

含测量误差的方差模型的多变点的估计

沙达克提·艾力, 董翠玲*

新疆师范大学数学科学学院, 新疆 乌鲁木齐

收稿日期: 2024年1月28日; 录用日期: 2024年2月22日; 发布日期: 2024年2月29日

摘要

本文讨论当已知含测量误差的方差模型存在变点时, 对方差变点给出了一个含有调节参数的“CUSUM型估计量”, 研究了方差变点统计量的弱(强)相合性, 得到收敛速度。结合“二元分割法”将其推广至多个方差变点的估计。模拟研究发现含调节参数 $\gamma \in (0.3, 0.7)$ 的CUSUM型估计量的精确度要优于无调节参数($\gamma = 0$)的CUSUM型估计量的精确度。进一步, 对原油价格涨跌幅进行实证分析验证了本文方法的有效性和可行性。

关键词

测量误差, 方差变点, 调节参数, 相合性, 二元分割法

Estimation of Variable Points of Variance Model with Measurement Error

Sadakat Ali, Cuiling Dong*

School of Mathematical Sciences, Xinjiang Normal University, Urumqi Xinjiang

Received: Jan. 28th, 2024; accepted: Feb. 22nd, 2024; published: Feb. 29th, 2024

Abstract

This article discusses when it is known that variance models with measurement errors have change points, a “CUSUM-type estimator” with adjustment parameters is given for the variance change point. The weak (strong) consistency of variance change point estimator is studied, and the convergence rate is obtained. It is extended to the estimation of multiple variance change points by using binary segmentation method. The simulation results show that the accuracy of CUSUM

*通讯作者。

type estimator with adjustment parameter $\gamma \in (0.3, 0.7)$ is better than that of CUSUM type estimator without adjustment parameter $\gamma = 0$. Furthermore, the validity and feasibility of this method are verified by an empirical analysis of the rise and fall of crude oil prices.

Keywords

Measurement Error, Variance Change Point, Adjustment Parameters, Consistency, Binary Segmentation Method

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

方差在统计应用中常常被理解为风险，估计方差变点的发生能够有效地监测风险的变化，避免不必要的损失。目前关于方差变点模型问题研究已获得一批重要理论成果。例如 Gombay (1996) 在独立随机误差条件下，构造了方差变点累积和(Cumulative Sum, 简记 CUSUM)型估计量，并获得变点估计量的极限分布等重要理论结果[1]；Inclán 和 Tiao (1994) 在独立正态随机误差条件下，基于迭代的累积平方和(ICSS) 算法，提出了关于方差变点的检测步骤[2]；Chen 和 Gupta (1997) 利用 Schwarz 信息准则(SIC) 构造方差变点 SIC 估计量，并获得估计量的极限分布[3]；Lee 和 Park (2001)、Lee (2003) 等将 Inclán 和 Tiao 的结果扩展到误差由独立随机变量构造的无限移动平均过程[4] [5]；Qin (2017) 等同样考虑误差由独立随机变量构造的无限移动平均过程，获得方差变点 CUSUM 型估计量的强收敛性和迭代算法[6]。陈璐(2019)建立简单的独立序列模型，构造均值方差变点位置的 CUSUM 型估计量，研究了独立序列均值方差变点估计问题[7]。房炜杰(2020)研究了方差变点的加权累积和估计的一些极限性质[8]。朱慧敏(2022)基于混合误差序列，研究方差变点模型 CUSUM 型估计量的相合性问题[9]。

有关测量误差的方差多变点的估计问题研究较少。因此估计测量误差模型中方差多变点的位置是一个重要且具有挑战性的统计问题。

测量误差模型中关于变点的研究，Chang 和 Huang (1997) 应用了似然比方法对测量误差模型中变量 X_j 的均值是否存在变点进行了检验，并给出了均值变点的极大似然估计[10]。王黎明(2002)应用了似然比方法对测量误差回归模型中回归系数是否存在变点进行了检验，并给出了回归系数变点的极大似估计[11]。You (2006) 对测量误差模型中回归系数的变点进行了研究[12]。而关于测量误差模型中方差变点的研究很少，Dong (2013, 2016) 分别用特征函数法[13]和非参数的 U 统计量法[14]研究了测量误差模型中的方差变点的估计量及其收敛速度。

Horváth 和 Kokoszka (1997) 应用 CUSUM 方法对单变量序列中的变点进行估计时，在数值模拟过程中发现调节参数对变点估计的精确度有显著影响[15]。Fremdt (2013) 对单变量序列是否存在变点提出了改进的 CUSUM 型检验统计量，数据模拟过程发现当调节参数 $\gamma \in (0, 0.5)$ 时变点检验效果更稳健[16]。谭常春和江敏(2020)利用蒙特卡洛方法研究了 CUSUM 型统计量中调节参数对单变量序列均值变点估计效果的影响[17]。沙达克提·艾力(2023)对方差变点用特征函数构造了一个含有调节参数的“CUSUM 型估计量”，研究了方差变点估计量的弱(强)相合性以及收敛速度[18]。本文在 Gombay (1994) 的基础上构造含有调节参数的 CUSUM 型估计量，研究方差变点估计量的相合性以及收敛速度。结合“二元分割法”推

广至多个方差变点的估计。最后, 对 2008 年 1 月~2024 年 1 月原油价格的涨跌幅周数据进行实证分析。

2. 模型与主要结果

考虑含有测量误差的方差变点模型: $1 \leq k_0 < n$, 假设 $\{X_1, X_2, \dots, X_n\}$ 是一组含有一个方差变点的测量误差模型, 观测值满足,

$$X_i = \begin{cases} \mu + \sigma_1 e_i, & 1 \leq i \leq k_0; \\ \mu + \sigma_2 e_i, & k_0 + 1 \leq i \leq n. \end{cases} \quad (1)$$

其中: μ 为常数, 若 $\sigma_1^2 \neq \sigma_2^2$, 则未知时刻 k_0 称为含测量误差模型中的方差变点。当 $k_0 = n$ 时, 表明不存在方差变点。 e_1, \dots, e_n 为独立随机误差序列, 满足 $Ee_i = 0$, $Ee_i^2 = 1$, $1 \leq i \leq n$, 方差变点 k_0 的 CUSUM 型估计量 \hat{k} 定义为

$$\hat{k} = \arg \max_{1 \leq k \leq n} |U_k| \quad (2)$$

$$U_k = \left(\frac{k(n-k)}{n^2} \right)^{1-\gamma} \left(\frac{1}{k} \sum_{i=1}^k (X_i - \bar{X})^2 - \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \bar{X})^2 \right), \quad (3)$$

其中 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, 且必存在 L , 使得 $|X_j| \leq L$ 。

注: 当 $\gamma = 0$ 时, (3)式为 Gombay (1996) CUSUM 型估计量。

变点位置 k_0 和变点估计量 \hat{k} 取值介于 1 和 n 之间, 为研究 k_0 与 \hat{k} 之间的准确性问题, 将其转化为研究 $\frac{\hat{k}}{n}$ 与 $\frac{k_0}{n}$ 之间的相合性问题。不失一般性, 假设存在一个整数 $\tau_0 \in (0, 1)$ 满足 $k_0 = [\tau_0 n]$, 其中 $[x]$ 为 x 取整符号。则方差变点位置 τ_0 的 CUSUM 型估计量 $\hat{\tau}$ 定为

$$\hat{\tau} = \frac{\hat{k}}{n} = \frac{1}{n} \arg \max_{1 \leq k \leq n} |U_k|. \quad (4)$$

本节将讨论方差变点 CUSUM 型估计量 $\hat{\tau} = \frac{\hat{k}}{n}$ 与 $\tau_0 = \frac{k_0}{n}$ 的相合性问题。记 $\delta = \sigma_1^2 - \sigma_2^2$, 且不是一般性假设 $\delta > 0$ 。

定理 1 考虑方差变点模型(1), 设 $X_j, j = 1, \dots, n$ 是模型(1)的独立的观测值, 且该模型中的测量误差 $\{e_j\}$ 的方差在未知时刻 k_0 发生改变, 则

(i) 由(4)定义的方差变点的估计量 $\hat{\tau}$ 是 τ_0 的弱相合估计, 且 $\lim_{n \rightarrow \infty} |\hat{\tau} - \tau_0| g_1(n) = 0$ 依概率成立, 其中:

$$g_1(n) = n^{\frac{1}{2}} \ell^{-1}(n).$$

(ii) $\hat{\tau}$ 是 τ_0 的强相合估计, 且 $\lim_{n \rightarrow \infty} |\hat{\tau} - \tau_0| g_2(n) = 0$ a.s. 其中: $g_2(n) = n^{\frac{1}{2}} (\log n)^{-\frac{1}{2}} \ell^{-1}(n)$ 。

其中: $\ell(n)$ 是满足 $\lim_{n \rightarrow \infty} \ell(n) = \infty$ 的慢变函数, $g_1(n)$ 与 $g_2(n)$ 是满足 $\lim_{n \rightarrow \infty} g_j(n) = \infty$ 的非负函数。

引理 1 [19] (Hoeffding 不等式) 假设 $0 \leq X_j \leq 1$, $\mu_j = E(X_j)$, $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$, $\mu = E(\bar{X})$,

$S_j = \sum_{i=1}^j X_i$, 则对于 $0 < x < 1 - \mu$,

$$P(\bar{X} - \mu \geq x) \leq P\left(\max_{1 \leq j \leq n} (S_j - E S_j) \geq nx\right) \leq \exp\{-2nx^2\},$$

如果存在着 $a_j \leq b_j$, 使得 $a_j \leq X_j \leq b_j$, $j = 1, 2, \dots, n$, 那么对任意的 $x > 0$,

$$P(\bar{X} - \mu \geq x) \leq P\left(\max_{1 \leq j \leq n}(S_j - E(S_j)) \geq nx\right) \leq \exp\left\{-\frac{2n^2x^2}{\sum_{j=1}^n (b_j - a_j)^2}\right\}.$$

特别地, 若当 $|X_j| \leq L$, 则 $b_j - a_j = 2L$, 那么,

$$P\left(\max_{1 \leq j \leq n}|S_j - ES_j| > nx\right) \leq \exp\left\{-\frac{2n^2x^2}{\sum_{k=1}^n (2L)^2}\right\} = \exp\left\{-\frac{nx^2}{2L^2}\right\}.$$

定理 1(i)的证明:

$$\begin{aligned} U_k &= \left(\frac{k(n-k)}{n^2}\right)^{1-\gamma} \left(\frac{1}{k} \sum_{i=1}^k (X_i - \bar{X})^2 - \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \bar{X})^2 \right) \\ &= \left(\frac{k(n-k)}{n^2}\right)^{1-\gamma} \left(\frac{1}{k} \sum_{i=1}^k ((X_i - \mu) + (\mu - \bar{X}))^2 - \frac{1}{n-k} \sum_{i=k+1}^n ((X_i - \mu) + (\mu - \bar{X}))^2 \right) \\ &= \left(\frac{k(n-k)}{n^2}\right)^{1-\gamma} \left[\left(\frac{1}{k} \sum_{i=1}^k (X_i - \mu)^2 + \frac{2(\mu - \bar{X})}{k} \sum_{i=1}^k (X_i - \mu) + (\mu - \bar{X})^2 \right) \right. \\ &\quad \left. - \left(\frac{1}{n-k} \sum_{i=k+1}^n (X_i - \mu)^2 + \frac{2(\mu - \bar{X})}{n-k} \sum_{i=k+1}^n (X_i - \mu) + (\mu - \bar{X})^2 \right) \right] \\ &= \left(\frac{k(n-k)}{n^2}\right)^{1-\gamma} \left(\frac{1}{k} \sum_{i=1}^k (X_i - \mu)^2 - \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \mu)^2 \right) \\ &\quad - \left(\frac{k(n-k)}{n^2}\right)^{1-\gamma} 2(\mu - \bar{X}) \left(\frac{1}{k} \sum_{i=1}^k (X_i - \mu) + \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \mu) \right) \\ &= U_{k,1} + U_{k,2}. \end{aligned} \tag{5}$$

由切比雪夫不等式, 知

$$P(|\bar{X} - \mu| > \varepsilon) \leq \frac{\text{var}(\bar{X})}{\varepsilon^2} \leq \frac{\alpha \max\left\{\frac{\sigma_1^2}{n}, \frac{\sigma_2^2}{n}\right\}}{\varepsilon^2} \rightarrow 0, \quad n \rightarrow \infty.$$

由中心极限定理知:

$$\frac{1}{k} \sum_{i=1}^k (X_i - \mu) = O(1), \quad \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \mu) = O(1)$$

在(5)式中 $U_{k,1}$ 和 $U_{k,2}$ 项相比, $U_{k,1}$ 是主项。即: $U_k = U_{k,1} + o(1)$ 。

当 $k \leq k_0$ 时, 有:

$$\begin{aligned} EU_{k,1} &= \left(\frac{k(n-k)}{n^2}\right)^{1-\gamma} E\left(\frac{1}{k} \sum_{i=1}^k (X_i - \mu)^2 - \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \mu)^2\right) \\ &= \left(\frac{k(n-k)}{n^2}\right)^{1-\gamma} \left(\frac{1}{k} \sum_{i=1}^k E(X_i - \mu)^2 - \frac{1}{n-k} \sum_{i=k+1}^{k_0} E(X_i - \mu)^2 - \frac{1}{n-k} \sum_{i=k_0+1}^n E(X_i - \mu)^2 \right) \\ &= \left(\frac{k(n-k)}{n^2}\right)^{1-\gamma} \left(\sigma_1^2 - \frac{k_0 - k}{n - k} \sigma_1^2 - \frac{n - k_0}{n - k} \sigma_2^2 \right) \\ &= \tau(1 - \tau_0) [\tau(1 - \tau)]^{-\gamma} \delta. \end{aligned} \tag{6}$$

当 $k > k_0$ 时, 有:

$$\begin{aligned}
 EU_{k,1} &= \left(\frac{k(n-k)}{n^2} \right)^{1-\gamma} E \left(\frac{1}{k} \sum_{i=1}^k (X_i - \mu)^2 - \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \mu)^2 \right) \\
 &= \left(\frac{k(n-k)}{n^2} \right)^{1-\gamma} \left(\frac{1}{k} \sum_{i=1}^{k_0} E(X_i - \mu)^2 + \frac{1}{k} \sum_{i=k_0+1}^k E(X_i - \mu)^2 - \frac{1}{n-k} \sum_{i=k+1}^n E(X_i - \mu)^2 \right) \\
 &= \left(\frac{k(n-k)}{n^2} \right)^{1-\gamma} \left(\frac{k_0}{k} \sigma_1^2 + \frac{k-k_0}{k} \sigma_2^2 - \sigma_2^2 \right) \\
 &= \tau_0 (1-\tau) [\tau (1-\tau)]^{-\gamma} \delta.
 \end{aligned} \tag{7}$$

由(6)式有:

$$EU_{k_0,1} = [\tau_0 (1-\tau_0)]^{1-\gamma} \delta. \tag{8}$$

对于 $U_{k,2}$, 有:

$$\begin{aligned}
 EU_{k,2} &= E \left[\left(\frac{k(n-k)}{n^2} \right)^{1-\gamma} 2(\mu - \bar{X}) \left(\frac{1}{k} \sum_{i=1}^k (X_i - \mu) - \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \mu) \right) \right] \\
 &= [\tau (1-\tau)]^{1-\gamma} E \left[2(\mu - \bar{X}) \left(\frac{1}{k} \sum_{i=1}^k (X_i - \mu) - \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \mu) \right) \right] \\
 &= o(1).
 \end{aligned} \tag{9}$$

$$EU_k = EU_{k,1} + EU_{k,2} = \begin{cases} \tau (1-\tau_0) [\tau (1-\tau)]^{-\gamma} \delta + o(1), & 1 \leq i \leq k_0, \\ \tau_0 (1-\tau) [\tau (1-\tau)]^{-\gamma} \delta + o(1), & k_0 + 1 \leq i \leq n, \end{cases}$$

从(6)~(9)式知, 当 n 充分大时,

$$\begin{aligned}
 &|EU_{k_0}| - |EU_k| \\
 &= \begin{cases} [\tau_0 (1-\tau_0)]^{1-\gamma} \delta - \tau (1-\tau_0) [\tau (1-\tau)]^{-\gamma} \delta, & 1 \leq i \leq k_0, \\ [\tau_0 (1-\tau_0)]^{1-\gamma} \delta - \tau_0 (1-\tau) [\tau (1-\tau)]^{-\gamma} \delta, & k_0 + 1 \leq i \leq n, \end{cases} \\
 &= \begin{cases} (1-\tau_0)^{1-\gamma} \delta [\tau_0^{1-\gamma} - \tau^{1-\gamma} (1-\tau_0)(1-\tau)^{-\gamma}], & 1 \leq i \leq k_0, \\ (\tau_0)^{1-\gamma} \delta [(1-\tau_0)^{1-\gamma} - \tau_0^\gamma \tau^{-\gamma} (1-\tau)^{1-\gamma}], & k_0 + 1 \leq i \leq n, \end{cases} \\
 &= \begin{cases} (1-\tau_0)^{1-\gamma} \delta \left[\tau_0^{1-\gamma} - \tau^{1-\gamma} \left(\frac{1-\tau_0}{1-\tau} \right)^\gamma \right], & 1 \leq i \leq k_0, \\ (\tau_0)^{1-\gamma} \delta \left[(1-\tau_0)^{1-\gamma} - (1-\tau)^{1-\gamma} \left(\frac{\tau_0}{\tau} \right)^\gamma \right], & k_0 + 1 \leq i \leq n, \end{cases} \\
 &\geq \begin{cases} (1-\tau_0)^{1-\gamma} \delta [\tau_0^{1-\gamma} - \tau^{1-\gamma}], & 1 \leq i \leq k_0, \\ (\tau_0)^{1-\gamma} \delta [(1-\tau_0)^{1-\gamma} - (1-\tau)^{1-\gamma}], & k_0 + 1 \leq i \leq n, \end{cases}
 \end{aligned}$$

若令 $f_1(\tau) = \tau^{1-\gamma}$, ($0 < \gamma < 1$), $f_2(\tau) = (1-\tau)^{1-\gamma}$, ($0 < \gamma < 1$) 由拉格朗日中值定理知,

$$|EU_{k_0}| - |EU_k| \geq \begin{cases} (1-\tau_0)^{1-\gamma} (1-\gamma) \tau_0 (\tau_0 - \tau) \delta, & k < k_0, \\ (\tau_0)^{1-\gamma} (1-\gamma) (1-\tau_0)^{-\gamma} (\tau - \tau_0) \delta, & k > k_0, \end{cases}$$

记: $\bar{\tau} = (1-\gamma) [\tau_0 (1-\tau_0)]^{-\gamma}$, $\alpha = \min \{\tau_0, 1-\tau_0\}$.

$$|EU_{k_0}| - |EU_k| \geq \bar{\tau} |\tau - \tau_0| \alpha \delta \quad (10)$$

由三角不等式,

$$\begin{aligned} |U_{k_0}| - |U_k| &\leq |U_k - EU_k| + |EU_k| + |U_{k_0} - EU_{k_0}| - |EU_{k_0}| \\ &\leq 2 \max_{1 \leq k \leq n} |U_k - EU_k| - |EU_{k_0}| + |EU_k|, \end{aligned} \quad (11)$$

结合(10)式, 并由 \hat{k} 的定义知 $|U_{\hat{k}}| \geq |U_{k_0}|$, 则有

$$\bar{\tau} |\hat{\tau} - \tau_0| \alpha \delta \leq |EU_{k_0}| - |EU_{\hat{k}}| \leq 2 \max_{1 \leq k \leq n} |U_k - EU_k| \quad (12)$$

即 $\bar{\tau} |\hat{\tau} - \tau_0| \alpha \delta \leq 2 \max_{1 \leq k \leq n} |U_k - EU_k|$.

结合(1)和(4)式有,

$$\begin{aligned} &\max_{1 \leq k \leq n} |U_k - EU_k| \\ &\leq \max_{1 \leq k \leq n} |U_{k,1} - EU_{k,1} + U_{k,2} - EU_{k,2}| \\ &\leq \max_{1 \leq k \leq n} |U_{k,1} - EU_{k,1}| + \max_{1 \leq k \leq n} |U_{k,2} - EU_{k,2}| \\ &= \max_{1 \leq k \leq n} \left[\left(\frac{k(n-k)}{n^2} \right)^{1-\gamma} \left| \frac{1}{k} \sum_{i=1}^k ((X_i - \mu)^2 - E(X_i - \mu)^2) - \frac{1}{n-k} \sum_{i=k+1}^n ((X_i - \mu)^2 - E(X_i - \mu)^2) \right| \right] \\ &\quad + \max_{1 \leq k \leq n} \left[2 |\mu - \bar{X}| \left(\frac{k(n-k)}{n^2} \right)^{1-\gamma} \left| \frac{1}{k} \sum_{i=1}^k ((X_i - \mu) - E(X_i - \mu)) - \frac{1}{n-k} \sum_{i=k+1}^n ((X_i - \mu) - E(X_i - \mu)) \right| \right] \\ &\leq \max_{1 \leq k \leq n} \frac{1}{k} \left| \sum_{i=1}^k ((X_i - \mu)^2 - E(X_i - \mu)^2) \right| + \max_{1 \leq k \leq n} \frac{1}{n-k} \left| \sum_{i=k+1}^n ((X_i - \mu)^2 - E(X_i - \mu)^2) \right| \\ &\quad + 2(\mu - \bar{X}) \max_{1 \leq k \leq n} \frac{1}{k} \left| \sum_{i=1}^k (X_i - E(X_i)) \right| + 2(\mu - \bar{X}) \max_{1 \leq k \leq n} \frac{1}{n-k} \left| \sum_{i=k+1}^n (X_i - E(X_i)) \right| \\ &\triangleq I_1 + I_2 + I_3 + I_4 \end{aligned} \quad (13)$$

由于 $\delta \neq 0$ 以及 $\bar{\tau} \neq 0$, 所以

$$|\hat{\tau} - \tau_0| \leq \frac{2 \max_{1 \leq k \leq n} |U_k - EU_k|}{\bar{\tau} \alpha \delta}$$

因此,

$$\begin{aligned} &P(|\hat{\tau} - \tau_0| > \varepsilon) \\ &\leq P\left(\frac{2 \max_{1 \leq k \leq n} |U_k - EU_k|}{\bar{\tau} \alpha \delta} > \varepsilon\right) \\ &\leq P\left(I_1 > \frac{\varepsilon \bar{\tau} \alpha \delta}{8}\right) + P\left(I_2 > \frac{\varepsilon \bar{\tau} \alpha \delta}{8}\right) + P\left(I_3 > \frac{\varepsilon \bar{\tau} \alpha \delta}{8}\right) + P\left(I_4 > \frac{\varepsilon \bar{\tau} \alpha \delta}{8}\right) \end{aligned} \quad (14)$$

令 $Y_i = X_i - \mu$, 记 $T_k = \sum_{i=1}^k Y_i^2$, $T_{n-k} = \sum_{i=k+1}^n Y_i^2$, 由引理 1, 我们有,

$$\begin{aligned}
I_1 &= \max_{1 \leq k \leq n} \frac{1}{k} \left| \sum_{i=1}^k (Y_i^2 - EY_i^2) \right| = \max_{1 \leq k \leq n} \frac{1}{k} |T_k - ET_k| \\
P\left\{ I_1 > \frac{\varepsilon \bar{\tau} \alpha \delta}{8} \right\} &= P\left\{ \max_{1 \leq k \leq n} \frac{1}{k} |T_k - ET_k| > \frac{\varepsilon \bar{\tau} \alpha \delta}{8} \right\} \\
&\leq P\left\{ \max_{1 \leq k \leq n} |T_k - ET_k| > \frac{n \varepsilon \bar{\tau} \alpha \delta}{8} \right\} \\
&\leq \exp\left\{ -\frac{n}{2} \left(\frac{\varepsilon \bar{\tau} \alpha \delta}{8L} \right)^2 \right\} \rightarrow 0, \quad n \rightarrow \infty.
\end{aligned} \tag{15}$$

同理,

$$\begin{aligned}
P\left\{ I_2 > \frac{\varepsilon \bar{\tau} \alpha \delta}{8} \right\} &= P\left\{ \max_{1 \leq k \leq n} \frac{1}{n-k} |T_{n-k} - ET_{n-k}| > \frac{\varepsilon \bar{\tau} \alpha \delta}{8} \right\} \\
&\leq P\left\{ \max_{1 \leq k \leq n} |T_{n-k} - ET_{n-k}| > \frac{n \varepsilon \bar{\tau} \alpha \delta}{8} \right\} \\
&\leq \exp\left\{ -\frac{n}{2} \left(\frac{\varepsilon \bar{\tau} \alpha \delta}{8L} \right)^2 \right\} \rightarrow 0, \quad n \rightarrow \infty.
\end{aligned} \tag{16}$$

对于 I_3, I_4 , 由于 $|X| \leq L$, 必存在 $M > 0$, 使得 $|\mu - \bar{X}| \leq M$,

$$\begin{aligned}
P\left\{ I_3 > \frac{\varepsilon \bar{\tau} \alpha \delta}{8} \right\} &= P\left(2 \left| (\mu - \bar{X}) \right| \max_{1 \leq k \leq n} \frac{1}{k} \left| \sum_{i=1}^k (X_i - E(X_i)) \right| > \frac{\varepsilon \bar{\tau} \alpha \delta}{8} \right) \\
&\leq P\left(\max_{1 \leq k \leq n} \frac{1}{k} \left| \sum_{i=1}^k (X_i - E(X_i)) \right| > \frac{\varepsilon \bar{\tau} \alpha \delta}{16M} \right) \\
&\leq \exp\left\{ -\frac{n}{2} \left(\frac{\varepsilon \bar{\tau} \alpha \delta}{16M} \right)^2 \right\} \rightarrow 0, \quad n \rightarrow \infty.
\end{aligned} \tag{17}$$

同理,

$$\begin{aligned}
P\left\{ I_4 > \frac{\varepsilon \bar{\tau} \alpha \delta}{8} \right\} &= P\left(2 \left| (\mu - \bar{X}) \right| \max_{1 \leq k \leq n} \frac{1}{n-k} \left| \sum_{i=1}^k (X_i - E(X_i)) \right| > \frac{\varepsilon \bar{\tau} \alpha \delta}{8} \right) \\
&\leq P\left(\max_{1 \leq k \leq n} \frac{1}{n-k} \left| \sum_{i=1}^k (X_i - E(X_i)) \right| > \frac{\varepsilon \bar{\tau} \alpha \delta}{16M} \right) \\
&\leq \exp\left\{ -\frac{n}{2} \left(\frac{\varepsilon \bar{\tau} \alpha \delta}{16M} \right)^2 \right\} \rightarrow 0, \quad n \rightarrow \infty.
\end{aligned} \tag{18}$$

因此必存在常数 C , 有,

$$P(|\hat{\tau} - \tau_0| > \varepsilon) \leq 4 \exp\left\{ -\frac{n}{2} \left(\frac{\varepsilon \bar{\tau} \alpha \delta}{C} \right)^2 \right\} \rightarrow 0, \quad n \rightarrow \infty.$$

这意味着 $\hat{\tau}$ 是 τ_0 的弱相合估计。由式(15)~(18)进一步可得,

$$P(g_1(n)|\hat{\tau} - \tau_0| > \varepsilon) \leq 4 \exp\left\{ -\frac{n}{2} \left(\frac{\varepsilon \bar{\tau} \alpha \delta}{Cg_1(n)} \right)^2 \right\}. \tag{19}$$

若 $g_1(n) = n^{\frac{1}{2}}\ell^{-1}(n)$ 时, 则有: $\lim_{x \rightarrow \infty} |\hat{\tau} - \tau_0| g_1(n) = 0$ 依概率成立, 其中: $\ell(n)$ 是一个趋于 ∞ 的慢变函数。

定理 1 (ii) 的证明: 由(19)式知,

$$\sum_{i=1}^n P(g_2(n)|\hat{\tau} - \tau_0| > \varepsilon) \leq 4 \sum_{i=1}^n \exp \left\{ -\frac{n}{2} \left(\frac{\varepsilon \bar{\tau} \alpha \delta}{C g_2(n)} \right)^2 \right\},$$

若取 $g_2(n) = n^{\frac{1}{2}}(\log n)^{-\frac{1}{2}}\ell^{-1}(n)$, 则有 $\sum_{i=1}^n \exp \left\{ -\frac{n}{2} \left(\frac{\varepsilon \bar{\tau} \alpha \delta}{C g_2(n)} \right)^2 \right\} < \infty$, 由 Borel-Cantelli 引理知 $\hat{\tau}$ 几乎

处处收敛到 τ_0 。且 $\lim_{n \rightarrow \infty} |\hat{\tau} - \tau_0| g_2(n) = 0$ a.s. 成立。

3. 含测量误差的方差多变点的估计

若模型(1)中存在 m 个变点, 且变点个数 m 已知, 则模型(1)转化为

$$X_i = \begin{cases} \mu + \sigma_1 e_i, & i = 1, 2, \dots, k_1; \\ \mu + \sigma_2 e_i, & i = k_1 + 1, \dots, k_2; \\ \vdots \\ \mu + \sigma_m e_i, & i = k_{m-1} + 1, \dots, k_m; \\ \mu + \sigma_{m+1} e_i, & i = k_m + 1, \dots, n. \end{cases} \quad (20)$$

再结合二元分割法将上述方法推广到多变点的情形, 则模型(20)中的方差多变点估计具体步骤如下:

步骤一: 对全体数据利用(4)式估计出第一个变点 \hat{k}_1 ;

步骤二: 变点 \hat{k}_1 将数据分成两个子序列, 第一部分为 $Y_1, Y_2, \dots, Y_{\hat{k}_1}$, 第二部分为 $Y_{\hat{k}_1+1}, Y_{\hat{k}_1+2}, \dots, Y_n$, 在这两段样本内分别计算 $\max_{1 \leq k \leq \hat{k}_1} |U_k(t)|$ 和 $\max_{\hat{k}_1 \leq k \leq n} |U_k(t)|$;

步骤三: 比较 $\max_{1 \leq k \leq \hat{k}_1} |U_k(t)|$ 和 $\max_{\hat{k}_1 \leq k \leq n} |U_k(t)|$ 的大小, 选择两者中较大的那个, 并选择较大的所对应的 k 值, 记为 \hat{k}_2 ;

步骤四: 重复步骤2和步骤3, 直到 m 个变点被估计出来。

4. 调节参数 γ 的取值对应变点位置估计的影响分析

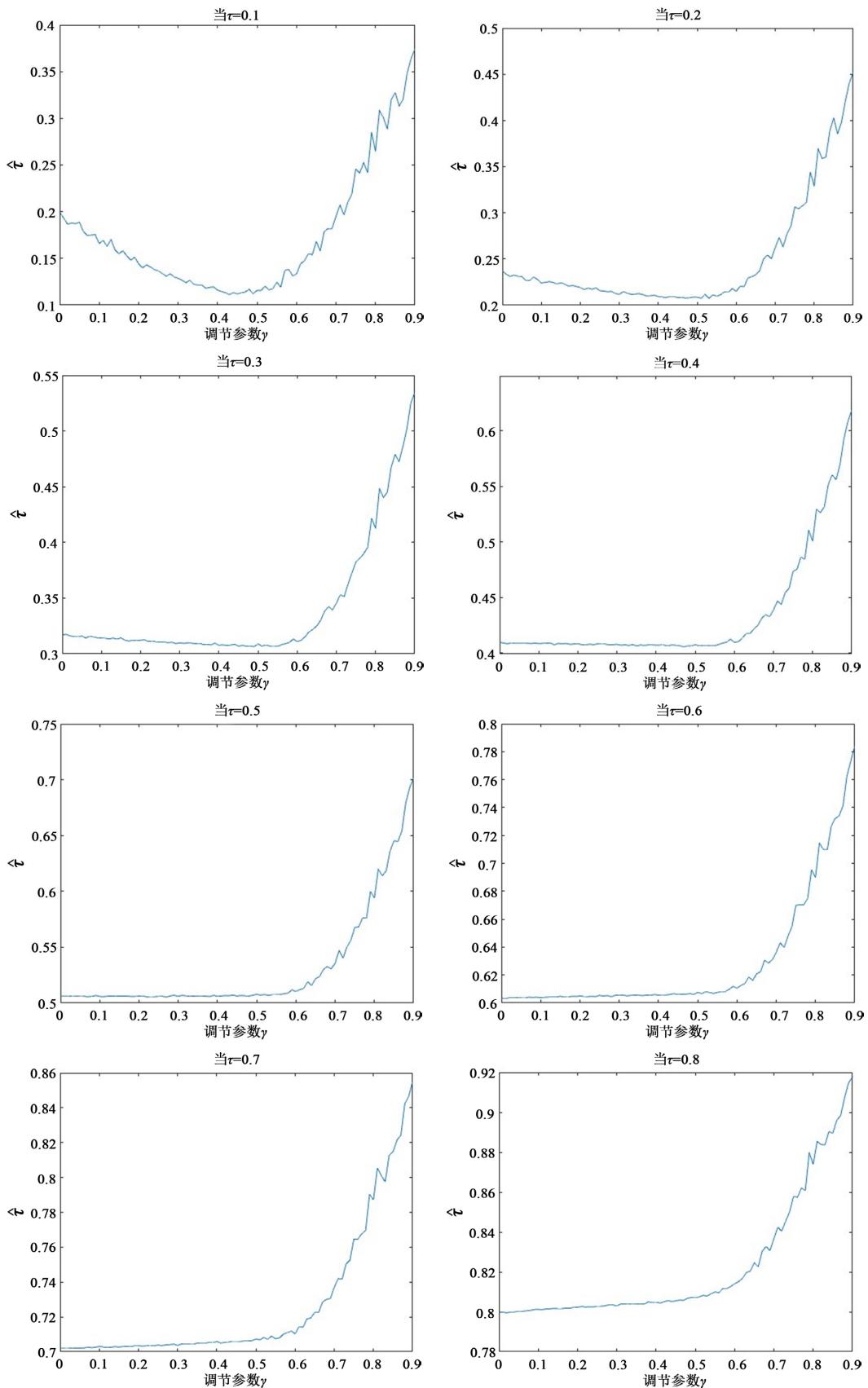
为了研究调节参数 γ 对变点估计量的影响, 用 Matlab 软件进行模拟分析, 假设变点真实位置 τ_0 分别为 $0.1, 0.2, \dots, 0.9$, 共九种情形, 取 $\mu = 1$, $\sigma_1 = 1$, $\sigma_2 = 2$, 调节参数 γ 取值 $(0, 1)$ 上的格子点, 步长为 0.01; 重复模拟次数 $m = 1000$; 样本容量 n 选取为 1000。

假设测量误差 $X_i, i = 1, 2, \dots, n$, i.i.d. $\sim N(0, 1)$ 此处为了方便起见, 在给定调节参数 γ 时, 记第 i 次重复的所得到的变点估计量为 $\hat{\tau}(i, \gamma)$, $i = 1, 2, \dots, m$; 取 $\hat{\tau}(i, \gamma)$ 的均值作为对应调节参数下的变点估计量, 并记为 $\hat{\tau}(\gamma)$, 即

$$\hat{\tau}(\gamma) = \frac{1}{m} \sum_{i=1}^m \hat{\tau}(i, \gamma) \quad (21)$$

图1给出了真实变点位置 τ_0 分别在 $0.1, 0.2, \dots, 0.9$ 时, 调节参数 γ 与变点估计量 $\hat{\tau}(\gamma)$ 的关系。

表1给出了真实变点位置 τ_0 分别在 $0.1, 0.2, \dots, 0.9$ 时, 不同调节参数 γ 下得到的变点估计量的均值 (mean), 标准差 (std), 众数 (mode), 中位数 (median)。



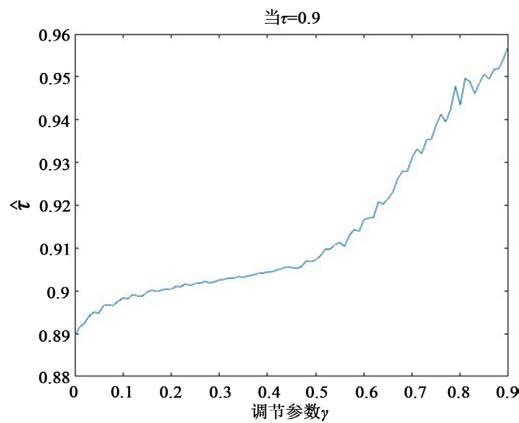


Figure 1. When $\sigma_1 = 1$, $\sigma_2 = 2$, the relationship between the adjustment parameter γ and the change point estimator $\hat{\tau}(\gamma)$ under the change point position τ_0 of different true values

图1. $\sigma_1 = 1$, $\sigma_2 = 2$ 时, 不同真实值变点位置 τ_0 下调节参数 γ 与变点估计量 $\hat{\tau}(\gamma)$ 之间的关系

Table 1. The mean, standard deviation, mode and median of the change point estimator given by (21) when the adjustment parameter $\gamma \in (0, 0.9)$ and step size is 0.1

表1. 调节参数 $\gamma \in (0, 0.9)$, 步长为 0.1 时, 由(21)给定的变点估计量的均值、标准差、众数、中位数

$\tau_0 = 0.1$										
	γ									
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.1994	0.1696	0.1427	0.1278	0.1149	0.11500	0.1337	0.1867	0.2786	0.3881
std	0.1187	0.1008	0.0782	0.0633	0.0520	0.0800	0.1577	0.2615	0.3570	0.4209
mode	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
median	0.1530	0.1270	0.1120	0.1080	0.1040	0.1030	0.1020	0.1010	0.1020	0.1010
$\tau_0 = 0.2$										
	γ									
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.2371	0.2269	0.2176	0.2137	0.2095	0.2088	0.2192	0.2558	0.3438	0.4537
std	0.0488	0.0431	0.0295	0.0205	0.0152	0.0389	0.1059	0.1975	0.3039	0.3713
mode	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000	0.2000
median	0.2160	0.2100	0.2070	0.2050	0.2040	0.2020	0.2020	0.2020	0.2020	0.2030
$\tau_0 = 0.3$										
	γ									
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.3168	0.3147	0.3113	0.3087	0.3076	0.3076	0.3121	0.3445	0.4242	0.5369
std	0.0244	0.0225	0.0179	0.0143	0.0133	0.0329	0.0698	0.1611	0.2629	0.3298
mode	0.3000	0.3000	0.3000	0.3000	0.3000	0.3000	0.3000	0.3000	0.3000	0.3000
median	0.3070	0.3060	0.3040	0.3040	0.303	0.3030	0.3020	0.3020	0.3030	0.3040

续表

$\tau_0 = 0.4$										
γ										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.4097	0.4086	0.4076	0.4076	0.4068	0.4068	0.4124	0.4378	0.5098	0.6224
std	0.0149	0.0143	0.0112	0.0122	0.0112	0.0222	0.0628	0.1343	0.2267	0.2889
mode	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.4000	0.9990
median	0.4040	0.4040	0.4030	0.4030	0.4030	0.4020	0.4020	0.4030	0.4030	0.4080
$\tau_0 = 0.5$										
γ										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.5060	0.5061	0.5062	0.5067	0.5067	0.5076	0.5126	0.5358	0.603	0.7026
std	0.0093	0.0091	0.0105	0.0104	0.0102	0.0247	0.0533	0.1156	0.1947	0.2462
mode	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.9990
median	0.5030	0.5030	0.5030	0.5030	0.5030	0.5030	0.5030	0.5030	0.5050	0.5160
$\tau_0 = 0.6$										
γ										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.6035	0.6035	0.6046	0.6051	0.6056	0.6071	0.6115	0.6350	0.6940	0.7793
std	0.0065	0.0063	0.0082	0.0091	0.0093	0.0203	0.0426	0.0987	0.1589	0.2030
mode	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.9990
median	0.6020	0.6010	0.6020	0.6020	0.6020	0.6030	0.6030	0.6040	0.6070	0.6350
$\tau_0 = 0.7$										
γ										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.7020	0.7026	0.7034	0.7040	0.7054	0.7069	0.7123	0.7349	0.7864	0.8568
std	0.0053	0.0062	0.0065	0.0073	0.0088	0.0174	0.0370	0.0798	0.1233	0.1588
mode	0.7000	0.7000	0.7000	0.7000	0.7000	0.7000	0.7000	0.7000	0.7000	0.9990
median	0.7010	0.7010	0.7010	0.7020	0.7020	0.7020	0.7030	0.7055	0.7120	0.9600
$\tau_0 = 0.8$										
γ										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.7997	0.8010	0.8022	0.8031	0.8044	0.8074	0.8138	0.8355	0.8744	0.9136
std	0.0070	0.0057	0.0064	0.0066	0.0082	0.0140	0.0320	0.0627	0.0836	0.1267
mode	0.8000	0.8000	0.8000	0.8000	0.8000	0.8000	0.8000	0.8000	0.9990	0.9990
median	0.8000	0.8010	0.8010	0.8010	0.8020	0.8030	0.8040	0.8060	0.8210	0.9780

续表

$\tau_0 = 0.9$										
γ										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
mean	0.8894	0.8983	0.9009	0.9026	0.9042	0.9077	0.9146	0.9300	0.9435	0.9500
std	0.0240	0.0092	0.0059	0.0065	0.0081	0.0143	0.0386	0.0357	0.0718	0.1264
mode	0.900	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9990	0.9990
median	0.899	0.9000	0.9010	0.9010	0.9020	0.9030	0.9040	0.9110	0.9375	0.9835

注1：从图1和表1中我们都可以发现，当变点位置靠近端点($\tau_0 = 0.1, 0.9$)时，变点估计值与真实值之间有一定的偏差；

注2：当 $\gamma \in (0.3, 0.7)$ 时，变点估计量与真实值最接近，都优于无调节参数($\gamma = 0$)时的情况；

注3：当 $\sigma_1 = 1, \sigma_2 = 3$ 时的模拟，跟上述模拟效果相差不大，由于篇幅限制，不再列出图表解释了。

5. 实证分析

基于上述模拟结果，选取调节参数 $\gamma \in (0.3, 0.7)$ ，以2008年1月~2024年1月原油价格的涨跌幅周数据(数据来源于<https://cn.investing.com>)作为研究对象，结合“二元分割法”并应用含有调节参数的“CUSUM型估计量”对原油价格涨跌幅的方差变点进行估计，样本量 $n = 832$ 。首先应用(4)式含有调节参数的“CUSUM型估计量”对这832个原油价格涨跌幅数据进行方差变点估计，得到变点的估计值见表2所示，得到的变点估计值都是60，所对应的实际时间是2009年2月28日，出现变点的原因时美国次贷危机出现，宏观经济拉低了原油的需求，导致2008年油价暴跌。OPEC于2008年9、11、12月均宣布减产，这个减产幅度尚不足以抵消过剩，加之刺激经济的政策在短时内还未能显现作用，宏观经济依旧疲软，因此至2009年1~2月，油价延续了这一周期的下行趋势，跌入该周期低位，因此价格出现波动。

Table 2. The first variance change point estimated by CUSUM estimators under different adjustment parameters γ and corresponding actual time

表2. 不同的调节参数 γ 下 CUSUM型估计量估计的第一个方差变点的估计值及对应的实际时间

变点对应的实际时间	调节参数 γ 的取值	变点估计值 \hat{k}
2009年2月22日	0.3	60
2009年2月22日	0.4	60
2009年2月22日	0.5	60
2009年2月22日	0.6	60
2009年2月22日	0.7	60

应用“二元分割法”，变点60(即2009年2月22日)将2008年1月~2024年1月原油价格涨跌幅的832个周数据一分为二，得到2008年1月~2009年2月和2009年3月~2024年1月两个子序列，由于2008年1月~2009年2月的样本量少，不作为研究对象。应用(4)式含有调节参数的“CUSUM型估计量”对2009年3月~2024年1月进行方差变点估计，样本量为771，得到的方差变点估计值见表3所示。对应的实际时间为2020年2月23日，出现变点的原因主要是2020年初，全球疫情的爆发使得原油再度暴跌。

Table 3. The second variance change point estimated by CUSUM estimators under different adjustment parameters γ and corresponding actual time

表3. 不同的调节参数 γ 下 CUSUM 型估计量估计的第二个方差变点估计值及对应的实际时间

变点对应的实际时间	调节参数 γ 的取值	变点估计值 \hat{k}
2020年2月23日	0.3	575
2020年2月23日	0.4	575
2020年2月23日	0.5	575
2020年2月23日	0.6	575
2020年2月23日	0.7	575

图2给出了2008年1月~2024年1月原油价格的涨跌幅周数据的方差变点(红线所示), 这与实际情况相符。

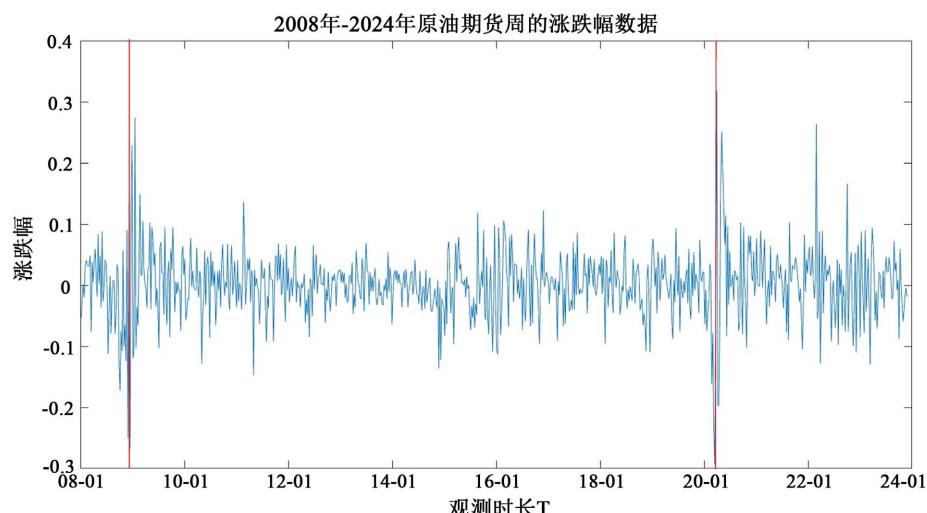


Figure 2. Variance change point of weekly data on the rise and fall of crude oil prices from January 2008 to January 2024
图2. 2008年1月~2024年1月原油价格涨跌幅周数据的方差变点

6. 结论

本文对含有测量误差的方差变点模型构造了含有调节参数的 CUSUM 型估计量, 并研究了方差变点估计量的相合性和收敛速度, 结合“二元分割法”推广至多个方差变点的情况。模拟发现, 当 $\gamma \in (0.3, 0.7)$ 时变点估计量与真实值最接近而且明显优于无调节参数($\gamma = 0$)时的情况。并进一步应用含有调节参数的 CUSUM 型估计量对原油价格涨跌幅的方差变点进行了实证分析, 结果表明基于调节参数 CUSUM 型估计量得到的方差变点与实际相符。

基金项目

新疆维吾尔自治区自然科学基金项目(2023D01A37, 2022D01A219)。

参考文献

- [1] Gombay, E., Horváth, L. and Husková, M. (1996) Estimators and Tests for Change in Variances. *Statistics & Risk Modeling*, **14**, 145-160. <https://doi.org/10.1524/strm.1996.14.2.145>

-
- [2] Inclan, C. and Tiao, G.C. (1994) Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance. *Journal of the American Statistical Association*, **89**, 913-923. <https://doi.org/10.1080/01621459.1994.10476824>
 - [3] Chen, J. and Gupta, A.K. (1997) Testing and Locating Variance Change Points with Application to Stock Prices. *Journal of the American Statistical Association*, **92**, 739-747. <https://doi.org/10.1080/01621459.1997.10474026>
 - [4] Lee, S. and Park, S. (2001) The Cusum of Squares Test for Scale Changes in Infinite Order Moving Average Processes. *Scandinavian Journal of Statistics*, **28**, 625-644. <https://doi.org/10.1111/1467-9469.00259>
 - [5] Lee, S., Ha, J., Na, O., et al. (2003) The Cusum Test for Parameter Change Time Series Models. *Scandinavian Journal of Statistics*, **30**, 781-796. <https://doi.org/10.1111/1467-9469.00364>
 - [6] Qin, R., Liu, W. and Tian, Z. (2017) A Strong Convergence Rate of Estimator of Variance Change in Linear Processes and Its Applications. *Statistics*, **51**, 314-330. <https://doi.org/10.1080/02331888.2016.1268614>
 - [7] 陈璐, 魏岳嵩, 尉梦珂. 独立序列均值与方差变点的估计[J]. 电脑知识与技术, 2019, 15(4): 218-220.
 - [8] 房炜杰, 金百锁. 方差变点的加权累积和估计的极限性质[J]. 中国科学技术大学学报, 2020, 50(4): 389-395.
 - [9] 朱慧敏, 王梓楠, 高敏, 等. 方差变点模型 CUSUM 型估计量的相合性[J]. 山东大学学报(理学版), 2023, 58(7): 106-114.
 - [10] Chang, Y.P. and Huang, W.T. (1997) Inferences for the Linear Errors-in-Variables with Change-Point Models. *Journal of the American Statistical Association*, **92**, 171-178. <https://doi.org/10.1080/01621459.1997.10473614>
 - [11] 王黎明. 测量误差模型只有一个变点的检验和估计[J]. 应用概率统计, 2002, 18(4): 385-392.
 - [12] You, J., Zhou, Y. and Chen, G. (2006) Corrected Local Polynomial Estimation in Varying-Coefficient Models with Measurement Errors. *Canadian Journal of Statistics*, **34**, 391-410. <https://doi.org/10.1002/cjs.5550340303>
 - [13] Dong, C., Miao, B., Tan, C., Wei, D. and Wu, Y. (2013) An Estimate of a Change Point in Variance of Measurement Errors and Its Convergence Rate. *Communications in Statistics—Theory and Methods*, **44**, 790-797. <https://doi.org/10.1080/03610926.2012.762395>
 - [14] Dong, C., Tan, C., Jin, B. and Miao, B. (2016) Inference on the Change Point Estimator of Variance in Measurement Error Models. *Lithuanian Mathematical Journal*, **56**, 474-491. <https://doi.org/10.1007/s10986-016-9330-3>
 - [15] Horváth, L. and Kokoszka, P. (1997) The Effect of Long-Range Dependence on Change-Point Estimators. *Journal of Statistical Planning and Inference*, **64**, 51-81. [https://doi.org/10.1016/S0378-3758\(96\)00208-X](https://doi.org/10.1016/S0378-3758(96)00208-X)
 - [16] Fremdt, S. (2014) Asymptotic Distribution of the Delay Time in Page's Sequential Procedure. *Journal of Statistical Planning & Inference*, **145**, 74-79. <https://doi.org/10.1016/j.jspi.2013.09.001>
 - [17] 谭常春, 江敏. CUSUM 型统计量中调节参数对变点估计效果的影响分[J]. 中国科学技术大学学报, 2020, 50(7): 920-928.
 - [18] 沙达克提·艾力, 董翠玲. 测量误差模型方差多变点的估计及收敛速度[J]. 理论数学, 2023, 13(11): 3262-3271.
 - [19] Hoeffding, W. (1963) Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, **58**, 13-30. <https://doi.org/10.1080/01621459.1963.10500830>