Change-point problems in nonlinear regression estimation with dependent observations

Ken-ichi Yoshihara iYokohama National University)
Shuya Kanagawa (Musashi Institute of Technology)

November 18, 2008

Let $\{X_i,\ i \geq 1\}$ and $\{Y_i,\ i \geq 1\}$ be stationary sequences of random variables observed from SDE's driven by processes with stationary increments, respectively. Assume that $\{Y_i\}$ can be represented by

$$Y_i = m\left(X_i\right) + \sigma\left(X_i\right)\varepsilon_i,$$

where $\{\varepsilon_i\}$ is a sequence of i.i.d. random variables which are independent of $\{X_i\}$. The regression function $m\left(\cdot\right)$ is smooth except for at τ defined by

$$m\left(x\right) = m_0\left(x\right) + \gamma I_{[\tau,1]}\left(x\right)$$

where $\gamma > 0$, $m_0 \in C^2$ and τ is called a change-point of the regression function. We consider an estimation of the change-point τ for the regression function $m\left(\cdot\right)$ in the case when the sequence of random variables $\{\left(X_i, Y_i\right),\ i \geq 1\}$ satisfies the strong mixing condition from the view points of time series analysis for mathematical finance.

Let $\{(X_i, Y_i)\}$ be a strictly stationary sequence of random vectors satisfying the strong mixing condition. As an estimator of $m(\cdot)$, we consider

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} w_i(x)Y_i}{\sum_{i=1}^{n} w_i(x)}$$

where

$$\begin{aligned} w_i(x) &= K\Big(\frac{x - X_i}{h_n}\Big)(S_2(x) - (x - X_i)S_1(x)), \quad i = 1, \cdots, n, \\ S_l(x) &= \sum_{i=1}^{n}(x - X_i)^l K\Big(\frac{x - X_i}{h_n}\Big), \qquad l = 0, 1, 2, \end{aligned}$$

K is a kernel and $\{h_n\}$ is a sequence of positive numbers such that $h_n \downarrow 0$ and $nh_n \to \infty$.

Define the estimate of τ as a value of t that maximizes $\hat{\gamma}(t)$ over $\kappa = [h_n, 1-h_n]$:

$$\hat{\tau} = \inf\{t \in \kappa; \hat{\gamma}(t) = \sup_{x\in\kappa} \hat{\gamma}(x)\}. \tag{1}$$

We exclude right and left edges of κ since 0 and 1 are themselves discontinuous points.

To investigate the asymptotic behaviour of $\hat{\tau}$ define

$$\mathcal{L}_n(z) = \alpha(n, h_n)\Big\{\hat{\gamma}\Big(\tau + \frac{h_n}{\beta(n, h_n)} z\Big) - \hat{\gamma}(\tau)\Big\}, \quad z \in [-M, M].$$

For M large enough, we have

$$\hat{\tau} = \arg \sup_{z\in[-M,M]} \mathcal{L}_n(z). \tag{2}$$

Let $p = p(n)$, $q = q(n)$ and $k = k(n)$ be integer valued functions of n such that $p(n) \uparrow \infty$, $q(n) = o(p(n)) \uparrow \infty$, and $k(n) = [n/(p(n) + q(n))] \uparrow \infty$.

The next theorem is an extension of Gregoire-Hamrouni (2001) to dependent random vectors.

Theorem 1 *Let $\{(X_i, Y_i)\}$ be a strictly stationary strong mixing random vectors such that for some $\delta > 0$*

$$E|Y_i|^{2+\delta} < \infty \tag{3}$$

and

$$\sum_{n=1}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(n) < \infty. \tag{4}$$

Assume further that $h_n \to 0, kh_n \to \infty$ and conditions H1-H3 are satisfied. Then, we have

$$\mathcal{L}_n \ \Rightarrow \mathcal{L} \quad \text{on } \mathcal{D}[-M, M],$$

where

$$\mathcal{L}(z) = -\frac{\lambda_3}{f_X(\tau)}\gamma(\tau)|z| + \frac{\lambda_1}{f_X(\tau)}\mathcal{N}(z),$$

with $\mathcal{N}(z)$ defined by

$$\mathcal{N}(z) = \begin{cases} \displaystyle\sum_{i=1}^{N_z^+}(-\gamma(\tau) - 2\sigma(\tau)\epsilon_i^+) & (z \geq 0) \\ \displaystyle\sum_{i=1}^{N_z^-}(-\gamma(\tau) + 2\sigma(\tau)\epsilon_i^-) & (z < 0). \end{cases}$$

The sequences (ϵ_i^+) and (ϵ_i^-) are independent and built with i.i.d. sequences as ϵ_i). N_z^+ and N_z^- are independent homogeneous Poisson processes with λ_2 as parameter and are independent of the sequences (ϵ_i^+) and (ϵ_i^-).