

Estimation non paramétrique pénalisée de la dérivée de la densité d'un processus de diffusion

Emeline Schmitter

► **To cite this version:**

Emeline Schmitter. Estimation non paramétrique pénalisée de la dérivée de la densité d'un processus de diffusion. 42èmes Journées de Statistique, 2010, Marseille, France, France. 2010. <inria-00494685>

HAL Id: inria-00494685

<https://hal.inria.fr/inria-00494685>

Submitted on 24 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ESTIMATION NON PARAMÉTRIQUE PÉNALISÉE DE LA DÉRIVÉE DE LA DENSITÉ D'UN PROCESSUS DE DIFFUSION

Emeline Schmisser

Laboratoire MAP5 Université Paris Descartes 45, rue des Saints-Pères 75006 Paris

Abstract

Nous considérons un processus de diffusion $(X_t)_{t \geq 0}$ de fonction de dérive b et de coefficient de diffusion σ . Nous supposons que le processus est strictement stationnaire et β -mélangeant. Ce processus est observé à des instants discrets $t_k = k\Delta$ pour k variant de 0 à $n + 1$. Nous calculons deux estimateurs non paramétriques de la dérivée f' de la densité f de ce processus. Pour cela, nous considérons une familles de sous-espaces vectoriels et, sur chacun de ces sous-espaces, nous calculons deux estimateurs de f' , $\hat{f}'_m^{(d)}$ et $\hat{f}'_m^{(b)}$. Nous choisissons ensuite les deux meilleurs estimateurs possibles, $\hat{f}'_{\hat{m}}^{(d)}$ et $\hat{f}'_{\hat{m}}^{(b)}$ en introduisant deux fonctions de pénalité $pen^{(d)}(m)$ et $pen^{(b)}(m)$. Le premier estimateur, $\hat{f}'_m^{(d)}$, est obtenu en dérivant un estimateur de f . Cet estimateur minimise une fonction de contraste. L'estimateur adaptatif $\hat{f}'_{\hat{m}}^{(d)}$ est convergent si $n\Delta \rightarrow \infty$ (que le pas Δ soit fixé ou qu'il tende vers 0). Pour calculer le second estimateur, nous supposons que le coefficient de diffusion σ est constant égal à 1. Dans ce cas, la fonction f' vérifie l'égalité $f'(x) = 2b(x)f(x)$. L'estimateur $\hat{f}'_m^{(b)}$ minimise une fonction de contraste basé sur cette égalité. Lorsque le pas Δ est petit, le risque de l'estimateur $\hat{f}'_{\hat{m}}^{(b)}$ est plus petit que le risque de $\hat{f}'_{\hat{m}}^{(d)}$. Cependant, l'estimateur $\hat{f}'_{\hat{m}}^{(b)}$ ne converge que si $\Delta \rightarrow 0$ et $n\Delta \rightarrow \infty$.

Let us consider a diffusion process $(X_t)_{t \geq 0}$ of drift function b and diffusion coefficient σ . This process is assumed to be strictly stationary and β -mixing. It is observed at discrete times $t_k = k\Delta$ for k from 0 to $n + 1$. Two non parametric estimators of the derivative f' of the density function f are computed. For this purpose, we consider a family of vectorial subspaces and compute a collection of estimators of f' on every subspace, $\hat{f}'_m^{(d)}$ and $\hat{f}'_m^{(b)}$. Then, the best estimators $\hat{f}'_{\hat{m}}^{(d)}$ and $\hat{f}'_{\hat{m}}^{(b)}$ are selected by introducing two penalty functions, $pen^{(d)}(m)$ and $pen^{(b)}(m)$. The first estimator, $\hat{f}'_m^{(d)}$, is the derivative of an estimator of f . This estimator minimises a contrast function. The adaptive estimator $\hat{f}'_{\hat{m}}^{(d)}$ converges if $n\Delta \rightarrow \infty$. In order to compute the second estimator, we assume that the diffusion coefficient σ is constant ($\sigma = 1$). In that case, the function f' satisfies the equality $f'(x) = 2b(x)f(x)$. The estimator $\hat{f}'_m^{(b)}$ minimises a contrast function based on this equality. If Δ is small, the risk of the estimator of $\hat{f}'_{\hat{m}}^{(b)}$ is smaller than the risk of $\hat{f}'_{\hat{m}}^{(d)}$. Nevertheless, this estimator converges only if $\Delta \rightarrow 0$ and $n\Delta \rightarrow \infty$.

Mots-clés: Estimation semi-paramétrique et non paramétrique, Processus

1 Notations

Considérons l'équation différentielle stochastique

$$dX_t = b(X_t)dt + dW_t, \quad X_0 = \eta$$

Le processus $(X_t)_{t \geq 0}$ est supposé strictement stationnaire et β -mélangeant de coefficient de β -mélange $\beta_X(t) \leq ce^{-\theta t}$. Notons f la densité stationnaire de (X_t) . Nous observons ce processus à des instants discrets $t = 0, \Delta, \dots, n\Delta$. Nous voulons estimer la dérive de la fonction de densité f' sur l'intervalle $[0, \pi]$.

2 Estimation par dérivée d'un estimateur

Dans cette partie, nous construisons un estimateur $\hat{f}'_m^{(d)}$ de f' en dérivant un estimateur de f . Considérons le contraste

$$\gamma_n(t) = \|t\|_{L^2}^2 - \frac{2}{n} \sum_{i=1}^n t(X_i)$$

et définissons l'estimateur

$$\hat{f}_m = \inf_{t \in S_m} \{\gamma_n(t)\}$$

où S_m est un sous-espace vectoriel engendré par les polynômes trigonométriques:

$$S_m = \text{Vect}(\cos(kx)1_{[0,\pi]}(x), 0 \leq k \leq m).$$

Notons f_m la projection de f sur S_m . Notons

$$V_m = \text{Vect}(\sin(kx)1_{[0,\pi]}(x), 1 \leq k \leq m).$$

La projection $f_m^{(d)}$ de f' sur $V_m = \text{Vect}(\sin(kx)1_{[0,\pi]}(x), 1 \leq k \leq m)$ vérifie

$$f_m^{(d)} = (f_m)'$$

Notons $\hat{f}'_m^{(d)} = (\hat{f}_m)'$. Cet estimateur minimise, pour $t \in V_m$, le contraste

$$\gamma_n^{(d)}(t) = \|t\|_{L^2}^2 + \frac{2}{n} \sum_{i=1}^n t'(X_i).$$

2.1 Risque de l'estimateur à m fixé.

Notons $\mathcal{M}_n = \{m, D_m \leq \mathcal{D}_n\}$. La dimension maximale \mathcal{D}_n sera précisée plus loin. Pour tout $m \in \mathcal{M}_n$, nous calculons l'estimateur $\hat{f}'_m^{(d)}$ de f' . D'après Pythagore,

$$\|\hat{f}'_m^{(d)} - f'1_{[0,\pi]}\|_{L^2}^2 = \|f_m^{(d)} - f'1_{[0,\pi]}\|_{L^2}^2 + \|\hat{f}'_m^{(d)} - f_m^{(d)}\|_{L^2}^2.$$

De plus, nous avons que

$$\|\hat{f}'_m^{(d)} - f_m^{(d)}\|_{L^2}^2 \leq D_m^2 \|\hat{f}_m - f_m\|_{L^2}^2.$$

Pour majorer cette quantité, nous utilisons des propriétés de β -mélange (voir Viennet (1997)).

Si le pas Δ est fixe, alors le risque de l'estimateur vérifie

$$\mathbb{E} \left(\left\| \hat{f}'_m^{(d)} - f'1_{[0,\pi]} \right\|_{L^2}^2 \right) \leq \left\| f_m^{(d)} - f'1_{[0,\pi]} \right\|_{L^2}^2 + c \frac{D_m^3}{n}.$$

Si $\Delta \rightarrow 0$ et $n\Delta \rightarrow \infty$, nous avons que

$$\mathbb{E} \left(\left\| \hat{f}'_m^{(d)} - f'_A \right\|_{L^2}^2 \right) \leq \left\| f_m^{(d)} - f'_A \right\|_{L^2}^2 + c \frac{D_m^3}{n\theta\Delta}.$$

2.2 Optimisation du choix de m

Supposons que f appartienne à un espace de Besov $B_{2,\infty}^\alpha$ et que $\|f'\|_{B_{2,\infty}^\alpha}^2 \leq 1$. Dans ce cas, nous avons que

$$\|f - f_m^{(d)}\|_{L^2}^2 \leq C(\alpha) D_m^{-2\alpha},$$

donc

$$\left\| f'1_{[0,\pi]} - f_m^{(d)} \right\|_{L^2}^2 \leq D_m^2 \|f - f_m^{(d)}\|_{L^2}^2 \leq C(\alpha) D_m^{-2(\alpha-1)}.$$

Pour que le compromis biais-variance soit minimal, nous devons choisir, si le pas Δ est fixe,

$$D_m \sim n^{1/(1+2\alpha)}$$

et dans ce cas, le risque de l'estimateur vérifie

$$\mathbb{E} \left(\left\| \hat{f}'_m^{(d)} - f'1_{[0,\pi]} \right\|_{L^2}^2 \right) \leq n^{-2(\alpha-1)/(1+2\alpha)}.$$

Si $\Delta \rightarrow 0$ et $n\Delta \rightarrow \infty$, nous devons prendre

$$D_m \sim (n\theta\Delta)^{1/(1+2\alpha)}$$

pour minimiser le compromis biais-variance. Dans ce cas, l'estimateur vérifie

$$\mathbb{E} \left(\left\| \hat{f}'_m^{(d)} - f'1_{[0,\pi]} \right\|_{L^2}^2 \right) \leq (n\theta\Delta)^{-2(\alpha-1)/(1+2\alpha)}.$$

La vitesse de convergence est toujours inférieure à $(n\Delta)$.

2.3 Risque de l'estimateur adaptatif

Introduisons une fonction de pénalité $pen^{(d)}(m)$ dépendant de D_m et de n . Choisissons \hat{m} tel que

$$\hat{m} = \arg \min_{m \in \mathcal{M}_n} \left[\gamma_n^{(d)} \left(\hat{f}'_m^{(d)} \right) + pen^{(d)}(m) \right]$$

et notons $\hat{f}'_{\hat{m}}$ l'estimateur résultant.

Si le pas est fixé, que

$$pen^{(d)}(m) \geq c \frac{D_m^3}{n} \quad \text{et} \quad \mathcal{D}_n^4 \leq c \ln^2(n) n \exp \left(c' \frac{\sqrt{n}}{\ln(n)} \right)$$

alors

$$\mathbb{E} \left(\left\| \hat{f}'_{\hat{m}}^{(d)} - f' 1_{[0,\pi]} \right\|_{L^2}^2 \right) \leq C \inf_{m \in \mathcal{M}_n} \left(\left\| f'_m^{(d)} - f' 1_{[0,\pi]} \right\|_{L^2}^2 + pen^{(d)}(m) + c \frac{\ln^4(n)}{n} \right).$$

Si $\Delta \rightarrow 0$ et $n\Delta \rightarrow \infty$, que

$$pen^{(d)}(m) \geq c \frac{D_m^3}{n\theta\Delta} \quad \text{et} \quad \mathcal{D}_n^4 \leq cn\Delta \ln^2(n) \exp \left(c' \frac{\sqrt{n\Delta}}{\ln(n)} \right),$$

alors

$$\mathbb{E} \left(\left\| \hat{f}'_{\hat{m}}^{(d)} - \hat{f}' 1_{[0,\pi]} \right\|_{L^2}^2 \right) \leq C \inf_{m \in \mathcal{M}_n} \left(\left\| f'_m^{(d)} - f' 1_{[0,\pi]} \right\|_{L^2}^2 + pen^{(d)}(m) + \frac{c \ln^4(n)}{n\Delta} + \frac{1}{n} \right).$$

3 Si le coefficient de diffusion est constant

3.1 Modèle et hypothèses

Supposons que $\sigma = 1$. Dans ce cas, la densité stationnaire vérifie

$$f(x) \propto \exp \left(2 \int_0^x b(x) dx \right)$$

et $f'(x) = 2b(x)f(x)$. Dans cette partie, nous supposons que la fonction b est lipschitzienne et qu'il existe des constantes $r > 0$ et $\alpha \geq 1$ telles que

$$\exists M_0 \in \mathbb{R}^+, \quad \forall x, |x| \geq M_0, \quad xb(x) \leq -r |x|^\alpha.$$

3.2 Risque de l'estimateur à m fixé

Considérons les espaces vectoriels $S_m = \text{Vect}(1, (\cos(kx))_{1 \leq k \leq m})$. Pour tout $m \in \mathcal{M}_n$, avec $\mathcal{M}_n = \{m, D_m \leq \mathcal{D}_n\}$, nous allons calculer un estimateur $\hat{f}'_m^{(b)}$ de f' . La dimension maximale \mathcal{D}_n sera précisée plus loin. Notons $f'_m{}^{(b)}$ la projection orthogonale de f' sur S_m . Considérons le contraste

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

et l'estimateur

$$\hat{f}'_m^{(b)} = \arg \min_{t \in S_m} \gamma_n^{(b)}(t).$$

Nous savons que

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

avec

$$I_{k\Delta} = \frac{1}{\Delta} \int_{k\Delta}^{(k+1)\Delta} (b(X_s) - b(X_{k\Delta})) ds \quad \text{et} \quad Z_{k\Delta} = \frac{1}{\Delta} \int_{k\Delta}^{(k+1)\Delta} dW_s.$$

Donc

$$\gamma_n(t) - \gamma_n(s) = \|t\|_{L^2}^2 - \|s\|_{L^2}^2 - \frac{4}{n} \sum_{k=1}^n (I_{k\Delta} + Z_{k\Delta} + b(X_{k\Delta})) (t(X_{k\Delta}) - s(X_{k\Delta})).$$

De plus, nous avons que

$$\|t - f'\|_{L^2}^2 = \|t\|_{L^2}^2 + \|f'\|_{L^2}^2 - 2 \int t(x) f'(x) dx = \|t\|_{L^2}^2 + \|f'\|_{L^2}^2 - 4 \int t(x) b(x) f(x) dx.$$

Notons

$$\begin{aligned} \nu_n(t) &= \frac{1}{n} \sum_{k=1}^n Z_{k\Delta} t(X_{k\Delta}) \\ \rho_n(t) &= \frac{1}{n} \sum_{k=1}^n \mathbb{E} (I_{k\Delta} t(X_{k\Delta})) \\ \xi_n(t) &= \frac{1}{n} \sum_{k=1}^n (I_{k\Delta} + b(X_{k\Delta})) t(X_{k\Delta}) - \mathbb{E} ((I_{k\Delta} + b(X_{k\Delta})) t(X_{k\Delta})) \end{aligned}$$

Nous avons que

$$\left\| \hat{f}'_m^{(b)} - f' 1_{[0, \pi]} \right\|_{L^2}^2 = \|t - f'\|_{L^2}^2 - \left\| \hat{f}'_m^{(b)} - f' \right\|_{L^2}^2 + 2(\nu_n + \rho_n + \xi_n) \left(\hat{f}'_m^{(b)} - f'_m{}^{(b)} \right)$$

La majoration des termes ν_n et ρ_n utilise les propriétés de la diffusion et est réalisée dans Comte et al (2007). La majoration du dernier terme, ξ_n , utilise le β -mélange et est réalisée dans Viennet(1997).

$$\mathbb{E} \left(\left\| \hat{f}'_m^{(b)} - f'1_{[0,\pi]} \right\|_{L^2}^2 \right) \leq \left\| f'_m^{(b)} - f'1_{[0,\pi]} \right\|_{L^2}^2 + c\Delta + c\frac{D_m}{n\Delta}$$

3.3 Optimisation du choix de m

Supposons que f' appartienne à la boule unité de l'espace de Besov $B_{2,\infty}^\alpha$. Dans ce cas, nous savons que $\left\| f' - f'_m^{(b)} \right\|_{L^2}^2 \leq D_m^{-2\alpha}$. Pour minimiser le compromis biais-variance, nous devons choisir

$$D_m \sim (n\Delta)^{1/(1+2\alpha)}$$

et dans ce cas le risque de l'estimateur vérifie:

$$\mathbb{E} \left(\left\| \hat{f}'_m^{(b)} - f' \right\|_{L^2}^2 \right) \leq C(n\Delta)^{-2\alpha/(1+2\alpha)} + c\Delta.$$

Dalayan et Kutoyants (2003) estiment la dérivée de la densité d'une diffusion dans le cas où on a des données continues. La borne minimax du risque qu'ils trouvent est $c(n\Delta)^{-2\alpha/(1+2\alpha)}$.

3.4 Risque de l'estimateur adaptatif

Posons

$$pen^{(b)}(m) = c\frac{D_m}{n\Delta}$$

et notons

$$\hat{m} = \inf_{m \in \mathcal{M}_n} \left\{ \gamma_n \left(\hat{f}'_m^{(b)} \right) + pen^{(b)}(m) \right\}$$

et notons $\hat{f}'_{\hat{m}}$ l'estimateur résultant. Si $\mathcal{D}_n^2 \leq n\theta\Delta$, alors

$$\mathbb{E} \left(\left\| \hat{f}'_{\hat{m}}^{(b)} - f' \right\|_{L^2}^2 \right) \leq C \inf_{m \in \mathcal{M}_n} \left\{ \left\| f'_m^{(b)} - f' \right\|_{L^2}^2 + pen(m) \right\} + \Delta + c\frac{\ln^2(n)}{n\Delta}$$

Bibliographie

- [1] Comte, F. genon-Catalot, V. et Rozenholc, Y. (2007) Penalized nonparametric mean square estimation of the coefficients of diffusion processes, *Berbouilli*, 13(2) 514–543.
- [2] Viennet, G. (1997) Inequalities for absolutely regular sequences: application to density estimation. *Probab. Theory Related Fields*, 107(4), 467–492.
- [3] Dalayan, A.S. et Kutoyants, A.Y. (2003) Asymptotically Efficient Estimation of the Derivative of the Invariant Density *Statistical Inference for Stochastic Processes*, 6(1), 89–107