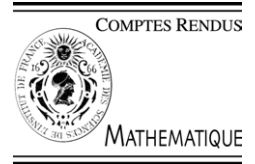




Available online at [www.sciencedirect.com](http://www.sciencedirect.com)



C. R. Acad. Sci. Paris, Ser. I 340 (2005) 389–392



<http://france.elsevier.com/direct/CRASS1/>

Statistics/Probability Theory

# Functional time series prediction via conditional mode estimation

Frédéric Ferraty<sup>a</sup>, Ali Laksaci<sup>b</sup>, Philippe Vieu<sup>a</sup>

<sup>a</sup> *Laboratoire de statistique et probabilités, Université Paul Sabatier, 31062 Toulouse, France*

<sup>b</sup> *Université Djillali Liabes, Sidi Bel Abbès, Algeria*

Received 9 July 2004; accepted after revision 18 January 2005

Presented by Paul Deheuvels

---

## Abstract

This Note focuses on an estimator of the conditional mode of a scalar response  $Y$  given a functional random variable  $X$ . We start by building a kernel estimator of the conditional density of  $Y$  given  $X$ ; the conditional mode is defined as the value which maximizes this conditional density. We establish the almost complete convergence for this estimate under  $\alpha$ -mixing assumption. *To cite this article: F. Ferraty et al., C. R. Acad. Sci. Paris, Ser. I 340 (2005).*

© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

## Résumé

**Prédiction de séries chronologiques par l'estimation en mode conditionnel.** On établit la convergence presque-complète de l'estimateur du mode de la distribution d'une variable réelle  $Y$  conditionnée par une variable fonctionnelle  $X$ . Le mode conditionnel est estimé par la valeur qui maximise l'estimateur à noyau de la densité conditionnelle de  $Y$  sachant  $X$ . Des résultats asymptotiques concernant cet estimateur sont établis sous l'hypothèse  $\alpha$ -mélangeante, rendant nos résultats opérationnels en prédiction de séries chronologiques. *Pour citer cet article : F. Ferraty et al., C. R. Acad. Sci. Paris, Ser. I 340 (2005).*

© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

---

## 1. Introduction

Often, prediction of a scalar response  $Y$  knowing an explained multivariate variable  $X$  is obtained by estimating the conditional expectation of  $Y$  given  $X$  (which is the standard regression function). However, this method is not efficient in some pathological situations. For instance, this is the case when the conditional density of  $Y$  given  $X$  is either unsymmetric or has several modes. In these latter cases, a pertinent predictor of  $Y$  is obtained by estimating the conditional mode (see Collomb et al. [2]). Recent advances concerning the conditional mode can be found for instance in Berlinet et al. [1] and Louani and Ould-Saïd [8].

---

*E-mail addresses:* [ferraty@cict.fr](mailto:ferraty@cict.fr) (F. Ferraty), [alilak@yahoo.fr](mailto:alilak@yahoo.fr) (A. Laksaci), [vieu@cict.fr](mailto:vieu@cict.fr) (P. Vieu).

1631-073X/\$ – see front matter © 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.  
doi:10.1016/j.crma.2005.01.016

On the other hand, more and more works deal with nonparametric approaches for variables valued in an infinite dimensional space. Gasser et al. [6] proposed an estimate of the mode of a distribution of random curves whereas Ferraty et al. [3] studied a kernel estimator in the functional regression setting. Note also that a kernel estimator for functional conditional distribution has been introduced by Ferraty et al. [5]. An overview on nonparametric methods for functional data can be found in Ferraty and Vieu [4]. The aim of this Note is to predict a scalar response  $Y$  given an explanatory functional variable  $X$  via the conditional mode estimate under  $\alpha$ -mixing assumption.

## 2. Model

Let  $(X_i, Y_i)_{i=1, \dots, n}$  be a stationary  $\alpha$ -mixing process valued in  $\mathfrak{S} \times \mathbb{R}$  where  $\mathfrak{S}$  is a semi-metric vector space,  $d(\cdot, \cdot)$  denoting the semi-metric. Assume there exists a regular version of the conditional probability of  $Y$  given  $X$ . Assume that for a given  $x$  there is some compact subset  $S = (\theta - \xi, \theta + \xi)$ ,  $\xi > 0$ , such that the conditional density of  $Y$  given  $X = x$  has a unique mode  $\theta$  on  $S$ . In the remaining of the paper  $x$  is fixed in  $\mathfrak{S}$  and  $N_x$  denotes a neighborhood of  $x$ . Let  $f^x$  (resp.  $f^{x(j)}$ ) be the conditional density (resp. the  $j$ th order derivative of the conditional density) of the variable  $Y$  given  $X = x$ . We define the kernel estimator  $\hat{f}^x$  of  $f^x$  as follows:

$$\hat{f}^x(y) = \frac{h_H^{-1} \sum_{i=1}^n K(h_K^{-1}d(x, X_i))H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^n K(h_K^{-1}d(x, X_i))}, \quad (1)$$

where  $K$  and  $H$  are kernels and  $h_K = h_{K,n}$  (resp.  $h_H = h_{H,n}$ ) is a sequence of positive real numbers. Note that a similar estimate was already introduced in the special case when  $X$  is a real random variable by Rosenblatt [9] and widely studied until now (see for instance Hassani and Youndjé [7], for several asymptotic results and references). A natural extension of the kernel estimator  $\hat{\theta}$  of the conditional mode  $\theta$  to the functional framework is given by:

$$\hat{\theta} = \arg \sup_{y \in S} \hat{f}^x(y). \quad (2)$$

The estimate  $\hat{\theta}$  is not necessarily unique, and if this is the case, all the remaining of our Note concerns any value  $\hat{\theta}$  satisfying (2). This work establishes the almost complete convergence of the kernel estimate  $\hat{\theta}$  of  $\theta$  under  $\alpha$ -mixing hypothesis. As we will see next Section, the flatness of the function  $f^x$  around the mode  $\theta$  plays a major role. One way to control this flatness is to consider the number of vanishing derivatives at  $\theta$ .

## 3. Main result

We introduce now some assumptions that are needed to state our results:

(H1)  $P(X \in B(x, r)) = \phi_x(r) > 0$ ,

(H2)  $\sup_{i \neq j} P((X_i, X_j) \in B(x, r) \times B(x, r)) = \phi_x(r)\psi_x(r) > 0$ ,

(H3) The coefficients of the  $\alpha$ -mixing sequence  $(X_i, Y_i)$  satisfy the following arithmetic condition:

$$\exists a > (5 + \sqrt{17})/2, \exists c > 0, \forall n, \alpha_n \leq cn^{-a},$$

(H4)  $f^x$  is  $j$ -times continuously differentiable with respect to  $y$  on  $(\theta - \xi, \theta + \xi)$ ,

(H5)  $\forall (y_1, y_2) \in S \times S, \forall (x_1, x_2) \in N_x \times N_x, |f^{x_1}(y_1) - f^{x_2}(y_2)| \leq C_x(d(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2})$ ,

(H6)  $f^x$  is strictly increasing on  $(\theta - \xi, \theta)$  and strictly decreasing on  $(\theta, \theta + \xi)$ ,

(H7)  $f^{x(l)}(\theta) = 0$ , if  $1 \leq l < j$ , and  $0 < |f^{x(j)}(\theta)| < \infty$ ,

(H8)  $K$  is a function with support  $(0, 1)$  such that  $0 < C_1 < K(t) < C_2 < \infty$ ,

(H9) 
$$\begin{cases} \forall (y_1, y_2) \in \mathbb{R}^2 & |H(y_1) - H(y_2)| \leq C|y_1 - y_2|, & \int |t|^{b_2} H(t) dt < \infty \\ \text{and } \exists \nu > 0, & \lim_{y \rightarrow \infty} |y|^{1+\nu} |H(y)| = 0. \end{cases}$$

(H10)  $\lim_{n \rightarrow \infty} h_H = 0$  and  $\exists \beta_1 \leq \frac{4}{(a+1)(a-2)}$  such that  $\lim_{n \rightarrow \infty} n^{\beta_1} h_H = \infty$ ,

(H11) 
$$\begin{cases} \lim_{n \rightarrow \infty} h_K = 0, & \lim_{n \rightarrow \infty} \frac{\log n}{nh_H \chi_x(h_K)} = 0, \\ \text{and } \exists \beta_2 > 0, c_1 > 0, c_2 > 0, & c_2 n^{(3-a)/(a+1)+\beta_2} \leq \chi_x(h_K) \leq c_1 n^{1/(1-a)}. \end{cases}$$

In this last hypothesis, the function  $\chi_x$  was defined by  $\chi_x(h) = \max\{\phi_x(h), \psi_x(h)\}$ .

**Theorem 3.1.** *Under the hypotheses (H1)–(H11) we have*

$$\hat{\theta} - \theta = O(h_K^{b_1/j} + h_H^{b_2/j}) + O\left(\left(\frac{\log n}{nh_H \chi_x(h_K)}\right)^{1/2j}\right), \quad a.co. \tag{3}$$

**Sketch of the proof.** Write a Taylor expansion of  $f^x$  in a neighbourhood of  $\theta$ . Because of (H4) and (H7), and because  $|f^x(\hat{\theta}) - f^x(\theta)| \leq 2 \sup_{y \in S} |\hat{f}^x(y) - f^x(y)|$ , we have for some  $\theta^*$  between  $\theta$  and  $\hat{\theta}$ :

$$|\theta - \hat{\theta}|^j \leq \frac{j!}{f^{x(j)}(\theta^*)} \sup_{y \in S} |\hat{f}^x(y) - f^x(y)|.$$

Theorem 3.1 is deduced from the lemmas and corollaries below, and from the decomposition:

$$\forall y \in S, \quad \hat{f}^x(y) - f^x(y) = \frac{1}{\hat{f}_D^x} \{(\hat{f}_N^x(y) - E \hat{f}_N^x(y)) - (f^x(y) - E \hat{f}_N^x(y))\} + \frac{f^x(y)}{\hat{f}_D^x} \{E \hat{f}_D^x - \hat{f}_D^x\},$$

where

$$\begin{aligned} \hat{f}_N^x(y) &= \frac{1}{nh_H EK(h_K^{-1}d(x, X_1))} \sum_{i=1}^n K(h_K^{-1}d(x, X_i)) H(h_H^{-1}(y - Y_i)) \quad \text{and} \\ \hat{f}_D^x &= \frac{1}{nEK(h_K^{-1}d(x, X_1))} \sum_{i=1}^n K(h_K^{-1}d(x, X_i)). \end{aligned}$$

**Lemma 3.2.** *Under the hypotheses (H1)–(H3), (H8) and (H11) we have*

$$\hat{f}_D^x - E \hat{f}_D^x = O\left(\sqrt{(\log n)/(n \chi_x(h_K))}\right), \quad a.co.$$

**Corollary 3.3.** *Under the hypotheses of Lemma 3.2, we have  $\sum_{n=1}^{\infty} P(\hat{f}_D^x < 1/2) < \infty$ .*

**Lemma 3.4.** *Under the hypotheses (H1), (H5), (H8), (H9) and (H11) we have*

$$\frac{1}{\hat{f}_D^x} \sup_{y \in S} |f^x(y) - E \hat{f}_N^x(y)| = O(h_K^{b_1}) + O(h_H^{b_2}), \quad a.co.$$

**Lemma 3.5.** *Under the hypotheses (H1)–(H3), (H5), and (H8)–(H11) we have*

$$\frac{1}{\hat{f}_D^x} \sup_{y \in S} |\hat{f}_N^x(y) - E \hat{f}_N^x(y)| = O\left(\sqrt{(\log n)/(nh_H \chi_x(h_K))}\right), \quad a.co.$$

**Corollary 3.6.** *Under the hypotheses of Lemma 3.5, we have  $\sup_{y \in S} |\hat{f}^x(y) - f^x(y)| \rightarrow 0$ , a.co.*

**Lemma 3.7.** *Under the hypotheses of Lemma 3.5 and (H6)–(H7), we have*

$$\exists c > 0, \quad \sum_{n=1}^{\infty} P(f^{X^{(j)}}(\theta^*) > c) < \infty.$$

**Remarks on the proof.** The complete proofs of these lemmas are available on request. Let us just note that the results linked with the dependence (Lemmas (3.2) and (3.5)) need a special attention and use exponential inequalities for dependent variables (see Rio [10]). For peoples wishing to apply these lemmas in other settings, note that the last part of condition (H11) is not necessary to get Lemmas 3.4 and 3.7.

**Remarks on key hypothesis.** The main novelty in our functional approach can be seen through expressions (H1) and (H2), that are not really restrictive. Indeed, as pointed out in [3] the expression of the function  $\phi_x$  appearing in (H1) can be specified for many usual continuous time processes, and is linked with small ball probability theory. To see the role of condition (H2) it suffices to think on the special case when  $X$  is real and has a density with respect to Lebesgues measure on  $\mathbb{R}$ . In this case, (H2) is true with  $\phi = \psi$  as long as the pairs  $(X_i, X_j)$  have a density with respect to the Lebesgues measure on  $\mathbb{R}^2$  (which is a quite often used assumption in the classical finite dimensional literature).

*Perspective for applications.* The result of this Note can be applied to the prediction of time series, just by cutting the past of some time series in continuous paths. Details can be found in [3] in which this is done but with other functional prediction technique (based on conditional expectation estimation).

## Acknowledgements

The authors would like to thank both referees whose comments and suggestions have improved the presentation of this work. All the participants of the working group STAPH on Functional and Operatorial Statistics in Toulouse are also gratified for their continuous supports and comments.

## References

- [1] A. Berlinet, A. Gannoun, E. Matzner-Lober, Normalité asymptotique d'estimateurs convergents du mode conditionnel, *Rev. Canad. Statist.* 26 (1998) 365–380.
- [2] G. Collomb, W. Härdle, S. Hassani, A note on prediction via conditional mode estimation, *J. Statist. Plann. Infer.* 15 (1987) 227–236.
- [3] F. Ferraty, A. Goia, P. Vieu, Functional nonparametric model for time series: a fractal approach to dimension reduction, *TEST* 11 (2) (2002) 317–344.
- [4] F. Ferraty, P. Vieu, Functional nonparametric statistics: a double infinite dimensional framework, in: M. Akritas, D. Politis (Eds.), *Recent Advances and Trends in Nonparametric Statistics*, Elsevier, Amsterdam, 2003.
- [5] F. Ferraty, A. Laksaci, P. Vieu, Estimating some characteristics of the conditional distribution in nonparametric functional models, *Statist. Infer. Stoch. Proc.*, (2004), in press.
- [6] T. Gasser, P. Hall, B. Presnell, Nonparametric estimation of the mode of a distribution of random curves, *J. Roy. Statist. Soc. Ser. B* 60 (1998) 681–691.
- [7] S. Hassani, E. Youndjé, Propriétés de convergence de l'estimateur noyau de la densité conditionnelle, *Rev. Roum. Math. Pures Appl.* 41 (1996) 535–566.
- [8] D. Louani, E. Ould-Saïd, Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis, *J. Nonparametric Statist.* 11 (4) (1999) 413–442.
- [9] M. Rosenblatt, Conditional probability density and regression estimators, in: P.R. Krishnaiah (Ed.), *Multivariate Analysis II*, Academic Press, New York, 1969.
- [10] E. Rio, Théorie asymptotique des processus faiblement dépendants, *Mathématiques & Applications*, vol. 31, Springer-SMAI, 1999.