Statistique et analyse des données

Edurne Biritxinaga

Consistance forte de l'estimation de la moyenne d'un processus stochastique par interpolation spline

Statistique et analyse des données, tome 15, nº 3 (1990), p. 47-59

<http://www.numdam.org/item?id=SAD_1990__15_3_47_0>

Statistique et Analyse des Données
1990 - Vol. 15 n° 3 p. 47 - 59

CONSISTANCE FORTE DE L'ESTIMATION DE LA MOYENNE D'UN PROCESSUS STOCHASTIQUE PAR INTERPOLATION SPLINE

Edurne BIRITXINAGA

Laboratoire de Statistique et Probabilités
U.A. 1204 CNRS
Université de Pau, Av. de l'Université.
64000 Pau, France.

Résumé

On montre dans cet article la consistance forte de l'estimation par fonction spline de la fonction moyenne m d'un processus stochastique du second ordre vérifiant la loi du logarithme itéré dans C[0,T]. La fonction moyenne est supposée régulière ($m \in H^k[0,T]$). Les estimations considérées sont alors des fonctions splines d'interpolation de la moyenne empirique issue de n réalisations indépendantes du processus connues en un nombre fini de points de discrétisation. On étudie aussi le cas où les trajectoires du processus sont non seulement continues mais appartiennent elles-mêmes à l'espace de Sobolev $H^k[0,T]$.

Mots clés : Espace de Sobolev, interpolation spline, estimation, consistance.

Classification AMS : 62 G 05, 60 G 12, 65 D 07

Abstract

In this paper we prove the strong consistency of the spline estimation of the mean function m of a second order stochastic process satisfying the iterated logarithm law in C[0,T]. We assume that the mean function is smooth (that is $m \in H^k[0,T]$). The estimations considered are then obtained by spline interpolation of the empirical mean defined from n independent realizations of the process at a finite set of discretization points. We also study the case when the paths of the process are not only continuous but in the Sobolev space $H^k[0,T]$.

Keywords : Sobolev space, spline interpolation, estimation, consistency.

Manuscrit reçu le 18 mai 1990, révisé le 6 novembre 1990.

1. INTRODUCTION

Les fonctions splines sont utilisées dans des domaines divers et de plus en plus nombreux en Statistique, par exemple, en régression non-paramétrique (cf. Wahba (1978), Nussbaum (1985), Silverman (1985),...), séries chronologiques (cf. Kimeldorf et Wahba (1970), Peele et Kimeldorf (1977),...), estimation de densité (cf. Cogburn et Davis (1974), De Montricher et al. (1975),...), et plus récemment pour l'approximation spline de l'Analyse en Composantes Principales d'un processus (cf. Besse et Ramsay (1986)).

Dans ce dernier cas, on suppose systématiquement le processus centré. Or ce n'est généralement pas le cas : la moyenne est inconnue et estimée point par point par la moyenne empirique.

Lorsque les trajectoires du processus sont "lisses", la moyenne possède la même propriété. Inversement, on rencontre des processus dont les trajectoires sont peu régulières (ou même ne le sont pas du tout : processus de naissance et de mort,...) mais où la moyenne est cependant lisse (cf. par exemple Ycart (1988)). Il parait alors souhaitable d'estimer cette moyenne par une fonction ayant le même degré de régularité.

La question qu'on se propose d'étudier dans cet article a pour origine un problème statistique : la consistance de l'estimation par fonction spline de la fonction moyenne m(.) d'un processus stochastique observé en un nombre fini d'instants. La construction de fonctions de prédiction et d'estimation de la fonction moyenne d'une série temporelle par des fonctions splines a déjà été traitée par Kimeldorf et Wahba (1970) et par Peele et Kimeldorf (1977), (1979), mais les problèmes de consistance n'ont pas été abordés à notre connaissance.

Soit X un processus stochastique défini sur un espace de probabilité $(\Omega',\mathcal{A}',P')$ et à trajectoires dans C[0,T]. On considère un échantillon $(X_i)_{i\in\mathbb{N}}$ de X, suite de variables aléatoires indépendantes, identiquement distribuées et de même loi que X. L'espace de probabilité considéré est donc $(\Omega',\mathcal{A}',P')^{\otimes\mathbb{N}}$, noté plus simplement $(\Omega,\mathcal{A},P)$. On pose alors $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, moyenne empirique issue de l'échantillon, qui ne sera supposée connue qu'aux instants t_i d'un sous-ensemble fini T_q de cardinal fini q de l'intervalle [0,T].

On va s'intéresser à des processus possédant une fonction moyenne m "régulière" et, plus précisement, appartenant à un espace de Sobolev $H^k[0,T]$. L'idée est donc d'estimer la moyenne (et éventuellement ses dérivées, qui ont dans certains cas une interprétation physique) en interpolant par splines la moyenne empirique de l'échantillon en un nombre fini d'instants. Ainsi, $n \in \mathbb{N}^*$, T_q et $\omega \in \Omega$ étant fixés, on note $H^{n,q}(\omega)$ le sous-espace affine de $H^k[0,T]$ défini par :

$$H^{n,q}(\omega) = \{u \in H^k[0,1] : \forall t_i \in T_q,\ u(t_i)=\overline{X}^n(\omega,t_i)\},$$

et on note $\hat{m}^{n,q}(\omega,.)$ l'élément de $H^k[0,T]$ tel que :

$$(P)\begin{cases} \hat{m}^{n,q}(\omega,.) \in H^{n,q}(\omega) \\ \forall v \in H^{n,q}(\omega),\ \| D^k \hat{m}^{n,q}(\omega,.) \|_{L^2[0,T]} \leq \| D^k v \|_{L^2[0,T]}. \end{cases}$$

L'interpolante $\hat{m}^{n,q}$ possède la même régularité que la moyenne théorique et conduit ainsi à une estimation a priori raisonnable. Mais une estimation de ce type ne sera vraiment justifiée que si elle est consistante. On se propose donc de donner des conditions suffisantes de convergence presque sûre de l'estimateur spline vers m, si possible dans $H^k[0,T]$. On va considérer pour cela des processus stochastiques vérifiant la loi du logarithme itéré dans C[0,T], celle-ci étant notre principale hypothèse d'un point de vue statistique.

A l'aide de quelques résultats d'approximation fonctionnelle regroupés dans le §3, on montre (Théorème 4.1) que dans les conditions ci-dessus, la suite $\{\hat{m}^{n,q}\}$ converge presque sûrement vers m dans $H^k[0,T]$ si la taille n de l'échantillon et le nombre q de points de discrétisation tendent vers l'infini de telle manière que $q^k n^{-1/2}(\log\log n)^{1/2}$ tende vers zéro, cette dernière condition étant exprimée de façon rigoureuse à travers une base de filtre sur $\mathbb{N}^* \times \mathbb{N}^*$.

On montre aussi (Théorème 4.2) que si les trajectoires du processus sont non seulement continues mais appartiennent en plus à l'espace de Sobolev $H^k[0,T]$, alors la suite $\{\hat{m}^{n,q}\}$ converge presque sûrement vers m(.) dans $H^k[0,T]$ lorsque n et q tendent vers l'infini indépendamment, même si le processus ne satisfait pas la loi du logarithme itéré.

2. NOTATIONS

Soit $k \in \mathbb{N}^*$ et soit $[0,T]$ un intervalle borné non vide de $\mathbb{R}$. Pour tout $\xi=(\xi_1,\ldots,\xi_q)$ et $\eta=(\eta_1,\ldots,\eta_q)$ de $\mathbb{R}^q$, on note :

$$<\xi \ , \ \eta>_q \ = \sum_{i=1}^{q} \xi_i \ \eta_i \quad \text{le produit scalaire usuel dans } \mathbb{R}^q ,$$

et

$$<\xi>_q \ = \ <\xi \ , \ \xi>_q^{1/2} \quad \text{la norme correspondante} .$$

Soit $H^k[0,T]$ l'espace de Sobolev des (classes de) fonctions de $L^2[0,T]$ dont toutes les dérivées (au sens des distributions) d'ordre inférieur ou égal à k sont des éléments de $L^2[0,T]$.

Si on note $(.,.)_0$ le produit scalaire usuel de $L^2[0,T]$, $H^k[0,T]$ muni du produit scalaire :

$$\forall u,v \in H^k[0,T], \quad ((u , v))_k \ = \sum_{i=0}^{k} (D^i u , D^i v)_0 ,$$

et de la norme associée :

$$\forall u \in H^k[0,T], \qquad \| u \|_k = ((u , u))_k^{1/2} ,$$

est un espace de Hilbert réel séparable. On va utiliser aussi le semi-produit scalaire sur $H^k[0,T]$ défini par $(u,v)_k = (D^k u , D^k v)_0$ et la semi-norme associée $| u |_k = (u , u)_k^{1/2}$.

Soit $T_q = \{t_1 ,\ldots, t_q\}$ un ensemble de $q \geq k$ points distincts de $[0,T]$. On introduit l'opérateur A_q de $\mathcal{L}(H^k[0,T], \mathbb{R}^q)$, défini par :

$$\forall u \in H^k[0,T], \quad A_q u = (u(t_1),\ldots, u(t_q)) .$$

La continuité de A_q découle de l'injection de Sobolev de $H^k[0,T]$ dans l'espace des fonctions continues sur $[0,T]$ (cf. par exemple Adams, 1975).

Soit u une fonction réelle définie sur $[0,T]$. Pour un T_q donné, on appelle fonction spline d'interpolation $\Sigma_q u$ la fonction "la plus lisse" rencontrant u aux points t_i de T_q :

$$\hat{u}^q = \Sigma_q u = \arg\min \ \{ \ | v |_k : v \in H^k[0,T] \ , A_q v = A_q u \ \}.$$

L'existence et unicité de $\hat{u}^q$ est garantie dès que $q \geq k$ (cf. Laurent, 1972). De plus, $\hat{u}^q$ appartient à un sous-espace S_q (de dimension finie q) de $H^k[0,T]$ appelé sous-espace des fonctions splines. Cet espace est engendré par la base $\{ \sigma_i^q \; ; \; i = 1,\ldots,q\}$ où σ_i^q est la fonction spline d'interpolation telle que, pour tout t_j de T_q $(j = 1,\ldots,q)$:

$$\sigma_i^q (t_j) = \delta_{ij}$$

Il est clair que:

$$\forall \; s \in S_q \; , \; s = \sum_{i=1}^{q} s(t_i)\sigma_i^q \, ,$$

et par conséquent, en notant W_q la matrice symétrique semi-définie positive de terme général $(\sigma_i^q, \sigma_j^q)_k$, on a :

$$\forall \; s_1, s_2 \in S_q \; , \; (s_1,s_2)_k = \langle W_q A_q s_1 \; , \; A_q s_2 \rangle_q \, .$$

3. QUELQUES RÉSULTATS D'APPROXIMATION FONCTIONNELLE

On regroupe dans ce paragraphe les résultats techniques qu'on va utiliser dans le §4 pour montrer la consistance forte des $\hat{m}^{n,q}$ comme estimateurs de la moyenne m d'un processus stochastique.

La convergence des fonctions splines d'interpolation pour des données non bruitées a été étudiée par de nombreux auteurs (cf. par exemple Atteia (1966), Schultz (1973), Prenter (1975), Schumaker (1981)). On va énoncer ici (Théorème3.4) la version unidimensionnelle d'un résultat de convergence plus général dû à Arcangeli (1986), qui montre que si la distance de Hausdorff de T_q à $[0,T]$ tend vers zéro, alors la suite des fonctions splines d'interpolation de f converge vers f dans $H^k[0,T]$. On va utiliser aussi (Théorème 3.6) un résultat important montré par Utreras (1983), concernant le comportement en fonction de q des valeurs propres associées aux fonctions splines construites sur une partition régulière de l'intervalle [0,T].

On définit h_{max} et h_{min} par :

$$h_{max} = \sup_{t\in[0,T]} \; \inf_{t_i \in T_q} | t - t_i | \qquad \text{et} \qquad h_{min} = \min_{t_i \neq t_j} | t_i - t_j | \, .$$

On considère les hypothèses suivantes concernant la régularité asymptotique des T_q :

Il existe une constante $C > 0$ telle que :

(H1) $\dfrac{h_{max}}{h_{min}} \leq C$

(H2) $\lim\limits_{q \to \infty} h_{max} = 0$.

Soit T_0 un sous-ensemble fixé d'au moins k points distincts de l'intervalle [0,T]. Compte tenu de l'hypothèse (H2), pour chaque $q \geq k$, on peut associer à chaque point t de T_0 un point t^q de T_q tel que $t = \lim\limits_{q \to \infty} t^q$. Pour tout q, on désigne par T_0^q l'ensemble des points t^q de T_q ainsi associés à T_0 et on considère la norme sur $H^k[0,T]$ définie par :

$$\forall u \in H^k[0,T], \quad |||\, u \,|||_q^2 = \sum_{t_0^q \in T_0^q} (u(t_0^q))^2 + |\, u\,|_k^2 .$$

D'après (Necas, 1967), $|||\,.\,|||_q$ est équivalente à la norme $\|\,.\,\|_k$ sur $H^k[0,T]$. On peut énoncer alors les deux lemmes suivants :

Lemme 3.2 (ARCANGELI, 1986)

Sous l'hypothèse (H2), il existe un $q_0 \geq k$ tel que pour tout $q \geq q_0$, $|||\,.\,|||_q$ est une norme sur $H^k[0,T]$ <u>uniformément</u> équivalente par rapport à q à la norme $\|\,.\,\|_k$.

Lemme 3.3

Sous l'hypothèse (H2), il existe un $q_0 \geq k$ tel que la suite d'opérateurs d'interpolation spline $\{\Sigma_q : q \geq q_0\}$ est uniformément bornée dans $\mathcal{L}(H^k[0,T])$.

Démonstration : Compte tenu de la définition de l'opérateur d'interpolation spline Σ_q ,

$\forall q \geq k$, $\forall u \in H^k[0,T]$,

$$|||\, \Sigma_q u \,|||_q^2 = \sum_{t_0^q \in T_0^q} (\Sigma_q u(t_0^q))^2 + |\, \Sigma_q u\,|_k^2 \leq \sum_{t_0^q \in T_0^q} (u(t_0^q))^2 + |\, u\,|_k^2 = |||\, u \,|||_q^2 ,$$

et on obtient le résultat par application du Lemme 3.2. ∎

Théorème 3.4 (ARCANGELI, 1986)

Si $\{T_q\ ;\ q\geq k\}$ est une suite de noeuds vérifiant l'hypothèse (H2), alors :

$$\forall\ u \in H^k[0,T],\quad \lim_{q\to\infty} \| \hat{u}^q - u \|_k = 0\ .$$

On signale que tous les résultats ci-dessus sont établis sous la seule hypothèse (H2). Cependant, on a besoin de l'hypothèse (H1) pour énoncer le théorème suivant, dû à Utreras (1988). Ce théorème montre le comportement en fonction de q des valeurs propres de la matrice W_q introduite dans le §2 et généralise au cas d'un découpage irrégulier le résultat précédemment démontré par le même auteur (Utreras, 1983) pour des noeuds équidistants.

Théorème 3.6 (UTRERAS, 1988)

Notons γ_i^q, i=1,...,q, les valeurs propres de la matrice W_q, rangées par ordre croissant. Alors, sous les hypothèses (3.1) et (3.2),

$$\gamma_1^q = \ldots = \gamma_k^q = 0$$

et il existe deux constantes positives C_1 et C_2 indépendantes de q telles que :

$$\exists\ q_0 \in \mathbb{N}^* : \forall\ q \geq q_0,\ C_1\, q^{-1}\, i^{2k} \leq \gamma_i^q \leq C_2\, q^{-1}\, i^{2k},\quad i = k+1,\ldots, q$$

4. APPLICATION À L'ESTIMATION DE LA MOYENNE D'UN PROCESSUS STOCHASTIQUE

Dans tout le paragraphe, on suppose que X est un processus du second ordre à trajectoires dans C[0,T]; ceci s'écrit :

(H3) $\quad E[\| X\|_\infty^2] < \infty.$

On suppose aussi que la fonction moyenne m du processus est régulière et, plus précisement, que :

(H4) $\quad m \in H^k[0,T]\ .$

On suppose enfin que le processus vérifie la loi du logarithme itéré dans C[0,T], explicitée ci-dessous. Avec les notations de l'introduction, posons $S_n = X_1 + ... X_n - nm$. On suppose donc qu'il existe un ensemble borné symétrique K dans C[0,T] tel que, en notant, pour toute suite $\{x_n\}$, $C(\{x_n\})$ l'ensemble de tous les points limites de la suite et en notant $LLn = \log \log n$ pour $n \geq 3$ et 1 pour $n = 1, 2$:

$$\text{(H5)} \qquad P\left\{ C\left(\left\{\frac{S_n}{(2nLLn)^{1/2}}\right\}\right) = K \right\} = 1,$$

En fait, K peut être caractérisé comme la boule unité fermée de l'espace reproduisant associé à la loi de X.

La loi du logarithme itéré pour des processus à trajectoires dans un espace de Banach a été étudiée par de nombreux auteurs. Des conditions suffisantes dans notre cas (dans C[0,T]) ont été établies par exemple par Kuelbs (1976a) et par Carmona et Kôno (1976) dans le cas Gaussien. De plus, sous l'hypothèse (H3), K est nécessairement compact dans C[0,T] (Cf. Kuelbs, 1976b).

Il existe alors une constante positive C telle que :

$$(4.1) \qquad P\left\{ \limsup_n \left\| \frac{\overline{X}^n - m}{(2n^{-1}LLn)^{1/2}} \right\|_\infty \leq C \right\} = 1.$$

Pour énoncer le principal théorème de ce paragraphe, on a besoin d'introduire encore quelques notations.

On définit :

$$\forall q \in \mathbb{N}^*, \ \forall n \in \mathbb{N}^*, \ \psi(q,n) = q^k n^{-1/2} (LL\, n)^{1/2}.$$

On pose alors :

$$\forall i \in \mathbb{N}^*, \ \forall j \in \mathbb{N}^*, \ \forall \theta > 0, \ B(i,j,\theta) = \{(q,n) \in \mathbb{N}^* \times \mathbb{N}^* ;\ q \geq i,\ n \geq j,\ \psi(q,n) < \theta\},$$

et

$$\mathfrak{B} = \{ B(i,j,\theta) ;\ i \in \mathbb{N}^*, j \in \mathbb{N}^*, \theta > 0 \}.$$

On vérifie aisément que $\mathfrak{B}$ est une base de filtre sur $\mathbb{N}^*\times\mathbb{N}^*$. De plus, pour une fonction Φ quelconque sur $\mathbb{N}^*\times\mathbb{N}^*$, notons $\lim_{\mathfrak{B}} \Phi$ la limite de Φ suivant la base de filtre $\mathfrak{B}$; il est clair que :

$$\lim_{\mathfrak{B}} \psi = 0 .$$

Soit $\hat{m}^{n,q}$ l'estimation spline solution du problème (P) énoncé dans l'introduction de cet article. Le théorème suivant montre que sous les hypothèses (H1) et (H2) assurant la régularité asymptotique de la partition induite par les T_q et sous les hypothèses (H3) à (H5) concernant le processus, la suite $\{\hat{m}^{n,q}\}$ converge presque sûrement vers m dans $H^k[0,T]$ si n et q tendent vers l'infini de telle manière que $q^k n^{-1/2}(LL\ n)^{1/2}$ tende vers zéro .

Théorème 4.1

Sous les hypothèses (H1) à (H5),

$$\lim_{\mathfrak{B}} \hat{m}^{n,q} = m \quad \text{dans } H^k[0,T], \text{ presque sûrement.}$$

Démonstration :

$$\text{(4.2)} \qquad \forall q\geq k,\ \forall n\in\mathbb{N}^*,\ \forall \omega\in\Omega,\ \| \hat{m}^{n,q}(\omega,.) - m \|_k \leq \| \hat{m}^{n,q}(\omega,.) - \hat{m}^q \|_k + \| \hat{m}^q - m \|_k.$$

Etant donné que m appartient à $H^k[0,T]$, par application du Théorème 3.4, on a :

$$\text{(4.3)} \qquad \lim_{q\to\infty} \| \hat{m}^q - m \|_k = 0.$$

De plus, par application du Lemme 3.2 :

$$\text{(4.4)} \qquad \begin{cases} \exists C>0,\ \exists q_1\in\mathbb{N}^* : \forall q\geq q_1, \forall n\in\mathbb{N}^*,\ \forall \omega\in\Omega, \\ \| \hat{m}^{n,q}(\omega,.) - \hat{m}^q \|_k^2 \leq C \displaystyle\sum_{t_0^q\in T_0^q} (\hat{m}^{n,q}(\omega,t_0^q) - \hat{m}^q(t_0^q))^2 + |\hat{m}^{n,q}(\omega,.) - \hat{m}^q|_k^2 . \end{cases}$$

Mais, compte tenu de la définition de $\hat{m}^{n,q}(\omega,.)$ et $\hat{m}^q$:

$$\text{(4.5)} \qquad \forall q\geq k,\ \forall n\in\mathbb{N}^*,\ \forall \omega\in\Omega,\ \sum_{t_0^q\in T_0^q} (\hat{m}^{n,q}(\omega,t_0^q) - \hat{m}^q(t_0^q))^2 = \sum_{t_0^q\in T_0^q} (\overline{X}^n(\omega,t_0^q) - m(t_0^q))^2$$

et, puisque $\hat{m}^{n,q}(\omega,.)$ et $\hat{m}^q$ sont dans S_q, on a aussi (cf. §2) :

$$\begin{cases} \forall\, q\geq k,\ \forall\, n\in \mathbb{N}^*,\ \forall\, \omega\in\Omega, \\ |\hat{m}^{n,q}(\omega,.) - \hat{m}^q|_k^2 = \langle W_q A_q(\hat{m}^{n,q}(\omega,.) - \hat{m}^q),\ A_q(\hat{m}^{n,q}(\omega,.) - \hat{m}^q)\rangle_q , \end{cases}$$

soit encore :

$$\begin{cases} \forall\, q\geq k,\ \forall\, n\in \mathbb{N}^*,\ \forall\, \omega\in\Omega, \\ |\hat{m}^{n,q}(\omega,.) - \hat{m}^q|_k^2 \leq \gamma_q^q \sum\limits_{t_i\in T_q} (\overline{X}^n(\omega,t_i) - m(t_i))^2. \end{cases}$$

Alors, par application du Théorème 3.6 :

$$\begin{cases} \exists C>0,\ \exists q_2\in \mathbb{N}^* : \forall\, q\geq q_2,\ \forall\, n\in \mathbb{N}^*,\ \forall\, \omega\in\Omega, \\ |\hat{m}^{n,q}(\omega,.) - \hat{m}^q|_k^2 \leq C\, q^{2k-1} \sum\limits_{t_i\in T_q} (\overline{X}^n(\omega,t_i) - m(t_i))^2. \end{cases} \tag{4.6}$$

Compte tenu de (4.1), il existe un ensemble Ω^* de probabilité 1 tel que :

$$\begin{cases} \exists C>0,\ (\forall\, \omega\in\Omega^*,\ \exists n_0(\omega)\in \mathbb{N}^*) : \forall\, n\geq n_0(\omega), \\ \sum\limits_{t_i\in T_q} (\overline{X}^n(\omega,t_i) - m(t_i))^2 < C\, q\, n^{-1} LLn, \\ \sum\limits_{t_0^q\in T_0^q} (\overline{X}^n(\omega,t_0^q) - m(t_0^q))^2 < C\, n^{-1} LLn\ . \end{cases} \tag{4.7}$$

Posons $q_0 = \max(q_1,q_2)$; on déduit des expressions (4.4) à (4.7) :

$$\begin{cases} \exists C>0,\ (\forall\, \omega\in\Omega^*,\ \exists n_0(\omega)\in \mathbb{N}^*) : \forall\, q\geq q_0,\ \forall\, n\geq n_0(\omega), \\ \| \hat{m}^{n,q}(\omega,.) - \hat{m}^q\|_k^2 \leq C\, (n^{-1} LLn + (\psi(q,n))^2) \end{cases}$$

Le théorème découle alors de (4.2) et (4.3). ∎

De plus, le théorème 4.1 peut être considérablement amélioré si les trajectoires du processus, au lieu d'être seulement continues, appartiennent elles-mêmes à l'espace de Sobolev $H^k[0,T]$. Dans ce cas, remplaçons (H4) et (H5) par l'hypothèse suivante :

(H6) Le processus X(.) est à trajectoires dans $H^k[0,T]$.

Sous l'hypothèse (H6), la suite $\{\overline{X}_n\}$ et la fonction moyenne m sont dans $H^k[0,T]$. De plus, $H^k[0,T]$ étant séparable et compte tenu de l'hypothèse (H3), la loi forte des grands nombres dans $H^k[0,T]$ permet d'écrire :

$$\lim_{n\to\infty} \| \overline{X}_n - m \|_k = 0, \text{ presque sûrement.} \tag{4.8}$$

Le théorème ci-dessous montre que, dans ce cas, la suite $\{ \hat{m}^{n,q} \}$ converge presque sûrement vers m dans $H^k[0,T]$ lorsque n et q tendent vers l'infini indépendamment.

Théorème 4.2

Sous les hypothèses (H2), (H3) et (H6),

$$\lim_{n,q\to\infty} \hat{m}^{n,q} = m \quad \text{dans } H^k[0,T], \text{ presque sûrement.}$$

Démonstration :

$$\| \hat{m}^{n,q} - m \|_k \leq \| \Sigma_q (\overline{X}_n - m) \|_k + \| \Sigma_q m - m \|_k$$

$$\leq \| \Sigma_q \|_{\mathcal{L}(H^k[0,T])} \| \overline{X}_n - m \|_k + \| \Sigma_q m - m \|_k$$

et, compte tenu de (4.8), on obtient le résultat par application directe du Lemme 3.3 et du Théorème 3.4. ∎

5. CONCLUSION

Des résultats de convergence similaires à ceux obtenus ici pour l'estimation par spline d'interpolation sont nécessaires également dans le cas où la fonction moyenne m(.) est estimée par spline d'ajustement à partir des valeurs $\overline{X}_n(t_1),\ldots, \overline{X}_n(t_q)$. Si $\lambda_0 > 0$ est un réel fixé, on peut montrer que sous des hypothèses similaires à celles considérées dans cet article, la famille des splines d'ajustement $\{\hat{m}^{n,q}_\lambda; (n,q,\lambda)\in\mathbb{N}^*\times\mathbb{N}^*\times]0,\lambda_0]\}$ converge presque sûrement vers m dans $H^k[0,T]$ (cf. Biritxinaga, 1987 et 1988).

Les résultats de cet article sont le fruit de nombreuses discussions avec B. YCART et R. ARCANGELI, que l'auteur remercie vivement.

Références

Adams R.A. (1975) "Sobolev Spaces", Academic Press, New York.

Atteia M. (1966) "Etude de certains noyaux et théorie des fonctions spline en Analyse Numérique", thèse, Grenoble.

Arcangeli R. (1986) "D^m-splines sur un domaine borné de $\mathbb{R}^n$", U.A.1204 CNRS, Publ. n° 1986/2, Université de Pau.

Besse P., Ramsay J. (1986) "Principal Components analysis of sampled functions", Psychometrika 51, 2, 285-311.

Biritxinaga E. (1987) "Estimation spline de la moyenne d'une fonction aléatoire", Thèse, Université de Pau et des Pays de l'Adour.

Biritxinaga E. (1988) "Ajustement spline et estimation de la moyenne d'une fonction aléatoire", U.A. 1204 CNRS, Publ.n° 88/4, Université de Pau.

Carmona R., Kôno N. (1976) "Convergence en loi et lois du logarithme itéré pour les vecteurs gaussiens", Z. Wahrsch. Gebiete 36, 241-267.

Cogburn R., Davis H.T. (1974) "Periodic splines and spectral density estimation", Ann. Stat. 2, 1108-1126.

De Montricher G.F., Tapia R.A., Thompson J.R. (1975) Nonparametric maximum likelihood estimation of probability densities by penalty function methods", Ann. Stat. 3, 1329-1348.

Kimeldorf G., Wahba G. (1970) "Spline Functions and Stochastic Processes", Sankhya A, 32, 173-180.

Kuelbs J. (1976a) "The Law of the Iterated Logarithm in C[0,1]", Z. Wahrsch. Gebiete 3, 221-235.

Kuelbs J. (1976b) "A strong convergence theorem for Banach space valued random variables", Ann. Prob. 4, No. 5, 744-771

Laurent P.J. (1972) "Approximation et Optimisation", Hermann, Paris.

Necas J. (1967) "Les méthodes directes en théorie des équations elliptiques", Masson, Paris.

Nussbaum M. (1985) "Spline smoothing in regression models and asymptotic efficiency in L^2", Ann. Stat. 13, 3, 984-997.J.A.S.A. 79, 3,

Peele L., Kimeldorf G. (1977) "Prediction functions and mean-estimation functions for a time series", Ann. Statist. 5, No.4, 709-721.

Peele L., Kimeldorf G. (1979) "Time series prediction functions based on imprecise observations", Ann. Statist. 7, No.4 , 801-811.

Prenter P.M. (1975) "Splines and Variational Methods", Wiley, New York.

Schultz M.H.(1973) "Spline Analysis", Prentice Hall, Englewood Cliffs, N.J.

Schumaker L.L. (1981) "Spline Functions : Basic Theory", Wiley, New York.

Silverman B.W. (1985) "Some aspects of the spline smoothing approach to nonparametric regression in curve fitting", J. R. Stat. Soc. B, 47, 1, 1-52.

Utreras F. (1983) "Natural Spline Functions, their Associate Eigenvalue Problem", Numer. Math. Vol.42 , 107-117.

Utreras F.(1988) "Convergence Rates for Multivariate Smoothing Spline Functions", J. Approximation Theory Vol. 52, 1-27.

Wahba G. (1978) "Improper priors, spline smoothing and a problem of guarding againts model errors in regression, J. R. Stat. Soc. B, 3, 364-372.

Ycart B. (1988) "A characteristic property of linear growth birth and death processes" Sankhya A, vol.50, 184-189.