

函数型核加权估计法及其在经济学中的应用

涂云东^{1,2,3}, 汪思韦¹

(1. 北京大学 光华管理学院, 北京 100871; 2. 北京大学 统计科学中心, 北京 100871; 3. 北京大学 数量经济与数理金融教育部重点实验室, 北京 100871)

摘要 本文基于变系数模型提出了一个新的统计推断方法: 函数型核函数加权最小二乘法。该方法将变系数模型中经典的核函数加权最小二乘法和参数模型中的函数型最小二乘法巧妙结合, 通过条件特征函数构造损失函数进而定义了函数型核函数最小二乘估计量。该估计量既具有函数型最小二乘法的优势——在扰动项服从厚尾分布时也能够稳健估计参数, 又具有非参数核估计的特点——估计量的相合性不依赖于参数模型的正确设定。同时, 本文探讨了该估计量的大样本性质, 证明了其相合性和渐近正态性。进一步, 本文研究了该估计量的自适应估计, 即基于估计量渐近方差的相合估计量来选择最优估计。最后, 本文通过数值模拟来探究函数型核函数最小二乘法的有限样本性质, 并将该方法应用到我国 PM_{2.5} 和经济增长关系的研究中。

关键词 自适应估计; 环境库兹涅茨曲线; 变系数模型; 函数型最小二乘; 厚尾分布; 核估计

Functional kernel-weighted least square estimation and its applications in economics

TU Yundong^{1,2,3}, WANG Siwei¹

(1. Guanghai School of Management, Peking University, Beijing 100871, China; 2. Center for Statistical Science, Peking University, Beijing 100871, China; 3. The Ministry of Education Key Laboratory of Mathematical Economics and Quantitative Finance, Peking University, Beijing 100871, China)

Abstract In this article, we propose a new method called functional kernel-weighted least squares (FKLS) method to estimate the smooth coefficient function in semi-parametric smooth coefficient model. This novel proposal ingeniously combines the kernel-weighted least squares (KLS) and the functional least squares (FLS) methods. The corresponding FKLS estimator is defined based on the loss function constructed by the conditional characteristic function. It not only has the advantage of FLS method that can produce robust parameter estimation even if the disturbance is subject to heavy tailed distributions, but also has the characteristics of non-parametric kernel estimation that consistency can be achieved without the knowledge of the correct functional form. The consistency and asymptotic normality of the proposed estimator are established. Furthermore, adaptive estimation is investigated based on the consistent estimator of the asymptotic variance. Finally, superiority of the FKLS estimator in finite samples, compared to the KLS estimator, is demonstrated through simulated numerical examples and the study of PM_{2.5} and economic growth in China.

Keywords adaptive estimation; environmental Kutznets curve; functional coefficient model; functional least squares; heavy tailed distribution; kernel estimation

收稿日期: 2018-10-15

作者简介: 涂云东 (1982-), 男, 汉, 湖北人, 副教授, 博士生导师, 研究方向: 计量经济学理论与应用, E-mail: yundong.tu@gsm.pku.edu.cn; 汪思韦 (1995-), 女, 汉, 湖北人, 博士研究生, 研究方向: 计量经济学理论与应用, E-mail: siweiwang@pku.edu.cn.

基金项目: 国家自然科学基金 (71472007, 71532001, 71671002); 国家重点研发计划专项项目 (2016YFC0207705)

Foundation item: National Natural Science Foundation of China (71472007, 71532001, 71671002); National Key Research and Development Program of China (2016YFC0207705)

中文引用格式: 涂云东, 汪思韦. 函数型核加权估计法及其在经济学中的应用 [J]. 系统工程理论与实践, 2019, 39(4): 839-853.

英文引用格式: Tu Y D, Wang S W. Functional kernel-weighted least square estimation and its applications in economics[J]. Systems Engineering — Theory & Practice, 2019, 39(4): 839-853.

1 引言

变系数模型作为线性模型的推广经常被用来处理经济学问题。该模型较线性模型更加灵活且具有优良的解释性,能够减少模型的错误设定带来的偏差,同时又能很好地避免“维数诅咒”问题。因此,变系数模型自 Cleveland 等^[1]提出以来就引起了众多理论研究者 and 应用研究者极大的兴趣。Chen 和 Tsay^[2], Hastie 和 Tibshirani^[3]进一步对变系数模型做出了各种形式的拓展。Li 等^[4]研究了随机样本数据下该模型的核估计量性质以及模型设定检验等问题。Cai 等^[5]讨论了该模型在非线性平稳时间序列数据中应用。蔡宗武,陈琳娜和方颖^[6]也将该模型应用到分析人民币对美元汇率收益率与波动率的非线性时间序列特征上。进一步, Cai 等^[7], Xiao^[8], Tu 和 Wang^[9]讨论了该模型在非平稳时间序列数据中的理论结果。其中, Tu 和 Wang^[9]对文献中变系数模型的理论研究和应用研究进行了较全面的回顾。

变系数模型通常可表述为:

$$y_j = \mathbf{x}_j^\top \mathbf{b}_0(z_j) + \varepsilon_j, \quad j = 1, \dots, n, \quad (1)$$

其中, y_j 为因变量, \mathbf{x}_j 为 L 维自变量, z_j 为协变量, $\mathbf{b}_0(z)$ 是关于 z 的函数, ε_j 为扰动项, n 为观测样本量, \top 为转置符号。 $\mathbf{b}_0(\cdot)$ 为待估计的未知函数, 它刻画了协变量 z 是如何通过影响系数来改变自变量 \mathbf{x} 和因变量 y 之间的关系。

为了估计 $\mathbf{b}_0(z)$, Li 等^[4]提出了核函数加权最小二乘法 (kernel-weighted least squares method, 简记为 KLS), 其估计量为:

$$\hat{\mathbf{b}}_n^{\text{KLS}}(z) = \left[\sum_{j=1}^n \mathbf{x}_j \mathbf{x}_j^\top K_j(z, h) \right]^{-1} \sum_{j=1}^n \mathbf{x}_j y_j K_j(z, h), \quad (2)$$

其中 h 为与 n 有关的窗宽, $K(\cdot)$ 为核函数, $K_j(z, h) = K(\frac{z_j - z}{h})$ 。 Li 等^[4]同时证明了在一定条件下 (例如扰动项方差 σ_ε^2 有限等), KLS 估计量具有相合性和渐近正态分布。然而, 经济和金融数据常具有厚尾分布的特性, 其方差可能不存在, 这就使 KLS 估计法在实际数据分析中遇到挑战。

在模型 (1) 中, 若 $P(z = a) = 1$, 其中 $a \in \mathbb{R}$ 为某常数, 则有 $\mathbf{b}(z) = \mathbf{b}$ 为常函数。此时, 变系数模型就退化成普通线性模型。在线性模型中, Chambers 和 Heathcote^[10]受到 $\lim_{t \rightarrow 0} -t^{-2} \log |E(e^{it\varepsilon})|^2 = \sigma_\varepsilon^2$ 的启发, 提出了一种基于特征函数来构造损失函数的估计方法——函数型最小二乘法 (functional least squares method, 简记为 FLS)。FLS 的损失函数和对应的样本损失函数为:

$$L^{\text{FLS}}(\beta; t) = -t^{-2} \log |E e^{it(y - \mathbf{x}^\top \beta)}|^2, \quad (3)$$

$$L_n^{\text{FLS}}(\beta; t) = -t^{-2} \log \left| n^{-1} \sum_{j=1}^n \exp it(y_j - \mathbf{x}_j^\top \beta) \right|^2, \quad (4)$$

其中 $0 < |t| < T, T > 0$, i 为虚数单位。在使用 FLS 估计未知参数 \mathbf{b} 时, 首先求出一系列使得样本损失函数 (4) 达到最小值的估计量 $\{\hat{\mathbf{b}}_n(t), 0 < |t| < T\}$, 然后利用这一系列估计量的渐近性质选出使渐近方差达到最小值的 t^* , 那么 t^* 所对应的 $\hat{\mathbf{b}}_n(t^*)$ 即为 FLS 估计量。

由损失函数 (3) 得知, 当 $t \rightarrow 0$ 时, $L^{\text{FLS}}(\beta; t) = E(y - \mathbf{x}^\top \beta)^2 + o(t)$ 。即, 当 $t \rightarrow 0$ 时, FLS 的损失函数与最小二乘法 (ordinary least squares, 简记为 OLS) 的损失函数渐近等价。由此, OLS 可以看作 FLS 在 $t = 0$ 且 σ_ε^2 存在时的特例, 而 FLS 则可以看作 OLS 的一种拓展。值得强调的是, 由于 FLS 的损失函数 (3) 是通过特征函数构造的, 所以只要 $t \neq 0$, 损失函数都有意义。也就是说, FLS 估计对扰动项的矩没有任何要求。那么 FLS 可以用来处理一些厚尾分布的数据, 进而得到稳健的估计。

正因为 FLS 的稳健性, 自它被提出后就被广泛应用和研究。首先, Chambers 和 Heathcote^[10]用该方法来筛选离群值。Heathcote^[11]进一步说明了 FLS 可以看作是 OLS 的一种推广来处理厚尾分布的问题和效率问题, 同时将其与角随机变量理论联系起来得到一系列重要的性质。Csorgo^[12]证明了 FLS 估计量在一定条件下的相合性和渐近正态性。Heathcote 和 Welsh^[13]考虑了自回归模型下 FLS 的性质。Welsh 和 Nicholls^[14]考察了当解释变量具有序列相关性时 FLS 的一些性质。Heathcote 和 Welsh^[15]进一步研究了多元线性回归中 FLS 的应用。

在变系数模型中, 本文将 FLS 与 KLS 结合, 提出了一种新的统计推断方法 —— 函数型核函数加权最小二乘法 (functional kernel-weighted least squares method, 简记为 FKLS), 构造如下损失函数:

$$L(\mathbf{b}; t, z) = -t^{-2} \log |f(z)E(\exp it(y - \mathbf{x}^\top \mathbf{b})|Z = z)|^2, \tag{5}$$

其中 $0 < |t| < T$, $|\cdot|$ 为欧式范数, $f(\cdot)$ 为协变量 Z 的概率密度函数. 该估计量既具有 FLS 的优势, 即在扰动项服从厚尾分布的情形下能够稳健估计参数, 又具有 KLS 的特点, 即其相合性不依赖于待估系数函数的正确设定. 同时, 本文探讨了该估计量的大样本性质, 证明了其相合性和渐近正态性. 进一步, 本文研究了该估计量的自适应估计, 即基于估计量渐近方差的相合估计量来进行最优估计. 最后, 本文通过数值模拟来探究该估计量的有限样本性质, 并将其应用到我国 $PM_{2.5}$ 和经济增长关系的研究中.

本文的主要内容安排如下: 第 2 节将具体介绍变系数模型中的 FKLS 方法, 讨论在计算 FKLS 估计量时发现的一些重要的性质. 第 3 节将探讨 FKLS 估计量的大样本性质, 证明其在一定条件下的相合性和渐近正态性. 同时, 从渐近方差角度比较 FKLS 与 KLS 估计量. 在第 4 节中, 我们采用数值模拟的方式, 比较在扰动项服从不同分布的情况下, FKLS 估计量和 KLS 估计量的表现. 第 5 节将本文提出的估计方法应用到分析中国 $PM_{2.5}$ 和经济增长的关系中. 最后, 我们对本文进行了总结, 并且讨论了该课题的后续研究. 理论结果的证明见附录.

2 变参数模型下的函数型核函数加权最小二乘法

在变参数模型 (1) 下, KLS 估计量 (2) 是通过最小化以下样本损失函数得到的:

$$L_n^{KLS}(\mathbf{b}, z) = \frac{1}{nh} \sum_j (y_j - \mathbf{x}_j^\top \mathbf{b})^2 K_j(z, h),$$

其对应的损失函数为:

$$L^{KLS}(\mathbf{b}, z) = E((y - \mathbf{x}^\top \mathbf{b})^2 | Z = z) f(z).$$

而 FKLS 的样本损失函数则是通过样本条件特征函数构造的:

$$L_n(\mathbf{b}; t, z, h) = -\frac{1}{t^2} \log \left| \frac{1}{nh} \sum_{j=1}^n \exp\{it(y_j - \mathbf{x}_j^\top \mathbf{b})\} K_j(z, h) \right|^2. \tag{6}$$

利用该样本损失函数, 模型 (1) 中 $b_0(z)$ 的 FKLS 估计量可以通过以下两步得到:

1. 对于给定的 t , 求出使得样本损失函数 (6) 达到最小值的估计量 $\hat{\mathbf{b}}_n(t, z, h) = \arg \min_{\mathbf{b} \in \Theta} L_n(\mathbf{b}; t, z, h)$, 这里 Θ 为 \mathbb{R}^L 中的紧集;
 2. 选择最优的 t^* 使估计量的渐近方差达到最小, 那么对应的 $\hat{\mathbf{b}}_n(t^*, z, h)$ 就是 $b_0(z, h)$ 的 FKLS 估计量.
- 为了表述的简洁, 下面简记

$$U_n(\mathbf{b}; t, z, h) = \frac{1}{nh} \sum_{j=1}^n \cos[t(y_j - \mathbf{x}_j^\top \mathbf{b})] K_j(z, h),$$

$$V_n(\mathbf{b}; t, z, h) = \frac{1}{nh} \sum_{j=1}^n \sin[t(y_j - \mathbf{x}_j^\top \mathbf{b})] K_j(z, h),$$

$$\Phi_n(\mathbf{b}; t, z, h) = U_n(\mathbf{b}; t, z, h)^2 + V_n(\mathbf{b}; t, z, h)^2 = \frac{1}{n^2 h^2} \sum_{j=1}^n \sum_{s=1}^n K_j(z, h) K_s(z, h) \cos(t(\hat{\varepsilon}_j - \hat{\varepsilon}_s)),$$

其中 $\hat{\varepsilon}_j = y_j - \mathbf{x}_j^\top \mathbf{b}$. 那么 FKLS 的样本损失函数 (6) 式就可以写成如下形式:

$$L_n(\mathbf{b}; t, z, h) = -\frac{1}{t^2} \log(U_n(\mathbf{b}; t, z, h)^2 + V_n(\mathbf{b}; t, z, h)^2) = -\frac{1}{t^2} \log \Phi_n(\mathbf{b}; t, z, h). \tag{7}$$

对上式关于 \mathbf{b} 求一阶导数得:

$$l_n(\mathbf{b}; t, z, h) = -\frac{1}{t^2} \Phi_n(\mathbf{b}; t, z, h)^{-1} \frac{\partial \Phi_n(\mathbf{b}; t, z, h)}{\partial \mathbf{b}}, \tag{8}$$

其中

$$\frac{\partial \Phi_n(\mathbf{b}; t, z, h)}{\partial \mathbf{b}} = \frac{1}{n^2 h^2} \sum_{j=1}^n \sum_{s=1}^n t K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) \sin(t(\hat{\varepsilon}_j - \hat{\varepsilon}_s)). \tag{9}$$

步骤 1 中的 $\hat{\mathbf{b}}_n(t, z, h)$ 需要满足一阶条件 $l_n(\mathbf{b}; t, z, h) = 0$. 由于 t 和 $\Phi_n(\mathbf{b}; t, z, h)$ 恒不等于 0, 因此可以将一阶条件调整为 $\frac{\partial \Phi_n(\mathbf{b}; t, z, h)}{\partial \mathbf{b}} = 0$. 步骤 2 中提到的 $\hat{\mathbf{b}}_n(t, z, h)$ 的渐近方差以及 t^* 的求解方法将在第 3 节中具体说明.

注 1: 当 $t \rightarrow 0$ 时, $-t^{-2} \log |f(z)E(\exp it(y - \mathbf{x}^\top \mathbf{b})|Z = z)|^2 = f(z)E((y - \mathbf{x}^\top \mathbf{b})|Z = z) + o(t)$, 即 FKLS 的 (样本) 损失函数等价于 KLS 的 (样本) 损失函数. 下面将从样本损失函数的一阶条件来说明当 $t \rightarrow 0$ 时, FKLS 和 KLS 等价. 当 $t \rightarrow 0$ 时

$$\begin{aligned} l_n(\mathbf{b}; t, z, h) = 0 &\iff \sum_{j=1}^n \sum_{s=1}^n K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) (\hat{\varepsilon}_j - \hat{\varepsilon}_s) = 0 \\ &\iff \sum_{j=1}^n \sum_{s=1}^n K_j(z, h) K_s(z, h) \hat{\varepsilon}_s \mathbf{x}_s = \sum_{j=1}^n \sum_{s=1}^n K_j(z, h) K_s(z, h) \hat{\varepsilon}_j \mathbf{x}_s \\ &\iff \sum_{s=1}^n K_s(z, h) \hat{\varepsilon}_s \mathbf{x}_s = \sum_{j=1}^n K_j(z, h) \hat{\varepsilon}_j = 0. \end{aligned}$$

注 2: 在固定 z 时, 可以将损失函数 $L(\mathbf{b}; t, z)$ (5) 式写成

$$\tilde{L}(\mathbf{b}; t, z) = -t^{-2} \log |E(\exp it(y - \mathbf{x}^\top \mathbf{b})|Z = z)|^2.$$

类似地, KLS 的损失函数也可以写成如下形式:

$$\tilde{L}^{\text{KLS}}(\mathbf{b}; z) = E((y - \mathbf{x}^\top \mathbf{b})^2|Z = z).$$

那么当 $\varepsilon|z$ 服从均值为 0, 方差为 σ^2 的正态分布时, $\tilde{L}(\mathbf{b}_0(z); t, z) = -t^{-2} \log |E(\exp it\varepsilon|Z = z)|^2 = \sigma^2 = \tilde{L}^{\text{KLS}}(\mathbf{b}_0(z); z)$. 由此推出: 在扰动项关于 z 的条件分布服从正态分布时, 无论 t 取何值, $\tilde{L}(\mathbf{b}_0(z); t, z)$ 都与 $\tilde{L}^{\text{KLS}}(\mathbf{b}_0(z); z)$ 相等, 即 FKLS 与 KLS 等价.

3 函数型核函数加权最小二乘估计量的渐近性质

3.1 大样本性质

记 $f(\varepsilon, \mathbf{x}|z)$ 为 ε 和 \mathbf{X} 在给定 $Z = z$ 下的联合条件密度函数, $f(\varepsilon|z)$ 和 $f(\mathbf{x}|z)$ 分别为 ε 和 \mathbf{X} 在给定 $Z = z$ 下的条件密度函数, $f(\varepsilon, \mathbf{x}, z)$ 为 $(\varepsilon, \mathbf{X}, Z)$ 的联合密度函数.

假设 1 (i) 对于 $\forall z \in \mathfrak{R}$, $\mathbf{b}(z)$ 是 Θ 的内点, 且 Θ 为 \mathfrak{R}^L 中的紧集;

(ii) $\{\varepsilon_j, \mathbf{x}_j, z_j\}_{j=1}^n$ 为简单随机样本, 且 ε 与 \mathbf{x} 在给定 z 下条件独立, 即 $f(\varepsilon, \mathbf{x}|z) = f(\varepsilon|z)f(\mathbf{x}|z)$, 对于 $\forall z \in \mathfrak{R}$;

(iii) $f(z)$ 为恒大于 0 的有界函数, $f(\varepsilon, \mathbf{x}, z)$ 和 $\mathbf{b}_0(z)$ 关于 z 满足李普希兹条件, 即 $|g(z+h) - g(z)| = O(h)$, $g(z) = f(\varepsilon, \mathbf{x}, z)$ 或者 $\mathbf{b}_0(z)$, $f(\varepsilon, \mathbf{x}, z)$ 关于 z 的二阶导数存在且有界;

(iv) 当 $n \rightarrow \infty$, 有 $h \rightarrow 0$, 同时 $nh \rightarrow \infty$;

(v) $K(\cdot)$ 为对称核函数, 且满足 $\int K(v)dv = 1$, $\int K^2(v)dv < \infty$.

其中假设 (i) 要求 $\mathbf{b}(\cdot)$ 的取值范围为 \mathfrak{R}^L 上的紧集, 该条件在极值估计中经常用以简化估计量一致性的证明. 由假设 (ii) 知, 这里只考虑了数据是随机样本以及 x 是外生变量的情况. 但是本文的极限结果可以通过调整假设以及使用相关数据的中心极限定理推广到数据是平稳遍历的情况. 假设 (iii) 要求 $f(z)$ 是恒大于 0 的有界函数以及 $f(\varepsilon, \mathbf{x}, z)$ 和 $\mathbf{b}_0(z)$ 满足李普希兹条件, 同时要求 $f(\varepsilon, \mathbf{x}, z)$ 二阶可导. 该条件见 Li 等^[4]. 假设 (iv) 和 (v) 是非参数估计中常见的对核函数和窗宽的条件.

定理 1 (相合性) 若假设 1 满足, 那么步骤 1 中的 $\hat{\mathbf{b}}_n(t, z, h)$ 依概率收敛到 $\mathbf{b}_0(z)$, 即当 $n \rightarrow \infty$ 时, $\forall \delta > 0, \lim_{n \rightarrow \infty} P(|\hat{\mathbf{b}}_n(t, z, h) - \mathbf{b}_0(z)| > \delta) = 0$.

建立 FKLS 估计量的渐近正态性质时还需要以下假设:

假设 2 (i) $E[\mathbf{x}_1 \mathbf{x}_1^\top | Z = z]$ 正定; (ii) 当 $n \rightarrow \infty$, 有 $nh^5 \rightarrow 0$.

其中假设 (i) 保证了 $E[\mathbf{x}_1 \mathbf{x}_1^\top | Z = z]$ 的可逆性. 假设 (ii) 中对窗宽 h 的条件 (undersmoothing) 使得估计量的渐近偏差为零. 这两个假设见 Li 等^[4].

定理 2 (渐近正态性) 若假设 1 和 2 满足, 那么当 $n \rightarrow \infty$ 时, 步骤 1 中的 $\hat{\mathbf{b}}_n(t, z, h)$ 具有如下渐近正态分布:

$$\sqrt{nh}(\hat{\mathbf{b}}_n(t, z, h) - \mathbf{b}_0(z)) \xrightarrow{d} N(0, \Omega(t, z)\mathbf{A}(z)^{-1}). \tag{10}$$

其中 $\Omega(t, z) = \frac{\int K^2(v)dv\Delta(t, z)}{2t^2 f(z)(u^2(t, z) + v^2(t, z))^2}$, $\mathbf{A}(z) = E(\mathbf{x}_1 \mathbf{x}_1^\top | Z = z) - E(\mathbf{x}_1 | Z = z)E(\mathbf{x}_1 | Z = z)^\top$, $\Delta(t, z) = u^2(t, z)(1 - u(2t, z)) + v^2(t, z)(1 + u(2t, z)) - 2v(t, z)u(t, z)v(2t, z)$, $u(t, z) = E(\cos t\varepsilon | Z = z)$, $v(t, z) = E(\sin t\varepsilon | Z = z)$.

注 3: 如果 $\varepsilon|z$ 的分布关于常数 c 对称, 那么 $E(e^{it\varepsilon} | Z = z) = e^{itc}u(t, z)$, 其中 $u(t, z) = E(\cos(t\varepsilon - c) | Z = z)$, 进而得到:

$$\Omega(t, z) = \frac{\int K^2(v)dv(1 - u(2t, z))}{2t^2 f(z)u^2(t, z)}.$$

注 4: 如果 $\varepsilon|z$ 服从均值为 0, 方差为 σ 的正态分布, 即 $u(t, z) = \exp(-\frac{1}{2}\sigma^2 t^2)$, 那么

$$\Omega(t, z) = \frac{\int K^2(v)dv \sinh(\sigma^2 t^2)}{2t^2 f(z)}.$$

在这种情形下, 对于 $\forall t \neq 0$, 都有 $\Omega(0, z) < \Omega(t, z)$. 结合注 2 可知在 $\varepsilon|z$ 服从正态分布时, KLS 就是最优的 FKLS.

注 5: 易证, 当 $t \rightarrow 0$ 时,

$$\frac{\Delta(t, z)}{2t^2(u^2(t, z) + v^2(t, z))^2} \rightarrow E[\varepsilon^2 | Z = z].$$

当假设 1 中条件 (ii) 满足时, KLS 估计量 $\hat{\mathbf{b}}_n^{KLS}(z)$ 的渐近方差可以写成 $f(z)^{-1}\mathbf{A}(z)^{-1}E[\varepsilon^2 | Z = z] \int K^2(v)dv$. 由此亦可知: 当 $t \rightarrow 0$ 时, FKLS 的渐近方差与 KLS 的渐近方差相同.

3.2 自适应估计

对于不同的 t 和 z , FKLS 估计量 $\hat{\mathbf{b}}_n(t, z, h)$ 具有不同的渐近方差. 从估计有效性的角度, 可以考虑选取 t 使渐近方差 $\Omega(t, z)\mathbf{A}(z)^{-1}$ 达到最小. 在实际数据分析中, 由于 $\varepsilon|z$ 的真实分布未知, 所以需要估计 $\Omega(t, z) \times \mathbf{A}(z)^{-1}$ 进行估计. 不妨将 $\Omega(t, z)$ 写成如下形式:

$$\Omega(t, z) = \frac{\int K^2(v)dv f^3(z)\Delta(t, z)}{2t^2 f^4(z)(u^2(t, z) + v^2(t, z))^2}.$$

下面我们用 $\hat{\mathbf{A}}_n(z, h) = \frac{1}{nh} \sum_j \mathbf{x}_j \mathbf{x}_j^\top K_j(z, h) / \frac{1}{nh} \sum_j K_j(z, h)$ 来估计 $\mathbf{A}(z)$, 用 $U_n(\mathbf{b}; t, z, h)$ 和 $V_n(\mathbf{b}; t, z, h)$ 来估计 $u(t, z)f(z)$ 和 $v(t, z)f(z)$, 用 $\underline{U}_n(\mathbf{b}; 2t, z, h) = \frac{1}{nh} \sum_j \cos 2t(y_j - \mathbf{x}_j^\top \mathbf{b})K_j^2(z, h)$ 来估计 $u(2t, z)f(z) \int K^2(v)dv$, 用 $\underline{V}_n(\mathbf{b}; 2t, z, h) = \frac{1}{nh} \sum_j \sin 2t(y_j - \mathbf{x}_j^\top \mathbf{b})K_j^2(z, h)$ 来估计 $v(2t, z)f(z) \int K^2(v)dv$, 用 $\hat{\mu}_2(K) = \frac{1}{nh} \sum_j K_j(z, h)^2$ 来估计 $\int K^2(v)dv f(z)$, 进而得到 $\Delta(t, z)$ 的估计:

$$\begin{aligned} \hat{\Delta}_n(\mathbf{b}; t, z, h) &= (\hat{\mu}_2(K) - \underline{U}_n(\mathbf{b}; 2t, z, h))U_n^2(\mathbf{b}; t, z, h) + (\hat{\mu}_2(K) + \underline{U}_n(\mathbf{b}; 2t, z, h))V_n^2(\mathbf{b}; t, z, h) \\ &\quad - 2\underline{V}_n(\mathbf{b}; 2t, z, h)U_n(\mathbf{b}; t, z, h)V_n(\mathbf{b}; t, z, h). \end{aligned}$$

进而可证

$$\hat{\Omega}_n(\mathbf{b}; t, z, h) = \frac{\hat{\Delta}_n(\mathbf{b}; t, z, h)}{2t^2(U_n^2(\mathbf{b}; t, z, h) + V_n^2(\mathbf{b}; t, z, h))^2} \tag{11}$$

是 $\Omega(t, z)$ 的相合估计. 由此得到 FKLS 估计量渐近方差的相合估计量为 $\hat{\Omega}_n(\mathbf{b}; t, z, h)\hat{\mathbf{A}}_n^{-1}(z, h)$.

对一系列给定的点 $z_s, s = 1, \dots, m$, FKLS 估计对应的 $\mathbf{b}(z_s)$ 时, 选择 t^* 有两种方法: 逐点选择和全局选择. 首先考虑逐点选择. 对于每一个 z_s , 将 t 和步骤 1 中得到的估计量 $\hat{\mathbf{b}}_n(t, z_s, h)$ 代入 $\hat{\Omega}_n(\mathbf{b}; t, z_s, h)$, 选择 $t^*(s)$:

$$t^*(s) = \arg \min \hat{\Omega}_n(\hat{\mathbf{b}}_n(t, z_s, h); t, z_s, h).$$

该方法选出来的 $t^*(s)$ 可能会因 z_s 的大小而不同. 若希望得到唯一最优的 t^* , 那么可以考虑全局选择.

在全球选择 t^* 时, 可考虑使得 $\hat{\mathbf{b}}(z_s, h)$ 的渐近方差加权求和达到最小, 即

$$t^* = \arg \min \sum_s \hat{\Omega}_n(\hat{\mathbf{b}}_n(t, z_s, h); t, z_s, h)\mathbf{A}(z_s)^{-1}W_s,$$

其中 W_s 是权重函数. 常见的权重函数可以是 z 的密度函数 (的估计) 或等权重函数. 由于 $A(s)^{-1}$ 为 $L \times L$ 的矩阵, 实际应用中可以考虑取其行列式或者迹. 即全局选择为

$$t^* = \arg \min_s \sum_s \hat{\Omega}_n(\hat{\mathbf{b}}_n(t, z_s, h); t, z_s, h) |A(z_s)^{-1}|,$$

或

$$t^* = \arg \min_s \sum_s \hat{\Omega}_n(\hat{\mathbf{b}}_n(t, z_s, h); t, z_s, h) \text{tr}(A(z_s)^{-1}).$$

最后, 将选择出来的 t^* 代入步骤 1 中, 就可以得到 FKLS 估计量 $\hat{\mathbf{b}}_n(z_s, h) = \hat{\mathbf{b}}_n(t^*, z_s, h)$. 需要注意的是, 选择出来的 t^* 满足使得估计量渐近方差达到最小, 即只要 $t^* \neq 0$ (或者不足够接近零), FKLS 的方差就比 KLS 的方差小.

3.3 选择窗宽 h

在实际应用中, 可以简便地使用拇指法则来选择窗宽. 根据 Silverman^[16], 可取 $h = 1.06 \times sd(z) \times n^{-1/5}$, 其中 $sd(z)$ 为 z 的标准差. 使用 FKLS 时, 也可以选择 $h = c \times sd(z) \times n^{-1/5}$, 其中 c 为常数. 根据假设 2(ii), 可以选择较小的窗宽来减小估计量的偏差, 但是此时估计量的方差反而会有所增加.

另外一种常见的窗宽选择方法是交叉验证法. 下面介绍该方法应用到求解 FKLS 估计量的步骤.

1. 对于固定的 t , 计算

$$\hat{\mathbf{b}}_{-i}(t, z_i, h) = \arg \min \left[-\frac{1}{t^2} \log \left| \frac{1}{nh} \sum_{j \neq i} \exp\{it(y_j - \mathbf{x}_j^\top \mathbf{b})\} K\left(\frac{z_j - z_i}{h}\right) \right|^2 \right];$$

2. 将 t 和对应的 $\hat{\mathbf{b}}_{-i}(t, z_i, h)$ 代入 $\hat{\Omega}_n(\mathbf{b}; t, z_i, h)$ ((11) 式) 计算出 $t^*(i) = \arg \min \hat{\Omega}_n(\hat{\mathbf{b}}_{-i}(t, z_i, h); t, z_i, h)$, 记 $\hat{\mathbf{b}}_{-i}(z_i, h) = \hat{\mathbf{b}}_{-i}(t^*(i), z_i, h)$;
3. 计算最优窗宽

$$h_{CV} = \arg \min \sum_{i=1}^n (y_i - \mathbf{x}_i^\top \hat{\mathbf{b}}_{-i}(z_i, h))^2.$$

4 数值模拟

本节使用蒙特卡罗数值模拟的方法求解模型 $y_j = x_j b(z_j) + \varepsilon_j, j = 1, \dots, n$, 中 $b(z)$ 的 FKLS 和 KLS 估计量, 进而比较 FKLS 和 KLS 在有限样本下的表现. 我们不妨考虑 $x_j \sim \text{i.i.d.} N(0, 1), z_j \sim \text{i.i.d.} N(0, 1), b(z) = e^z / (1 + e^z)$. 对于扰动项 ε_j 的分布, 考虑以下 6 种设定:

- (E1) $\varepsilon_j \sim \text{i.i.d.} N(0, 1), j = 1, \dots, n$;
- (E2) $\varepsilon_j \sim \text{i.i.d.} 0.4N(-6, 1) + 0.6N(4, 10), j = 1, \dots, n$;
- (E3) $\varepsilon_j \sim \text{i.i.d.} t(3), j = 1, \dots, n$;
- (E4) $\varepsilon_j \sim \text{i.i.d.} Cauchy(0, 1), j = 1, \dots, n$;
- (E5) $\varepsilon_j \sim \text{i.i.d.} Gamma(6, 1) - 6, j = 1, \dots, n$;
- (E6) $\varepsilon_j \sim \text{i.i.d.} Laplace(0, 2), j = 1, \dots, n$.

在估计 $b(z)$ 时, 我们使用拇指法则选择窗宽 $h = 1.06sd(z)n^{-1/4}$, 并采用标准正态核函数 $K(v) = \frac{1}{\sqrt{2\pi}} \times e^{-v^2/2}$. 首先, 我们在 $(-1, 1)$ 上等间隔地取 100 个格点: $z_s, s = 1, \dots, 100$, 对每一组参数生成 $n = 100$ 或 400 个样本, 分别求出 $b(z_s)$ 的 FKLS 和 KLS 估计量. 其中 FKLS 估计量的 t^* 由逐点最优法选择得到. 为了比较两种方法的估计精度, 重复上述过程 100 次, 并分别计算两个估计量的逐点样本均方误差 (MSE)、偏差的平方 (Bias²) 和方差 (VAR):

$$MSE(z_s) = \frac{1}{100} \sum_{r=1}^{100} [\hat{b}(z_s)_r - b(z_s)]^2; Bias^2(z_s) = \left[\frac{1}{100} \sum_{r=1}^{100} \hat{b}(z_s)_r - b(z_s) \right]^2;$$

$$VAR(z_s) = \frac{1}{100} \sum_{r=1}^{100} \left[\hat{b}(z_s)_r - \frac{1}{100} \sum_{r=1}^{100} \hat{b}(z_s)_r \right]^2,$$

这里 $\hat{b}(\cdot)$ 表示 FKLS 或 KLS 估计, $\hat{b}(z_s)_r$ 表示基于第 r 组样本得到的对 $b(z_s)$ 的估计. 图 1~6 画出了扰动项 6 种分布设定下三个估计精度度量随 z 的变化曲线. 其中, 细 (粗) 实线对应 $n = 100$ (400) 时 FKLS 的计算结果, 细 (粗) 虚线对应 $n = 100$ (400) 时 KLS 的计算结果.

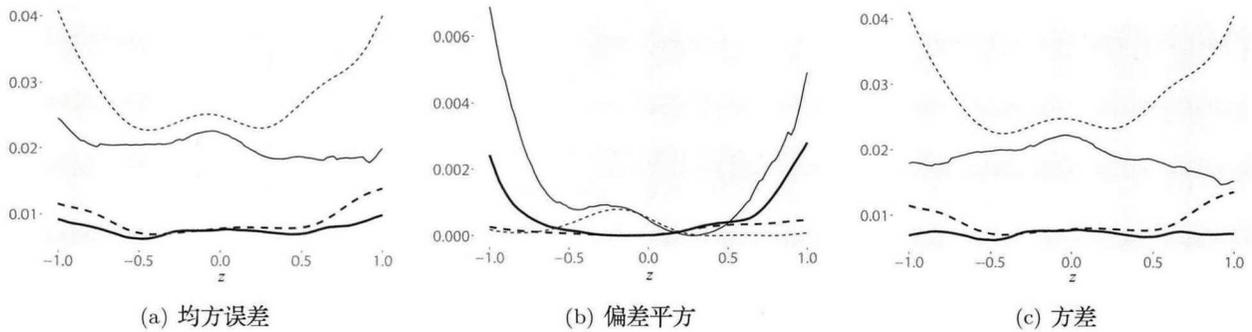


图 1 FKLS 与 KLS 估计的均方误差、偏差平方和方差比较: E1

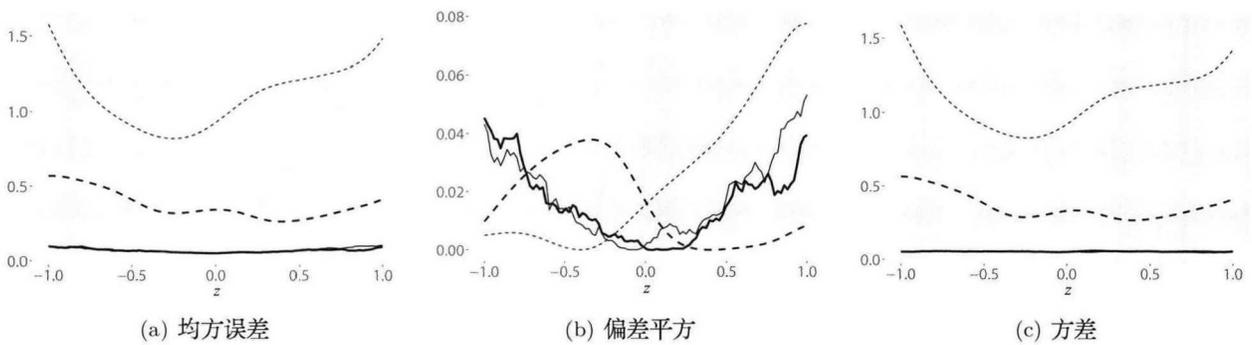


图 2 FKLS 与 KLS 估计的均方误差、偏差平方和方差比较: E2

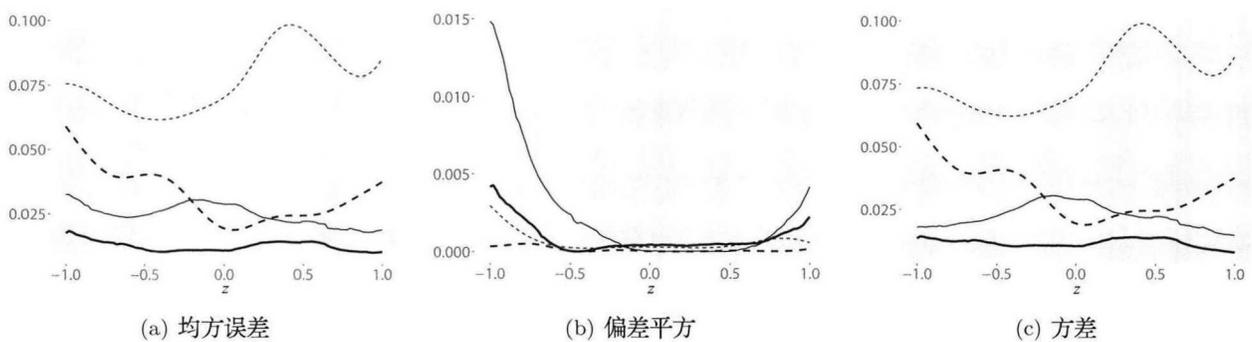


图 3 FKLS 与 KLS 估计的均方误差、偏差平方和方差比较: E3

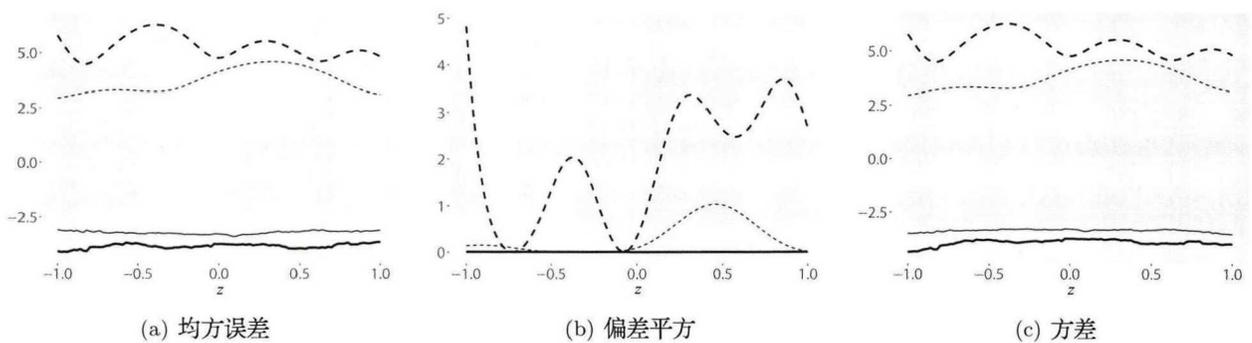


图 4 FKLS 与 KLS 估计的 (对数) 均方误差、偏差平方和 (对数) 方差比较: E4

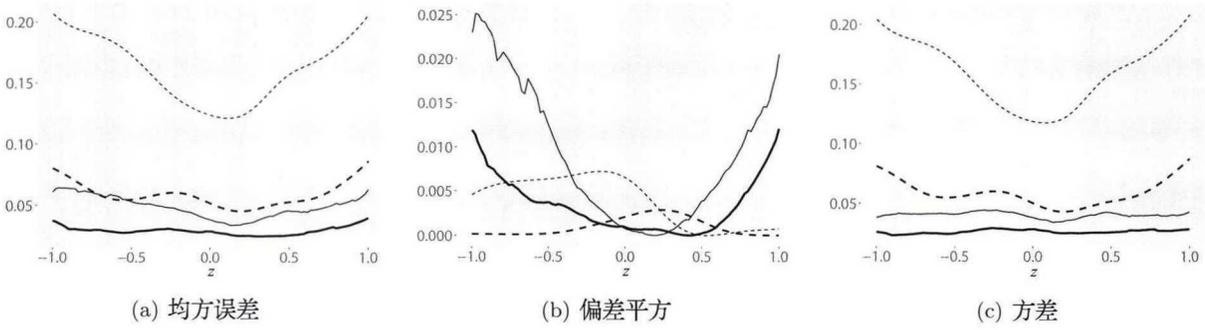


图 5 FKLS 与 KLS 估计的均方误差、偏差平方和方差比较: E5

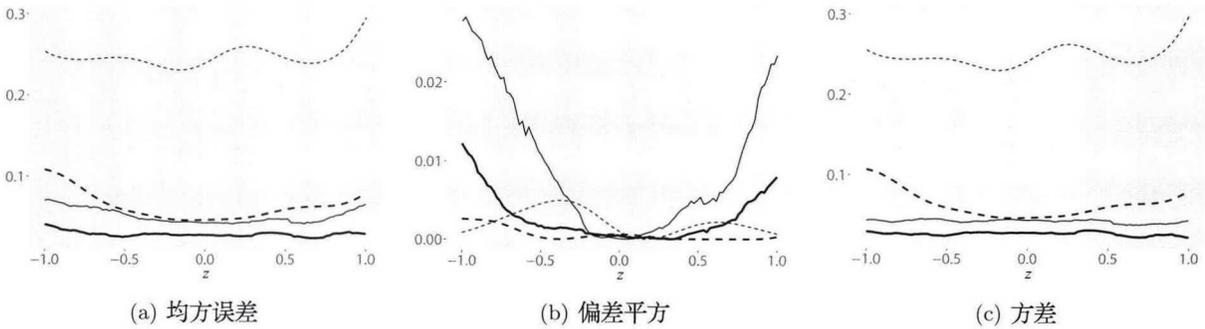


图 6 FKLS 与 KLS 估计的均方误差、偏差平方和方差比较: E6

如图 1 所示,当扰动项服从正态分布时,FKLS 估计的 MSE 和 VAR 与 KLS 相差不多.从 $Bias^2$ 角度观察可以发现由于端点附近的样本较少,FKLS 和 KLS 的估计有一定的偏差.但是观察纵轴可以发现两者的偏差平方都十分的小,最大只有 0.0015.观察图 2 可以发现,当扰动项服从混合正态分布(分布不对称)时,FKLS 估计表现明显优于 KLS 估计.在样本量较小的时候 ($n = 100$),KLS 的最大均方误差为 3,而 FKLS 的均方误差均小于 0.5. FKLS 在端点位置估计偏差的问题同样存在,但是较之 KLS 已经有了一定的优化.图 3 与图 1 类似,当扰动项服从自由度为 3 的 t 分布时,FKLS 在端点附近的估计偏差较大,但总体表现优于 KLS.图 4 展示了情形 4 (扰动项服从柯西分布)下 FKLS 和 KLS 的估计效果.由于 KLS 估计的 MSE 和 VAR 过大,我们对纵轴做了取对数处理.观察发现,当扰动项方差不存在时,KLS 的估计效果明显劣于 FKLS,且随样本量增加 ($n = 100$ 增加到 $n = 400$),KLS 估计的 VAR 和 MSE 并没有减小.但是 FKLS 估计量在该情形下表现良好.图 5,图 6 (扰动项服从伽马分布,拉普拉斯分布)与图 1 的情形类似,即 FKLS 在端点附近存在一定的估计偏差,但是其方差和均方误差均小于 KLS.而当样本量较小时 ($n = 100$) KLS 估计的均方误差甚至大于 0.25,即此时估计不稳定.

表 1 中列出了两种方法估计的加总均方误差 (IMSE, 即 $MSE(z_s)$ 对 s 求和), 加总偏差的平方 ($IBias^2$) 以及加总方差 (IVAR). 由表 1 可知,在 E1 (正态分布) 情形下,FKLS 的 IVAR 比 KLS 的相对大一点,但是其 IMSE 以及 $IBias^2$ 都较小.同时,随着样本量的增大 ($n = 100$ 增加到 $n = 800$),FKLS 的 IVAR 有明显减小 (从 2.65 降低到了 0.59). 在 E3 (t 分布), E5 (伽马分布), E6 (拉普拉斯分布) 情形下两种估计量的表现与 E1 类似.同时,随着样本量的增加,FKLS 的 IVAR 和 IMSE 都有明显减小趋势.当扰动项服从混合正态分布 (E2 情形) 时,KLS 的 IMSE 和 IVAR 都较大,超过了 200,同时其 $IBias^2$ 也明显大于 FKLS.观察 E4 (柯西分布) 情形发现,KLS 估计量的 IMSE 和 IVAR 都很大,甚至超过了 999.但是 FKLS 的 IMSE, IVAR 和 $IBias^2$ 都随着样本量的增大而减小.当样本量增加到 800 时,FKLS 的 $IBias^2$ 只有 0.09,而 KLS 的 $IBias^2$ 却达到了 46.7.

综上所述,在以上六种情形下,FKLS 的 VAR 均小于 KLS.当扰动项不存在二阶矩 (柯西分布) 或分布不对称 (混合正态分布) 时,FKLS 的估计效果明显优于 KLS; 而当扰动项服从正态分布时,FKLS 与 KLS 估计结果相近; 当扰动项的二阶矩存在且为厚尾分布时,FKLS 存在一定的估计偏差,但随着样本量的增大偏差逐渐减小.总体来看,FKLS 的估计效果优于 KLS.

表 1 FKLS 与 KLS 估计的加总均方误差, 加总偏差的平方和加总方差的比较

	n = 100		n = 400		n = 800		n = 100		n = 400		n = 800	
	\hat{b}_{KLS}	\hat{b}_{FKLS}										
E1						E4						
均方误差	2.62	2.69	0.92	0.92	0.59	0.60	-	6.58	-	3.02	-	1.81
偏差平方	0.05	0.04	0.02	0.01	0.01	0.01	389	0.03	98.8	0.01	46.7	0.02
方差	2.57	2.65	0.90	0.91	0.58	0.59	-	6.55	-	3.01	-	1.79
E2						E5						
均方误差	267	11.9	76.3	7.27	50.7	5.05	18.7	10.6	5.57	4.82	3.32	3.18
偏差平方	3.69	0.29	0.66	0.31	0.19	0.11	0.50	0.33	0.16	0.17	0.02	0.02
方差	263	11.6	75.4	6.96	50.5	4.94	18.2	10.3	5.41	4.65	3.30	3.16
E3						E6						
均方误差	6.63	4.83	2.79	1.61	1.51	0.90	25.1	10.8	7.03	5.19	5.25	3.50
偏差平方	0.08	0.02	0.07	0.02	0.00	0.01	0.42	0.34	0.02	0.04	0.04	0.03
方差	6.55	4.81	2.72	1.59	1.51	0.89	24.7	10.5	7.01	5.15	5.21	3.47

注: 其中 - 表示数值大于 999.

5 中国 PM_{2.5} 和经济增长关系探究

5.1 背景介绍

下面, 我们将变系数模型和 FKLS 方法应用到中国的环境与经济数据以探究环境库兹涅茨曲线 (environmental Kutznets curve, 简记为 EKC). 近年来, PM_{2.5} (细颗粒物) 排放导致的污染问题引发越来越多的关注. 与较粗的大气颗粒物相比, PM_{2.5} 粒径小、面积大、活性强、易附带有毒有害物质 (例如, 重金属、微生物等). 同时, 它能在空气中悬浮较长时间, 对人体健康和大气环境质量有很大影响. 为了研究环境污染与经济发展之间的关系, Grossman 和 Krueger^[17] 参照经济学中的库兹涅茨曲线提出了 EKC, 并且指出在经济发展的初期环境污染会随经济的发展而加重, 当经济发展到某一水平时环境污染程度达到最大, 而后随着经济继续发展, 环境质量却逐渐改善. EKC 被提出之后, 国内外众多学者对其做出了讨论. 彭水军和包群^[18] 对我国经济增长与 6 类环境污染指标之间的关系进行了实证检验; 虞义华等^[19] 利用可行的广义最小二乘法 (FGLS) 模型分析了二氧化碳排放强度同经济发展水平及产业结构之间的关系; 王志华等^[20] 以北京为例验证了 EKC 的假设及其条件, 解析十类环境指标的 EKC 演变轨迹和阶段特征. 本文则从经济与环境互相影响的角度出发, 引入第二产业占比作为协变量构造一个变系数模型. 下面我们将利用 FKLS 和 KLS 方法探讨 PM_{2.5} 的排放和 GDP 以及第二产业占比之间的关系.

PM_{2.5} 排放的数据来源于北京大学地表过程实验室. 原始数据是在 0.1 × 0.1 的空间分辨率 (空间分辨率是指遥感图像上能够详细区分的最小单元的尺寸或大小, 是用来表征影像分辨地面目标细节的指标) 上的月度总排放数据. 为了得到各个省份的季度数据, 我们对月度格点数据按照季度和中国省级行政区划进行加总. 最后将加总数据除以每个省份的人口数目, 得到了每个省份的人均季度 PM_{2.5} 排放数据, 单位为 10³ kg 每人. 其中人口数据来自国家统计局, 原始数据为年度数据. 我们假设每一年人口季度增长率不变, 对原数据进行了对数线性插值得到人口的季度数据.

国民生产总值 (gross domestic product, 简记 GDP) 是衡量经济发展的重要数据. 各个省份名义 GDP (用生产物品和劳务的当年价格计算的全部最终产品的市场价值) 的数据来自 WIND 数据库中的国家统计局六十年统计资料汇编, 原始数据为季度累计 GDP 并且有部分数据缺失. 我们先利用迭代插值的方法对数据进行了插补, 然后将相邻两个季度的 GDP 做差得到各个省份的季度名义 GDP. 对从国家统计局六十年统计资料汇编中得到的每年 GDP 平减指数 (没有剔除物价变动前的 GDP 增长与剔除了物价变动后的 GDP 增长之商) 进行迭代插补之后, 我们计算了以 2000 年为基年的各个省份季度实际 GDP (按基年价格计算的最终产品的价值), 单位为万元. 将各个省份的季度实际 GDP 除以对应的人口数目之后得到各个省份的人均季度实际 GDP, 单位为万元每人. 我们在模型中采用第二产业占比 (z) 作为协变量. 因此, 我们对来自国家统计局六十年统计资料汇编的季度第二产业占比数据也做了类似的迭代插值. 上述变量均含有较明显的季节性,

我们都做了去季节化处理。

下面,我们以 2000 年第一季度到 2014 年第四季度一共 60 个季度,北京、天津、河北、山西四个地区一共 $n = 240$ 个数据为例,分析人均季度实际 GDP(x),人均季度 PM_{2.5} 排放量 (y) 以及季度第二产业占比 (z) 之间的关系。

5.2 模型建立和参数估计

由于 z 对 x, y 的影响未知,我们考虑如下含有未知函数 $b(z) = (b_0(z), b_1(z), b_2(z))$ 的变系数模型:

$$y_i = b_0(z_i) + b_1(z_i)x_i + b_2(z_i)x_i^2 + \varepsilon_i, \quad i = 1, \dots, n. \quad (12)$$

下面分别使用 KLS 和 FKLS 两种方法来估计 $b(z)$ 。首先选择 $K(v) = \frac{1}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}$ 作为核函数。再使用交叉验证法对两种方法进行最优窗宽选择,进而得到 KLS 的最优窗宽为 $h^{\text{KLS}} = 0.0168$, FKLS 的最优窗宽为 $h^{\text{FKLS}} = 0.0167$ 。接下来,我们使用 FKLS 和 KLS 估计一系列 $b(z_m), m = 1, \dots, 50$, 其中 z_m 为 (0.43, 0.56) 上的等距离格点, 0.43 和 0.56 分别为 z 的 1/4 和 3/4 分位点。最后,利用两种估计量的渐近方差的相合估计计算出置信度为 95% 的逐点的置信区间。估计结果如图 7 所示,其中细实(虚)线为 FKLS (KLS) 估计值,粗实(虚)线为 FKLS (KLS) 的置信区间。

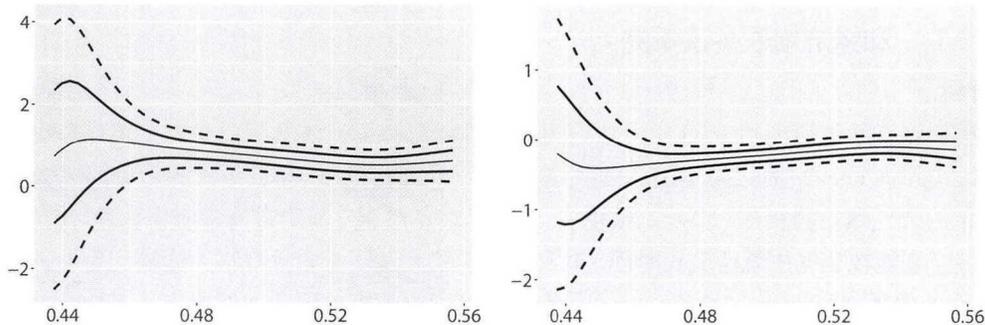


图 7 FKLS 和 KLS 估计的 $b_1(z)$ (左) 和 $b_2(z)$ (右) 曲线以及置信区间

由图 7 可知, FKLS 和 KLS 两种方法所得到的估计结果十分接近,即两条估计曲线几乎重合,但是两种方法所得到的置信区间却相差甚远。首先观察 $b_1(z)$ 和 $b_2(z)$ 估计曲线发现, $b_1(z)$ 恒大于 0, $b_2(z)$ 恒小于 0。由此可得 x 与 y 之间的关系呈现倒 U 型,与 EKC 理论吻合。即在经济发展初期,经济发展伴随着环境问题加剧;当经济发展到一定阶段,经济发展有助于缓解污染问题加剧。随着第二产业占比 (z) 的增大, $b_1(z)$ 先增大后减小,末端又存在着平稳上升的态势;而 $b_2(z)$ 则先减小后增大,末端又缓慢减小。进而发现季度第二产业占比的变化影响着人均季度实际 GDP 与人均季度 PM_{2.5} 排放量之间的关系。当 $z \in (0.54, 0.56)$ 时, $b_2(z)$ 的绝对值较小,即此时 y 与 x 之间近乎斜率大于 0 的线性关系,随着人均季度实际 GDP 的增长,人均季度 PM_{2.5} 排放量也显著提升。当 $z \in (0.46, 0.48)$ 时, $b_2(z)$ 的绝对值较大,即此时 x 的增长不会对于 y 有很强烈的影响。由以上分析可知,当第二产业占比较少时,由于经济发展不主要依赖于第二产业,其发展对于 PM_{2.5} 的排放问题没有太大的影响;而当第二产业占比较大时,经济发展依赖第二产业的发展,由此就造成了环境污染的问题。

如图 7 所示, FKLS 的置信区间宽度明显小于 KLS。当 $z \in (0.44, 0.48)$ 时, KLS 估计量的 95% 置信区间包含了 0,即认为 $b_1(z)$ 和 $b_2(z)$ 不显著, x 与 y 之间没有相关性;而此时 FKLS 则能给出更准确的置信区间。观察 KLS 的置信区间发现,其包含了常数值 $b_1(z) \equiv 0.5, b_2(z) \equiv -0.2$,即使用 KLS 方法会让我们忽视 z 对于 x 和 y 的影响,而不能准确刻画两者的关系。

综上所述,在实际应用数据中, FKLS 不仅能给出准确估计,还能给出较 KLS 更为精确的置信区间。同时,我们利用这两种方法说明了环境与经济发展之间相互影响的关系:经济发展初期,随着人均 GDP 的提高,污染问题会愈演愈烈;但是当经济发展到一定程度之后,人均 GDP 的提高反而能帮助减缓污染问题。当经济发展主要依赖于第二产业时,人均 GDP 的提高会加速污染问题;但是当第二产业在经济发展中占比较小时,人均 GDP 的提高反而会减缓环境污染。

6 总结

本文在变系数模型中提出了新的估计方法: 函数型核加权最小二乘法 (FKLS). 与传统的 KLS 方法进行比较之后发现, FKLS 的估计更为稳健, 且当 $t \rightarrow 0$ 时, 两种估计方法等价 (扰动项方差存在时). 通过理论推导, 证明了在一定条件下 FKLS 的相合性和渐近正态性, 并且提出了其自适应估计. 基于不同情形的蒙特卡罗试验发现: 扰动项不论服从何种分布, FKLS 的估计方差均小于 KLS; 当扰动项分布厚尾或不对称时, FKLS 的估计结果明显优于 KLS. 将 FKLS 方法应用到环境与经济的问题上发现国内四个地区的人均 PM_{2.5} 排放与人均 GDP 之间的关系与 EKC 理论吻合, 同时第二产业占比又会对两者的关系造成一定影响. 在实际数据中, FKLS 估计出来的置信区间明显优于 KLS.

本文的研究只局限于数据是随机样本的情形, 没有考虑自变量的相关性、非平稳性以及内生性. 若能将 FKLS 推广到面板数据中, 那么在实际应用中将有更大的发挥空间. 在分析环境与经济二者关系时, 我们没有考虑地理因素, 若将地理因素 (空间相关性) 也引入模型中, 那么会得到更加准确的结论. 这些拓展会带来新的理论挑战, 将会是重要的研究课题.

参考文献

- [1] Cleveland W, Mallows C, Mcrae J. ATS methods: Nonparametric regression for non-gaussian data[J]. Journal of the American Statistical Association, 1993, 88(423): 821-835.
- [2] Chen R, Tsay R S. Functional-coefficient autoregressive models[J]. Journal of the American Statistical Association, 1993, 88(421): 298-308.
- [3] Hastie T, Tibshirani R. Varying-coefficient models[J]. Journal of the Royal Statistical Society, 1993, 55(4): 757-796.
- [4] Li Q, Huang C J, Li D, et al. Semiparametric smooth coefficient models[J]. Journal of Business and Economic Statistics, 2002, 20(3): 412-422.
- [5] Cai Z, Fan J, Yao Q. Functional-coefficient regression models for nonlinear time series[J]. Journal of the American Statistical Association, 2000, 95: 941-956.
- [6] 蔡宗武, 陈琳娜, 方颖. 人民币汇率的半参数预测模型 [J]. 系统工程理论与实践, 2012, 32(4): 685-692.
Cai Z W, Chen L N, Fang Y. Semi-parametric forecasting model for USD/CNY exchange rate[J]. Systems Engineering — Theory & Practice, 2012, 32(4): 685-692.
- [7] Cai Z, Li Q, Park J Y. Functional-coefficient models for nonstationary time series data[J]. Journal of Econometrics, 2009, 148(2): 101-113.
- [8] Xiao Z. Functional-coefficient cointegration models[J]. Journal of Econometrics, 2009, 152(2): 81-92.
- [9] Tu Y, Wang Y. Adaptive estimation of functional-coefficient cointegration models with nonstationary volatility[R]. Technical Report, Peking University, 2017.
- [10] Chambers R L, Heathcote C R. On the estimation of slope and the identification of outliers in linear regression[J]. Biometrika, 1981, 68(1): 21-33.
- [11] Heathcote C R. Linear regression by functional least squares[J]. Journal of Applied Probability, 1982, 19(A): 225-239.
- [12] Csorgo S. The theory of functional least squares[J]. Journal of the Australian Mathematical Society, 1983, 34(3): 336-355.
- [13] Heathcote C R, Welsh A H. The robust estimation of autoregressive processes by functional least squares[J]. Journal of Applied Probability, 1983, 20(4): 737-753.
- [14] Welsh A H, Nicholls D F. Robust estimation of regression models with dependent regressors: The functional least squares approach[J]. Econometric Theory, 1986, 2(1): 132-150.
- [15] Heathcote C R, Welsh A H. Multivariate functional least squares[J]. Journal of Multivariate Analysis, 1988, 25(1): 45-64.
- [16] Silverman B W. Density estimation for statistics and data analysis[M]. Chapman and Hall, 1986.
- [17] Grossman G M, Krueger A B. Environmental impacts of a North American free trade agreement[J]. Social Science Electronic Publishing, 1991, 8(2): 223-250.
- [18] 彭水军, 包群. 经济增长与环境污染 —— 环境库兹涅茨曲线假说的中国检验 [J]. 财经问题研究, 2006(8): 3-17.
Peng S J, Bao Q. Economic growth and environmental pollution: An empirical test for the environmental Kuznets curve hypothesis in China[J]. Research on Financial and Economic Issues, 2006(8): 3-17.
- [19] 虞义华, 郑新业, 张莉. 经济发展水平、产业结构与碳排放强度 —— 中国省级面板数据分析 [J]. 经济理论与经济管理, 2011(3): 72-81.

- Yu Y H, Zheng X Y, Zhang L. Carbon dioxide emission and economic development: A panel data analysis[J]. *Economic Theory and Business Management*, 2011(3): 72-81.
- [20] 王志华, 温宗国, 闫芳, 等. 北京环境库兹涅茨曲线假设的验证 [J]. *中国人口·资源与环境*, 2007(2): 40-47.
Wang Z H, Wen Z G, Yan F, et al. Verifying the environmental Kuznets curve hypothesis and its conditions in Beijing[J]. *China Population, Resources and Environment*, 2007(2): 40-47.
- [21] Bierens H J. *Introduction to the mathematical and statistical foundations of econometrics*[M]. Cambridge University Press, 2004.
- [22] van Der Vaart A W. *Asymptotic statistics*[M]. Cambridge University Press, 1998.

附录

定理 1 的证明 首先定义 $\phi_j(\mathbf{b}; t, z, h) = \frac{1}{h} \exp\{it(y_j - \mathbf{x}_j^\top \mathbf{b})\} K_j(z, h)$. 对 $\phi_j(\mathbf{b}; t, z, h)$ 求期望得到:

$$\begin{aligned} & E[\phi_1(\mathbf{b}; t, z, h)] \\ &= E\left[\frac{1}{h} \exp\{it(y_j - \mathbf{x}_j^\top \mathbf{b})\} K_j(z, h)\right] \\ &= \frac{1}{h} \int \exp[it(\varepsilon_1 + \mathbf{x}_1^\top (\mathbf{b}_0(z_1) - \mathbf{b}))] K\left(\frac{z_1 - z}{h}\right) f(\varepsilon_1, \mathbf{x}_1, z_1) d\varepsilon_1 d\mathbf{x}_1 dz_1 \\ &= \int \exp[it(\varepsilon_1 + \mathbf{x}_1^\top (\mathbf{b}_0(z + hv) - \mathbf{b}))] K(v) f(\varepsilon_1, \mathbf{x}_1, z + hv) d\varepsilon_1 d\mathbf{x}_1 dv \\ &= \int \exp[it(\varepsilon_1 + \mathbf{x}_1^\top (\mathbf{b}_0(z) - \mathbf{b}))][1 + O(h)] K(v) [f(\varepsilon_1, \mathbf{x}_1, z) + O(h)] d\varepsilon_1 d\mathbf{x}_1 dv \\ &= \int \exp[it(\varepsilon_1 + \mathbf{x}_1^\top (\mathbf{b}_0(z) - \mathbf{b}))] f(\varepsilon_1, \mathbf{x}_1, z) d\varepsilon_1 d\mathbf{x}_1 + O(h) \\ &= f(z) \int \exp[it(\varepsilon_1 + \mathbf{x}_1^\top (\mathbf{b}_0(z) - \mathbf{b}))] f(\varepsilon_1, \mathbf{x}_1 | z) d\varepsilon_1 d\mathbf{x}_1 + O(h) \\ &= f(z) E[\exp(it\varepsilon_1 + i\mathbf{x}_1^\top (\mathbf{b}_0(z) - \mathbf{b})) | z] + O(h) \\ &= f(z) E[\exp(it\varepsilon + i\mathbf{x}^\top (\mathbf{b}_0(z) - \mathbf{b})) | z] + O(h) \\ &= f(z) E(\exp it(y - \mathbf{x}^\top \mathbf{b}) | z) + O(h). \end{aligned}$$

其中第二个等式由模型 (1) 的假设得到, 第三个等式做了变量代换 $v = \frac{z_1 - z}{h}$, 第四个等号由假设 1. (iii) 中 $b(z)$, $f(\varepsilon, x, z)$ 的李普希兹条件得到, 第五个等号来自于假设 1. (v) $\int K(v) dv = 1$. 类似可证

$$VAR[\phi_1(\mathbf{b}; t, z, h)] = o(1).$$

由此可得当 $n \rightarrow \infty$ 时,

$$E[\phi_1(\mathbf{b}; t, z, h)] \xrightarrow{P} f(z) E(\exp it(y - \mathbf{x}^\top \mathbf{b}) | z).$$

又因为 $E \sup_{\mathbf{b} \in \beta} |\phi_1(\mathbf{b}; t, z, h)| \leq E \frac{1}{h} K_1(z, h) = 1$, 由 Bierens^[21] 定理 6.10 可知, $\forall \delta > 0$,

$$\lim_{n \rightarrow \infty} P \left[\sup_{\mathbf{b} \in \beta} |\Phi_n(\mathbf{b}; t, z, h) - |f(z) E(\exp it(y - \mathbf{x}^\top \mathbf{b}) | z)|^2| > \delta \right] = 0.$$

下面, 我们将证明 $\forall \delta > 0$, $\sup_{|\mathbf{b}_0(z) - \mathbf{b}| \geq \delta} |f(z) E(\exp it(y - \mathbf{x}^\top \mathbf{b}) | z)|^2 < |f(z) E(\exp it\varepsilon | z)|^2$

$$\begin{aligned} & |E(\exp it(y - \mathbf{x}^\top \mathbf{b}) | z)|^2 \\ &= |E(\exp\{it(\varepsilon + \mathbf{x}^\top (\mathbf{b}_0(z) - \mathbf{b}))\} | z)|^2 \\ &= \left| \iint \exp\{it(\varepsilon + \mathbf{x}^\top (\mathbf{b}_0(z) - \mathbf{b}))\} f(\varepsilon, \mathbf{x} | z) d\varepsilon d\mathbf{x} \right|^2 \\ &= \left| \int \exp\{it\varepsilon\} f(\varepsilon | z) d\varepsilon \int \exp\{it\mathbf{x}^\top (\mathbf{b}_0(z) - \mathbf{b})\} f(\mathbf{x} | z) d\mathbf{x} \right|^2 \\ &= |E(\exp\{it\varepsilon\} | z) E(\exp\{it\mathbf{x}^\top (\mathbf{b}_0(z) - \mathbf{b})\} | z)|^2 \\ &< |E(\exp\{it\varepsilon\} | z)|^2, \end{aligned}$$

其中第四个等号由假设 1. (ii) 中 ε 和 \mathbf{x} 关于 z 条件独立得到, 而最后一个不等号由 $|E(\exp\{it\mathbf{x}^\top \delta\} | z)|^2 < 1$ 得到.

由于 $\hat{\mathbf{b}}_n(t, z, h) = \arg \max_{\mathbf{b} \in \Theta} \Phi_n(\mathbf{b}; t, z, h)$, 综合上述所证结论及 van Der Vaart^[22] 定理 5.7 可得 $\hat{\mathbf{b}}_n(t, z, h)$ 的相合性.

定理 2 的证明 对于固定的 t, z , FKLS 估计量 $\hat{\mathbf{b}}_n(t, z, h)$ 满足以下一阶条件:

$$\frac{\partial \Phi_n(\mathbf{b}; t, z, h)}{\partial \mathbf{b}} = 0,$$

其中

$$\frac{\partial \Phi_n(\mathbf{b}; t, z, h)}{\partial \mathbf{b}} = \frac{1}{n^2 h^2} \sum_{j=1}^n \sum_{s=1}^n t K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) \sin(t(\hat{\varepsilon}_j - \hat{\varepsilon}_s)).$$

下面简记 $\Psi(\mathbf{b}; t, z, h) = \frac{\partial \Phi_n(\mathbf{b}; t, z, h)}{\partial \mathbf{b}}$, $\Upsilon(\mathbf{b}; t, z, h) = \frac{\partial^2 \Phi_n(\mathbf{b}; t, z, h)}{\partial \mathbf{b}^2}$. 将 $\Psi(\hat{\mathbf{b}}_n(t, z, h); t, z, h)$ 在 $\mathbf{b}_0(z)$ 处进行泰勒展开得到:

$$\Psi(\hat{\mathbf{b}}_n(t, z, h); t, z, h) = \Psi(\mathbf{b}_0(z); t, z, h) + \Upsilon(\mathbf{b}_0(z); t, z, h)(\hat{\mathbf{b}}_n(t, z, h) - \mathbf{b}_0(z)) + o(|\hat{\mathbf{b}}_n(t, z, h) - \mathbf{b}_0(z)|).$$

又由 $\Psi(\hat{\mathbf{b}}_n(t, z, h); t, z, h) = 0$ 以及 $\hat{\mathbf{b}}_n(t, z, h)$ 的相合性, 计算得到: 当 $n \rightarrow \infty$ 时,

$$-\Upsilon(\mathbf{b}_0(z); t, z, h)(\hat{\mathbf{b}}_n(t, z, h) - \mathbf{b}_0(z)) = \Psi(\mathbf{b}_0(z); t, z, h) + o_p(1).$$

若下面我们能证明 $\sqrt{nh}\Psi(\mathbf{b}_0(z); t, z, h) \xrightarrow{d} N(0, \Omega)$, $-\Upsilon(\mathbf{b}_0(z); t, z, h) \xrightarrow{p} \mathbf{A}$, 则由 Slutsky 定理可得

$$\sqrt{nh}(\hat{\mathbf{b}}_n(t, z, h) - \mathbf{b}_0(z)) \xrightarrow{d} N(0, \mathbf{A}^{-1}\Omega\mathbf{A}^{-1}).$$

首先不妨简记 $d_j = \mathbf{b}_0(z_j) - \mathbf{b}_0(z)$, 那么

$$h^2 \Psi(\mathbf{b}_0(z); t, z, h) = \frac{1}{n^2} \sum_{j=1}^n \sum_{s=1}^n t K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) \sin [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s]$$

为一个 U 统计量, 其核为 $\zeta_{js}(\mathbf{b}_0(z); t, z, h) = t K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) \sin [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s]$.

下面我们将求解

$$E\zeta_{js}(\mathbf{b}_0(z); t, z, h), E(\zeta_{js}(\mathbf{b}; t, z, h)^\top \zeta_{js}(\mathbf{b}; t, z, h))$$

以及

$$COV(\zeta_{12}(\mathbf{b}_0(z); t, z, h), \zeta_{13}(\mathbf{b}_0(z); t, z, h)),$$

进而得到 U 统计量的渐近正态分布.

$$\begin{aligned} & E\zeta_{js}(\mathbf{b}_0(z); t, z, h) \\ &= E\{t K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) \sin [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s]\} \\ &= \int t K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) \sin [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s] f(\varepsilon_j, \mathbf{x}_j, z_j) f(\varepsilon_s, \mathbf{x}_s, z_s) d\varepsilon_j d\mathbf{x}_j dz_j d\varepsilon_s d\mathbf{x}_s dz_s \\ &= \int t h^2 K(v) K(u) (\mathbf{x}_j - \mathbf{x}_s) \sin [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top (\mathbf{b}_0(z + hv) - \mathbf{b}_0(z)) - t\mathbf{x}_s^\top (\mathbf{b}_0(z + hu) - \mathbf{b}_0(z))] \\ &\quad \times f(\varepsilon_j, \mathbf{x}_j, z + hv) f(\varepsilon_s, \mathbf{x}_s, z + hu) d\varepsilon_j d\mathbf{x}_j dv d\varepsilon_s d\mathbf{x}_s du \\ &= \int t h^2 K(v) K(u) (\mathbf{x}_j - \mathbf{x}_s) [\sin(t(\varepsilon_j - \varepsilon_s)) + O(h)] \left[f(\varepsilon_j, \mathbf{x}_j, z) + f_z(\varepsilon_j, \mathbf{x}_j, z) h v + \frac{1}{2} f_{zz}(\varepsilon_j, \mathbf{x}_j, z) h^2 v^2 + O(h^3) \right] \\ &\quad \times [f(\varepsilon_s, \mathbf{x}_s, z) + f_z(\varepsilon_s, \mathbf{x}_s, z) h u + \frac{1}{2} f_{zz}(\varepsilon_s, \mathbf{x}_s, z) h^2 u^2 + O(h^3)] d\varepsilon_j d\mathbf{x}_j dv d\varepsilon_s d\mathbf{x}_s du \\ &= \int t h^2 K(v) K(u) (\mathbf{x}_j - \mathbf{x}_s) \sin(t(\varepsilon_j - \varepsilon_s)) \left[\frac{1}{2} f_{zz}(\varepsilon_s, \mathbf{x}_s, z) u^2 f(\varepsilon_j, \mathbf{x}_j, z) \right. \\ &\quad \left. + \frac{1}{2} f_{zz}(\varepsilon_j, \mathbf{x}_j, z) u^2 f(\varepsilon_s, \mathbf{x}_s, z) \right] d\varepsilon_j d\mathbf{x}_j dv d\varepsilon_s d\mathbf{x}_s du + O(h^5) \\ &= h^4 c_1 + O(h^5), \end{aligned}$$

这里, 第三个等式进行了变量代换 $v = \frac{z_j - z}{h}$, $u = \frac{z_s - z}{h}$, 第四个等式利用了假设 1. (iii) 中 $\mathbf{b}_0(z)$ 的李普希兹条件, 同时对 $f(\varepsilon_j, \mathbf{x}_j, z_j + hv)$ 和 $f(\varepsilon_s, \mathbf{x}_s, z_s + hu)$ 进行泰勒展开, 其中 $c_1 = \int t K(v) K(u) (\mathbf{x}_j - \mathbf{x}_s) \sin(t(\varepsilon_j - \varepsilon_s))$

× [½f_{zz}(ε_s, x_s, z)u²f(ε_j, x_j, z) + ½f_{zz}(ε_j, x_j, z)u²f(ε_s, x_s, z)]dε_jd \mathbf{x}_j dv_sdε_sd \mathbf{x}_s du, f_z(ε_s, x_s, z) 和 f_{zz}(ε_s, x_s, z) 表示 f(ε_s, x_s, z) 关于 z 的一阶和二阶导数. 由假设 2. (ii) 的条件可得 $E\frac{\sqrt{nh}}{h^2}\zeta_{js}(\mathbf{b}_0(z); t, z, h) \rightarrow 0$.

$$\begin{aligned} & E(\zeta_{js}(\mathbf{b}; t, z, h))^\top \zeta_{js}(\mathbf{b}; t, z, h) \\ &= E\{t^2 K_j^2(z, h) K_s^2(z, h) (\mathbf{x}_j - \mathbf{x}_s)^\top (\mathbf{x}_j - \mathbf{x}_s) \sin^2 [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s]\} \\ &= \int t^2 K_j^2(z, h) K_s^2(z, h) (\mathbf{x}_j - \mathbf{x}_s)^\top (\mathbf{x}_j - \mathbf{x}_s) \sin^2 [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s] \times \\ & \quad f(\varepsilon_j, \mathbf{x}_j, z_j) f(\varepsilon_s, \mathbf{x}_s, z_s) d\varepsilon_j d\mathbf{x}_j dz_j d\varepsilon_s d\mathbf{x}_s dz_s \\ &= \int t^2 h^2 K^2(v) K^2(u) (\mathbf{x}_j - \mathbf{x}_s)^\top (\mathbf{x}_j - \mathbf{x}_s) \sin^2 [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top (\mathbf{b}_0(z + hv) - \mathbf{b}_0(z)) - t\mathbf{x}_s^\top (\mathbf{b}_0(z + hu) - \mathbf{b}_0(z))] \times \\ & \quad f(\varepsilon_j, \mathbf{x}_j, z + hv) f(\varepsilon_s, \mathbf{x}_s, z + hu) d\varepsilon_j d\mathbf{x}_j dv d\varepsilon_s d\mathbf{x}_s du \\ &< \int t^2 h^2 K^2(v) K^2(u) (\mathbf{x}_j - \mathbf{x}_s)^\top (\mathbf{x}_j - \mathbf{x}_s) f(\varepsilon_j, \mathbf{x}_j, z + hv) f(\varepsilon_s, \mathbf{x}_s, z + hu) d\varepsilon_j d\mathbf{x}_j dv d\varepsilon_s d\mathbf{x}_s du \\ &= t^2 h^2 \left[\int K^2(v) dv \right]^2 f(z)^2 E[(\mathbf{x}_j - \mathbf{x}_s)^\top (\mathbf{x}_j - \mathbf{x}_s) | z] + O(h^3). \end{aligned}$$

这里, 第三个等式进行了变量代换 $v = \frac{z_j - z}{h}$, $u = \frac{z_s - z}{h}$, 同时利用了假设 1. (iii) 中 $\mathbf{b}_0(z)$ 的李普希兹条件, 第四个不等号成立是由于 $|\sin^2 [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top (\mathbf{b}_0(z_j) - \mathbf{b}_0(z)) - t\mathbf{x}_s^\top (\mathbf{b}_0(z_s) - \mathbf{b}_0(z))]| < 1$. 由假设 1. (v) 得到 $\int K^2(v) dv < \infty$, 假设 2. (ii) 得到 $E[(\mathbf{x}_j - \mathbf{x}_s)^\top (\mathbf{x}_j - \mathbf{x}_s) | z] < \infty$, 进而得到 $E(\zeta_{js}(\mathbf{b}; t, z, h))^\top \zeta_{js}(\mathbf{b}; t, z, h) < \infty$.

$$\begin{aligned} & COV(\zeta_{12}(\mathbf{b}_0(z); t, z, h), \zeta_{13}(\mathbf{b}_0(z); t, z, h)) \\ &= E\{t^2 K_1^2(z, h) K_2(z, h) K_3(z, h) (\mathbf{x}_1 - \mathbf{x}_2) (\mathbf{x}_1 - \mathbf{x}_3)^\top \sin [t(\varepsilon_1 - \varepsilon_2) + t\mathbf{x}_1^\top \mathbf{d}_1 - t\mathbf{x}_2^\top \mathbf{d}_2] \\ & \quad \times \sin [t(\varepsilon_1 - \varepsilon_3) + t\mathbf{x}_1^\top \mathbf{d}_1 - t\mathbf{x}_3^\top \mathbf{d}_2]\} \\ &= \int t^2 K_1^2(z, h) K_2(z, h) K_3(z, h) (\mathbf{x}_1 - \mathbf{x}_2) (\mathbf{x}_1 - \mathbf{x}_3)^\top \sin [t(\varepsilon_1 - \varepsilon_2) + t\mathbf{x}_1^\top \mathbf{d}_1 - t\mathbf{x}_2^\top \mathbf{d}_2] \\ & \quad \times \sin [t(\varepsilon_1 - \varepsilon_3) + t\mathbf{x}_1^\top \mathbf{d}_1 - t\mathbf{x}_3^\top \mathbf{d}_2] f(\varepsilon_1, \mathbf{x}_1, z_1) f(\varepsilon_2, \mathbf{x}_2, z_2) f(\varepsilon_3, \mathbf{x}_3, z_3) d\varepsilon_1 d\mathbf{x}_1 dz_1 d\varepsilon_2 d\mathbf{x}_2 dz_2 d\varepsilon_3 d\mathbf{x}_3 dz_3 \\ &= \int t^2 h^3 K^2(v) K(u) K(r) (\mathbf{x}_1 - \mathbf{x}_2) (\mathbf{x}_1 - \mathbf{x}_3)^\top \sin [t(\varepsilon_1 - \varepsilon_2) + O(h)] \\ & \quad \times \sin [t(\varepsilon_1 - \varepsilon_3) + O(h)] f(\varepsilon_1, \mathbf{x}_1, z + hv) f(\varepsilon_2, \mathbf{x}_2, z + hu) f(\varepsilon_3, \mathbf{x}_3, z + hr) d\varepsilon_1 d\mathbf{x}_1 dv d\varepsilon_2 d\mathbf{x}_2 du d\varepsilon_3 d\mathbf{x}_3 dr \\ &= \int t^2 h^3 K^2(v) K(u) K(r) (\mathbf{x}_1 - \mathbf{x}_2) (\mathbf{x}_1 - \mathbf{x}_3)^\top \sin [t(\varepsilon_1 - \varepsilon_2)] \sin [t(\varepsilon_1 - \varepsilon_3)] \\ & \quad \times f(\varepsilon_1, \mathbf{x}_1, z) f(\varepsilon_2, \mathbf{x}_2, z) f(\varepsilon_3, \mathbf{x}_3, z) d\varepsilon_1 d\mathbf{x}_1 dv d\varepsilon_2 d\mathbf{x}_2 du d\varepsilon_3 d\mathbf{x}_3 dr + O(1) \\ &= t^2 h^3 f^3(z) \int K^2(v) dv E[(\mathbf{x}_1 - \mathbf{x}_2) (\mathbf{x}_1 - \mathbf{x}_3)^\top | z] E\{\sin [t(\varepsilon_1 - \varepsilon_2)] \sin [t(\varepsilon_1 - \varepsilon_3)] | z\} + O(h^4) \\ &= \frac{t^2 h^3}{2} f^3(z) \int K^2(v) dv \mathbf{A}(z) \Delta(t, z) + O(h^4). \end{aligned}$$

这里, 第三个等式进行了变量代换 $v = \frac{z_j - z}{h}$, $u = \frac{z_s - z}{h}$, 同时利用了假设 1. (iii) 中 $\mathbf{b}_0(z)$ 的李普希兹条件, 第五个等式是由假设 2. (i) 确保的, 最后的结果 $\frac{t^2 h^3}{2} f^3(z) \int K^2(v) dv \mathbf{A}(z) \Delta(t, z)$ 有界利用了条件 1. (iii). 其中 $\Delta(t, z) = u^2(t, z)(1 - u(2t, z)) + v^2(t, z)(1 + u(2t, z)) - 2v(t, z)u(t, z)v(2t, z)$, $\mathbf{A}(z) = E(\mathbf{x}_1 \mathbf{x}_1^\top | z) - E(\mathbf{x}_1 | z)E(\mathbf{x}_1 | z)^\top$, $u(t, z) = E(\cos t\varepsilon | z)$, $v(t, z) = E(\sin t\varepsilon | z)$.

结合上面三个等式以及 van Der Vaart^[21] 定理 12.3 可得:

$$\sqrt{nh^2} \Psi(\mathbf{b}_0; t, z, h) \xrightarrow{d} N\left(0, 2t^2 h^3 f^3(z) \int K^2(v) dv \mathbf{A}(z) \Delta(t, z)\right).$$

进一步可以得到:

$$\sqrt{nh} \Psi(\mathbf{b}_0; t, z, h) \xrightarrow{d} N\left(0, 2t^2 f^3(z) \int K^2(v) dv \mathbf{A}(z) \Delta(t, z)\right).$$

下面我们将证明 $\Upsilon(\mathbf{b}_0(z); t, z, h)$ 的渐近性质. 首先 $\Upsilon(\mathbf{b}_0(z); t, z, h)$ 的形式如下:

$$\Upsilon(\mathbf{b}_0(z); t, z, h) = -\frac{1}{n^2 h^2} \sum_{j=1}^n \sum_{s=1}^n t^2 K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) (\mathbf{x}_j - \mathbf{x}_s)^\top \cos [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s].$$

对 $\Upsilon(\mathbf{b}_0(z); t, z, h)$ 求期望得到

$$\begin{aligned}
 & E[-\Upsilon(\mathbf{b}_0(z); t, z, h)] \\
 &= E \left\{ \frac{t^2}{h^2} K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) (\mathbf{x}_j - \mathbf{x}_s)^\top \cos [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s] \right\} \\
 &= \int \frac{t^2}{h^2} K_j(z, h) K_s(z, h) (\mathbf{x}_j - \mathbf{x}_s) (\mathbf{x}_j - \mathbf{x}_s)^\top \cos [t(\varepsilon_j - \varepsilon_s) + t\mathbf{x}_j^\top \mathbf{d}_j - t\mathbf{x}_s^\top \mathbf{d}_s] \\
 &\quad \times f(\varepsilon_j, \mathbf{x}_j, z_j) f(\varepsilon_s, \mathbf{x}_s, z_s) d\varepsilon_j d\mathbf{x}_j dz_j d\varepsilon_s d\mathbf{x}_s dz_s \\
 &= \int t^2 K(v) K(u) (\mathbf{x}_j - \mathbf{x}_s) (\mathbf{x}_j - \mathbf{x}_s)^\top \cos [t(\varepsilon_j - \varepsilon_s) + O(h)] \\
 &\quad \times f(\varepsilon_j, \mathbf{x}_j, z + hv) f(\varepsilon_s, \mathbf{x}_s, z + hu) d\varepsilon_j d\mathbf{x}_j dv d\varepsilon_s d\mathbf{x}_s du \\
 &= t^2 f^2(z) E[(\mathbf{x}_j - \mathbf{x}_s) (\mathbf{x}_j - \mathbf{x}_s)^\top | z] E[\cos t(\varepsilon_j - \varepsilon_s) | z] + O(h) \\
 &= t^2 f^2(z) E[(\mathbf{x}_j - \mathbf{x}_s) (\mathbf{x}_j - \mathbf{x}_s)^\top | z] (u^2(t, z) + v^2(t, z)) + O(h) \\
 &= 2t^2 f^2(z) \mathbf{A}(z) (u^2(t, z) + v^2(t, z)) + O(h).
 \end{aligned}$$

其中第三个等式进行了变量代换 $v = \frac{z_j - z}{h}$, $u = \frac{z_s - z}{h}$, 同时利用了假设 1. (iii) 中 $\mathbf{b}_0(z)$ 的李普希兹条件, 第四个等式利用了假设 1. (iii) 中 $f(\varepsilon, \mathbf{x}, z)$ 的李普希兹条件. 类似可证 $\text{VAR}[\Upsilon(\mathbf{b}_0(z); t, z, h)] = o(1)$, 进而得到 $-\Upsilon(\mathbf{b}_0(z); t, z, h) \xrightarrow{P} 2t^2 f^2(z) \mathbf{A}(z) (u^2(t, z) + v^2(t, z))$.

结合 $\Psi(\mathbf{b}_0; t, z, h)$ 和 $-\Upsilon(\mathbf{b}_0(z); t, z, h)$ 的渐近性质以及 Slutsky 定理可得: 当假设 1, 2 成立时, 有

$$\sqrt{nh}(\hat{\mathbf{b}}_n(t, z, h) - \mathbf{b}_0(z)) \xrightarrow{d} N \left(0, \frac{\int K^2(v) dv \Delta(t, z)}{2t^2 f(z) (u^2(t, z) + v^2(t, z))^2} \mathbf{A}(z)^{-1} \right).$$

定理证毕.