications*, 1996, 65(1):81–101.
- [19] Kreiss J. On adaptive estimation in stationary ARMA process [J]. *The Annals of Statistics*, 1987, 15(1):112–133.
- [20] Xia Y C. Asymptotic distributions for two estimators of the single-index model [J]. *Econometric Theory*, 2006, 22(6):1112–1137.
- [21] Wang H X, Wang J D, Huang B. Prediction for spatio-temporal models with autoregression in errors [J]. *Journal of Nonparametric Statistics*, 2012, 24(1–2):217–244.
- [22] Cressie N A C, Wikle C K. Statistics for spatio-temporal data [M]. New York: Wiley, 2011.
- [23] Cressie N A C. Statistics for spatial data [M]. New York: Wiley, 1993.
- [24] Ibragimov I A, Linnik Y V. Independent and stationary sequences of random variables [M]. Groningen: Wolters-Nordhoff, 1971.
- [25] Deo C M. A note on empirical process of strong mixing sequences [J]. *The Annals of Statistics*, 1973, 1(5):870–875.

Adaptive Nonparametric Estimation of Spatio-Temporal Models with Unknown Error Distributions

WANG Hongxia¹ LUO Xuehong¹ LIN Jinguan² TANG Xing¹

¹School of Statistics and Mathematics, Nanjing Audit University, Nanjing 211815, China.

E-mail: hxwang@nau.edu.cn; xueh_luo@163.com; xtang_tj@163.com

²Corresponding author. School of Statistics and Mathematics, Nanjing Audit University, Nanjing 211815, China. E-mail: jglin@nau.edu.cn

Abstract Maximum likelihood is an effective method for parameter estimation, but it can't work with unknown error distribution. On the other hand, when the data contain outliers or come from population with heavy-tailed distributions, which appear very often in spatiotemporal data, the estimation methods based on least-squares method will not perform well. Considering the spatio-temporal heterogeneity and correlation, in this paper, the authors propose an adaptive nonparametric estimator for spatio-temporal model with unknown error distributions based on kernel density estimation. Asymptotic theory properties show that the proposed estimator is as asymptotically efficient as local maximum likelihood estimator with known error distributions, and more efficient than local polynomial estimator based on least-square method under some mild conditions. Simulation and case study are conducted to investigate the finite-sample performance of our procedure, especially for the existence of singularities, for boundary points, the performance of our procedure is significantly better than the method based on least squares.

Keywords Spatio-temporal model, Kernel density estimation, Local polynomial method, Local maximum likelihood method

2000 MR Subject Classification 62M30, 62G08

The English translation of this paper will be published in

Chinese Journal of Contemporary Mathematics, Vol. 42 No. 2, 2021

by ALLERTON PRESS, INC., USA