Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 346 (2008) 999–1002
COMPTES RENDUS MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Statistique/Probabilités

Estimation suroptimale de la densité spectrale par projection adaptative : Application à l'estimation de l'ordre d'une moyenne mobile

Mory Souare

L.S.T.A., Université Paris 6, 175, rue du Chevaleret, 75013 Paris, France

Reçu le 15 août 2007 ; accepté après révision le 2 juillet 2008

Disponible sur Internet le 23 août 2008

Présenté par Paul Deheuvels

Résumé

Dans cette Note on étudie un estimateur par projection adaptative de la densité spectrale. On montre que cet estimateur atteint une vitesse suroptimale sur un ensemble dense dans la classe des densités spectrales et une vitesse quasi-optimale ailleurs. Cet ensemble peut être choisi par le statisticien et la vitesse suroptimale est atteinte pour l'erreur quadratique intégrée et la convergence uniforme presque sûre. On en déduit un estimateur de l'ordre d'une moyenne mobile. ***Pour citer cet article : M. Souare, C. R. Acad. Sci. Paris, Ser. I 346 (2008).***

Abstract

Superoptimal estimator of the spectral density by adaptive projection: an application to the estimator of a moving average order. We study an adaptive estimator of the spectral density by projection. We show that this estimator reaches a superoptimal rate on a dense set in the spectral densities class, and a quasi-optimal rate elsewhere. This set can be chosen by the Statistician, and the superoptimal speed is reached for integrated quadratic error and almost sure uniform convergence. As an application we obtain a consistent estimator of a moving average order. ***To cite this article: M. Souare, C. R. Acad. Sci. Paris, Ser. I 346 (2008).***

1. Estimation de la densité spectrale par projection adaptative

Nous nous intéressons dans cette Note à un processus $X = (X_t)_{t\in\mathbb{Z}}$ défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Le processus X est du second ordre, strictement stationnaire, centré, et fortement mélangeant, de coefficient de mélange vérifiant $\alpha(k) \leqslant a\exp(-bk)$ où $(a > 0,\ b > 0,\ k \geqslant 0)$. Sa fonction d'autocovariance

$$\gamma(j) = \mathbb{E}(X_t X_{t+j}), \quad j \in \mathbb{Z},\ \text{satisfaisant la condition} \sum_{j\in\mathbb{Z}} \bigl|\gamma(j)\bigr| = \sum_{j\in\mathbb{Z}} \bigl|\mathrm{Cov}(X_0, X_j)\bigr| < \infty.$$

Adresse e-mail : souaremory@yahoo.fr.

1631-073X/$ – see front matter
doi:10.1016/j.crma.2008.07.019

Alors la densité spectrale associée à X appartient à l'espace $\mathcal{H} = L^2([-\pi, \pi], \mathcal{B}_{[-\pi,\pi]}, l)$, où l est la mesure de Lebesgue sur $[-\pi, \pi]$, et s'écrit

$$f(\lambda) = \frac{1}{2\pi} \sum_{j \in \mathbb{Z}} \gamma(j) \cos \lambda j, \quad \forall |\lambda| \leqslant \pi.$$

On construit un estimateur adaptatif en définissant

$$\hat{k}_n = \max\bigl\{j\colon\ 0 \leqslant j \leqslant k_n \text{ tel que } \bigl|\hat{\gamma}_n(j)\bigr| \geqslant \theta_n\bigr\}, \tag{1}$$

avec la convention

$$\hat{k}_n = k_n \quad \text{si } \bigl\{j\colon\ 0 \leqslant j \leqslant k_n \text{ tel que } \bigl|\hat{\gamma}_n(j)\bigr| \geqslant \theta_n\bigr\} = \emptyset,$$

$(k_n)_{n\in\mathbb{N}}$, et $(\theta_n)_{n\in\mathbb{N}}$ étant des suites positives choisies par le statisticien. Par ailleurs, on suppose que

$$k_n < n, \quad k_n \underset{n\to\infty}{\longrightarrow} \infty, \quad n^{-1}k_n \underset{n\to\infty}{\longrightarrow} 0, \quad \text{et} \quad \theta_n = n^{-1/2}(\log\log n)^{1/2} \log^{\Gamma} n, \quad \text{avec } \Gamma > 1.$$

L'estimateur de la densité spectrale s'écrit alors

$$\hat{f}_n(\lambda) = \frac{1}{2\pi} \sum_{|j| \leqslant \hat{k}_n} \hat{\gamma}_n(j) \cos \lambda j, \quad \forall |\lambda| \leqslant \pi, \tag{2}$$

où

$$\begin{cases} \hat{\gamma}_n(j) = \dfrac{1}{n-j} \displaystyle\sum_{t=1}^{n-j} X_t X_{t+j} & \text{si } 0 \leqslant j < n, \\ \hat{\gamma}_n(-j) = \hat{\gamma}_n(j) & \\ \hat{\gamma}_n(j) = 0 & \text{si } j \geqslant n. \end{cases}$$

Si $\theta_n = 0$, on retrouve l'estimateur par projection non adaptatif de la densité spectrale

$$f_n(\lambda) = \frac{1}{2\pi} \sum_{|j| \leqslant k_n} \hat{\gamma}_n(j) \cos \lambda j \quad (k_n < n). \tag{3}$$

L'estimateur f_n a été étudié par de nombreux auteurs, voir par exemple [6], et [7].

Maintenant nous distinguons le cas où f admet un développement fini, du cas où celui-ci est infini.

$$\mathcal{F}_0(K) := \bigl\{f \in \mathcal{H}\colon\ \gamma(K) \neq 0 \text{ et } \gamma(j) = 0 \text{ pour } j > K\bigr\},$$
$$\mathcal{F}_0 := \bigcup_{K=0}^{\infty} \mathcal{F}_0(K),$$
$$\mathcal{F}_1 := \mathcal{H} - \mathcal{F}_0.$$

Notons que la famille $\mathcal{F}_0$ des densités spectrales dont le développement de Fourier n'a qu'un nombre fini de termes non nuls correspond à la classe des processus à moyenne mobile finie, c'est-à-dire les processus de la forme

$$X_t = a_0\varepsilon_t + a_1\varepsilon_{t-1} + \cdots + a_K\varepsilon_{t-K}, \quad \forall t \in \mathbb{Z}, \text{ avec } a_0 = 1 \text{ et } a_K \neq 0, \tag{4}$$

où les a_j pour $j = 1, \ldots, K$ sont des réels tels que le polynôme $1 + a_1 z + \cdots + a_K z^K$ n'a pas de zéro à l'intérieur du disque unité voir [5, page 89] ou [8, page 135], et $(\varepsilon_t)_{t\in\mathbb{Z}}$ est un bruit blanc faible, de variance σ^2.

1.1. Les hypothèses

(i) Le processus $(X_t)_{t\in\mathbb{Z}}$ est supposé borné

$$A = \|X_t\|_\infty < \infty.$$

(ii) Dans toute la suite de cette Note, nous supposons que $(\varepsilon_t)_{t\in\mathbb{Z}}$ est un bruit blanc fort de variance σ^2.

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

Proposition 2.1. *Si $f \in \mathcal{F}_0$, alors*

$$\hat{k}_n = K \quad \textit{p.s. pour n assez grand.}$$

Si $f \in \mathcal{F}_1$, alors

$$\hat{k}_n \underset{n\to\infty}{\longrightarrow} \infty \quad \textit{p.s.}$$

Ce résultat montre que, sur $\mathcal{F}_0$, $\hat{k}_n$ est un estimateur convergent de l'ordre d'une moyenne mobile.
Pour préciser le comportement asymptotique de $\hat{k}_n$ quand $f \in \mathcal{F}_1$ on pose

$$q(\eta) = \min\bigl\{q \in \mathbb{N}\colon \bigl|\gamma(j)\bigr| \leqslant \eta \text{ pour tout } j > q\bigr\} \quad \text{avec } \eta > 0.$$

Comme $\gamma(j) \to 0$, alors $q(\eta)$ est toujours bien définie. Dans la suite on posera

$$q_n(\varepsilon) = q\bigl((1+\varepsilon)\theta_n\bigr) \quad \text{avec } \varepsilon > 0 \quad \text{et} \quad q_n'(\varepsilon') = q\bigl((1-\varepsilon')\theta_n\bigr) \quad \text{avec } 0 < \varepsilon' < 1.$$

La proposition suivante donne un encadrement asymptotique presque sûr de $\hat{k}_n$:

Proposition 2.2. *Si $f \in \mathcal{F}_1$, alors $\forall \varepsilon > 0$, $\forall \varepsilon' \in]0,1[$, et $k_n \geqslant q_n(\varepsilon)$, on a*

$$q_n(\varepsilon) \leqslant \hat{k}_n \leqslant \min\bigl(q_n'(\varepsilon'), k_n\bigr) \quad \textit{p.s. pour n assez grand.}$$

3. Comportement asymptotique de $\hat{f}_n$

3.1. Vitesses suroptimales

Sur $\mathcal{F}_0$, $\hat{f}_n$ a les propriétés d'un estimateur paramétrique. En fait, il se comporte comme

$$f_{n,K}(\lambda) = \frac{1}{2\pi} \sum_{|j|\leqslant K} \hat{\gamma}_n(j) \cos \lambda j, \quad \forall |\lambda| \leqslant \pi.$$

L'énoncé suivant résume ses propriétés asymptotiques :

Proposition 3.1. *Si $f \in \mathcal{F}_0$, alors*

(i) $n\,\mathbb{E}\|\hat{f}_n - f\|^2 = \mathcal{O}(1)$,
(ii) $\overline{\lim}_{n\to\infty} \sup_{f\in\mathcal{F}_0(K)} n\,\mathbb{E}\|\hat{f}_n - f\|^2 \leqslant \frac{64a(K+1)A^4}{1-\exp(-b)}$,
(iii) $\overline{\lim}_{n\to\infty} (\frac{n}{\log\log n})^{1/2} \sup_{\lambda\in[-\pi,\pi]} |\hat{f}_n(\lambda) - f(\lambda)| \leqslant \sqrt{2}(1+K)$, *p.s.*

3.2. Normalité asymptotique

Dans cette partie on étudie la normalité asymptotique de $\hat{f}_n$. Remarquons par ailleurs que l'estimateur de la densité spectrale est liée aux autocovariances empiriques, il est donc naturel d'utiliser la normalité asymptotique des autocovariances empiriques, on pourra se servir du corollaire 8.3.1 dans [1] page 467.

Proposition 3.2. *Si $f \in \mathcal{F}_0$, et si $(\varepsilon_t)_{t\in\mathbb{Z}}$ est tel que $\mathbb{E}\varepsilon_t^4 = 3\sigma^4 + \kappa_4 < \infty$, où κ_4 est le cumulant d'ordre 4 de ε_t, alors*

$$\sqrt{n}\bigl(\hat{f}_n(\lambda) - f(\lambda)\bigr) \xrightarrow{\mathcal{L}} \frac{1}{\pi} \sum_{j=1}^{K} N_j \cos \lambda j,$$

où $(N_1, \ldots, N_K)$ est un vecteur gaussien centré de matrice de covariance Σ telle que

$$\Sigma_{j,j'} = 4\pi \int_{-\pi}^{\pi} \cos\lambda j \;\cos\lambda j' f^2(\lambda)\,\mathrm{d}\lambda + \frac{\kappa_4}{\sigma^4} \int_{-\pi}^{\pi} \cos\lambda j f(\lambda)\,\mathrm{d}\lambda \int_{-\pi}^{\pi} \cos\lambda j' f(\lambda)\,\mathrm{d}\lambda, \quad 0 \leqslant j \leqslant j' \leqslant K.$$

3.3. Vitesses quasi-optimales

Nous donnons ci-dessous un encadrement asymptotique de l'erreur quadratique intégrée.

Proposition 3.3. *Si $f \in \mathcal{F}_1$, et si $k_n \geqslant q_n(\varepsilon)$ avec $\varepsilon > 0$, $0 < \varepsilon' < 1$, on a*

(i) $\mathbb{E}\|\hat{f}_n - f\|^2 \leqslant \frac{64aA^4 \min(k_n, q_n'(\varepsilon'))}{n\pi(1-\exp(-b))} + \frac{1}{\pi}\sum_{j \geqslant q_n(\varepsilon)} \gamma^2(j) + \mathcal{O}\big(n^{1-\frac{\log\log n}{64A^4}}\big)$,
(ii) $\mathbb{E}\|\hat{f}_n - f\|^2 \geqslant \frac{64A^4 a q_n(\varepsilon)}{n\pi(1-\exp(-b))} + \frac{1}{\pi}\sum_{j \geqslant \min(q_n'(\varepsilon'), k_n)} \gamma^2(j) + \mathcal{O}\big(n^{1-\frac{\log\log n}{64A^4}}\big)$.

Corollaire 3.4. *Si $f \in \mathcal{F}_1$, et si $k_n > q_n(\varepsilon)$ avec $\varepsilon > 0$, et $\Gamma > 1$, on a*

$$\mathbb{E}\|\hat{f}_n - f\|^2 = \mathcal{O}\big(n^{-1}k_n(\log\log n)\log^{2\Gamma} n\big) + \frac{1}{\pi}\sum_{j>k_n} \gamma^2(j).$$

Ainsi, ce résultat nous permet de faire une comparaison avec l'estimateur non adaptatif de la densité spectrale f_n, en effet

$$\mathbb{E}\|f_n - f\|^2 = \mathcal{O}\big(n^{-1}k_n\big) + \frac{1}{\pi}\sum_{j>k_n} \gamma^2(j).$$

On peut donc espérer que la perte de vitesse de $\hat{f}_n$ par rapport à f_n est en $\big(\log\log n\big)\log^{2\Gamma} n$.

4. Commentaires

Pour les preuves de nos résultats, nous utilisons principalement les inégalités de [9, page 9], et [3, page 25] ainsi que le théorème 4.1 dans [4]. Notons par ailleurs qu'un estimateur analogue pour la densité figure dans [2].

Remerciements

Je tiens sincèrement à remercier R. Ignaccolo et J.B. Aubin pour les divers échanges très instructifs que nous avons eu ensemble et dont les critiques ont permis d'améliorer le contenu de cette Note. Je remercie également le professeur D. Bosq et les référés anonymes pour leurs suggestions et remarques.

Références

[1] T.W. Anderson, The Statistical Analyse of Time Series, Wiley, New York, 1971.
[2] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.
[3] D. Bosq, Estimation suroptimale de la densité par projection, Canad. J. Statist. 33 (1) (2005) 21–37.
[4] D. Bosq, Nonparametric Statistics for Stochastic Processes. Estimation and Prediction, second ed., Lecture Notes in Statist., vol. 110, Springer-Verlag, New York, 1998.
[5] P.J. Brokwell, R.A. Davis, Time Series: Theory and Methods, Springer-Verlag, New York, 1991.
[6] M. Carbon, Sur la convergence uniforme presque sûre des estimateurs de la densité spectrale des processus stationnaires et mélangeants, C. R. Acad. Sci. Paris, Sér. I 292 (1981) 95–98.
[7] V.F. Gaposhkin, Almost sure convergence of estimates for the spectral density of a stationary process, Teor. Veroyatnost. i Primenen. 25 (1980) 172–178.
[8] M.B. Priestley, Spectral Analysis and Times Series, vol. 2, Academic Press, 1981.
[9] E. Rio, Théorie asymptotique des processus aléatoires faiblement dépendants, Mathematics & Applications, vol. 31, Springer-Verlag, Berlin, 2000.