Nonparametric adaptive estimation of the derivatives of the stationary density of a stationary process

SCHMISSER Emeline

Université Paris Descartes
Laboratoire MAP5
Emeline.Schmisser@mi.parisdescartes.fr

19 june 2010

Table of contents

Table of contents

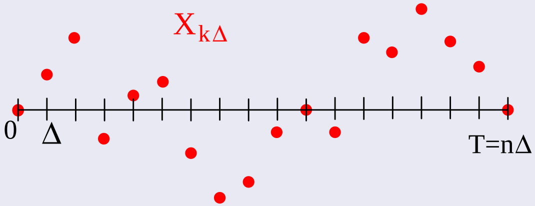
Table of contents

Table of contents

Assumptions

Model

- process (X_t) stricly stationary and exponentially β-mixing



$X_0, X_\Delta, \ldots, X_{n\Delta}$
f stationary density $\mathscr{C}^k$

- β-mixing coefficient: $\beta_X(t) = \frac{1}{2}\|P_{(X_0,X_t)} - P_{X_0} \otimes P_{X_0}\|_{TV}$
- (X_t) is exponentially β-mixing if $\beta_X(t) \leq c\exp(-\theta t)$

Aim:

Estimation of the derivatives of the stationary density:
$f^{(j)}$ $(0 \leq j \leq k)$ over a compact set $[0,1]$

Example of exponentially β-mixing processes

Diffusions

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \qquad X_0 = \eta$$

Assumptions

- functions b and σ are Lipschitz
- σ is bounded and positive: $\exists \sigma_0, \sigma_1, \quad 0 < \sigma_0^2 \leq \sigma^2(x) \leq \sigma_1^2$
- b is "elastic":
  $\exists \alpha \geq 1, r > 0$ such that for $|x|$ large, $xb(x) \leq -r|x|^\alpha$

Example of exponentially β-mixing processes

Diffusions

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \qquad X_0 = \eta$$

Assumptions

- functions b and σ are Lipschitz
- σ is bounded and positive: $\exists \sigma_0, \sigma_1, \quad 0 < \sigma_0^2 \leq \sigma^2(x) \leq \sigma_1^2$
- b is "elastic":
  $\exists \alpha \geq 1, r > 0$ such that for $|x|$ large, $xb(x) \leq -r|x|^\alpha$
- (X_t) is stationary

Property

(X_t) is exponentially β-mixing: $\exists c, \theta > 0, \quad \beta_X(t) \leq c\exp(-\theta t)$

Adaptive estimation: Model selection

- g function to estimate on $[0,1]$
- $(S_m)_{m\leq M}$: $\nearrow$ sequence, linear subspaces, $\subseteq L^2([0,1])$, $D_m = \dim(S_m)$
- contrast function $\gamma_n : \ t \in L^2([0,1]) \ \rightarrow \ \gamma_n(t) \ \in \mathbb{R}$

Adaptive estimation: Model selection

- g function to estimate on $[0,1]$
- $(S_m)_{m\leq M}$: $\nearrow$ sequence, linear subspaces, $\subseteq L^2([0,1])$, $D_m = \dim(S_m)$
- contrast function $\gamma_n : \ t \in L^2([0,1]) \ \rightarrow \ \gamma_n(t) \ \in \mathbb{R}$

For each $m \leq M$

- estimator $\hat{g}_m = \arg\inf_{t\in S_m} \gamma_n(t)$
- Risk computation: $R_m = \mathbb{E}\left(\|\hat{g}_m - g\|^2_{L^2}\right)$ where $\|.\|_{L^2} = \|.\|_{L^2([0,1])}$

$$R_m \leq \text{Bias} + \text{Variance} + \text{Remainder}$$

- bias: $\|g_m - g\|^2_{L^2} \searrow$ with D_m
- variance: $\nearrow$ with D_m

Adaptive estimation: Model selection

- g function to estimate on $[0,1]$
- $(S_m)_{m\leq M}$: $\nearrow$ sequence, linear subspaces, $\subseteq L^2([0,1])$, $D_m = \dim(S_m)$
- contrast function $\gamma_n : \ t \in L^2([0,1]) \ \to \ \gamma_n(t) \ \in \mathbb{R}$

For each $m \leq M$

- estimator $\hat{g}_m = \arg\inf_{t\in S_m} \gamma_n(t)$
- Risk computation: $R_m = \mathbb{E}\left(\|\hat{g}_m - g\|_{L^2}^2\right)$ where $\|.\|_{L^2} = \|.\|_{L^2([0,1])}$

$$R_m \leq \mathrm{Bias} + \mathrm{Variance} + \mathrm{Remainder}$$

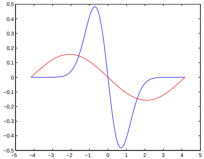
- bias: $\|g_m - g\|_{L^2}^2 \searrow$ with D_m
- variance: $\nearrow$ with D_m

Aim

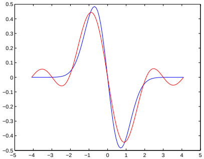
Find compromise between bias and variance terms

Importance of the choice of D_m

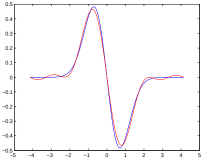
$D_m = 2$



$D_m = 6$

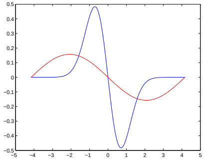


$D_m = 8$

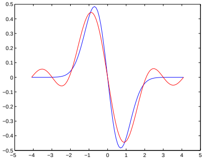


Importance of the choice of D_m

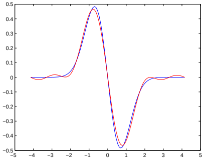
$D_m = 2$



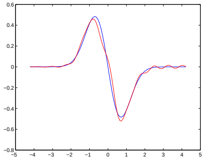
$D_m = 6$



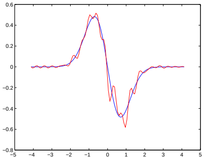
$D_m = 8$



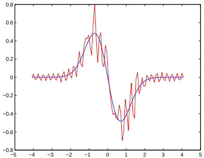
$D_m = 22$



$D_m = 40$

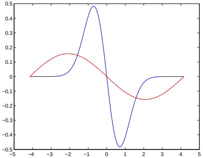


$D_m = 60$

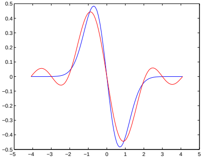


Importance of the choice of D_m

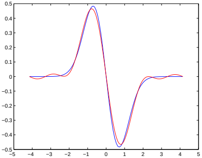
$D_m = 2$



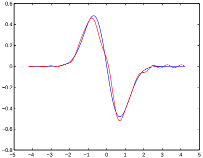
$D_m = 6$



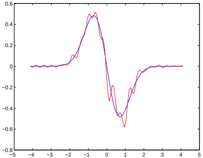
$D_m = 8$



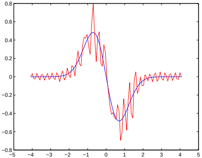
$D_m = 22$



$D_m = 40$



$D_m = 60$



Find the best D_m

- Simple is regularity of g is known ($\Rightarrow$ bias known)
- If not: penalty function

Adaptive estimator

Choice of m

- penalty function $pen \sim \text{Variance}$
- adaptive estimator:

$$\hat{m} = \arg\inf_{m} \{\gamma_n(\hat{g}_m) + pen(m)\} \qquad \hat{g}_{\hat{m}} \text{ adaptive estimator}$$

Risk of the adaptive estimator

- risk of the adaptive estimator $\hat{g}_{\hat{m}}$: $R = \mathbb{E}\left(\|\hat{g}_{\hat{m}} - g\|_{L^2}^2\right)$

$$R \leq C \inf_{m \leq M} (\text{bias} + pen(m)) + \text{Remainders}$$

- Adaptive estimator $\hat{g}_{\hat{m}}$ automatically realises the bias/variance compromise

First estimator: Differentiating an estimator of f

Aim

Estimate $(f^{(j)})_{0\leq j\leq k}$ for stationary exponentially β-mixing processes

Model

- Asymptotic frame: $n \to \infty$, $T = n\Delta \to \infty$, Δ fixed or $\Delta \to 0$
- spaces S_m: special conditions $(t^{(j)}(0) = t^{(j)}(1) = 0$ for $0 \leq j \leq k-1)$

Contrast function

$$\gamma_{j,n}(t) = \|t\|_{L^2}^2 - (-1)^j \frac{2}{n} \sum_{k=1}^{n} t^{(j)}(X_{k\Delta})$$

Justification: least-square contrast function:

$$\mathbb{E}\left(\gamma_{j,n}(t)\right) = \|t\|_{L^2}^2 - 2*(-1)^j < t^{(j)}, f >= \|t\|_{L^2}^2 - 2 < t, f^{(j)} >$$

Results

Risk of the estimator with fixed m:

$$R_m := \mathbb{E}\left(\|\widehat{f^{(j)}}_m - f^{(j)}\|_{L^2}^2\right) \leq \|f_m^{(j)} - f^{(j)}\|_{L^2}^2 + D_m^{2j+1}/(n\theta\Delta)$$

Adaptive estimator

- penalty: $pen(m) \propto D_m^{2j+1}/(n\theta\Delta)$
- risk: $R \leq C min_{m \leq M}\left(\|f_m^{(j)} - f^{(j)}\|_{L^2}^2 + pen(m)\right) + \frac{\ln^{2j+2}(n\Delta)}{n\Delta}$

Convergence's speed in Besov's spaces $B_{2,\infty}^{\alpha}$

$R \leq (n\Delta)^{-2\alpha/(1+2j+2\alpha)}$

Rao: independent variables:
$R \leq n^{-2\alpha/(1+2j+2\alpha)}$

Second estimator: Estimation for diffusions

Aim

Estimation of f' for diffusions with $\sigma = 1$

Model

- Estimation of $f' = 2bf$
- Asymptotic frame: $n \to \infty$, $T = n\Delta \to \infty$, $\Delta \to 0$

Contrast function: (least square contrast)

$$\gamma_n(t) = \|t\|_{L^2}^2 - \frac{4}{n\Delta} \sum_{k=1}^{n} \left(X_{(k+1)\Delta} - X_{k\Delta}\right) t(X_{k\Delta})$$

Justification: $\Delta^{-1} \left(X_{(k+1)\Delta} - X_{k\Delta}\right) = b(X_{k\Delta}) + I_{k\Delta} + Z_{k\Delta}$

$$\mathbb{E}\left(\gamma_n(t)\right) = \|t\|_{L^2}^2 - 2 < t, 2bf > + O\left(\Delta^{1/2}\right) \simeq \|t\|_{L^2}^2 - 2 < t, f' >$$

Results

Risk of the estimator with fixed m

$$R_m := \mathbb{E}\left(\|\widehat{f'}_m - f'\|_{L^2}^2\right) \leq \|f'_m - f'\|_{L^2}^2 + D_m/(n\Delta)\left(1 + \frac{1}{\theta}\right) + \Delta$$

Adaptive estimation

- penalty $pen(m) \propto D_m/(n\Delta)\left(1 + \frac{1}{\theta}\right)$
- Risk: $R \leq C \min_{m \leq M}\left(\|f'_m - f'\|_{L^2}^2 + pen(m)\right) + \Delta + \frac{\ln^2(n\Delta)}{n\Delta}$

Convergence's speed: if $f' = g$ in Besov's space $B^{\alpha}_{2,\infty}$

$$R \leq (n\Delta)^{-2\alpha/(1+2\alpha)} + \Delta$$

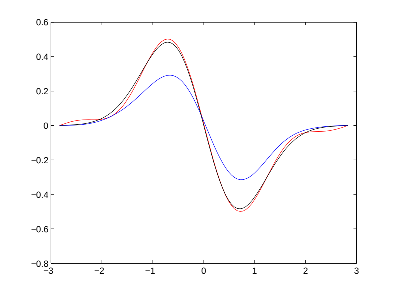
Dalayan and Kutoyants (2003): minimax speed with continuous time ($\Delta = 0$): $R \simeq T^{-2\alpha/(1+2\alpha)}$

Ornstein Uhlenbeck: estimation of the first derivative

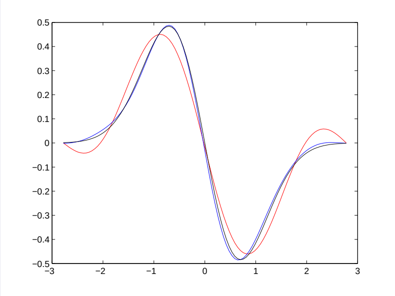
- $b(x) = -x$ and $\sigma(x) = 1$ $\qquad f'(x) = -2x\exp(-x^2)/\sqrt{\pi}$

Graphics

$N = 10^4$ and $\Delta = 1$



$N = 10^5$ and $\Delta = 10^{-2}$



True function

Adaptive estimator of f': first method

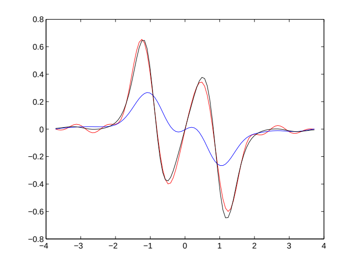
Adaptive estimator of f': second method

Sine function: estimation of the first derivative

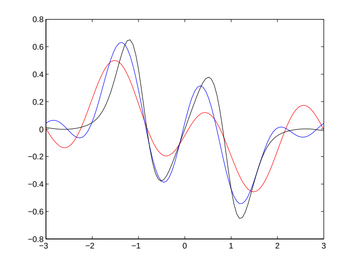
- $b(x) = sin(x) - x/\sqrt{1+x^2}$, $\sigma(x) = 1$ and
  $f'(x) \propto 2b(x)\exp(-2\cos(x) - 2\sqrt{1+x^2})$
- Simulation of $(X_0, \dots, X_{n\Delta})$: retrospective algorithm of Beskos et al

Graphics

$N = 10^4$ and $\Delta = 1$



$N = 10^5$ and $\Delta = 10^{-3}$



Adaptive estimator of f': first method

Adaptive estimator of f': second method

Application: estimation of b by quotient

Application: estimation of b by quotient

Equation

diffusion $X_t = b(X_t)dt + dW_t \quad \Longrightarrow \quad f' = 2bf$

Estimator of b

- $\tilde{f}$ adaptive estimator of f
- $\tilde{f}'$ adaptive estimator of f'
- estimator of b:

$$\tilde{b} = \frac{\tilde{f}'}{2\tilde{f}} \quad \text{if} \quad \tilde{f}' \leq 2n\tilde{f} \quad \text{and} \quad \tilde{b} = 0 \quad \text{if not}$$

Risk for $\Delta = 1$: if b in Besov's space $B^{\alpha}_{2,\infty}$ with $\alpha \geq 1$

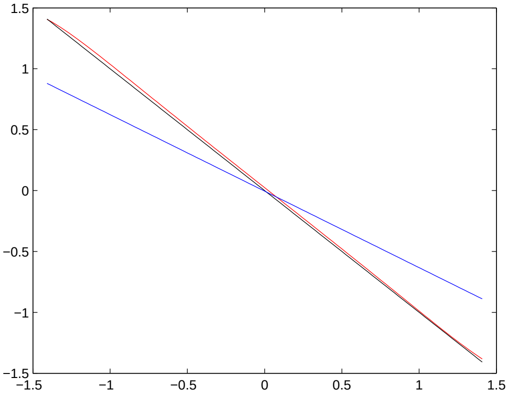
$$R(\tilde{b}) \leq R(\tilde{f}) + R(\tilde{f}') + \frac{1}{n} \leq cn^{-2\alpha/(2\alpha+3)}$$

Minimax speed defined by Gobet et al (2004) (Δ fixed)

Ornstein-Uhlenbeck: drift estimation

- Parameters : $b(x) = -x$ and $\sigma(x) = 1$

$n = 10000$ and $\Delta = 1$

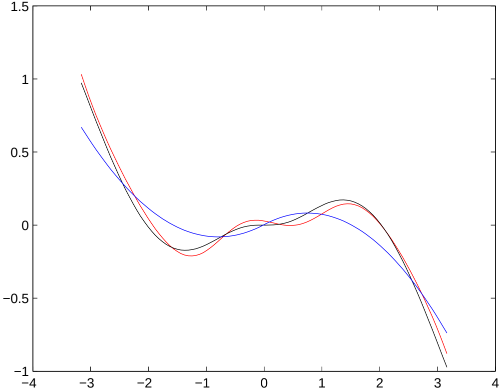


- True drift b
- estimator of b by quotient
  f and f' are estimated by the first method
- adaptive estimator of b by least-square contrast: see Comte, Genon-Catalot et Rozenholc

Function sine: drift estimation

- Parameters : $b(x) = sin(x) - x/\sqrt{1+x^2}$ and $\sigma(x) = 1$

$n = 10000$ and $\Delta = 1$



- True drift b
- estimator of b by quotient
  f and f' are estimated by the first method
- adaptive estimator of b by least-square contrast