Detecting Smooth Changes in Locally Stationary Processes

Michael Vogt
University of Konstanz

Holger Dette
Ruhr-Universität Bochum

Workshop on
“Inference for Change-Point and Related Processes”

Newton Institute, Cambridge
January 14, 2014
Introduction

- In a wide range of time series applications, the stochastic properties of the data change over time.
- It is often natural to assume that the properties change gradually rather than abruptly.
- In a number of applications, the properties can be expected to be constant for some time before they start to vary.
- In such situations, it is often of interest to locate the time point where the properties start to vary.
Introduction

Example: \( X_{t,T} = \mu\left(\frac{t}{T}\right) + \varepsilon_t \).

Aim: Estimate the region \([0, u_0]\) where the function \( \mu \) is constant, i.e., estimate the smooth change point \( u_0 \).
Example: $X_{t,T} = \sigma^2 \left( \frac{t}{T} \right) \epsilon_t$.

Aim: Estimate the region $[u_0, 1]$ where the function $\sigma^2$ is constant.
Model Setting

- Consider a locally stationary process \( \{X_t, T\} \) which is univariate for simplicity.
- Let \( \lambda_{t,T} \) be some time-varying feature of the process, e.g.,
  \[
  \lambda_{t,T} = \mathbb{E}[X_{t,T}]
  \]
  \[
  \lambda_{t,T} = \mathbb{E}[X_{t,T}^p] \text{ for some } p
  \]
  \[
  \lambda_{t,T} = (\gamma_{\ell,t,T})_{\ell=1}^p \text{ with } \gamma_{\ell,t,T} = \text{Cov}(X_{t,T}, X_{t-\ell,T})
  \]
  \[
  \lambda_{t,T} = F_{t,T}(\cdot) = \mathbb{P}(X_{t,T} \leq \cdot)
  \]
- We assume that \( \lambda_{t,T} \) is constant on the rescaled time interval \([u_0, 1]\) and varies smoothly prior to \(u_0\).
- We want to construct a statistical procedure to estimate the time point \(u_0\).
Model Setting

We consider a generic feature $\lambda_{t,T}$ of the process $\{X_{t,T}\}$ with the following property:

**(P)** $\lambda_{t,T}$ is fully characterized by the set of moments 
$\{\mathbb{E}[f(X_{t,T})]: f \in \mathcal{F}\}$, where $\mathcal{F}$ is a family of measurable functions $f$.

Examples:

1. $\lambda_{t,T} = \mathbb{E}[X_{t,T}] = \mathbb{E}[\text{id}(X_{t,T})] \rightarrow \mathcal{F} = \{\text{id}\}$
2. $\lambda_{t,T} = \mathbb{E}[X_{t,T}^p] \rightarrow \mathcal{F} = \{f\}$ with $f(x) = x^p$
3. $\lambda_{t,T} = F_{t,T}(\cdot) = \mathbb{P}(X_{t,T} \leq \cdot) \rightarrow \mathcal{F} = \{1(\cdot \leq x) : x \in \mathbb{R}\}$
   
   $= \mathbb{E}[\mathbb{1}(X_{t,T} \leq \cdot)]$
Measures of Time-Variation

In order to estimate $u_0$, we first construct a measure of time-variation.

**Definition:** A function $D : [0, 1] \rightarrow \mathbb{R}_{\geq 0}$ is called a *measure of time-variation* if it has the following property:

$$D(u) \begin{cases} = 0 & \text{if } \lambda_{t,T} \text{ does not vary on } [u, 1] \\ > 0 & \text{if } \lambda_{t,T} \text{ varies on } [u, 1]. \end{cases}$$

This implies that

$$D(u) \begin{cases} = 0 & \text{for } u \geq u_0 \\ > 0 & \text{for } u < u_0. \end{cases}$$
Measures of Time-Variation

- For each $f \in \mathcal{F}$, write $\mathbb{E}[f(X_{t,T})] = \mu_{t,T,f}$.

- For simplicity of exposition, assume that $\mu_{t,T,f} = \mu_f(t/T)$ for some function $\mu_f$.

- By (P), $\lambda_{t,T}$ is fully characterized by the set of moments $\{\mu_{t,T,f} : f \in \mathcal{F}\}$.

- Hence, time-variation in $\lambda_{t,T}$ is equivalent to time-variation in $\mu_{t,T,f}$ for some $f \in \mathcal{F}$.

- Put differently, time-variation in $\lambda_{t,T}$ is equivalent to variation of $\mu_f(t/T)$ in the time argument $t/T$ for some $f$. 
Measures of Time-Variation

If $\mu_f$ is constant on $[u, 1]$, then for any $v \in [u, 1]$,

$$\frac{1}{1 - v} \int_v^1 \mu_f(w)dw = \frac{1}{1 - u} \int_u^1 \mu_f(w)dw.$$ 

Hence, if $\mu_f$ is constant on $[u, 1]$, then for any $v \in [u, 1]$,

$$D(u, v, f) = \int_v^1 \mu_f(w)dw - \left(\frac{1 - v}{1 - u}\right) \int_u^1 \mu_f(w)dw = 0.$$ 

The measure of time-variation may now be defined as

$$\mathcal{D}(u) = \sup_{f \in \mathcal{F}, v \in [u, 1]} |D(u, v, f)|.$$
Measures of Time-Variation

- It holds that
  \[ \int_{\nu}^{1} \mu_f(w)dw \approx \frac{1}{T} \sum_{t=\lceil \nu T \rceil}^{T} \mu_f \left( \frac{t}{T} \right) = \frac{1}{T} \sum_{t=\lceil \nu T \rceil}^{T} \mathbb{E}[f(X_{t,T})] \approx \frac{1}{T} \sum_{t=\lceil \nu T \rceil}^{T} f(X_{t,T}). \]

- Estimate \( D(u, \nu, f) \) by
  \[
  \hat{D}_T(u, \nu, f) = \frac{1}{T} \sum_{t=\lceil \nu T \rceil}^{T} f(X_{t,T}) - \left( \frac{1 - \nu}{1 - u} \right) \frac{1}{T} \sum_{t=\lceil u T \rceil}^{T} f(X_{t,T}).
  \]

- Finally, define the estimator of \( D(u) \) by
  \[
  \hat{D}_T(u) = \sup_{f \in \mathcal{F}, \nu \in [u,1]} |\hat{D}_T(u, \nu, f)|.
  \]
Estimation Method

Our estimation method is based on the observation that

$$\sqrt{T}\cap(u)\begin{cases} = 0 & \text{for } u \geq u_0 \\ \rightarrow \infty & \text{for } u < u_0 \end{cases}$$

as $T \rightarrow \infty$. As $\hat{\cap}_T$ estimates $\cap$, its scaled version $\sqrt{T}\hat{\cap}_T$ should behave similarly. Indeed,

$$\sqrt{T}\hat{\cap}_T(u)\begin{cases} = O_p(1) & \text{for } u \geq u_0 \\ \overset{P}{\rightarrow} \infty & \text{for } u < u_0. \end{cases}$$

The main idea of our method is to exploit this dichotomous behaviour of the process $\sqrt{T}\hat{\cap}_T$. 
Estimation Method

**Step 1:** We first transform the statistic $\sqrt{T} \hat{D}_T$ to behave approximately like a function that has a jump at $u_0$. To do so, define

$$\hat{q}_T(u) = \Phi(\rho_T \sqrt{T} \hat{D}_T(u)),$$

where $\Phi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is strictly increasing with $\Phi(0) = 0$ and $\lim_{x \rightarrow \infty} \Phi(x) = 1$ and $\{\rho_T\}$ is a sequence of positive constants which slowly converges to zero. This yields

$$\hat{q}_T(u) \xrightarrow{P} \begin{cases} 
0 & \text{for } u \geq u_0 \\
1 & \text{for } u < u_0,
\end{cases}$$

i.e., $\hat{q}_T$ behaves approximately like the step function $1(\cdot < u_0)$. 
**Estimation Method**

**Step 2:** We now use $\hat{q}_T$ to construct a criterion function which is minimized approximately at $u_0$. To do so, we define

$$\hat{Q}_T(u) = u + (1 - u)\hat{q}_T(u)$$

and use

$$\hat{u}_0 := \arg\min_{u \in [0,1]} \hat{Q}_T(u)$$

as an estimator of $u_0$. 

![Diagram showing the estimation method](attachment:image.png)
Step 2: We now use $\hat{q}_T$ to construct a criterion function which is minimized approximately at $u_0$. To do so, we define

$$\hat{Q}_T(u) = u + (1 - u)\hat{q}_T(u)$$

and use

$$\hat{u}_0 := \arg\min_{u \in [0,1]} \hat{Q}_T(u)$$

as an estimator of $u_0$. 

![Diagram showing the function $\hat{Q}_T(u)$ and the estimation process.](image-url)
Step 2: We now use $\hat{q}_T$ to construct a criterion function which is minimized approximately at $u_0$. To do so, we define

$$\hat{Q}_T(u) = u + (1 - u)\hat{q}_T(u)$$

and use

$$\hat{u}_0 := \arg\min_{u \in [0,1]} \hat{Q}_T(u)$$

as an estimator of $u_0$. 
Step 2: We now use $\hat{q}_T$ to construct a criterion function which is minimized approximately at $u_0$. To do so, we define

$$\hat{Q}_T(u) = u + (1 - u)\hat{q}_T(u)$$

and use

$$\hat{u}_0 := \arg\min_{u \in [0,1]} \hat{Q}_T(u)$$

as an estimator of $u_0$. 
Alternative Estimation Approaches

The problem of estimating $u_0$ may be approached by alternative methods:

- In a time-varying mean setting, one can regard $u_0$ as a break point in a higher-order derivative and use change point methods; cp. for example Müller (1992).

- One may try to apply sequential testing methods; cp. Mercurio & Spokoiny (2004) who consider a time-varying volatility model.

- In a time-varying mean setting, a $p$-value based procedure has been proposed by Mallik et al. (2011) and Mallik et al. (2013).
Asymptotic Results

Assumptions:

(A1) \( \{X_{t,T}\} \) is locally stationary.

(A2) \( \{X_{t,T}\} \) is strongly mixing with a sufficiently fast rate.

(A3) The variables \( X_{t,T} \) fulfill appropriate moment conditions.

(A4) The family \( \mathcal{F} \) is not too complex.

(A5) \( D(u) \) is as smooth as a polynomial of order \( k > 0 \) at \( u_0 \), that is,

\[
\frac{D(u)}{(u_0 - u)^k} \to c_k > 0 \quad \text{as } u \to u_0
\]

for some constant \( c_k > 0 \).
Asymptotic Results

Theorem (Convergence rate of $\hat{u}_0$): Under the above conditions,

$$\hat{u}_0 - u_0 = O_p(\gamma_T),$$

where $\gamma_T = \max\{\rho_T, (\rho_T \sqrt{T})^{-1/k}\}$.

Remarks:

- The rate $\gamma_T$ depends on the degree of smoothness $k$. In particular, the larger $k$, the slower the rate.
- $\gamma_T$ also depends on the shrinkage parameter $\rho_T$.
- The “optimal” rate is $\gamma_T = T^{-\frac{1}{2(k+1)}}$ and obtained by setting $\rho_T = T^{-\frac{1}{2(k+1)}}$. 
Simulations

We consider the time-varying mean model

$$X_{t,T} = \mu\left(\frac{t}{T}\right) + \varepsilon_t$$

with AR errors $\varepsilon_t = 0.25\varepsilon_{t-1} + \eta_t$ and $\eta_t \sim \text{i.i.d. } N(0, 0.25)$. 

![Graphs showing \(\mu_1(u)\) and \(\mu_2(u)\)]
Simulations

Figure: A typical sample path of length 500 drawn from the design with $\mu_1$. 
Simulations

Figure: Simulation results for the design with $\mu_1$. 
Simulations

Figure: A typical sample path of length 500 drawn from the design with $\mu_2$. 
Simulations

Figure: Simulation results for the design with $\mu_2$. 

- For $T = 500$, the distribution of $\hat{u}_0$ is approximately normal with a peak around 0.5.
- For $T = 1000$, the distribution of $\hat{u}_0$ is also approximately normal, with a peak around 0.5.

Number of simulations: 0, 500, 1000, 1500.
Applications

![Graph of temperature anomalies](chart1)

- Temperature anomalies from 1850 to 2000
- Y-axis range from -0.4 to 0.4
- X-axis ranges from 1850 to 2000

![Graph of S&P 500 returns](chart2)

- S&P 500 returns from 2011 to 2013
- Y-axis range from -0.06 to 0.02
- X-axis ranges from 2011 to 2013
Summary

- We have investigated the question how to estimate a smooth change point in a locally stationary setup.
- We have derived the asymptotic properties of our estimator, in particular its convergence rate, and have studied its small sample performance in a simulation study.
- Future work: Our measure of time-variation may be used to construct tests for time-invariance/stationarity.
Literature

Fitting Time Series Models to Nonstationary Processes.
*Annals of Statistics.*

Threshold estimation based on a p-value framework in dose-response and regression settings.
*Biometrika.*

Asymptotics for p-value based threshold estimation in regression settings.
*Electronic Journal of Statistics.*

Statistical inference for time-inhomogeneous volatility models.
*Annals of Statistics.*

Change-points in nonparametric regression analysis.
*Annals of Statistics.*