Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Acad. Sci. Paris, Ser. I 337 (2003) 293–296
COMPTES RENDUS MATHEMATIQUE

Statistique/Probabilités

Comportement asymptotique d'un estimateur de la densité adaptatif par méthode d'ondelettes

Jean-Baptiste Aubin, Anne Massiani

Université Paris 6, LSTA, boîte 158, 175, rue du Chevaleret, 75013 Paris, France

Reçu le 25 mars 2003 ; accepté après révision le 17 juin 2003

Présenté par Paul Deheuvels

Résumé

Nous étudions une version tronquée de l'estimateur de la densité par méthode d'ondelettes qui consiste à introduire un niveau d'analyse multirésolution adaptatif $\hat{\jmath}_n$. Nous décrivons d'abord le comportement asymptotique de $\hat{\jmath}_n$. Nous montrons alors que l'estimateur basé sur $\hat{\jmath}_n$ atteint une vitesse suroptimale au sens de l'erreur quadratique intégrée sur un sous-ensemble dense de $L_2(\mathbb{R})$. De plus, cet estimateur atteint une vitesse quasi-optimale si la densité inconnue f appartient à l'espace de Sobolev W_2^p, où $p \geqslant 1$, et a un support compact. ***Pour citer cet article : J.-B. Aubin, A. Massiani, C. R. Acad. Sci. Paris, Ser. I 337 (2003).***

Abstract

Asymptotic behavior of an adaptative wavelet density estimator. We present a data-driven version of the wavelet density estimator, where the traditional multiresolution analysis level j_n is replaced by an adaptative multiresolution analysis level $\hat{\jmath}_n$. First, we describe the limiting behavior of $\hat{\jmath}_n$. Next, we show that the estimator based on $\hat{\jmath}_n$ reaches a superoptimal rate for the mean square error on a dense subset of $L_2(\mathbb{R})$. We finally state that this estimator reaches quasioptimal rate of convergence when the unknown density f belongs to a Sobolev class W_2^p, where $p \geqslant 1$, and has compact support. ***To cite this article: J.-B. Aubin, A. Massiani, C. R. Acad. Sci. Paris, Ser. I 337 (2003).***

1. Introduction

Au cours de la dernière décennie, les statisticiens ont introduit la théorie des ondelettes en estimation non paramétrique. Citons par exemple l'ouvrage de Härdle et al. [6] ainsi que leur bibliographie. En transposant le travail général de Bosq [1] sur l'estimation de la densité par projection, nous étudions ici une version tronquée de l'estimateur de la densité par méthode d'ondelettes.

Soit $X_1, X_2, \ldots$ une suite de variables aléatoires indépendantes et équidistribuées, de densité inconnue f appartenant à $L_2(\mathbb{R})$, que l'on cherche à estimer. Soient également φ une fonction d'échelle, et ψ l'ondelette

Adresses e-mail : jbaubin@ccr.jussieu.fr (J.-B. Aubin), amassia@ccr.jussieu.fr (A. Massiani).

1631-073X/$ – see front matter
doi:10.1016/S1631-073X(03)00310-8

mère associée, toutes deux supposées bornées et à support compact. Notons, pour $j, k \in \mathbb{Z}$:

$$\varphi_{jk} = 2^{j/2}\varphi\big(2^j \cdot -k\big), \quad \psi_{jk} = 2^{j/2}\psi\big(2^j \cdot -k\big), \quad \alpha_{jk} = \int_{-\infty}^{+\infty} \varphi_{jk}(u) f(u)\,\mathrm{d}u, \quad \beta_{jk} = \int_{-\infty}^{+\infty} \psi_{jk}(u) f(u)\,\mathrm{d}u.$$

On peut alors estimer f par la quantité $\sum_{k\in\mathbb{Z}} \hat{\alpha}_{j_0k}\varphi_{j_0k} + \sum_{j=j_0}^{j_n} \sum_{k\in\mathbb{Z}} \hat{\beta_{jk}}\psi_{jk}$, avec $\hat{\alpha}_{j_0k} = \frac{1}{n}\sum_{i=1}^n \varphi_{j_0k}(X_i)$, et $\hat{\beta}_{jk} = \frac{1}{n}\sum_{i=1}^n \psi_{jk}(X_i)$, et où $j_n \geqslant j_0$. Un problème important réside dans le choix du niveau d'analyse multirésolution j_n. Parmi les différentes manières de procéder, nous introduisons ici un niveau d'analyse multirésolution $\hat{j}_n$ adaptatif défini par :

$$\hat{j}_n = \max\big\{j_0 \leqslant j \leqslant j_n \mid \exists k \in \mathbb{Z},\ |\hat{\beta}_{jk}| > c_j\gamma_n\big\}, \quad \text{avec la convention } \max\{\emptyset\} = j_0, \tag{1}$$

les quantités j_n, c_j et γ_n restant à choisir. Ceci conduit ainsi à considérer l'estimateur $\hat{f}_{\hat{j}_n}$ de f défini par :

$$\hat{f}_{\hat{j}_n} = \sum_{k\in\mathbb{Z}} \hat{\alpha}_{j_0k}\varphi_{j_0k} + \sum_{j=j_0}^{\hat{j}_n} \sum_{k\in\mathbb{Z}} \hat{\beta}_{jk}\Psi_{jk}. \tag{2}$$

Remarquons que si $\gamma_n = 0$, $\hat{f}_{\hat{j}_n}$ est simplement l'estimateur dit « linéaire » par méthode d'ondelettes introduit par Doukhan et Leon [4]. Nous faisons ici le choix classique dans la littérature $\gamma_n = (\frac{\log n}{n})^{1/2}$. Nous ajustons alors ce choix pour des raisons techniques liées à l'inégalité de Höeffding dans nos preuves en fixant un c_j proportionnel à $\|\psi_{jk}\|_\infty$, pour $j_n \geqslant j_0$ et $k \in \mathbb{Z}$, i.e. $c_j = \delta 2^{j/2}$, où $\delta > 0$. Nous faisons dans la suite l'hypothèse naturelle : $j_n \to \infty$, lorsque $n \to \infty$. Nous imposons cependant la condition peu restrictive suivante : $2^{j_n} \leqslant n$, pour tout $n \geqslant 1$.

Signalons par ailleurs que $\hat{f}_{\hat{j}_n}$ peut se réécrire de la façon suivante :

$$\hat{f}_{\hat{j}_n}(x) = \frac{2^{\hat{j}_n}}{n} \sum_{i=1}^n K\big(2^{\hat{j}_n}X_i, 2^{\hat{j}_n}x\big), \quad \text{où } K(x,y) = \sum_{k\in\mathbb{Z}} \varphi(x-k)\varphi(y-k), \text{ pour } x, y \in \mathbb{Z}. \tag{3}$$

Notons que certains auteurs, notamment Donoho et Johnstone [3], Hall et al. [5], ont déjà étudié des méthodes de seuillage un peu différentes, qui consistent à supprimer certains coefficients ou blocs de coefficients jugés « trop petits », tandis que notre démarche repose sur l'introduction d'un $\hat{j}_n$ adaptatif et sur l'étude de son comportement asymptotique.

Nous abordons ici l'inévitable question de la convergence de l'estimateur $\hat{f}_{\hat{j}_n}$. Nous commençons donc par décrire dans la Section 2 le comportement asymptotique de $\hat{j}_n$, dont dépend fortement celui de $\hat{f}_{\hat{j}_n}$. Ceci nous amène à distinguer le cas où $f \in \mathcal{F}_0$ de celui où $f \in \mathcal{F}_1$, les sous-espaces $\mathcal{F}_0$ et $\mathcal{F}_1$ étant définis par :

$$\mathcal{F}_0 = \bigcup_{J\geqslant j_0} \mathcal{F}_0(J) \quad \text{où } \mathcal{F}_0(J) = \left\{ f = \sum_{k\in\mathbb{Z}} \alpha_{j_0k}\varphi_{j_0k} + \sum_{j=j_0}^{J} \sum_{k\in\mathbb{Z}} \beta_{jk}\psi_{jk} \,\middle|\, \exists k \in \mathbb{Z},\ \beta_{Jk} \neq 0 \right\}, \text{ pour } J \geqslant j_0,$$

$$\mathcal{F}_1 = L_2(\mathbb{R}) \setminus \left\{ \widetilde{\mathcal{F}}_0 \bigcup_{J\geqslant j_0} \mathcal{F}_0(J) \right\}, \quad \text{où } \widetilde{\mathcal{F}}_0 = \left\{ f = \sum_{k\in\mathbb{Z}} \alpha_{j_0k}\varphi_{j_0} \right\}.$$

Nous établissons alors dans la Section 3 que l'estimateur $\hat{f}_{\hat{j}_n}$ atteint une vitesse suroptimale, au sens de l'erreur quadratique intégrée sur le sous-espace $\mathcal{F}_0$, qui est dense dans $L_2(\mathbb{R})$. De plus, l'estimateur $\hat{f}_{\hat{j}_n}$ est quasi-optimal si la densité f appartient à l'espace de Sobolev W_2^P, où $p \geqslant 1$, et a un support compact (cf. Section 4). Rappelons que la vitesse optimale est donnée par Bretagnolle et Huber [2], qui se sont intéressés à la théorie minimax dans les espaces de Sobolev. Nous obtenons aussi dans un cas particulier une loi du logarithme itéré ponctuelle pour $\hat{f}_{\hat{j}_n}$, comparable à celle déjà établie pour l'estimateur linéaire par méthode d'ondelettes (cf. Massiani [7]).

Par souci de concision, la preuve de ces résultats n'est pas présentée ici.

2. Comportement asymptotique de $\hat{j}_n$

Nous supposons dans tout ce qui suit que $\int_{-\infty}^{+\infty} x^2 f(x)\,\mathrm{d}x < \infty$.

Proposition 2.1. *Choisissons δ tel que $\delta > 2\sqrt{2}\|\psi\|_\infty$, où $\|\psi\|_\infty = \sup_{x\in\mathbb{R}} |\psi(x)|$.*

(1) *Si $f \in \mathcal{F}_0(J)$, alors, presque sûrement, $\hat{j}_n \to J$ lorsque $n \to \infty$.*
(2) *Si $f \in \mathcal{F}_1$, alors, presque sûrement, $\hat{j}_n \to \infty$ lorsque $n \to \infty$.*

Nous allons maintenant étudier plus précisément le comportement asymptotique de $\hat{j}_n$, dans le cas particulier où les coefficients théoriques β_{jk} vérifient les conditions suivantes, pour un $s > 0$:

$$\forall j,k \in \mathbb{Z}, \quad |\beta_{jk}| \leqslant \alpha 2^{-js}, \quad \text{où } \alpha > 0 \tag{4}$$

$$\forall j \in \mathbb{Z},\ \exists k \in \mathbb{Z} \text{ tel que } \quad |\beta_{jk}| \geqslant \gamma 2^{-js}, \quad \text{où } \gamma > 0. \tag{5}$$

Proposition 2.2. *Soient $a,b > 0$ tels que $\frac{\alpha}{b^{(2s+1)/2}} + 2\sqrt{2}\|\Psi\|_\infty < \frac{\gamma}{a^{(2s+1)/2}} - \sqrt{2}\|\Psi\|_\infty$. Choisissons δ tel que $\frac{\alpha}{b^{(2s+1)/2}} + 2\sqrt{2}\|\Psi\|_\infty < \delta < \frac{\gamma}{a^{(2s+1)/2}} - \sqrt{2}\|\Psi\|_\infty$, et j_n tel que $2^{j_n} > b(\frac{n}{\log n})^{1/(2s+1)}$. Si la condition* (4) *est vérifiée, alors, presque sûrement, si n est assez grand* :

$$2^{\hat{j}_n} \leqslant b\left(\frac{n}{\log n}\right)^{1/(2s+1)}, \tag{6}$$

et si la condition (5) *est vérifiée, alors, presque sûrement, si n est assez grand* :

$$2^{\hat{j}_n} \geqslant a\left(\frac{n}{\log n}\right)^{1/(2s+1)}. \tag{7}$$

Remarque 1. Supposons que φ appartient à l'espace de Sobolev W_2^N, où $N \geqslant 0$. Si f appartient à l'espace de Sobolev W_2^p, où $1 \leqslant p \leqslant N+1$, alors la condition (4) est vérifée (pour $s = p$). Si de plus $f \in \mathrm{W}_2^{p+1}$, alors la condition (5) n'est pas vérifée (cf. Härdle et al. [6], p. 120, formule 3.37).

Remarque 2. Notons que si $f \in \mathrm{W}_2^p$, le choix $2^{j_n} \sim n^{1/(2p+1)}$ minimise le majorant de l'erreur quadratique intégrée de l'estimateur linéaire (cf. Härdle et al. [6], Théorème 10.1, p. 130). Si $f \in \mathrm{W}_2^p$ et vérifie les conditions (4) et (5), $\hat{j}_n$ est donc une approximation, à un facteur logarithmique près, de ce $2^{j_n} \sim n^{1/(2p+1)}$ « optimal ».

3. Suroptimalité sur $\mathcal{F}_0$

Une conséquence importante de la première partie de la Proposition 2.1 est que l'estimateur $\hat{f}_{\hat{j}_n}$ atteint des vitesses suroptimales sur $\mathcal{F}_0$ au sens de l'erreur quadratique intégrée. De plus, cette propriété permet également d'obtenir une vitesse en $(\frac{n}{\log\log n})^{1/2}$ pour la convergence uniforme presque sûre.

Proposition 3.1. *Si $\delta > \sqrt{10}\|\psi\|_\infty$ et si $f \in \mathcal{F}_0(J)$, alors*

$$n E_f \|\hat{f}_{\hat{j}_n} - f\|_2^2 \to 2^J \int_{-\infty}^{\infty} K\left(2^J u, 2^J u\right) f(u)\,\mathrm{d}u - \int_{-\infty}^{\infty} f^2(u)\,\mathrm{d}u, \quad \textit{quand } n \to \infty.$$

Proposition 3.2. *Si $f \in \mathcal{F}_0(J)$, alors, pour une certaine constante $C_1 \geqslant 0$, on a, presque sûrement* :

$$\limsup_{n\to\infty}\left(\frac{n}{\log\log n}\right)^{1/2} \|\hat{f}_{\hat{j}_n} - f\|_\infty \leqslant C_1.$$

4. Convergence dans $\mathcal{F}_1$

Proposition 4.1. *Supposons que la fonction d'échelle φ appartient à l'espace de Sobolev W_2^N, où $N \geqslant 0$. Soit $\widetilde{W}_p^2$ l'ensemble des densités f appartenant à l'espace de Sobolev W_p^2, où $1 \leqslant p \leqslant N+1$, ayant un support inclus dans l'intervalle $[-L, L]$, où $L \in \mathbb{R}$ et telles que $\|f\|_2 \leqslant L'$ où $L' \in \mathbb{R}$. Alors, en choisisssant $\delta > \sqrt{10}\|\psi\|_\infty$, il existe une constante $C_2 \geqslant 0$ telle que* :

$$\sup_{f\in\widetilde{W}_p^2} \|\hat{f}_{\hat{j}_n} - f\|^2 \leqslant C_2\left(\frac{\log n}{n}\right)^{2p/(2p+1)}. \tag{8}$$

Les encadrements (6) et (7) de la quantité $2^{\hat{j}_n}$ permettent maintenant d'étudier la vitesse de convergence ponctuelle presque sûre de $\hat{f}_{\hat{j}_n}$ dans le cas particulier où les coefficients β_{jk} vérifient les conditions (4) et (5).

Avant d'énoncer notre résultat, notons $f_{\hat{j}_n} = \sum_{k\in\mathbb{Z}} \alpha_{j_0 k}\varphi_{j_0 k} + \sum_{j=j_0}^{\hat{j}_n}\sum_{k\in\mathbb{Z}} \beta_{jk}\psi_{jk}$.

Proposition 4.2. *Supposons que φ soit une fonction continue à gauche, à support compact, et à variations bornées. Supposons également que f soit continue en un point x tel que la quantité $f(x)$ soit strictement positive. Si f vérifie les conditions* (4) *et* (5)*, alors, en choisissant δ comme dans la Proposition* 2.2*, l'estimateur $\hat{f}_{\hat{j}_n}$ vérifie* :

$$\limsup_{n\to\infty}\left(\frac{n}{2\times 2^{\hat{j}_n}\log\log n}\right)^{1/2} \frac{\hat{f}_{\hat{j}_n}(x) - f_{\hat{j}_n}(x)}{(f(x)\int_{-\infty}^{\infty} K^2(2^{\hat{j}_n}x, u)\,\mathrm{d}u)^{1/2}} \leqslant 1 \quad p.s. \tag{9}$$

De plus, l'inégalité précédente est une égalité si x est un point dyadique (*i.e., de la forme $x = p/2^q$, où $p, q \in \mathbb{Z}$*).

Notons qu'il est facile de vérifier que le terme de normalisation $\int_{-\infty}^{\infty} K^2(2^{\hat{j}_n}x, u)\,\mathrm{d}u$ intervenant dans la Proposition 4.2 vérifie presque sûrement, pour n assez grand :

$$B \leqslant \int_{-\infty}^{\infty} K^2\big(2^{\hat{j}_n}x, u\big)\,\mathrm{d}u \leqslant C, \quad \text{où } B \text{ et } C \text{ sont deux constantes strictement positives.}$$

Remerciements

Nous remercions pour leurs suggestions constructives Denis Bosq et Daniel Pierre-Loti-Viaud.

Références

[1] D. Bosq, Estimation localement suroptimale et adaptative de la densité, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 591–595.
[2] J. Bretagnolle, C. Huber, Estimation des densités : risque minimax, Z. Wahrscheinlichkeitstheorie Verw. Gebiete 47 (1979) 119–137.
[3] D. Donoho, I. Johnstone, Minimax risk over l_p-balls for L_p-error, Probab. Theory Related Fields 99 (1994) 277–303.
[4] P. Doukhan, J. Leon, Déviation quadratique d'estimateurs d'une densité par projection orthogonale, C. R. Acad. Sci. Paris, Ser. I 310 (1990) 425–430.
[5] P. Hall, G. Kerkyacharian, D. Picard, Block threshold rules for curve estimation using kernel and wavelet methods, Ann. Statist. 26 (1998) 922–942.
[6] W. Härdle, G. Kerkyacharian, D. Picard, A. Tsybakov, Wavelets, Approximation, and Statistical Applications, Springer-Verlag, New York, 1998.
[7] A. Massiani, Étude asymptotique locale de l'estimateur par méthode d'ondelettes, C. R. Math. Acad. Sci. Paris, Ser. I 335 (2002) 553–556.