

横截面独立的长相依面板数据的公共均值变点

习代青¹, 傅承德², 庞天晓³

(1. 中南财经政法大学 统计与数学学院, 湖北武汉 430073;

2. 中央大学 统计研究所, 台湾桃园 320317;

3. 浙江大学 数学科学学院, 浙江杭州 310058)

摘要: 研究横截面独立的长相依面板数据的公共均值变点的估计问题. 文中分别在强变点信号, 中等变点信号和弱变点信号这三种情形下研究了公共变点的估计量, 建立了估计量的渐近性质, 包括相合性, 收敛速度和极限分布. 理论结果表明在变点信号和长相依之间存在着一种权衡关系. 具体来说, 在强变点信号情形下, 长相依不影响估计量的渐近性质; 在中等变点信号情形下, 长相依不影响估计量的收敛速度, 但影响估计量的极限分布; 在弱变点信号情形下, 长相依既影响估计量的收敛速度, 也影响估计量的极限分布. Monte Carlo模拟评估了估计量在有限样本情形下的表现, 并支持文中的理论结果.

关键词: 公共变点; 极限分布; 长相依; 面板数据

中图分类号: O212.1

文献标识码: A **文章编号:** 1000-4424(2024)03-0253-20

§1 引言

许多经济和金融时间序列的数据都含有单变点或者多变点的特征, 并兼具长相依的性质. 因此本文在横截面独立的假设下研究长相依面板数据的公共均值变点. 在该面板数据中, 有 N 个时间序列, 每个时间序列有 T 个观测值, 假设在每个时间序列(或大部分的时间序列)上存在一个公共变点, 笔者感兴趣的是: 当 N 很大时公共变点估计量的渐近性质.

面板数据中的公共变点现象在实际生活中是很常见的. 例如, 税收政策的改变会影响大部分人的收入; 2008年的全球金融危机和之前几年的新冠病毒大流行明显影响了每个国家的GDP数据和经济增长率数据.

尽管生活中的公共变点现象很普遍, 但文献中关于面板数据的变点理论研究却还很少. 在上世纪九十年代, 时间序列中的变点研究才扩展到了面板数据. Joseph等人的[1-2]是关于面板数

据变点分析的最早文献,在文中他们提出了一个随机变点模型,即在 N 个时间序列中存在着 N 个独立同分布的(i.i.d.)变点,他们研究了变点分布的估计问题. Joseph等人^[3]随后将该随机变点模型推广到了自回归面板数据模型. 后来, Bai^[4]在横截面独立的假设下研究了弱相依面板数据中的公共均值变点问题,他发现当变点信号足够强时,公共变点估计量是相合的. 值得注意的是,在平稳时间序列模型中,变点估计量是不相合的. 因此Bai^[4]的工作表明,当同时利用许多时间序列数据时,可以相合地估计出变点. 这是一个崭新的结论. 自Bai^[4]的开创性工作以来,面板数据中公共变点的检测和估计在计量经济学和统计学领域中得到了进一步的研究. 例如, Horváth等人^[5]以及Chan等人^[6]研究了面板数据中的公共均值变点的假设检验, Kim^[7-8]研究了面板数据中的时间趋势的公共变点估计, Baltagi等人^[9]研究了异质性面板数据中回归系数的公共变点估计, Li等人^[10]研究了面板数据中的多变点估计, Horváth等人^[11]研究了面板数据中的公共均值变点的CUSUM统计量, Fisher等人^[12]用贝叶斯的方法研究了面板数据中变点的统计推断. 然而,以上提到的所有文献都假设面板数据在时间维度上是i.i.d.或弱相依的. 关于长相依面板数据中变点问题的研究还是非常少见的.

本文的目的是沿着Bai^[4]的工作,对横截面独立的长相依面板数据的公共均值变点的估计问题进行系统的研究. 讨论了所有情形的变点信号:强变点信号,中等变点信号和弱变点信号. 值得一提的是, Bai^[4]和其他文献均还未研究过弱变点信号的情形(即使是在弱相依的情形下). 本文建立了估计量的渐近性质,包括相合性,收敛速度和极限分布. 本文的理论结果表明:在变点信号和长相依之间存在着一种权衡关系. 具体来说,在强变点信号情形下,长相依不影响估计量的渐近性质;在中等变点信号情形下,长相依不影响估计量的收敛速度,但影响估计量的极限分布;在弱变点信号情形下,长相依既影响估计量的收敛速度,也影响估计量的极限分布.

本文结构如下. §2给出模型的具体形式以及相应的理论结果. §3汇报Monte Carlo模拟结果,用于评估公共变点估计量在有限样本情形下的表现. §4对文章进行了总结. §5是论文的理论证明.

§2 模型和结论

先介绍一些符号. 符号“ $\xrightarrow{p}$ ”表示依概率收敛,“ $\xrightarrow{d}$ ”表示依分布收敛,“ $\Rightarrow$ ”表示概率测度弱收敛,“ $\stackrel{d}{=}$ ”表示同分布,“ $[x]$ ”表示 x 的整数部分, $a_N \asymp b_N$ 表示存在着两个正常数 c_1 和 c_2 使得对所有的 $N \geq 1$, $c_1 \leq a_N/b_N \leq c_2$ 成立,其中 $\{a_N, N \geq 1\}$ 和 $\{b_N, N \geq 1\}$ 是两个正常数序列.

研究公共变点模型(Bai^[4]曾在强变点信号和中等变点信号情形下研究过此模型,但他假设所有时间序列都是弱相依的)

$$y_{it} = \begin{cases} \mu_{i1} + x_{it}, & t = 1, \dots, k^0, \\ \mu_{i2} + x_{it}, & t = k^0 + 1, \dots, T, \end{cases} \quad i = 1, 2, \dots, N, \quad (1)$$

其中 N 表示时间序列或个体的个数, T 表示每个时间序列的样本容量, k^0 表示未知的公共均值变点. 这里考虑两种关于 T 的情形. 第一种情形假设 T 是固定的, k^0 可以取1到 $T-1$ 之间的任意整数;另一种情形假设 T 是无界的,但 k^0 不能太靠近1和 T ,即存在 $\tau^0 \in (0, 1)$ 使得 $k^0 = \lfloor \tau^0 T \rfloor$. 这是变点分析领域的一个常规假设,可参考文献[13]. 为了避免技术上的细枝末节,假设 $k^0 = \tau^0 T$,并称 τ^0 为模型(1)的公共变点分数. 此外与Bai^[4]文章一样,假设面板是横截面独立的. 即 y_{it} 在个体维度上是相互独立的. μ_{i1} 和 μ_{i2} 分别表示第 i 个时间序列的变前均值和变后均值. μ_{i1} 和 μ_{i2} 的大小不起本质作用,起本质作用的是它们的差. 值得一提的是,对于某些时间序列,允许 $\mu_{i2} - \mu_{i1}$ 为0,

即允许这些时间序列没有发生结构变化.

记

$$\lambda_N = \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2.$$

众所周知 λ_N 越大, 越容易估计出公共变点. 因此 λ_N 的大小将对公共变点估计量的渐近性质有本质性的影响. 下文称 λ_N 为模型(1)的变点信号(λ_N 可理解为单个时间序列的平均变点信号). 将在变点信号的三种不同情形下研究公共变点估计量的渐近性质: (1) 强变点信号, 即 $\lambda_N \rightarrow \infty$ ($N \rightarrow \infty$); (2) 中等变点信号, 即 $\lambda_N \rightarrow \lambda \in (0, \infty)$ ($N \rightarrow \infty$); (3) 弱变点信号, 即 $\lambda_N \rightarrow 0$ ($N \rightarrow \infty$).

本文假设模型误差 $\{x_{it}, t \geq 1\}$ 是长相依的. 即对每一个 $i \geq 1$, 存在一个常数 $d \in (0, 0.5)$ 使得

$$(1 - B)^d x_{it} = u_{it}, \quad (2)$$

其中 B 是时间维度上的滞后算子, u_{it} 在个体维度和时间维度上均为i.i.d.随机变量, 具有零均值和有限方差 $0 < \sigma_u^2 < \infty$. 下文中, 称 d 为分数差分参数或记忆参数.

上述关于 d 的限制($0 < d < 0.5$)保证了 x_{it} 在时间维度上是平稳遍历的. x_{it} 的一些重要性质(参见Sowell^[14])如下, 它们在模型的渐近分析中扮演着重要作用.

P1 对每一个给定的 $i \geq 1$, $\text{Var}(\sum_{t=1}^T x_{it}) = O(T^{1+2d})$. (其中 $T \rightarrow \infty$)

P2 对每一个给定的 $i \geq 1$, $\text{Cov}(x_{it}, x_{i,t+s}) = O(|s|^{-1+2d})$. (其中 $|s| \rightarrow \infty$)

P3 $\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \xrightarrow{p} E(x_{it}^2) = \frac{\Gamma(1-2d)}{\Gamma^2(1-d)} \sigma_u^2$. (其中 $N \rightarrow \infty$ 或 $T \rightarrow \infty$)

由性质P2, 对每一个给定的 $i \geq 1$, 有 $\sum_{s=1}^{\infty} |\text{Cov}(x_{it}, x_{i,t+s})| = \infty$. 这就是称 $\{x_{it}, i \geq 1, t \geq 1\}$ 为长相依面板数据的原因. 接下来用最小二乘的方法估计模型(1)中的公共变点. 定义 k^0 的最小二乘估计(简称为LSE)为

$$\hat{k} = \arg \min_{1 \leq k \leq T-1} S_{NT}(k),$$

其中

$$S_{NT}(k) = \sum_{i=1}^N \left[\sum_{t=1}^k (y_{it} - \bar{y}_i(k))^2 + \sum_{t=k+1}^T (y_{it} - \bar{y}_i^*(k))^2 \right],$$

$\bar{y}_i(k) = \frac{1}{k} \sum_{t=1}^k y_{it}$, $\bar{y}_i^*(k) = \frac{1}{T-k} \sum_{t=k+1}^T y_{it}$. 公共变点分数 τ^0 的LSE定义为 $\hat{\tau} = \hat{k}/T$.

为了研究LSE的渐近性质, 对模型(1)作如下假设.

条件A1 对每一个 $i \geq 1$, $\{x_{it}, t \geq 1\}$ 是长相依的, 即(2)成立, 其中 $d \in (0, 0.5)$, u_{it} 在个体维度和时间维度上均为i.i.d.随机变量, 具有零均值和有限方差 $0 < \sigma_u^2 < \infty$.

条件A2 $0 < \tau^0 < 1$.

条件A3 $N \rightarrow \infty$ 时(在下文中, T 无界时“ $N \rightarrow \infty$ ”均理解成“ $N, T \rightarrow \infty$ ”),

$$\frac{1}{T^{0.5-d}} = o\left(\frac{\lambda_N}{N}\right).$$

注1 在条件A1中, 由 u_{it} 的有限二阶矩假设使得作者能对长相依随机变量序列应用Hájek-Rényi不等式^[15]. Hájek-Rényi类型的不等式在变点分析中扮演着关键作用. u_{it} 的有限二阶矩假设还可以让作者应用 $\{x_{it}, t \geq 1\}$ 的泛函中心极限定理, 即

$$\frac{1}{T^{0.5+d}} \sum_{t=1}^{\lfloor T\tau \rfloor} x_{it} \Rightarrow \kappa_d B_d(\tau), \quad T \rightarrow \infty, \quad 0 \leq \tau \leq 1,$$

其中 B_d 是一个记忆参数为 d 的分数Brownian运动^[16],

$$\kappa_d = \{\sigma_u^2 \Gamma(1-2d) / [(1+2d)\Gamma(1+d)\Gamma(1-d)]\}^{1/2}, \quad (3)$$

参考[17]. 也可以假设 u_{it} 在个体维度上是异方差的. 但为了表述上的方便, 这里假设它在个体维度上是方差齐性的. 条件A2是变点分析领域的一个标准假设, 这个假设意味着当 T 有界时, k^0 可以是 $[1, T-1]$ 中的任何一个整数; 而当 T 无界时, k^0 不能太靠近样本的左右两个端点, 这是为了保证变点的可识别性. 条件A3不要求每一个时间序列都有变点, 而是要求平均变点信号需要比 $\frac{1}{T^{0.5-d}}$ 大得多.

接下来继续介绍几个假设.

条件A4 (强变点信号情形) 对每一个 $i \geq 1$, $\mu_{i2} - \mu_{i1}$ 是固定的, 且 $\lim_{N \rightarrow \infty} \lambda_N = \infty$.

条件A5 (中等变点信号情形) 对每一个 $i \geq 1$, 存在非负常数 Δ_i 使得 $\mu_{i2} - \mu_{i1} = \Delta_i/\sqrt{N}$, 且存在某个 C_0 使得 $\sup_{i \geq 1} |\Delta_i| \leq C_0$ 成立. 此外假设

$$\lim_{N \rightarrow \infty} \lambda_N = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Delta_i^2 = \lambda \in (0, \infty).$$

条件A6 (弱变点信号情形) $\lim_{N \rightarrow \infty} \lambda_N = 0$.

注2 (1) 条件A4假设变点幅度 $\mu_{i2} - \mu_{i1}$ 是固定的, 且允许某些变点幅度是零, 即允许某些时间序列没有发生结构变化. 但是 $\lambda_N \rightarrow \infty$ 这个假设要求存在着足够多的时间序列, 它们具有非零的变点幅度. (2) 条件A5要求 $\mu_{i2} - \mu_{i1}$ 的大小与 i 有关, 且可以写成 $\mu_{i2} - \mu_{i1} = \Delta_i/\sqrt{N}$, 其中 $\{|\Delta_i|, i \geq 1\}$ 是一致有界的, 即所有的变点幅度 $|\mu_{i2} - \mu_{i1}|$ 以 $1/\sqrt{N}$ 的速度收缩到零, 但允许某些时间序列没有发生结构变化. 给定条件A5, 条件A3等价于 $N = o(T^{0.5-d})$, 这意味着 T 的发散速度要快于 N . (3) 条件A6要求所有的变点幅度都需以较快的速度收缩到零. 条件A6和A3意味着 λ_N 收敛到零的速度要慢于 $\frac{N}{T^{0.5-d}}$. 若 λ_N 以太快的速度收敛到零, 那么估计 k^0 就会变得非常困难, 任何估计方法都将无能为力.

定理2.1 对于模型(1), 假设条件A1-A3成立, 有下面的结论.

(1) 条件A4成立时有

$$\lim_{N \rightarrow \infty} P(\hat{k} = k^0) = 1. \quad (4)$$

(2) 条件A5成立时, 当 $N, T \rightarrow \infty$ 有

$$\hat{k} - k^0 = O_p(1). \quad (5)$$

此外

$$\hat{k} - k^0 \xrightarrow{d} \arg \min_{l \in \{\dots, -2, -1, 0, 1, 2, \dots\}} W(l), \quad (6)$$

其中

$$W(l) = \begin{cases} -l\sqrt{\lambda} + 2\sigma_u \sum_{t=l}^{-1} (1-B)^{-d} Z_t, & l = -1, -2, \dots, \\ 0, & l = 0, \\ l\sqrt{\lambda} + 2\sigma_u \sum_{t=1}^l (1-B)^{-d} Z_t, & l = 1, 2, \dots, \end{cases}$$

$\{Z_t, t = \dots, -2, -1, 0, 1, 2, \dots\}$ 是i.i.d.的标准正态随机变量序列.

(3) 条件A6成立时, 当 $N, T \rightarrow \infty$ 有

$$|\hat{k} - k^0| = O_p(\lambda_N^{-1/(1-2d)}). \quad (7)$$

此外

$$\lambda_N^{1/(1-2d)}(\hat{k} - k^0) \xrightarrow{d} \arg \min_{-\infty < s < \infty} \Upsilon(s), \quad (8)$$

其中

$$\Upsilon(s) = \begin{cases} -s + 2\kappa_d B_d(s), & s < 0, \\ 0, & s = 0, \\ s + 2\kappa_d B_d(s), & s > 0, \end{cases}$$

κ_d 的定义见(3), $B_d(\cdot)$ 是一个双边分数Brownian运动.

定理2.1说明: (1) 在强变点信号情形下, $\hat{k}$ 是 k^0 的相合估计, 且这一结论与 T 是否无界是无关的. (2) 在中等变点信号情形下, $\hat{k}$ 不是 k^0 的相合估计, 但估计误差是依概率有界的, 这使得 $\hat{\tau}$ 仍是 τ^0 的相合估计. (3) 在弱变点信号情形下, $\hat{k}$ 的估计误差发散到无穷大(在概率意义下). 但是, $\hat{\tau}$ 仍是 τ^0 的相合估计, 这是因为条件A3保证了 $T\lambda_N^{1/(1-2d)} \rightarrow \infty$ 成立($N, T \rightarrow \infty$). 注意到当 $\lambda_N \rightarrow 0$ 时, $\lambda_N^{-1/(1-2d)}$ 是 $d \in (0, 0.5)$ 的单调增函数, 所以 d 越大, 准确地估计出公共变点就越困难, 这与直觉相符. (4) 定理2.1中的两个极限分布看起来是非常规的(但都是变点分析领域的常规极限分布), 这主要是因为变点估计量没有显示表达式, 所以只能通过研究目标函数的极限性质来刻画变点估计量的极限分布; 参考Bai^[4]和Csörgő等人^[13].

定理2.1还揭示了一个有趣的现象: 在弱变点信号情形下, $\hat{k}$ 的极限分布与单个时间序列中变点的LSE的极限分布非常相似, 参考[15]的定理8, [18]的定理2.3和定理2.4. 一个可能的理由是: 尽管有 N 条时间序列, 但每条时间序列的变点信号非常微弱, 把 N 条时间序列的变点信号加总后得到 λ_N , 而 λ_N 依然趋于0, 这意味着面板数据所带来的信息和单个时间序列所带来的信息并没有本质性的差别.

值得一提的是, 由定理2.1还可知在变点信号和长相依之间存在着一一种权衡关系. 即在强变点信号情形下, 长相依不影响估计量的渐近性质; 在中等变点信号情形下, 长相依不影响估计量的收敛速度, 但影响估计量的极限分布; 在弱变点信号情形下, 长相依既影响估计量的收敛速度, 也影响估计量的极限分布.

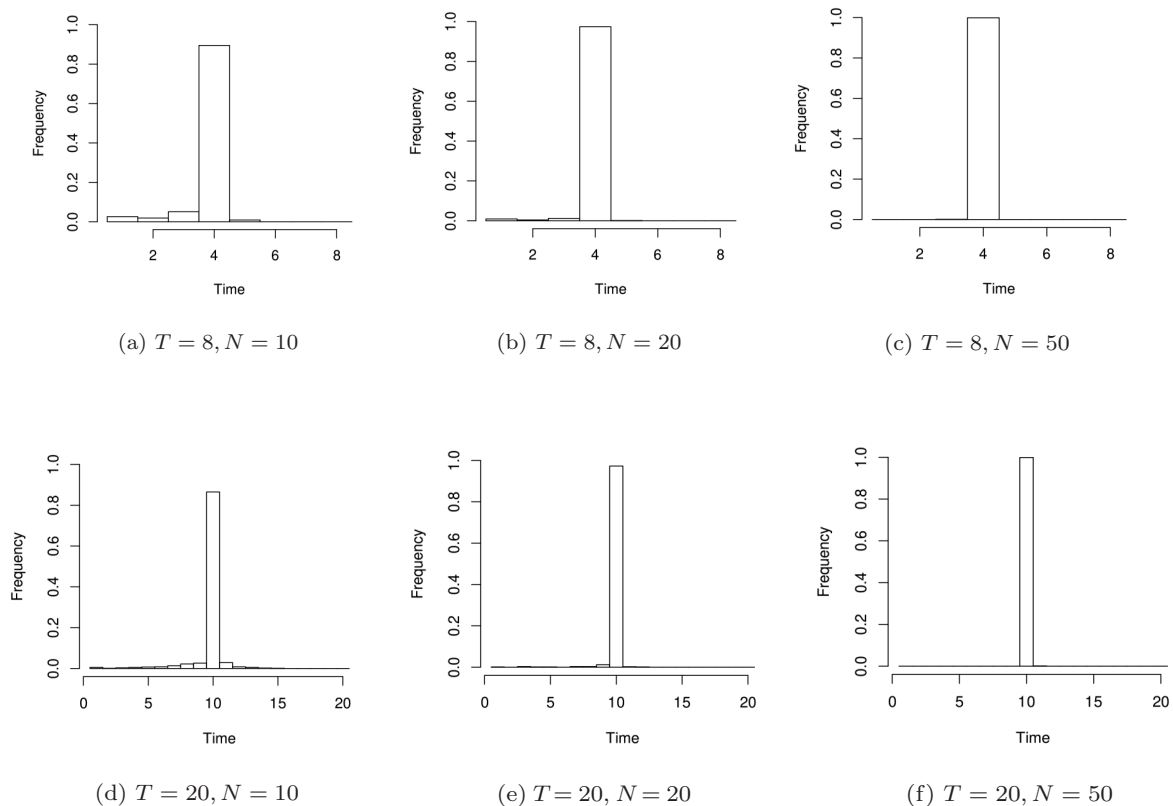
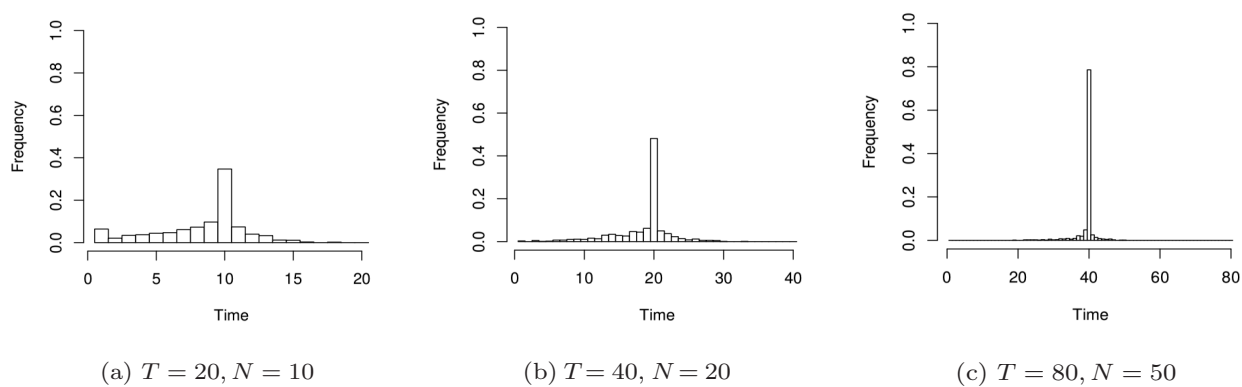
§3 模拟

为了评估公共变点估计量在有限样本情形下的表现, 做了一些模拟试验来考察估计量的估计误差以及估计量的有限样本分布与极限分布的吻合程度. 在所有的试验中, 长相依过程 $\{x_{it}, t \geq 1\}$ 的随机数是按照McLeod和Hipel^[19]以及Hosking^[20]的方式产生, 记忆参数 $d = 0.25$, 公共变点分数 $\tau^0 = 0.5$, $\hat{k}$ 在 $\{1, 2, \dots, T-1\}$ 中进行搜索.

3.1 估计误差

在这一节里, 试验的重复次数都是1000. 首先在强变点信号情形下评估 $\hat{k}$ 在有限样本情形下的表现. 由结论(4)可知 $\hat{k}$ 是 k^0 的相合估计, 且这一结论无要求 T 是无界的. 因此, 考虑 $T = 8$ 和20两种情形, 其中 $T = 8$ 代表小 T 这种情形, $T = 20$ 代表大 T 这种情形. 对于每一个 T , 考虑 $N \in \{10, 20, 50\}$. 变点幅度 $\mu_{i2} - \mu_{i1}$ 假设产生自 $U(-2, 2)$. 图1展示了模拟结果. 从图1可以看出, 即使 $N = 10$, 真实的公共变点(当 $T = 8$ 时它等于4, 当 $T = 20$ 时它等于10)都能够很精确地被估计出来, 并且随着 N 的增大, 估计精度逐步得到改进. 因此这里的模拟结果支持结论(4).

接下来在中等变点信号情形下评估 $\hat{k}$ 的有限样本表现. 由结论(5)可知 $\hat{k}$ 不是 k^0 的相合估计, 但是估计误差是依概率有界的. 注意到结论(5)的成立需要 $N, T \rightarrow \infty$, 此外回顾注2的第二部分, 考虑下列的 $(N, T) \in \{(10, 20), (20, 40), (50, 80)\}$. 变点幅度 $\mu_{i2} - \mu_{i1}$ 假设由 $U(-1, 1)$ 产生. 图2展

图1 $\tau^0 = 0.5$ 时强变点信号情形下 $\hat{k}$ 的直方图图2 $\tau^0 = 0.5$ 时中等变点信号情形下 $\hat{k}$ 的直方图

示了模拟结果. 可以看出, 总的来说 $\hat{k}$ 的估计误差是中等大小的, 且估计精度随着 (N, T) 的增大而逐步得到改善. 所以这里的模拟结果支持结论(5).

最后在弱变点信号情形下评估 $\hat{k}$ 的有限样本表现. 由结论(7)可知 $\hat{k}$ 不是 k^0 的相合估计, 且估计误差可以是无界. 除了变点幅度, 采用与中等变点信号情形的相同模拟设置; 这里假设 $\mu_{i2} - \mu_{i1}$ 由 $U(-0.5, 0.5)$ 产生. 图3展示了模拟结果. 可以看出, 即使 $N = 50$, $\hat{k}$ 的估计误差仍然很大, 但估计精度随着 (N, T) 的增大得到缓慢的改善. 因此这里的模拟结果支持结论(7).

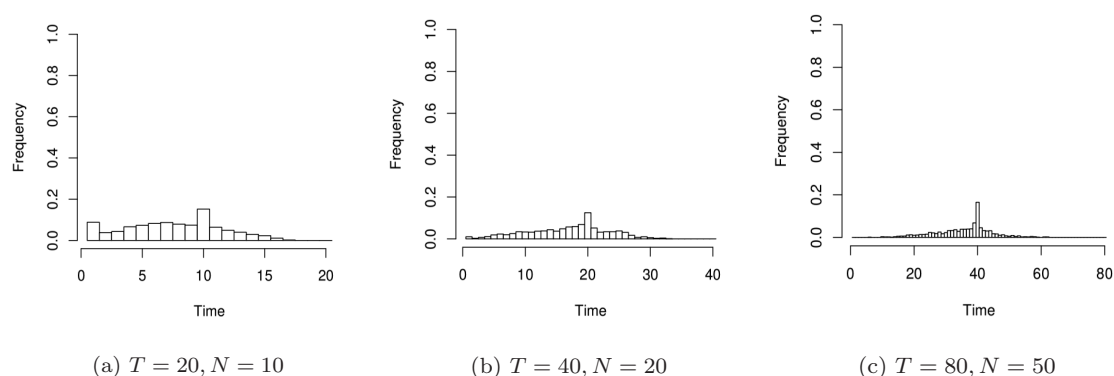


图3 $\tau^0 = 0.5$ 时弱变点信号情形下 $\hat{k}$ 的直方图

3.2 有限样本分布与极限分布

在实际问题中, 变点信号 λ_N 总是一个正常数(假设变点存在). 也就是说, 中等变点信号是现实生活中最常见的一种情形, 而强变点信号和弱变点信号可看作是两种极端情形. 因此, 接下来考察中等变点信号情形下变点估计量的有限样本分布与极限分布的吻合程度.

回忆 $d = 0.25$ 以及 $\tau^0 = 0.5$. 此外, 让 $\mu_{i2} - \mu_{i1} = \frac{\Delta_i}{\sqrt{N}}$ 中的 Δ_i 由 $U(-1, 4)$ 产生, 在(6)中取 $|l| \in \{0, 1, \dots, T/2\}$, $(T, N) \in \{(200, 15), (500, 40)\}$, 试验的重复次数是5000. 两个试验结果见图4. 从图4可以看出, $\hat{k} - k^0$ 的有限样本分布与极限分布的吻合度相当不错.

§4 总结

在本文中, 在横截面独立的假设下对长相依面板数据的公共均值变点估计提供了一个完整的理论分析. 具体来说, 分别在强, 中等和弱变点信号情形下, 研究了公共均值变点的最小二乘估计的渐近性质. 理论结果表明, 只有在强变点信号的情形下才能相合地估计出公共变点, 但公共变点分数的最小二乘估计在上述三种变点信号情形下总是相合的. 更重要的是, 本文发现在变点信号和长相依之间存在着一种权衡关系. 文中做了一些Monte Carlo模拟来评估估计量在有限样本情形下的表现, 模拟结果支持文中的理论发现.

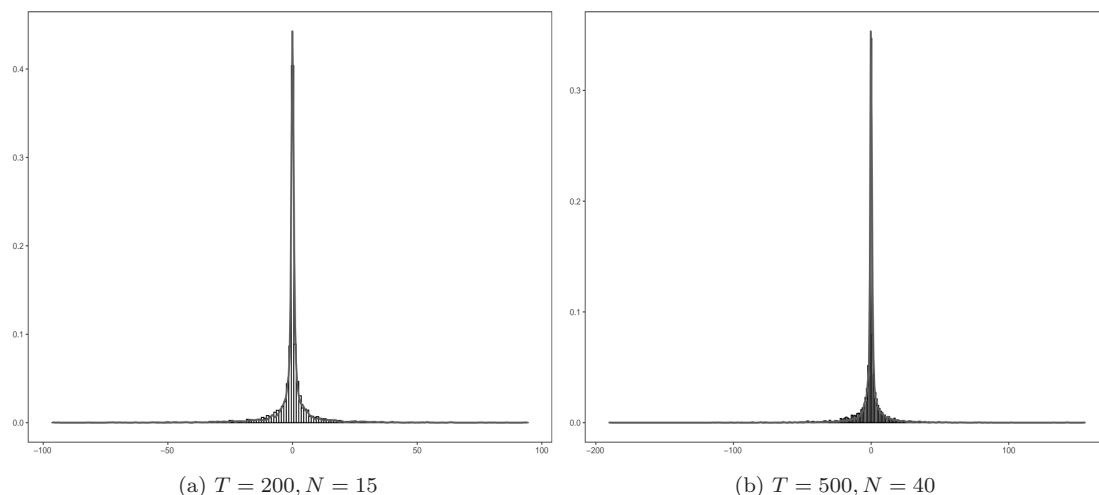


图4 中等变点信号情形下 $\hat{k} - k^0$ 的有限样本分布与极限分布(黑色曲线)

§5 证 明

引理5.1 对于记忆参数为 $d \in (0, 0.5)$ 的长相依过程 $\{\varphi_t, t \geq 1\}$ (即 $(1-B)^d \varphi_t = u_t$, 其中 $\{u_t, t \geq 1\}$ 是i.i.d.随机变量序列, 均值为0, 方差有限), 可以得到

$$\sup_{1 \leq k \leq n} \frac{1}{\sqrt{k}} \left| \sum_{t=1}^k \varphi_t \right| = O_p(n^d \sqrt{\ln n}), \quad \sup_{k \geq n} \frac{1}{k} \left| \sum_{t=1}^k \varphi_t \right| = O_p\left(\frac{1}{n^{0.5-d}}\right).$$

证 它们由Lavielle和Moulines^[15]中的Hájek-Rényi不等式推得.

引理5.2 条件A1成立时, 对于每一个 $i \geq 1$ 和正整数 $p < q$, 可以得到

$$\mathbb{E} \left[\frac{1}{q-p} \left(\sum_{t=1}^p x_{it} \right) \left(\sum_{t=p+1}^q x_{it} \right) \right] \leq O(p^{2d}).$$

证 根据性质P2, $\gamma_x(h) := \mathbb{E}(x_{it}x_{i,t+h}) = O(|h|^{2d-1})$. 因此

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{q-p} \left(\sum_{t=1}^p x_{it} \right) \left(\sum_{s=p+1}^q x_{is} \right) \right] \right| = \left| \frac{1}{q-p} \sum_{t=1}^p \sum_{s=p+1}^q \gamma_x(s-t) \right| \\ & \leq \frac{C}{q-p} \sum_{t=1}^p \sum_{s=p+1}^q (p+1-t)^{2d-1} = C \sum_{t=1}^p (p+1-t)^{2d-1} = O(p^{2d}). \end{aligned}$$

注意到 $S_{NT}(k)$ 的极小值点也是 $\frac{1}{\lambda_{NT}} \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right]$ 的极小值点, 因此证明思路如下: 首先证明 $\frac{1}{\lambda_{NT}} \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right]$ 以一定的速度接近它的数学期望, 然后把这个数学期望作为优化函数.

在给出定理2.1的证明之前, 先提供关于 $S_{NT}(k)$ 的两个分解式

$$\begin{aligned}
 S_{NT}(k) &= \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 - \frac{1}{k} \sum_{i=1}^N \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{1}{T-k} \sum_{i=1}^N \left(\sum_{t=k+1}^T x_{it} \right)^2 \\
 &\quad + 2 \frac{T-k^0}{T-k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} - 2 \frac{k^0-k}{T-k} \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it} \\
 &\quad + \frac{(k^0-k)(T-k^0)}{T-k} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2, \quad k \in [1, k^0],
 \end{aligned} \tag{9}$$

$$\begin{aligned}
 S_{NT}(k) &= \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 - \frac{1}{k} \sum_{i=1}^N \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{1}{T-k} \sum_{i=1}^N \left(\sum_{t=k+1}^T x_{it} \right)^2 \\
 &\quad + 2 \frac{k-k^0}{k} \sum_{i=1}^N \sum_{t=1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} - 2 \frac{k^0}{k} \sum_{i=1}^N \sum_{t=k^0+1}^k (\mu_{i1} - \mu_{i2}) x_{it} \\
 &\quad + \frac{(k-k^0)k^0}{k} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2, \quad k \in [k^0+1, T].
 \end{aligned} \tag{10}$$

以下引理5.3表明 $\frac{1}{\lambda_N T} \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right]$ 以速度 $\max \left\{ O_p \left(\frac{N \ln T}{\lambda_N T^{1-2d}} \right), O_p \left(\frac{1}{\sqrt{\lambda_N T^{0.5-d}}} \right) \right\}$ 接近它的数学期望 $\frac{1}{\lambda_N T} \mathbb{E} \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right]$.

引理5.3 对于模型(1), 假设条件A1成立, 那么对有界和无界的 T , 都有

$$\begin{aligned}
 &\sup_{1 \leq k \leq T} \left| \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] - \mathbb{E} \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] \right| \\
 &\leq \max \left\{ O_p \left(NT^{2d} \ln T \right), O_p \left(\sqrt{\lambda_N T^{0.5+d}} \right) \right\}.
 \end{aligned}$$

证 只给出 $k \leq k^0$ 情形下的证明, 因为 $k > k^0$ 情形下的证明是类似的. 对于 $k \leq k^0$, 由(9)可得

$$\begin{aligned}
 &\sup_{1 \leq k \leq k^0} \left| \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] - \mathbb{E} \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] \right| \\
 &= \sup_{1 \leq k \leq k^0} \left| \frac{1}{k} \sum_{i=1}^N \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{1}{k} \sum_{i=1}^N \mathbb{E} \left(\sum_{t=1}^k x_{it} \right)^2 \right. \\
 &\quad + \frac{1}{T-k} \sum_{i=1}^N \left(\sum_{t=k+1}^T x_{it} \right)^2 - \frac{1}{T-k} \sum_{i=1}^N \mathbb{E} \left(\sum_{t=k+1}^T x_{it} \right)^2 \\
 &\quad \left. - 2 \frac{T-k^0}{T-k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} + 2 \frac{k^0-k}{T-k} \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it} \right| \\
 &\leq \sup_{1 \leq k \leq k^0} \left| \frac{1}{k} \sum_{i=1}^N \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{1}{k} \sum_{i=1}^N \mathbb{E} \left(\sum_{t=1}^k x_{it} \right)^2 \right|
 \end{aligned}$$

$$\begin{aligned}
& + \sup_{1 \leq k \leq k^0} \left| \frac{1}{T-k} \sum_{i=1}^N \left(\sum_{t=k+1}^T x_{it} \right)^2 - \frac{1}{T-k} \sum_{i=1}^N \mathbb{E} \left(\sum_{t=k+1}^T x_{it} \right)^2 \right| \\
& + 2 \sup_{1 \leq k \leq k^0} \left| \frac{T-k^0}{T-k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} \right| + 2 \sup_{1 \leq k \leq k^0} \left| \frac{k^0-k}{T-k} \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it} \right|.
\end{aligned} \tag{11}$$

接下来分别研究(11)中的各项.

对于第一项, 由引理5.1和性质P1可知

$$\begin{aligned}
& \sup_{1 \leq k \leq k^0} \left| \frac{1}{k} \sum_{i=1}^N \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{1}{k} \sum_{i=1}^N \mathbb{E} \left(\sum_{t=1}^k x_{it} \right)^2 \right| \\
& \leq \sum_{i=1}^N \left(\sup_{1 \leq k \leq k^0} \frac{1}{\sqrt{k}} \left| \sum_{t=1}^k x_{it} \right| \right)^2 + \sum_{i=1}^N \sup_{1 \leq k \leq k^0} \frac{1}{k} \mathbb{E} \left(\sum_{t=1}^k x_{it} \right)^2 \leq O_p(NT^{2d} \ln T).
\end{aligned}$$

对于第二项, 应用泛函中心极限定理和性质P1可知

$$\begin{aligned}
& \sup_{1 \leq k \leq k^0} \left| \frac{1}{T-k} \sum_{i=1}^N \left(\sum_{t=k+1}^T x_{it} \right)^2 - \frac{1}{T-k} \sum_{i=1}^N \mathbb{E} \left(\sum_{t=k+1}^T x_{it} \right)^2 \right| \\
& \leq \sum_{i=1}^N \frac{1}{T-k^0} \left(\sup_{1 \leq k \leq k^0} \sum_{t=k+1}^T x_{it} \right)^2 + \sum_{i=1}^N \frac{1}{T-k^0} \sup_{1 \leq k \leq k^0} \mathbb{E} \left(\sum_{t=k+1}^T x_{it} \right)^2 = O_p(NT^{2d}).
\end{aligned}$$

研究第三项. 记

$$\psi_t = \frac{1}{\sqrt{\lambda_N}} \sum_{i=1}^N (\mu_{i1} - \mu_{i2}) x_{it}. \tag{12}$$

显然 $\{\psi_t, t \geq 1\}$ 仍是记忆参数为 d 的长相依过程, 且 $\mathbb{E}(\psi_t) = 0$, $\mathbb{E}(\psi_t^2) = \mathbb{E}(x_{1t}^2) < \infty$. 因此利用

泛函中心极限定理可知 $\sup_{1 \leq k \leq k^0} \left| \sum_{t=k+1}^{k^0} \psi_t \right| = O_p(T^{0.5+d})$, 这意味着

$$\begin{aligned}
& \sup_{1 \leq k \leq k^0} 2 \left| \frac{T-k^0}{T-k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} \right| \\
& \leq 2\sqrt{\lambda_N} \sup_{1 \leq k \leq k^0} \frac{T-k^0}{T-k} \cdot \sup_{1 \leq k \leq k^0} \left| \sum_{t=k+1}^{k^0} \psi_t \right| = O_p(\sqrt{\lambda_N} T^{0.5+d}).
\end{aligned}$$

研究最后一项. 首先根据性质P1可知

$$\mathbb{E} \left(\sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it} \right)^2 = \sum_{i=1}^N (\mu_{i1} - \mu_{i2})^2 \mathbb{E} \left(\sum_{t=k^0+1}^T x_{it} \right)^2 = O(\lambda_N T^{1+2d}),$$

那么有

$$\sup_{1 \leq k \leq k^0} \left| \frac{k^0-k}{T-k} \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it} \right| \leq \left| \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it} \right| = O_p(\sqrt{\lambda_N} T^{0.5+d}). \tag{13}$$

把上述的结论整理一下, 即可证明引理5.3.

接下来证明 $\frac{1}{\lambda_N T} \mathbb{E} \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right]$ 有唯一的极小值点 k^0 .

引理5.4 对于模型(1), 若条件A1和A2成立, 且当 $N \rightarrow \infty$ 时 $\frac{1}{T^{1-2d}} = o(\frac{\lambda_N}{N})$ 成立, 那么当 N 充分大后,

$$\begin{aligned} & \mathbb{E} \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] - \mathbb{E} \left[S_{NT}(k^0) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] \\ &= \mathbb{E} [S_{NT}(k)] - \mathbb{E} [S_{NT}(k^0)] \geq c\lambda_N |k - k^0| \end{aligned} \quad (14)$$

关于 $k \in [1, T-1]$ 一致成立, 这里 c 是一个仅与 τ^0 有关的正常数.

证 类似于引理5.3的证明, 只给出 $k \leq k^0$ 情形下的证明. 根据(9)可得

$$\begin{aligned} & S_{NT}(k) - S_{NT}(k^0) \\ &= \frac{(k^0 - k)(T - k^0)}{T - k} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2 - \sum_{i=1}^N \left[\frac{1}{k} \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{1}{k^0} \left(\sum_{t=1}^{k^0} x_{it} \right)^2 \right] \\ & \quad - \sum_{i=1}^N \left[\frac{1}{T - k} \left(\sum_{t=k+1}^T x_{it} \right)^2 - \frac{1}{T - k^0} \left(\sum_{t=k^0+1}^T x_{it} \right)^2 \right] \\ & \quad + 2 \frac{T - k^0}{T - k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} - 2 \frac{k^0 - k}{T - k} \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it}. \end{aligned} \quad (15)$$

显然上式最后两项的数学期望为零. 此外第一项是非随机的, 且注意到 $(T - k^0)/(T - k) \geq (T - k^0)/T = 1 - \tau^0$, 可知

$$\frac{(k^0 - k)(T - k^0)}{T - k} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2 \geq (1 - \tau^0)(k^0 - k)\lambda_N \quad (16)$$

对 $k \in [1, k^0]$ 一致成立.

考虑(15)中的第二项的数学期望. 由性质P1和引理5.2可知, 对于 $1 \leq k \leq k^0$,

$$\begin{aligned} & \mathbb{E} \left\{ \sum_{i=1}^N \left[\frac{1}{k} \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{1}{k^0} \left(\sum_{t=1}^{k^0} x_{it} \right)^2 \right] \right\} \\ &= N(k^0 - k) \left\{ \mathbb{E} \left[\frac{1}{kk^0} \left(\sum_{t=1}^k x_{it} \right)^2 \right] - \mathbb{E} \left[\frac{2}{k^0(k^0 - k)} \left(\sum_{t=1}^k x_{it} \right) \left(\sum_{t=k+1}^{k^0} x_{it} \right) \right] \right. \\ & \quad \left. - \mathbb{E} \left[\frac{1}{k^0(k^0 - k)} \left(\sum_{t=k+1}^{k^0} x_{it} \right)^2 \right] \right\} \\ &= N(k^0 - k) \cdot \left[O \left(\frac{k^{2d}}{k^0} \right) + O \left(\frac{k^{2d}}{k^0} \right) + O \left(\frac{(k^0 - k)^{2d}}{k^0} \right) \right] \\ &\leq N(k^0 - k) \cdot O(T^{2d-1}). \end{aligned}$$

类似地, 对于(15)中的第三项有

$$\mathbb{E} \left\{ \sum_{i=1}^N \left[\frac{1}{T - k} \left(\sum_{t=k+1}^T x_{it} \right)^2 - \frac{1}{T - k^0} \left(\sum_{t=k^0+1}^T x_{it} \right)^2 \right] \right\} \leq N(k^0 - k) \cdot O(T^{2d-1}).$$

因此当 $\frac{1}{T^{1-2d}} = o(\frac{\lambda_N}{N})$ ($N \rightarrow \infty$)成立时, (15)中的第二和第三项的数学期望都被第一项的数

学期望所控制, 这就证明了(14).

给定一个满足 $0 < \eta < 1$ 的常数 η , 记

$$\begin{cases} D = \{k : T\eta \leq k \leq T(1 - \eta)\}, \\ D(k^0) = \{k : T\eta \leq k \leq T(1 - \eta), k \neq k^0\}. \end{cases} \quad (17)$$

引理5.5 对于模型(1), 若条件A1-A4成立, 那么当 $N \rightarrow \infty$ 时, 有

$$P\left(\min_{k \in D(k^0)} S_{NT}(k) - S_{NT}(k^0) \leq 0\right) \rightarrow 0.$$

证 显然, 只需证明当 $N \rightarrow \infty$ 时, $P(\min_{k \in D(k^0)} S_{NT}(k) - S_{NT}(k^0) > 0) \rightarrow 1$. 与前面引理的证明类似, 只考虑 $k < k^0$ 的情形. 即, 证明当 $N \rightarrow \infty$ 时, 有

$$P\left(\min_{k \in D(k^0), k < k^0} S_{NT}(k) - S_{NT}(k^0) > 0\right) \rightarrow 1.$$

注意到(15)已提供了一个关于 $S_{NT}(k) - S_{NT}(k^0)$ 的表达式. 下面验证(15)中的第一项是主项.

对于(15)中的第一项, (16)已经证明这一项不小于 $(1 - \tau^0)(k^0 - k)\lambda_N$.

考虑(15)中的第二项.

$$\begin{aligned} & \sum_{i=1}^N \left[\frac{1}{k} \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{1}{k^0} \left(\sum_{t=1}^{k^0} x_{it} \right)^2 \right] = (k^0 - k) \times \\ & \sum_{i=1}^N \left[\frac{1}{kk^0} \left(\sum_{t=1}^k x_{it} \right)^2 - \frac{2}{k^0(k^0 - k)} \left(\sum_{t=1}^k x_{it} \right) \left(\sum_{t=k+1}^{k^0} x_{it} \right) - \frac{1}{k^0(k^0 - k)} \left(\sum_{t=k+1}^{k^0} x_{it} \right)^2 \right] \\ & = (k^0 - k) \left\{ \frac{N}{k^0} \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{\sqrt{k}} \sum_{t=1}^k x_{it} \right)^2 - \frac{N}{k^0} \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{\sqrt{k^0 - k}} \sum_{t=k+1}^{k^0} x_{it} \right)^2 \right. \\ & \quad \left. - 2 \frac{N}{(k^0)^{0.5-d}} \frac{1}{N} \sum_{i=1}^N \left[\left(\frac{1}{(k^0)^{0.5+d}} \sum_{t=1}^k x_{it} \right) \left(\frac{1}{k^0 - k} \sum_{t=k+1}^{k^0} x_{it} \right) \right] \right\}. \end{aligned}$$

接下来说明当 $k \in D(k^0)$ 且 $k < k^0$ 时, 花括号里的每一项都被 λ_N 控制. 注意到 $k \in D(k^0)$ 时有 $T\eta \leq k$. 应用泛函中心极限定理并利用条件A3, 有

$$\begin{aligned} \sup_{T\eta < k < k^0} \frac{N}{k^0} \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{\sqrt{k}} \sum_{t=1}^k x_{it} \right)^2 & \leq \frac{N}{k^0} \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{\sqrt{T\eta}} \sup_{T\eta < k < k^0} \left| \sum_{t=1}^k x_{it} \right| \right)^2 \\ & \leq O_p \left(\frac{NT^{2d}}{T} \right) = o_p(\lambda_N). \end{aligned}$$

类似地, 利用引理5.1有

$$\begin{aligned} \sup_{T\eta < k < k^0} \frac{N}{k^0} \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{\sqrt{k^0 - k}} \sum_{t=k+1}^{k^0} x_{it} \right)^2 & \leq \frac{N}{k^0} \frac{1}{N} \sum_{i=1}^N \left(\sup_{1 \leq k < k^0} \frac{1}{\sqrt{k^0 - k}} \left| \sum_{t=k+1}^{k^0} x_{it} \right| \right)^2 \\ & = O_p \left(\frac{NT^{2d} \ln T}{T} \right) = o_p(\lambda_N). \end{aligned}$$

对于最后一项, 同样根据泛函中心极限定理, 引理5.1和条件A3可得

$$\begin{aligned} & \sup_{T\eta < k < k^0} \left| \frac{N}{(k^0)^{0.5-d}} \frac{1}{N} \sum_{i=1}^N \left[\left(\frac{1}{(k^0)^{0.5+d}} \sum_{t=1}^k x_{it} \right) \left(\frac{1}{k^0 - k} \sum_{t=k+1}^{k^0} x_{it} \right) \right] \right| \\ & \leq \frac{N}{(k^0)^{0.5-d}} \frac{1}{N} \sum_{i=1}^N \sup_{1 \leq k < k^0} \left| \frac{1}{(k^0)^{0.5+d}} \sum_{t=1}^k x_{it} \right| \cdot \sup_{k^0 - k \geq 1} \left| \frac{1}{k^0 - k} \sum_{t=k+1}^{k^0} x_{it} \right| \\ & = O_p \left(\frac{N}{T^{0.5-d}} \right) = o_p(\lambda_N). \end{aligned}$$

综上所述, (15)中的第二项可被第一项控制.

类似地可以证明(15)中的第三项可被第一项控制.

对于(15)的第四项有

$$\frac{T - k^0}{T - k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} = \frac{T - k^0}{T - k} \sqrt{\lambda_N} \sum_{t=k+1}^{k^0} \psi_t,$$

其中 ψ_t 由(12)定义. 根据引理5.1有

$$\sup_{T\eta \leq k < k^0} \frac{1}{k^0 - k} \left| \sum_{t=k+1}^{k^0} \psi_t \right| \leq \sup_{k^0 - k \geq 1} \frac{1}{k^0 - k} \left| \sum_{t=k+1}^{k^0} \psi_t \right| = O_p(1).$$

注意到 $|(T - k^0)/(T - k)| < 1$, 且 $\lambda_N \rightarrow \infty$, 那么可进一步推出

$$\sup_{T\eta \leq k < k^0} \frac{1}{(k^0 - k)\lambda_N} \frac{T - k^0}{T - k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} = O_p \left(\frac{1}{\sqrt{\lambda_N}} \right) = o_p(1).$$

因此(15)中的第四项可被第一项控制.

对于(15)中的最后一项, 根据(13)有

$$\begin{aligned} & \sup_{T\eta \leq k < k^0} \left| \frac{1}{(k^0 - k)\lambda_N} \cdot \frac{k^0 - k}{T - k} \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it} \right| \\ & \leq \frac{C}{\lambda_N T} \left| \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2}) x_{it} \right| = O_p \left(\frac{\sqrt{\lambda_N} T^{0.5+d}}{\lambda_N T} \right) = o_p(1). \end{aligned}$$

因此(15)中的最后一项也被第一项控制.

基于以上的讨论, 可以得到这样的结论: 当 $k \in D(k^0)$ 且 $k < k^0$ 时, 有

$$S_{NT}(k) - S_{NT}(k^0) = (1 - \tau^0)(k^0 - k)\lambda_N(1 + o_p(1)).$$

注意到 $\min_{k \in D(k^0), k < k^0} S_{NT}(k) - S_{NT}(k^0) \geq (1 - \tau^0)\lambda_N(1 + o_p(1))$, $0 < \tau^0 < 1$ 且 $\lambda_N \rightarrow \infty$ (当 $N \rightarrow \infty$), 因此当 $N \rightarrow \infty$ 时, 有 $P(\min_{k \in D(k^0), k < k^0} S_{NT}(k) - S_{NT}(k^0) > 0) \rightarrow 1$.

定理2.1中结论(4)的证明 根据引理5.4(注意到条件A3可推出 $\frac{1}{T^{1-2d}} = o(\frac{\lambda_N}{N})$), 容易看出,

对充分大的 N ,

$$\begin{aligned} & S_{NT}(k) - S_{NT}(k^0) \\ &= \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] - E \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] \\ & \quad - \left\{ \left[S_{NT}(k^0) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] - E \left[S_{NT}(k^0) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] \right\} \\ & \quad + E[S_{NT}(k)] - E[S_{NT}(k^0)] \\ & \geq -2 \sup_{1 \leq k \leq T} \left| \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] - E \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] \right| + c\lambda_N |k - k^0| \quad (18) \end{aligned}$$

对 $k \in [1, T-1]$ 一致成立. 由于(18)对所有的 $k \in [1, T-1]$ 成立, 那么它对 $\hat{k}$ 自然成立; 此外注意到 $S_{NT}(\hat{k}) - S_{NT}(k^0) \leq 0$ (这是因为 $\hat{k} = \arg \min S_{NT}(k)$), 因此对充分大的 N ,

$$|\hat{k} - k^0| \leq \frac{2}{c\lambda_N} \sup_{1 \leq k \leq T} \left| \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] - E \left[S_{NT}(k) - \sum_{i=1}^N \sum_{t=1}^T x_{it}^2 \right] \right|$$

成立. 然后应用引理5.3和条件A3, 当 $N \rightarrow \infty$ 时有

$$|\hat{k} - k^0| = \frac{2}{c\lambda_N} \cdot \max \left\{ O_p(NT^{2d} \ln T), O_p(\sqrt{\lambda_N} T^{0.5+d}) \right\} = o_p(T). \quad (19)$$

由于当 T 有界时, (19)已可推出 $\lim_{N \rightarrow \infty} P(\hat{k} = k^0) = 1$, 因此在剩下的证明中只需考虑无界的 T .

对于无界的 T , 由(19)可知, 对于任意给定的 $\epsilon > 0$, $\lim_{N, T \rightarrow \infty} P(|\hat{k} - k^0| \leq T\epsilon) = 1$. 因此根据条件A2, 存在着一个满足 $\tau^0 \in (\eta, 1 - \eta)$ 的 $\eta > 0$ 使得

$$\lim_{N, T \rightarrow \infty} P(\hat{k} \in [T\eta, T(1 - \eta)]) = 1. \quad (20)$$

回忆(17)并注意

$$\begin{aligned} P(\hat{k} \neq k^0) &= P(\hat{k} \notin D, \hat{k} \neq k^0) + P(\hat{k} \in D, \hat{k} \neq k^0) \\ &\leq P(\hat{k} \notin D) + P(\hat{k} \in D(k^0)) \leq P(\hat{k} \notin D) + P\left(\min_{k \in D(k^0)} S_{NT}(k) - S_{NT}(k^0) \leq 0\right), \end{aligned}$$

那么根据(20)和引理5.5, 就可得到 $\lim_{N, T \rightarrow \infty} P(\hat{k} = k^0) = 1$.

为证明结论(5), 先证下面这个引理, 它的作用类似于引理5.5.

引理5.6 对于模型(1), 假设条件A1-A3以及A5成立, 那么存在着一个正常数 $M < \infty$ 使得当 $N, T \rightarrow \infty$ 时, 有

$$P\left(\min_{k \in D, |k - k^0| > M} S_{NT}(k) - S_{NT}(k^0) \leq 0\right) \rightarrow 0.$$

证 类似于引理5.5的证明, 只需证明当 $N, T \rightarrow \infty$ 时, 有

$$P\left(\min_{k \in D, k^0 - k > M} S_{NT}(k) - S_{NT}(k^0) > 0\right) \rightarrow 1.$$

$S_{NT}(k) - S_{NT}(k^0)$ 的表达式已在(15)给出.

对于(15)的第一项, 注意到

$$\frac{(k^0 - k)(T - k^0)}{T - k} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2 \geq (1 - \tau^0)(k^0 - k)\lambda_N$$

仍然成立. 接下来证明当 $k \in D$ 且 $k^0 - k > M$ 时, (15)的第一项控制着其它所有项. 如果这个结

论是正确的, 那么

$$\min_{k \in D, k^0 - k > M} S_{NT}(k) - S_{NT}(k^0) > (1 - \tau^0)M\lambda/2$$

发生的概率会趋向于1, 结论就可得证.

仅详细说明为什么(15)的第四项可以被第一项控制, 因为对于其它项的证明与引理5.5的证明是类似的. 引理5.1可推得

$$\sup_{k^0 - k > M} \frac{1}{k^0 - k} \left| \sum_{t=k+1}^{k^0} \psi_t \right| = O_p \left(\frac{1}{M^{0.5-d}} \right),$$

其中 ψ_t 的定义见(12). 这意味着

$$\begin{aligned} & \sup_{k^0 - k > M} \left| \frac{2}{(k^0 - k)\lambda_N} \frac{T - k^0}{T - k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2})x_{it} \right| \\ &= \sup_{k^0 - k > M} \left| \frac{2}{\sqrt{\lambda_N}} \frac{T - k^0}{T - k} \frac{1}{k^0 - k} \sum_{t=k+1}^{k^0} \psi_t \right| = O_p \left(\frac{1}{\sqrt{\lambda_N} M^{0.5-d}} \right) = O_p \left(\frac{1}{M^{0.5-d}} \right). \end{aligned}$$

因此只要让 M 充分大, 就可知(15)的第四项可以被第一项控制.

定理2.1中的结论(5)的证明 首先通过检查(19)的证明可以发现

$$\left| \frac{\hat{k}}{T} - \frac{k^0}{T} \right| = o_p(1)$$

仍然成立. 即当 $N, T \rightarrow \infty$ 时, $P(\hat{k} \in D) \rightarrow 1$ 成立. 然后, 沿着结论(4)的证明思路并利用引理5.6即可完成证明.

定理2.1中的结论(6)的证明 考虑过程

$$\Delta_{NT}(l) = S_{NT}(k^0 + l) - S_{NT}(k^0),$$

其中 l 是正整数, 满足 $|l| \leq M$, M 是任意给定的正常数. 记

$$\hat{l} = \arg \min_{|l| \leq M} \Delta_{NT}(l).$$

那么当 $M \rightarrow \infty$ 时有 $\hat{l} = \hat{k} - k^0$.

首先考虑 $-M \leq l < 0$ 的情形. 在(15)中取 $k = k^0 + l$, 有

$$\begin{aligned} \Delta_{NT}(l) &= \frac{-l(T - k^0)}{T - k^0 - l} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2 - \sum_{i=1}^N \left[\frac{1}{k^0 + l} \left(\sum_{t=1}^{k^0+l} x_{it} \right)^2 - \frac{1}{k^0} \left(\sum_{t=1}^{k^0} x_{it} \right)^2 \right] \\ &\quad - \sum_{i=1}^N \left[\frac{1}{T - k^0 - l} \left(\sum_{t=k^0+l+1}^T x_{it} \right)^2 - \frac{1}{T - k^0} \left(\sum_{t=k^0+1}^T x_{it} \right)^2 \right] \\ &\quad + 2 \frac{T - k^0}{T - k^0 - l} \sum_{i=1}^N \sum_{t=k^0+l+1}^{k^0} (\mu_{i1} - \mu_{i2})x_{it} - 2 \frac{-l}{T - k^0 - l} \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2})x_{it}. \end{aligned}$$

通过检查结论(5)的证明可以发现上面等式右边的第一和第四项是主项. 对于第一项, 当 $N, T \rightarrow \infty$ 时有

$$\frac{-l(T - k^0)}{T - k^0 - l} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2 \rightarrow -l\lambda.$$

考虑第四项,

$$\begin{aligned} & 2 \frac{T - k^0}{T - k^0 - l} \sum_{i=1}^N \sum_{t=k^0+l+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} \\ &= 2 \frac{T - k^0}{T - k^0 - l} \sum_{t=k^0+l+1}^{k^0} \sum_{i=1}^N (\mu_{i1} - \mu_{i2}) (1 - B)^{-d} u_{it} \\ &= 2 \frac{T - k^0}{T - k^0 - l} \sqrt{\sigma_u^2 \lambda_N} (1 - B)^{-d} \sum_{t=k^0+l+1}^{k^0} \left(\frac{1}{\sqrt{\sigma_u^2 \lambda_N}} \sum_{i=1}^N (\mu_{i1} - \mu_{i2}) u_{it} \right). \end{aligned}$$

下面证明当 $N, T \rightarrow \infty$ 时有

$$\frac{1}{\sqrt{\sigma_u^2 \lambda_N}} \sum_{i=1}^N (\mu_{i1} - \mu_{i2}) u_{it} \xrightarrow{d} Z_t, \quad (21)$$

其中 Z_t 是 $N(0, 1)$ 随机变量. 因为 $\sum_{i=1}^N (\mu_{i1} - \mu_{i2}) u_{it}$ 是零均值的独立随机变量的部分和, 且

$$\frac{1}{\sigma_u^2 \lambda_N} \sum_{i=1}^N \text{Var}((\mu_{i1} - \mu_{i2}) u_{it}) = 1,$$

所以只需验证Lindeberg条件

$$\sum_{i=1}^N \text{E}(\xi_{it}^2 I\{|\xi_{it}| > \eta\}) = o(1), \quad \forall \eta > 0$$

成立, 其中

$$\xi_{it} = \frac{(\mu_{i1} - \mu_{i2}) u_{it}}{\sqrt{\sigma_u^2 \lambda_N}}.$$

回忆条件 $\mu_{i1} - \mu_{i2} = \Delta_i / \sqrt{N}$ 和 $\sup_{i \geq 1} |\Delta_i| \leq C_0$. 根据 $\text{E}(u_{it}^2)$ 的有限性, 可知当 $N, T \rightarrow \infty$ 时,

$$\begin{aligned} \sum_{i=1}^N \text{E}(\xi_{it}^2 I\{|\xi_{it}| > \eta\}) &= \frac{1}{\sigma_u^2 \lambda_N} \sum_{i=1}^N (\mu_{i1} - \mu_{i2})^2 \text{E}(u_{it}^2 I\{|u_{it}| > \eta \sigma_u \sqrt{N \lambda_N} / C_0\}) \\ &\leq \frac{1}{\sigma_u^2} \text{E}(u_{1t}^2 I\{|u_{1t}| > \eta \sigma_u \sqrt{N \lambda_N} / C_0\}) \rightarrow 0 \end{aligned}$$

成立. 因此(21)得证. 注意到 $\{\frac{1}{\sqrt{\sigma_u^2 \lambda_N}} \sum_{i=1}^N (\mu_{i1} - \mu_{i2}) u_{it}, t = k^0 + l + 1, \dots, k^0\}$ 相互独立且

$$\sum_{t=k^0+l+1}^{k^0} \left(\frac{1}{\sqrt{\sigma_u^2 \lambda_N}} \sum_{i=1}^N (\mu_{i1} - \mu_{i2}) u_{it} \right) \stackrel{d}{=} \sum_{t=l}^{-1} \left(\frac{1}{\sqrt{\sigma_u^2 \lambda_N}} \sum_{i=1}^N (\mu_{i1} - \mu_{i2}) u_{it} \right).$$

因此

$$2 \frac{T - k^0}{T - k^0 - l} \sum_{i=1}^N \sum_{t=k^0+l+1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} \xrightarrow{d} 2 \sigma_u \sqrt{\lambda} \sum_{t=l}^{-1} (1 - B)^{-d} Z_t.$$

综合上述的分析可知, 对于 $-M \leq l < 0$ 情形, $N, T \rightarrow \infty$ 时有

$$\Delta_{NT}(l) \xrightarrow{d} -l\lambda + 2\sigma_u \sqrt{\lambda} \sum_{t=l}^{-1} (1 - B)^{-d} Z_t. \quad (22)$$

类似地, 对于 $0 < l \leq M$ 情形, 可知当 $N, T \rightarrow \infty$ 时有

$$\begin{aligned} \Delta_{NT}(l) &= \frac{lk^0}{k^0+l} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2 - \sum_{i=1}^N \left[\frac{1}{k^0+l} \left(\sum_{t=1}^{k^0+l} x_{it} \right)^2 - \frac{1}{k^0} \left(\sum_{t=1}^{k^0} x_{it} \right)^2 \right] \\ &\quad - \sum_{i=1}^N \left[\frac{1}{T-k^0-l} \left(\sum_{t=k^0+l+1}^T x_{it} \right)^2 - \frac{1}{T-k^0} \left(\sum_{t=k^0+1}^T x_{it} \right)^2 \right] \\ &\quad + 2 \frac{k^0}{k^0+l} \sum_{i=1}^N (\mu_{i2} - \mu_{i1}) \sum_{t=k^0+1}^{k^0+l} x_{it} + 2 \frac{l}{k^0+l} \sum_{i=1}^N \sum_{t=1}^{k^0} (\mu_{i1} - \mu_{i2}) x_{it} \\ &\xrightarrow{d} l\lambda + 2\sigma_u \sqrt{\lambda} \sum_{t=1}^l (1-B)^{-d} Z_t. \end{aligned} \quad (23)$$

此外, 注意到 $l=0$ 时有 $\Delta_{NT}(l)=0$, 且 M 可以任意大, 因此定义

$$W(l) = \begin{cases} -l\sqrt{\lambda} + 2\sigma_u \sum_{t=l}^{-1} (1-B)^{-d} Z_t, & l = -1, -2, \dots, \\ 0, & l = 0, \\ l\sqrt{\lambda} + 2\sigma_u \sum_{t=1}^l (1-B)^{-d} Z_t, & l = 1, 2, \dots, \end{cases}$$

就有

$$\Delta_{NT}(l) \xrightarrow{d} \sqrt{\lambda} W(l).$$

这意味着根据 $\arg \max / \arg \min$ 的连续映射定理^[21]有

$$\hat{k} - k^0 \xrightarrow{d} \arg \min_{l \in \{\dots, -2, -1, 0, 1, 2, \dots\}} \{\sqrt{\lambda} W(l)\} = \arg \min_{l \in \{\dots, -2, -1, 0, 1, 2, \dots\}} W(l).$$

为证结论(7), 先证明下面的引理, 它的作用类似于引理5.5和5.6.

引理5.7 对于模型(1), 假设条件A1-A3和A6成立, 那么存在一个正常数 $M < \infty$ 使得当 $N, T \rightarrow \infty$ 时, 有

$$P \left(\min_{k \in D, |k-k^0| > M\lambda_N^{-1/(1-2d)}} S_{NT}(k) - S_{NT}(k^0) \leq 0 \right) \rightarrow 0.$$

证 类似于引理5.5的证明, 只需证明当 $N, T \rightarrow \infty$ 时, 有

$$P \left(\min_{k \in D, k^0 - k > M\lambda_N^{-1/(1-2d)}} S_{NT}(k) - S_{NT}(k^0) > 0 \right) \rightarrow 1.$$

$S_{NT}(k) - S_{NT}(k^0)$ 的表达式见(15). 接下来, 证明当 $k \in D, k^0 - k > M\lambda_N^{-1/(1-2d)}$, 且 M 充分大时, (15)的第一项控制着其它项.

对于(15)的第一项, 容易看出(16)仍然成立, 且当 $N, T \rightarrow \infty$ 时有

$$\begin{aligned} &\min_{k \in D, k^0 - k > M\lambda_N^{-1/(1-2d)}} (1 - \tau^0)(k^0 - k)\lambda_N \geq (1 - \tau^0)M\lambda_N^{-1/(1-2d)}\lambda_N \\ &= (1 - \tau^0)M\lambda_N^{-2d/(1-2d)} \rightarrow \infty. \end{aligned}$$

考虑(15)的第四项. 应用引理5.1可得

$$\sup_{k^0 - k > M\lambda_N^{-1/(1-2d)}} \left| \frac{1}{k^0 - k} \sum_{t=k+1}^{k^0} \psi_t \right| = O_p \left(\frac{\sqrt{\lambda_N}}{M^{0.5-d}} \right),$$

其中 ψ_t 的定义见(12). 那么

$$\begin{aligned} & \sup_{k^0-k > M\lambda_N^{-1/(1-2d)}} \left| \frac{2}{(k^0-k)\lambda_N} \frac{T-k^0}{T-k} \sum_{i=1}^N \sum_{t=k+1}^{k^0} (\mu_{i1} - \mu_{i2})x_{it} \right| \\ & \leq \sup_{k^0-k > M\lambda_N^{-1/(1-2d)}} \frac{2}{\sqrt{\lambda_N}} \left| \frac{1}{k^0-k} \sum_{t=k+1}^{k^0} \psi_t \right| = O_p\left(\frac{1}{M^{0.5-d}}\right). \end{aligned}$$

因此, 当 $k \in D$, $k^0 - k > M\lambda_N^{-1/(1-2d)}$, 且 M 充分大时, 第四项被第一项控制.

剩下的证明与引理5.5的证明没有什么本质区别, 故省略.

定理2.1中结论(7)的证明 首先通过检查(19)的证明, 可以发现 $\left|\frac{\hat{k}}{T} - \frac{k^0}{T}\right| = o_p(1)$ 仍然成立. 然后应用引理5.7即可完成证明.

定理2.1中结论(8)的证明 类似于(6)的证明, 研究过程

$$A_T(s) = \lambda_N^{2d/(1-2d)} \left[S_{NT}(k^0 + \lfloor s\lambda_N^{-1/(1-2d)} \rfloor) - S_{NT}(k^0) \right],$$

其中 $s \in [-M, M]$, M 是任意给定的正常数. 记 $l = \lfloor s\lambda_N^{-1/(1-2d)} \rfloor$,

$$\Psi_{NT}(l) = \lambda_N^{2d/(1-2d)} [S_{NT}(k^0 + l) - S_{NT}(k^0)]$$

且

$$\hat{l} = \arg \min_{|l| \leq M\lambda_N^{-1/(1-2d)}} \Psi_{NT}(l).$$

显然当 $M \rightarrow \infty$ 时有 $\hat{l} = \hat{k} - k^0$. 因此只需要研究 $\Psi_{NT}(l)$. 由于篇幅的原因, 下面只提供 $[-M\lambda_N^{-1/(1-2d)}] \leq l < 0$ 情形下的详细证明.

首先容易看出条件A3可推出 $l = o(T)$.

$$\begin{aligned} & \Psi_{NT}(l) \\ &= \lambda_N^{\frac{2d}{1-2d}} \frac{-l(T-k^0)}{T-k^0-l} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2 - \lambda_N^{\frac{2d}{1-2d}} \sum_{i=1}^N \left[\frac{1}{k^0+l} \left(\sum_{t=1}^{k^0+l} x_{it} \right)^2 - \frac{1}{k^0} \left(\sum_{t=1}^{k^0} x_{it} \right)^2 \right] \\ & \quad - \lambda_N^{\frac{2d}{1-2d}} \sum_{i=1}^N \left[\frac{1}{T-k^0-l} \left(\sum_{t=k^0+l+1}^T x_{it} \right)^2 - \frac{1}{T-k^0} \left(\sum_{t=k^0+1}^T x_{it} \right)^2 \right] \\ & \quad + 2\lambda_N^{\frac{2d}{1-2d}} \frac{T-k^0}{T-k^0-l} \sum_{i=1}^N \sum_{t=k^0+l+1}^{k^0} (\mu_{i1} - \mu_{i2})x_{it} \\ & \quad - 2\lambda_N^{\frac{2d}{1-2d}} \frac{-l}{T-k^0-l} \sum_{i=1}^N \sum_{t=k^0+1}^T (\mu_{i1} - \mu_{i2})x_{it}. \end{aligned} \tag{24}$$

通过检查(5)的证明, 可以看出(24)中的第一和第四项是主项.

对于(24)中的第一项, 当 $N, T \rightarrow \infty$ 时有

$$\begin{aligned} & \lambda_N^{\frac{2d}{1-2d}} \frac{-l(T-k^0)}{T-k^0-l} \sum_{i=1}^N (\mu_{i2} - \mu_{i1})^2 = \lambda_N^{\frac{1}{1-2d}} \frac{-l(T-k^0)}{T-k^0-l} \\ &= -\lambda_N^{\frac{1}{1-2d}} \lfloor s\lambda_N^{-1/(1-2d)} \rfloor \cdot (1 + o(1)) \rightarrow -s. \end{aligned}$$

对于(24)中的第四项,

$$2\lambda_N^{\frac{2d}{1-2d}} \frac{T - k^0}{T - k^0 - l} \sum_{i=1}^N \sum_{t=k^0+l+1}^{k^0} (\mu_{i1} - \mu_{i2})x_{it} = 2\lambda_N^{\frac{1+2d}{2(1-2d)}} \sum_{t=k^0+l+1}^{k^0} \psi_t \cdot (1 + o(1)),$$

其中 ψ_t 的定义见(12). 由于 $\{\psi_t, t \geq 1\}$ 是长相依的, 均值为零, $\text{Var}(\psi_t) = \text{Var}(x_{1t}) < \infty$. 那么根据

据 $\sum_{t=k^0+l+1}^{k^0} \psi_t \stackrel{d}{=} \sum_{t=l}^{-1} \psi_t$ 和 $\lambda_N^{-1/(1-2d)} \rightarrow \infty$, $\left(\lambda_N^{-1/(1-2d)}\right)^{0.5+d} = \lambda_N^{-\frac{1+2d}{2(1-2d)}}$, 并应用泛函中心极限定理, 可知当 $N, T \rightarrow \infty$ 时有

$$2\lambda_N^{\frac{2d}{1-2d}} \frac{T - k^0}{T - k^0 - l} \sum_{i=1}^N \sum_{t=k^0+l+1}^{k^0} (\mu_{i1} - \mu_{i2})x_{it} \Rightarrow 2\kappa_d B_d(s),$$

其中 $B_d(\cdot)$ 是一个分数Brownian运动, κ_d 的定义见(3). 因此当 $N, T \rightarrow \infty$ 时有

$$\Lambda_T(s) = \Psi_{NT}(l) \Rightarrow -s + 2\kappa_d B_d(s).$$

类似的分析可以证明: 对于 $0 < l \leq \lfloor M\lambda_N^{-1/(1-2d)} \rfloor$ 的情形, 当 $N, T \rightarrow \infty$ 时有

$$\Lambda_T(s) = \Psi_{NT}(l) \Rightarrow s + 2\kappa_d B_d(s).$$

注意到 $\Lambda_T(0) = 0$, 且 M 可以任意大, 那么定义

$$\Upsilon(s) = \begin{cases} -s + 2\kappa_d B_d(s), & s < 0, \\ 0, & s = 0, \\ s + 2\kappa_d B_d(s), & s > 0, \end{cases}$$

即知当 $N, T \rightarrow \infty$ 时, 有 $\Lambda_T(s) \Rightarrow \Upsilon(s)$. 这意味着由 $\arg \max / \arg \min$ 的连续映射定理^[21], 当 $N, T \rightarrow \infty$ 时有

$$\lambda_N^{1/(1-2d)} (\hat{k} - k^0) \xrightarrow{d} \arg \min_{-\infty < s < \infty} \Upsilon(s).$$

参考文献:

- [1] Joseph L, Wolfson D. Estimation in the multi-path change-point problems[J]. Communications in Statistics—Theory and Methods, 1992, 21(4): 897-913.
- [2] Joseph L, Wolfson D. Maximum likelihood estimation in the multi-path change-point problem[J]. Annals of the Institute of Statistical Mathematics, 1993, 45(3): 511-530.
- [3] Joseph L, Vandal A, Wolfson D. Estimation in the multi-path change-point problem for correlated data[J]. The Canadian Journal of Statistics, 1996, 24(1): 37-54.
- [4] Bai Jushan. Common breaks in means and variances for panel data[J]. Journal of Econometrics, 2010, 157(1): 78-92.
- [5] Horváth L, Hušková M. Change-point detection in panel data[J]. Journal of Time Series Analysis, 2012, 33(4): 631-648.
- [6] Chan J, Horváth L, Hušková M. Darling-Erdős limit results for change-point detection in panel data[J]. Journal of Statistical Planning and Inference, 2013, 143(5): 955-970.
- [7] Kim D. Estimating a common deterministic time trend break in large panels with cross sectional dependence[J]. Journal of Econometrics, 2011, 164(2): 310-330.
- [8] Kim D. Common breaks in time trends for large panel data with a factor structure[J]. The Econometrics Journal, 2014, 17(3): 301-337.
- [9] Baltagi B, Feng Qu, Kao C. Estimation of heterogeneous panels with structural breaks[J]. Journal of Econometrics, 2016, 191(1): 176-195.

- [10] Li Degui, Qian Junhui, Su Liangjun. Panel data models with interactive fixed effects and multiple structural breaks[J]. *Journal of the American Statistical Association*, 2016, 111(516): 1804-1819.
- [11] Horváth L, Hušková M, Rice G, Wang J. Asymptotic properties of the CUSUM estimator for the time of change in linear panel data models[J]. *Econometric Theory*, 2017, 33(2): 366-412.
- [12] Fisher M, Jensen M. Bayesian inference and prediction of a multiple-change-point panel model with nonparametric priors[J]. *Journal of Econometrics*, 2019, 210(1): 187-202.
- [13] Csörgő M, Horváth L. *Limit Theorems in Change-Point Analysis*[M]. New York: Wiley, 1997.
- [14] Sowell F. The fractional unit root distribution[J]. *Econometrica*, 1990, 58(2): 495-505.
- [15] Lavielle M, Moulines E. Least-squares estimation of an unknown number of shifts in a time series[J]. *Journal of Time Series Analysis*, 2000, 21(1): 33-59.
- [16] Mandelbrot B B, Van Ness J W. Fractional Brownian motions, fractional noises and applications[J]. *SIAM Review*, 1968, 10(4): 422-437.
- [17] Wang Qiying, Lin Yan-Xia, Gulati C M. Asymptotics for general fractionally integrated processes with applications to unit root tests[J]. *Econometric Theory*, 2003, 19(1): 143-164.
- [18] Xi Daiqing, Pang Tianxiao. Estimating multiple breaks in mean sequentially with fractionally integrated errors[J]. *Statistical Papers*, 2021, 62(1): 451-494.
- [19] McLeod A I, Hipel K W. Preservation of the rescaled adjusted range. 1: A reassessment of the Hurst phenomenon[J]. *Water Resources Research*, 1978, 14(3): 491-508.
- [20] Hosking J R M. Modelling persistence in hydrological time series using fractional differencing[J]. *Water Resources Research*, 1984, 20(12): 1898-1908.
- [21] Kim J, Pollard D. Cube root asymptotics[J]. *The Annals of Statistics*, 1990, 18(1): 191-219.

A common break in means for long-range dependent panel data under cross-sectional independence

XI Dai-qing¹, FUH Cheng-Der², PANG Tian-xiao³

- (1. School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China;
- 2. Graduate Institute of Statistics, Central University, Taoyuan (Taiwan) 320317, China;
- 3. School of Mathematical Sciences, Zhejiang University, Hangzhou 310058, China)

Abstract: This paper focuses on estimating a common break point in means for long-range dependent panel data under cross-sectional independence. The common break-point estimator is examined under three scenarios: strong, moderate and weak break signals. Asymptotic properties, including consistency, rate of convergence and limiting distribution, of the estimator are established. The theoretical results reveal that there is a trade-off between the break signal and long-range dependence. To be more precise, the long-range dependence has no ability to influence the asymptotic behaviors of the estimator if the break signal is strong, it does not influence the rate of convergence but has an impact on the limiting distribution of the estimator when the break signal is moderate, and it influences both the rate of convergence and the limiting distribution of the estimator when the break signal is weak. Monte Carlo simulations are conducted to assess the finite-sample performance of the estimator, and the theoretical results are supported by the simulation results.

Keywords: common break; limiting distribution; long-range dependence; panel data

MR Subject Classification: 60F05; 62F12