Z-process method for statistical change point problems with applications to discretely observed diffusion processes

Ilia Negri

University of Bergamo (Italy)

(Joint work with Y. Nishiyama)

SAPS – XI, 17 – 21 July 2017, Peterhof, Russia

Outline of the talk

Overview

- Testing on structural change problems has been an important issue in statistics
- Horváth and Parzen (1994) are apparently the firsts to introduce a statistics based on the Fisher-score process to test a parameter change for independent data
- Berkes et al. (2004) successfully apply this method to GARCH(p, q) models
- N & N (2012) take the same approach to the change point problem for an ergodic diffusion process model based on the continuous observation

This work in two points

- The aim of this work is to develop a general approach to change point problems using the partial sum processes of estimating equations, which may be called the Z-process method
- These problems arise in statistics when the parameters governing the model have a jump or a change during the observation period

To introduce the problem and illustrate the method based on Z-process we give an overview of the case with independent data

Preliminaries

Let us give an illustration by the example of independent data:

- Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a measure space
- Let $f(\cdot;\theta)$ a parametric family of probability densities with respect to μ, where $\theta \in \Theta \subset \mathbb{R}^d$.
- Let $X_1, X_2, \ldots$ be an independent sequence of $\mathcal{X}$-valued random variables from this parametric model.

$\mathbb{M}$ and $\mathbb{Z}$ processes

- Define
$$\theta \mapsto \mathbb{M}_n(\theta) = \sum_{k=1}^{n} \log f(X_k;\theta),$$
and
$$\mathbb{Z}_n(\theta) = \dot{\mathbb{M}}_n(\theta)$$
where $\dot{\mathbb{M}}_n(\theta)$ is the gradient vector of $\mathbb{M}_n(\theta)$.
- Introduce the partial sum process
$$\mathbb{M}_n(u,\theta) = \sum_{k=1}^{[un]} \log f(X_k;\theta), \quad \forall u \in [0,1],$$
and denote its gradient vectors by
$$\mathbb{Z}_n(u,\theta) = \dot{\mathbb{M}}_n(u,\theta)$$

The maximum likelihood estimator

There are at least two ways to define the maximum likelihood estimator (MLE) in statistics.

- One way is to define it as the maximum point of the random function

$$\theta \mapsto \mathbb{M}_n(\theta) = \sum_{k=1}^{n} \log f(X_k;\theta),$$

- The other is to do it as the solution to the estimating equation

$$\mathbb{Z}_n(\theta) = \dot{\mathbb{M}}_n(\theta) = 0$$

The former is a special case of M-estimators, and the latter is that of Z-estimators

Test for change point in i.i.d. case

In the previous framework the following testing problem (change point problem) is considered:

H_0: the true value $\theta_0 \in \Theta$ does not change during $u \in [0,1]$

versus any alternative that we can state as H_1: "not H_0".
To deal with this change point problems the partial sum process

$$\mathbb{M}_n(u,\theta) = \sum_{k=1}^{[un]} \log f(X_k;\theta), \quad \forall u \in [0,1],$$

and its gradient vectors

$$\mathbb{Z}_n(u,\theta) = \dot{\mathbb{M}}_n(u,\theta)$$

play a crucial role.

Horváth and Parzen (1994) statistics

Let $\widehat{\theta}_n$ be the MLE for the full data $X_1, \ldots, X_n$ as a special case of Z-estimators, that is, $\widehat{\theta}_n$ is the solution to the estimating equation

$$\mathbb{Z}_n(1, \theta) = \dot{\mathbb{M}}_n(1, \theta) = 0.$$

To test a change in the value of the parameter, Horváth and Parzen (1994) introduce the test statistic

$$\mathcal{T}_n = n^{-1} \sup_{u \in [0,1]} \mathbb{Z}_n(u, \widehat{\theta}_n)^\top \widehat{I}_n^{-1} \mathbb{Z}_n(u, \widehat{\theta}_n),$$

where $\widehat{I}_n$ is a consistent estimator for the Fisher Information matrix $I(\theta_0)$.

Limit distribution

- From Donsker's theorem

$$u \rightsquigarrow \sqrt{n}\mathbb{Z}_n(u, \theta_0) \quad \text{converges weakly to} \quad u \rightsquigarrow I(\theta_0)^{1/2} B(u)$$

in the Skorohod space $D[0,1]$, where $u \rightsquigarrow B(u)$ is a vector of independent standard Brownian motions.

- It also holds

$$u \rightsquigarrow \sqrt{n}\mathbb{Z}_n(u, \widehat{\theta}_n) \quad \text{converges weakly to} \quad u \rightsquigarrow I(\theta_0)^{1/2} B^{\circ}(u)$$

in $D[0,1]$, where $u \rightsquigarrow B^{\circ}(u)$ is a vector of independent standard Brownian bridges.

- From the continuous mapping theorem

$$\mathcal{T}_n \to^d \sup_{u \in [0,1]} ||B^{\circ}(u)||^2.$$

Beyond the Horváth and Parzen approach

- The same approach of Horváth and Parzen has been applied:
  - Berkes et al. (2004) to GARCH(p, q) models
  - N-N (2012) to change point problem for ergodic diffusion process with continuous observations
- Horváth and Parzen did not discuss the asymptotic behavior of the test statistics under the alternative.
- In N.N (2012) also the consistency of the test under an alternative which has sufficient generality was proved
- The argument for alternatives in N-N (2012) works also for the case of independent data presented in Horváth and Parzen (1994)

The proposed Z-process method

The proposed Z-process method is not just a simple generalization of Fisher-score process method in the case of independent random sequences proposed by Horváth and Parzen for i.i.d model.

- It makes possible to treat new applications in broad spectrum of statistical change point problems including not only models for ergodic dependent data but also non-ergodic cases
- The proofs of main results are based on some asymptotic representations of Z-estimators that are new from the viewpoint of mathematical statistics
- An argument to prove the consistency of the test based on the proposed method under some specified alternatives is developed

Set up of the change point problem

All process $u \rightsquigarrow X(u)$, are assumed to take values in $D[0,1]$, the space of cad lag functions defined on $[0,1]$ equipped with the Skorohod metric.

- Let $\Theta \subset \mathbb{R}^d$ be bounded, open and convex
- Let $u \rightsquigarrow \mathbb{Z}_n(u,\theta)$ be an $\mathbb{R}^d$-valued random process indexed by $\theta \in \Theta \subset \mathbb{R}^d$, defined on a probability space $(\Omega, \mathcal{F}, P)$ that is common for all $n \in \mathbb{N}$
- Let $\dot{\mathbb{Z}}_n(u,\theta)$ denote the $d \times d$ random matrix whose (i,j)-component is $\dot{\mathbb{Z}}_n^{i,j}(u,\theta) = \frac{\partial}{\partial \theta_j}\mathbb{Z}_n^i(u,\theta)$.

We consider the following testing problem:
H_0: the true value $\theta_0 \in \Theta$ does not change during $u \in [0,1]$;
H_1: "not H_0".

Conditions N under H_0 for the limit process

Let be given the *limit* process

$$u \rightsquigarrow Z_{\theta_0}(u, \theta)$$

[N]. Under H_0, the following conditions have to be satisfied.

1. There exists a sequence of diagonal matrices Q_n such that

$$\sup_{\theta \in \Theta} ||Q_n^{-2}\mathbb{Z}_n(1, \theta) - Z_{\theta_0}(1, \theta)|| \to^p 0,$$

2. The limits $Z_{\theta_0}(1, \theta)$'s satisfy $\inf_{\theta : ||\theta - \theta_0|| > \varepsilon} ||Z_{\theta_0}(1, \theta)|| > 0, \quad \forall \varepsilon > 0$, and $||Z_{\theta_0}(1, \theta_0)|| = 0$, almost surely.

Consistency of $\mathbb{Z}$ estimators

The following Lemma prove the consistency of a sequence of Z-estimators. It can be proved exactly in the same way as Theorems 5.7 and 5.9 of van der Vaart (1998),

Lemma

Under [N], for any sequence of Θ-valued random vectors $\widehat{\theta}_n$ such that $||Q_n^{-2}\mathbb{Z}_n(1,\widehat{\theta}_n)|| = o_P(1)$, it holds that $\widehat{\theta}_n \to^p \theta_0$.

The test statistics

- Let $\widehat{\theta}_n$ be any sequence of Θ-valued random vectors such that $||Q_n^{-1}\mathbb{Z}_n(1,\widehat{\theta}_n)|| = o_P(1)$ under H_0
- Let $u \rightsquigarrow V(u,\theta_0)$ be a non-negative definite matrix valued random process such that $V(1,\theta_0)$ is positive definite almost surely, and let $u \rightsquigarrow B(u)$ be a vector of independent standard Brownian motions;
- Let $u \rightsquigarrow \widehat{V}_n(u)$ a uniformly consistent sequence of estimators for the non-negative definite matrix valued random process $u \rightsquigarrow V(u,\theta_0)$

Introduce the test statistic

$$\mathcal{T}_n = \sup_{u\in(0,1]} (Q_n^{-1}\mathbb{Z}_n(u,\widehat{\theta}_n))^\top (u\widehat{V}_n(u)^{-1}) Q_n^{-1}\mathbb{Z}_n(u,\widehat{\theta}_n).$$

Conditions under H_0 for the limit process

Under [N]

(i) There exist a sequence of diagonal matrices R_n and a sequence of matrix valued random processes $u \rightsquigarrow V_n(u,\theta_0)$ such that $V_n(1,\theta_0)$'s are non-singular almost surely and that for any sequence of Θ-valued random vectors $\widetilde{\theta}_n(u)$ indexed by $u \in [0,1]$ satisfying $\sup_{u\in[0,1]} ||\widetilde{\theta}_n(u) - \theta_0|| \to^p 0$,

$$\sup_{u\in[0,1]} ||Q_n^{-1}\dot{\mathbb{Z}}_n(u,\widetilde{\theta}_n(u))R_n^{-1} - (-V_n(u,\theta_0))|| \to^p 0.$$

(ii) It holds that, in $D[0,1]$,

$$(Q_n^{-1}\mathbb{Z}_n(u,\theta_0), V_n(u,\theta_0)) \to^d ((u^{-1}V(u,\theta_0))^{1/2}B(u), V(u,\theta_0)),$$

where $u \rightsquigarrow B(u)$ is a d-dimensional standard Brownian motion, and the value of the first vector of the limit at $u = 0$ should be read as zero.

(iii) It holds that $\sup_{u\in[0,1]} ||\widehat{V}_n(u) - V(u,\theta_0)|| \to^p 0$.

The main result under H_0

Theorem (The main result under H_0)

Under [N] with a sequence of diagonal matrices Q_n in the above set up, assume the conditions (i) – (iii) hold true. Moreover let assume that $u \rightsquigarrow V(u,\theta_0)$ and $u \rightsquigarrow B(u)$ are independent. Then, it holds that

$$\mathcal{T}_n \to^d \sup_{u\in[0,1]} ||B(u) - u^{1/2}V(u,\theta_0)^{1/2}V(1,\theta_0)^{-1/2}B(1)||^2.$$

Remark. If $V(u,\theta_0) = uV(1,\theta_0)$ for every $u \in [0,1]$, then the test is asymptotically distribution free. In this case the limit is reduced to $\sup_{u\in[0,1]} ||B^\circ(u)||^2$ where $u \rightsquigarrow B^\circ(u) = B(u) - uB(1)$ is a vector of independent standard Brownian bridges.

Remarks

- Since $u \rightsquigarrow V(u, \theta_0)$ and $u \rightsquigarrow B(u)$ are independent, then the limit is approximated by

$$\sup_{u \in [0,1]} ||B(u) - u^{1/2} \widehat{V}_n(u)^{1/2} \widehat{V}_n(1)^{-1/2} B(1)||^2$$

  whose approximate distribution can be computed by computer simulations

- In the examples the rate matrices Q_n and R_n are diagonal $\sqrt{n} I_d$, and they are such that $Q_n^{-1} \ddot{\mathbb{M}}_n(u, \theta) R_n^{-1}$ converges to a limit, where $\ddot{\mathbb{M}}_n(u, \theta)$ is a Hessian matrix of the partial sum process $\mathbb{M}_n(u, \theta)$ of a contrast like a log-likelihood function.
- The underlying probability spaces do not have to be common for $n \in \mathbb{N}$ if $V(u, \theta_0)$ appearing in the limit is non-random.

Conditions A under H_1 for the limit process

Under H_1, let be given the *limit* process $u \rightsquigarrow \mathcal{Z}(u, \theta)$ such that:

1. there exists a sequence of diagonal matrices Q_n such that

$$\sup_{u\in[0,1]} \sup_{\theta\in\Theta} ||Q_n^{-2}\mathbb{Z}_n(u,\theta) - \mathcal{Z}(u,\theta)|| \to^p 0,$$

2. There exists a Θ-valued random vector θ_* such that $\forall\varepsilon > 0$

$$\inf_{\theta:||\theta-\theta_*||>\varepsilon} ||\mathcal{Z}(1,\theta)|| > 0,$$

   and $||\mathcal{Z}(1,\theta_*)|| = 0$, almost surely.

3. For the same θ_* it holds that

$$\sup_{u\in(0,1)} ||\mathcal{Z}(u,\theta_*)|| > 0, \quad \text{almost surely.}$$

Remarks (1)

- Assuming conditions (1) and (2) under H_1 is natural. See e.g. Theorems 5.7 and 5.9 of van der Vaart (1998).
- In many cases the alternatives has the form:
  H_1': there exists a constant $u_* \in (0,1)$ such that the true value is $\theta_0 \in \Theta$ for $u \in [0, u_*]$, and $\theta_1 \in \Theta$ for $u \in (u_*, 1]$, where $\theta_0 \neq \theta_1$.
- Condition (1) under H_1' is satisfied with $\mathcal{Z}(u, \theta)$ such that

$$\mathcal{Z}(u_*, \theta) = u_* Z_{\theta_0}(1, \theta)$$

and

$$\mathcal{Z}(1, \theta) = u_* Z_{\theta_0}(1, \theta) + (1 - u_*) Z_{\theta_1}(1, \theta)$$

where $Z_{\theta_1}(1, \theta)$ are also assumed to satisfy (2) under H_0

Remarks (2)

- Condition (3) under H_1' is satisfied observing that

$$\mathcal{Z}(u_*, \theta_*) = u_*(1 - u_*)(Z_{\theta_0}(1, \theta_*) - Z_{\theta_1}(1, \theta_*));$$

  is greater than 0 with probability one.
- If this were zero with positive probability, then it should follow from $\mathcal{Z}(1, \theta_*) = 0$ that $Z_{\theta_0}(1, \theta_*) = Z_{\theta_1}(1, \theta_*) = 0$ with positive probability, and this contradicts (2) under H_0 and the assumption that $\theta_0 \neq \theta_1$
- Hence, it holds, almost surely,

$$\begin{aligned} \sup_{u \in (0,1)} ||\mathcal{Z}(u, \theta_*)|| &\geq ||\mathcal{Z}(u_*, \theta_*)|| \\ &= u_*(1 - u_*)||Z_{\theta_0}(1, \theta_*) - Z_{\theta_1}(1, \theta_*)|| > 0 \end{aligned}$$

- This positive value is closely related to the power of our test under H_1'

Consistency of $\mathbb{Z}$ estimators

The following Lemma prove the consistency of a sequence of Z-estimators under H_1.. It can be also proved exactly in the same way as Theorems 5.7 and 5.9 of van der Vaart (1998).

Lemma

Let conditions [A] under H_1 be satisfied. Then for any sequence of Θ-valued random vectors $\widehat{\theta}_n$ such that $||\mathbb{Z}_n(1, \widehat{\theta}_n)|| = o_P(1)$, it holds that $\widehat{\theta}_n \to^p \theta_*$.

The main result under H_1

Theorem

Under H_1 and related above conditions, it holds for any random point $\check{u}$ in $(0,1)$ that

$$\mathcal{T}_n \geq \lambda(\check{u} Q_n \widehat{V}_n(\check{u})^{-1} Q_n) \left\{ ||\mathcal{Z}(\check{u}, \theta_*)||^2 + o_P(1) \right\},$$

where $\lambda(A)$ denotes the smallest eigenvalue of the random matrix A. Hence, if there exists a random point $\check{u}$ in $(0,1)$ such that

$$||\mathcal{Z}(\check{u}, \theta_*)|| > 0$$

almost surely, and such that

$$\lambda(Q_n(\widehat{V}_n(\check{u})^{-1} Q_n) \to^p \infty$$

then the test is consistent.

Ergodic diffusion process: set up

Let us consider an $I = (l, r)$-valued diffusion process $t \rightsquigarrow X_t$ unique strong solution solution to the stochastic differential equation (SDE)

$$X_t = X_0 + \int_0^t S(X_s; \alpha) ds + \int_0^t \sigma(X_s; \beta) dW_s,$$

where $s \rightsquigarrow W_s$ is a standard Wiener process.

- The parameter is $\theta = (\alpha^\top, \beta^\top)^\top$ where $\alpha \in \Theta_A \subset \mathbb{R}^{d_A}$ and $\beta \in \Theta_B \subset \mathbb{R}^{d_B}$.

Ergodic diffusion process: sample scheme

We observe the process X at discrete time grids $0 = t_0^n < t_1^n < \cdots < t_n^n$, and the asymptotic scheme is $n\Delta_n^2 \to 0$ and $t_n^n \to \infty$ as $n \to \infty$, where

$$\Delta_n = \max_{1 \le k \le n} |t_k^n - t_{k-1}^n|,$$

and

$$\sum_{k=1}^{n} \left| \frac{|t_k^n - t_{k-1}^n|}{t_n^n} - \frac{1}{n} \right| \to 0, \quad \text{as } n \to \infty. \tag{1}$$

Ergodic diffusion process: $\mathbb{Z}$ process

We introduce the $(d_A + d_B)$-dimensional random vectors

$$\mathbb{Z}_n(u,\theta) = \dot{\mathbb{M}}_n(u,\theta)$$

where

$$\mathbb{M}_n(u,\theta) = -\sum_{k:t^n_{k-1}\le ut^n_n} \left\{ \log\sigma(X_{t^n_{k-1}};\beta) + \frac{|X_{t^n_k} - X_{t^n_{k-1}} - S(X_{t^n_{k-1}};\alpha)|t^n_k - t^n_{k-1}||^2}{2\sigma(X_{t^n_{k-1}};\beta)^2|t^n_k - t^n_{k-1}|} \right\}$$

and $R_n = Q_n$ are the diagonal matrix such that $R_n^{(i,i)}$ is $\sqrt{t^n_n}$ for $i = 1, ..., d_A$ and $\sqrt{n}$ for $i = d_A + 1, ..., d$ with $d = d_A + d_B$.

Ergodic diffusion process: $\mathbb{Z}$ process

We re-write with natural notation, the $(d_A + d_B)$-dimensional random vectors

$$\mathbb{Z}_n(u,\theta) = \dot{\mathbb{M}}_n(u,\theta) = (\dot{\mathbb{M}}_n^A(u,\theta)^\top, \dot{\mathbb{M}}_n^B(u,\theta)^\top)^\top,$$

We consider the $(d_A + d_B) \times (d_A + d_B)$-random matrices

$$\dot{\mathbb{Z}}_n(u,\theta) = \ddot{\mathbb{M}}_n(u,\theta) = \begin{pmatrix} \ddot{\mathbb{M}}_n^A(u,\theta) & \ddot{\mathbb{M}}_n^C(u,\theta) \\ \ddot{\mathbb{M}}_n^C(u,\theta)^\top & \ddot{\mathbb{M}}_n^B(u,\theta) \end{pmatrix}.$$

Ergodic diffusion process: the limit processes

- The conditions [N]-(1) and [N]-(2) under H_0 are satisfied with $Z_{\theta_0}(1,\theta) = (Z^A_{\theta_0}(1,\theta)^\top, Z^B_{\theta_0}(1,\theta)^\top)^\top$, given by

$$Z^A_{\theta_0}(1,\theta) = \int_I \frac{\dot{S}(x;\alpha)}{\sigma(x;\beta)}(S(x;\alpha_0) - S(x;\alpha))\mu_{\theta_0}(dx)$$

and

$$Z^B_{\theta_0}(1,\theta) = \int_I \frac{\dot{\sigma}(x;\beta)}{\sigma(x;\beta)^3}(\sigma(x;\beta_0)^2 - \sigma(x;\beta)^2)\mu_{\theta_0}(dx),$$

where μ_{θ_0} denotes the invariant distribution of X when the true value is θ_0.

- Moreover the conditions [A]-(1) and [A]-(2) under H_1' are satisfied with

$$\mathcal{Z}(u,\theta) = (u \wedge u_*)Z_{\theta_0}(1,\theta) + ((u-u_*)\vee 0)Z_{\theta_1}(1,\theta).$$

Ergodic diffusion process: auxiliary results

Under some regularity conditions it can be proved that

- $$\sup_{u\in[0,1]}\left\|\frac{1}{t_n^n}\dot{\mathbb{M}}_n^A(u,\theta_0)-\frac{1}{t_n^n}\sum_{k:t_{k-1}^n\le ut_n^n}\frac{\dot{S}(X_{t_{k-1}^n};\alpha_0)}{\sigma(X_{t_{k-1}^n};\beta_0)}(W_{t_k^n}-W_{t_{k-1}^n})\right\|=o_P((t_n^n)^{-1/2})$$

- $$\sup_{u\in[0,1]}\left\|\frac{1}{n}\dot{\mathbb{M}}_n^B(u,\theta_0)-\frac{1}{n}\sum_{k:t_{k-1}^n\le ut_n^n}\frac{\dot{\sigma}(X_{t_{k-1}^n};\beta_0)}{\sigma(X_{t_{k-1}^n};\beta_0)}\left\{\frac{|W_{t_k^n}-W_{t_{k-1}^n}|^2}{|t_k^n-t_{k-1}^n|}-1\right\}\right\|=o_P(n^{-1/2})$$

- $$\sup_{u\in[0,1]}\sup_{\theta\in\Theta}\left\|\frac{1}{t_n^n}\ddot{\mathbb{M}}_n^A(u,\theta)-\frac{1}{t_n^n}\sum_{k:t_{k-1}^n\le ut_n^n}H^A(X_{t_{k-1}^n};\theta_0,\theta)|t_k^n-t_{k-1}^n|\right\|=o_P((t_n^n)^{-1/2})$$

- $$\sup_{u\in[0,1]}\sup_{\theta\in\Theta}\left\|\frac{1}{n}\ddot{\mathbb{M}}_n^B(u,\theta)-\frac{1}{n}\sum_{k:t_{k-1}^n\le ut_n^n}H^B(X_{t_{k-1}^n};\theta_0,\theta)\right\|=o_P(n^{-1/2})$$

- $$\sup_{u\in[0,1]}\sup_{\theta\in\Theta}\left\|\frac{1}{\sqrt{t_n^n n}}\ddot{\mathbb{M}}_n^C(u,\theta)\right\|=o_P(n^{-1/4})$$

Ergodic diffusion process: auxiliary results

Where

$$
\begin{aligned}
H^A(x;\theta_0,\theta) &= \frac{\ddot{S}(x;\alpha)(S(x;\alpha_0)-S(x;\alpha))-\dot{S}(x;\alpha)\dot{S}(x;\alpha)^\top}{\sigma(x;\beta)^2}, \\
H^B(x;\theta_0,\theta) &= \left\{\frac{\ddot{\sigma}(x;\beta)}{\sigma(x;\beta)^3}-3\frac{\dot{\sigma}(x;\beta)\dot{\sigma}(x;\beta)^\top}{\sigma(x;\beta)^4}\right\}\left(\sigma(x;\beta_0)^2-\sigma(x;\beta)^2\right) \\
&\quad -2\frac{\dot{\sigma}(x;\beta)\dot{\sigma}(x;\beta)^\top}{\sigma(x;\beta)^2}.
\end{aligned}
$$

Ergodic diffusion process: conditions under H_0

Using these facts and the usual martingale central limit theorem, we can see that the conditions (i) and (ii) in the main theorem under H_0 hold for

$$V_n(u,\theta_0) = \begin{pmatrix} V_n^A(u,\theta_0) & 0 \\ 0 & V_n^B(u,\theta_0) \end{pmatrix},$$

where

$$\begin{aligned} V_n^A(u,\theta_0) &= \frac{1}{t_n^n} \sum_{k:t_{k-1}^n \le u t_n^n} \frac{\dot{S}(X_{t_{k-1}^n};\alpha_0)\dot{S}(X_{t_{k-1}^n};\alpha_0)^\top}{\sigma(X_{t_{k-1}^n};\beta_0)^2} |t_k^n - t_{k-1}^n|, \\ V_n^B(u,\theta_0) &= \frac{2}{n} \sum_{k:t_{k-1}^n \le u t_n^n} \frac{\dot{\sigma}(X_{t_{k-1}^n};\beta_0)\dot{\sigma}(X_{t_{k-1}^n};\beta_0)^\top}{\sigma(X_{t_{k-1}^n};\beta_0)^2}. \end{aligned}$$

Ergodic diffusion process: limit process

The limit of $V_n(u, \theta_0)$ is $V(u, \theta_0) = uI_{\theta_0}(\theta_0)$, where

$$I_{\theta_0}(\theta) = \begin{pmatrix} I^A_{\theta_0}(\theta) & 0 \\ 0 & I^B_{\theta_0}(\theta) \end{pmatrix}$$

with

$$I^A_{\theta_0}(\theta) = \int_I \frac{\dot{S}(x;\alpha)\dot{S}(x;\alpha)^\top}{\sigma(x;\beta)^2} \mu_{\theta_0}(dx),$$

and

$$I^B_{\theta_0}(\theta) = 2\int_I \frac{\dot{\sigma}(x;\beta)\dot{\sigma}(x;\beta)^\top}{\sigma(x;\beta)^2} \mu_{\theta_0}(dx).$$

We suppose that $I_{\theta_0}(\theta)$'s are positive definite.

Ergodic diffusion process: consistent estimator

Condition (iii) of the main theorem under H_0 is satisfied if as consistent estimator $\widehat{V}_n(u)$ for $V(u,\theta_0)$, we introduce

$$\widehat{V}_n(u) = \begin{pmatrix} u\widehat{I}_n^A & 0 \\ 0 & u\widehat{I}_n^B \end{pmatrix},$$

where

$$\widehat{I}_n^A = \frac{1}{n}\sum_{k=1}^n \frac{\dot{S}(X_{t_{k-1}^n};\widehat{\alpha}_n)\dot{S}(X_{t_{k-1}^n};\widehat{\alpha}_n)^\top}{\sigma(X_{t_{k-1}^n};\widehat{\beta}_n)^2},$$

and

$$\widehat{I}_n^B = \frac{2}{n}\sum_{k=1}^n \frac{\dot{\sigma}(X_{t_{k-1}^n};\widehat{\beta}_n)\dot{\sigma}(X_{t_{k-1}^n};\widehat{\beta}_n)^\top}{\sigma(X_{t_{k-1}^n};\widehat{\beta}_n)^2}.$$

Ergodic diffusion process: test's consistency

Condition [A]-(3) under H_1' is automatically satisfied as soon as the natural conditions [N]-(2) and [A]-(2) are satisfied.
Under H_1', since the matrix

$$u_* I_{\theta_0}(\theta_*) + (1 - u_*) I_{\theta_1}(\theta_*)$$

is positive definite, and

$$\widehat{V}_n(u_*) \to u_* I_{\theta_0}(\theta_*) + (1 - u_*) I_{\theta_1}(\theta_*)$$

then

$$\lambda(Q_n \widehat{V}_n(u_*) Q_n) \to^p \infty$$

Hence the test is consistent.

Numerical study: Ornstein-Uhlenbeck process

We consider the Ornstein-Uhlenbeck process starting from $x_0 = 0$ for the true (data-generating) process:

$$X_t = x_0 - \int_0^t \alpha X_s ds + \beta W_t, \quad t \in [0, T].$$

- We treat the equidistant sampling case, that is, $\Delta_n = |t_k^n - t_{k-1}^n|$ for every $k = 1, ..., n$.
- We observe the trajectory of the process for different time horizons $t_n^n = T$
- The number n of observations for each trajectory is such that $t_n^n = n^{1/3}$, so $\Delta_n = n^{-2/3}$.
- We simulate $M = 10^4$ trajectories

The limit distribution of test statistics

- For any fixed level $\varepsilon > 0$ the critical value c_ε is given by

$$P\left(\sup_{u\in[0,1]} \sum_{i=1}^{d_A+d_B} |B^{\circ,(i)}(u)|^2 > c_\varepsilon\right) = \varepsilon.$$

- Table 1 of Lee *et al.* (2003) gives a table of the critical values for the significance levels $\varepsilon = 0.01, 0.05, 0.10$ and for different values of the dimension $d = d_A + d_B$ computed by Monte Carlo simulation for the limit distribution.
- Throughout we take the significance level to be $\varepsilon = 0.05$. For two parameters ($d = 2$) the critical value is $c_\varepsilon = 2.408$.

H_0 true

Empirical size for different time horizons.

parameter		T	5	10	15	20	25
α	β	n	125	1000	3375	8000	15625
1	1		0.044	0.054	0.050	0.052	0.053
0.25	0.02		0.047	0.061	0.058	0.064	0.054
1	8		0.37	0.049	0.052	0.051	0.052

The empirical size gains along with increasing terminal time $T = t_n^n$, attaining at 0.05, but also for small terminal T. In the second example, the values of the parameter are the maximum likelihood estimate for the mostly federal funds data 1963-1998 in Aït-Sahalia (1999).

H_0 not true

Regarding the alternative hypothesis we study the behavior of the test statistic in three different situations and for different change point $u_* T$ of the parameters, as follows:

- The drift coefficient changes from α_0 to α_1, but the diffusion coefficient does not change.
- The drift coefficient does not change, but the diffusion coefficient changes from β_0 to β_1.
- Both coefficients change.

For each of the above scenarios we consider the following change points, $u_* = \frac{1}{2}, \frac{3}{4}, \frac{9}{10}$.

H_0 not true: 1st scenario

parameter			T	5	10	15	20	25
α_0	α_1	β	n	125	1000	3375	8000	15625
0.25	0.50	0.02	$u_* = \frac{1}{2}$	0.31	0.52	0.73	0.79	0.88
0.25	1.25	0.02		0.35	0.60	0.78	0.88	0.94
1	2	1		0.10	0.13	0.16	0.20	0.54
1	4	1		0.27	0.45	0.64	0.78	1.00
0.25	0.50	0.02	$u_* = \frac{3}{4}$	0.12	0.17	0.23	0.26	0.35
0.25	1.25	0.02		0.13	0.20	0.28	0.31	0.38
1	2	1		0.07	0.08	0.08	0.09	0.19
1	4	1		0.10	0.14	0.16	0.19	0.63
0.25	0.50	0.02	$u_* = \frac{9}{10}$	0.05	0.07	0.08	0.08	0.09
0.25	1.25	0.02		0.06	0.08	0.09	0.11	0.11
1	2	1		0.05	0.05	0.06	0.06	0.07
1	4	1		0.06	0.07	0.07	0.08	0.12

H_0 not true: 2nd scenario

parameter			T	5	10	15	20	25
α	β_0	β_1	n	125	1000	3375	8000	15625
0.25	0.020	0.025	$u_* = \frac{1}{2}$	0.87	1	1	1	1
0.25	0.020	0.030		0.99	1	1	1	1
1	1	0.95		0.04	0.11	0.35	0.74	0.97
1	1	0.70		0.47	1	1	1	1
0.25	0.020	0.025	$u_* = \frac{3}{4}$	0.52	0.99	1	1	1
0.25	0.020	0.030		0.86	1	1	1	1
1	1	0.95		0.02	0.06	0.17	0.43	0.78
1	1	0.70		0.11	0.99	1	1	1
0.25	0.020	0.025	$u_* = \frac{9}{10}$	0.14	0.62	0.99	1	1
0.25	0.020	0.030		0.36	0.99	1	1	1
1	1	0.95		0.03	0.05	0.06	0.09	0.16
1	1	0.70		0.04	0.23	0.99	1	1

Numerical study: Cox-Ingersoll-Ross process

In the same framework of Ornstein-Uhlenbeck process, let us consider the Cox-Ingersoll-Ross (CIR) process

$$X_t = x_0 + \int_0^t (\alpha - X_s) ds + \int_0^t \beta \sqrt{X_s} dW_s, \quad t \in [0, T].$$

- This model is widely used in mathematical financial modeling for return asset prices.
- To have stationarity the model parameters have to satisfy $2\alpha > \beta^2$.

Numerical study: the results under H_0

Cox-Ingersoll-Ross process: empirical size based on $M = 10^4$ independent statistics, for different time horizons.

T	5	10	15	20	25
n	125	1000	3375	8000	15625
$\alpha_0 = 1,\ \beta_0 = 1$	0.175	0.136	0.113	0.093	0.053
$\alpha_0 = 0.25,\ \beta_0 = 0.02$	0.134	0.126	0.126	0.128	0.118

- For small T the performance of the test on this model is worse than for the OU model.
- A simulation for increasing $T = 75$ with $M = 1000$ independent statistics show that we have convergence to the theoretical value 0.05.
- For financial data $T = 75$ may be too long. For small T the fact that the significance level could be greater than what is expected asymptotically should be taken in account.

Volatility of diffusion process: set up

Let us consider an $I = (l, r)$-valued diffusion process $t \rightsquigarrow X_t$ unique strong solution solution to the SDE

$$X_t = X_0 + \int_0^t S(X_s)ds + \int_0^t \sigma(X_s;\theta)dW_s,$$

where $s \rightsquigarrow W_s$ is a standard Wiener process.

- The drift coefficient $S(\cdot)$ is treated as an unknown nuisance function
- The parameter is $\theta \in \Theta \subset \mathbb{R}^d$.

Volatility of diffusion process: sample scheme

We observe the process X at discrete time grids :

$$0 = t_0^n < t_1^n < \cdots < t_n^n = T < \infty$$

and the asymptotic scheme is

$$\sum_{k=1}^{n} \left| \frac{|t_k^n - t_{k-1}^n|}{t_n^n} - \frac{1}{n} \right| \to 0, \quad \text{as } n \to \infty.$$

Volatility of diffusion process: $\mathbb{Z}$ process

We introduce

$$\mathbb{Z}_n(u,\theta) = \dot{\mathbb{M}}_n(u,\theta)$$

where

$$\mathbb{M}_n(u,\theta) = -\sum_{k:t^n_{k-1}\le ut^n_n} \left\{ \log \sigma(X_{t^n_{k-1}};\theta) + \frac{|X_{t^n_k} - X_{t^n_{k-1}}|^2}{2\sigma(X_{t^n_{k-1}};\theta)^2 |t^n_k - t^n_{k-1}|} \right\}.$$

The rate matrices are given by $R_n = Q_n = \sqrt{n} I_d$.
An interesting point of this example is that the limit of $-\ddot{\mathbb{M}}_n(u,\widetilde{\theta}_n(u))$ is random and depend on $u \in [0,1]$ in a complex way.

Volatility of diffusion process: limit process

Under some standard conditions on the parametric family for the diffusion coefficient, we can show that the conditions [N]-(1) and [N]-(2) under H_0 are satisfied with

$$Z_{\theta_0}(1,\theta) = \int_0^T \frac{\dot{\sigma}(X_t;\theta)}{\sigma(X_t;\theta)^3}(\sigma(X_t;\theta_0)^2 - \sigma(X_t;\theta)^2)dt$$

Volatility of diffusion process: auxiliary results

Under some regularity conditions, it can be proved that

$$\sup_{u\in[0,1]}\left\|\dot{\mathbb{M}}_n(u,\theta_0) - \frac{1}{n}\sum_{k:t_{k-1}^n\le ut_n^n} \frac{\dot\sigma(X_{t_{k-1}^n};\theta_0)}{\sigma(X_{t_{k-1}^n};\theta_0)}\left\{\frac{|\Delta W_k^n|^2}{\Delta t_k^n}-1\right\}\right\| = o_P(n^{-1/2}),$$

$$\sup_{u\in[0,1]}\sup_{\theta\in\Theta}\left\|\ddot{\mathbb{M}}_n(u,\theta) - \frac{1}{n}\sum_{k:t_{k-1}^n\le ut_n^n} H(X_{t_{k-1}^n};\theta_0,\theta)\right\| = o_P(n^{-1/2}),$$

where

$$\begin{aligned} H(x;\theta_0,\theta) &= \left\{\frac{\ddot\sigma(x;\theta)}{\sigma(x;\theta)^3} - 3\frac{\dot\sigma(x;\theta)\dot\sigma(x;\theta)^\top}{\sigma(x;\theta)^4}\right\}(\sigma(x;\theta_0)^2-\sigma(x;\theta)^2) \\ &\quad -2\frac{\dot\sigma(x;\theta)\dot\sigma(x;\theta)^\top}{\sigma(x;\theta)^2} \end{aligned}$$

Volatility of diffusion process: conditions under H_0

To check (i) of main theorem under H_0 , observe that it holds that for any sequence of random vectors $\widetilde{\theta}_n(u)$ indexed by $u \in [0,1]$ such that $\sup_{u\in[0,1]} ||\widetilde{\theta}_n(u) - \theta_0|| \to^p 0$,

$$\sup_{u\in[0,1]} \left|\left| \frac{1}{n}\ddot{\mathbb{M}}_n(u, \widetilde{\theta}_n(u)) - (-V_n(u,\theta_0)) \right|\right| \to^p 0,$$

where

$$V_n(u,\theta_0) = \frac{2}{n} \sum_{k:t^n_{k-1}\le ut^n_n} \frac{\dot{\sigma}(X_{t^n_{k-1}};\theta_0)\dot{\sigma}(X_{t^n_{k-1}};\theta_0)^\top}{\sigma(X_{t^n_{k-1}};\theta_0)^2}, \quad \forall u \in [0,1]$$

Volatility of diffusion process: conditions under H_0

Condition (ii) of main theorem under H_0 follows from the well known theory of martingales. Indeed in $D[0,1]$

$$(\sqrt{n}\dot{\mathbb{M}}_n(u,\theta_0), V_n(u,\theta_0)) \to^d ((u^{-1}V(u,\theta_0))^{1/2}B(u), V(u,\theta_0))$$

where $u \rightsquigarrow B(u)$ is a vector of independent standard Brownian motions which is independent of the matrix valued random process $u \rightsquigarrow V(u,\theta_0)$ given by

$$V(u,\theta_0) = 2\int_0^{uT} \frac{\dot{\sigma}(X_s;\theta_0)\dot{\sigma}(X_s;\theta_0)^\top}{\sigma(X_s;\theta_0)^2} ds, \quad \forall u \in [0,1]$$

Volatility of diffusion process: conditions under H_0

To prove assumption (iii) in the main theorem under H_0 define the matrices

$$V(u,\theta) = 2\int_0^{uT} \frac{\dot\sigma(X_t;\theta)\dot\sigma(X_t;\theta)^\top}{\sigma(X_t;\theta)^2}dt, \quad \forall u \in [0,1]$$

A consistent estimator $\widehat{V}_n(u)$ for $V(u,\theta_0)$ is given by

$$\widehat{V}_n(u) = \frac{2}{n}\sum_{k:t^n_{k-1}\le ut^n_n} \frac{\dot\sigma(X_{t^n_{k-1}};\widehat{\theta}_n)\dot\sigma(X_{t^n_{k-1}};\widehat{\theta}_n)^\top}{\sigma(X_{t^n_{k-1}};\widehat{\theta}_n)^2}, \quad \forall u \in [0,1]$$

Volatility of diffusion process: numerical results

The data-generating process is the following:

$$X_t = 4 - \int_0^t (X_s - 4)ds + \int_0^t \exp\left(\theta \frac{X_s^2}{1+X_s^2}\right) dW_s, \quad t \in [0,1],$$

- The sample scheme is the equidistant time grid $t_k^n = \frac{k}{n}$, $k = 0, 1, \ldots, n$.
- $M = 10^4$ trajectories are simulated
- The test statistic can be computed, estimating

$$V(u, \theta_0) = 2 \int_0^u \left| \frac{X_s^2}{1+X_s^2} \right|^2 ds$$

by the natural estimator

$$\widehat{V}_n(u) = \frac{2}{n} \sum_{k=1}^{[un]} \left| \frac{X_{t_{k-1}^n}^2}{1+X_{t_{k-1}^n}^2} \right|^2 .$$

H_0 true

Empirical size for different n. The critical value is $c_\varepsilon = 1.820$. The true value of the parameter is set as $\theta_0 = 1.0$ or 1.5.

n	20	40	100	200
$\theta_0 = 1.0$	0.013	0.024	0.031	0.034
$\theta_0 = 1.5$	0.028	0.029	0.035	0.036

Simulations for the model with $\theta_0 = 1.0$ seem to confirm the theoretical result: with $M = 1000$ simulated trajectories, for $n = 1000$ and $n = 5000$ the empirical size is 0.039 and 0.049 respectively.

H_0 not true

The empirical power under H_1'.
The true values of the parameter change from $\theta_0 = 1.0$ to $\theta_1 = 1.5$ at time point $u_* = \frac{1}{2}, \frac{3}{4}$ or $\frac{9}{10}$.

n	20	40	100	200
$u_* = \frac{1}{2}$	0.288	0.793	0.974	0.994
$u_* = \frac{3}{4}$	0.457	0.768	0.951	0.979
$u_* = \frac{9}{10}$	0.251	0.466	0.805	0.935

Thank you for your attention !!

References

1. Horváth, L. and Parzen, E. (1994). Limit theorems for Fisher-score change processes. IMS Lecture Notes
2. N & N (2012). Asymptotically distribution free test for parameter change in a diffusion process model. *Ann. Inst. Statist. Math.*
3. N & N (2016). Z-process method for change point problems with applications to discretely observed diffusion processes. *Stat. Methods Appl.*