REPUBLIQUE ALGERIENNE Démocratique & Populaire
Ministére de l'Enseignement Supérieur et la Recherche Scientifique

Université Djillali Liabés de Sidi Bel Abbés

Faculté des Sciences Exactes

Département des probabilites-Statistiques

THÈSE DE DOCTORAT

Spécialité : Mathématiques
Option : Statistiques

Présentée par

ARDJOUN FATIMA ZOHRA

Soutenue le , devant le jury composé de

A la mémoire de ma défunte mére ...

Prière et Salut soient sur Notre Cher Maître & Prophète " Mohammed " et sur sa famille et ses fidèles compagnons.

Remerciements

Je tiens à remercier vivement tout d'abord mon directeur de recherches, monsieur Laksaci Ali, Professeur à l'université de Sidi Bel Abbés pour tout le temps qu'ils m'ont consacré et pour sa patience, son suivie et ses conseils durant l'évolution de ce travail.

Que Kadi Mohammed Attouch trouve ici l'expression de mes remerciements pour l'intérêt qu'il a porté à cette thèse. Je le remercie chaleuresement d'avoir accepté d'être le président du jury de ce travail.

Je voudrais également remercier les membres de jury pour l'honneur qu'ils m'a fait pour avoir accepter d'examiner cette thèse : Monsieur Benchikh Toufik, Maître de conférence à l'université de Sidi Bel-Abbés, Monsieur Guendouzi Toufik, Professeur à l'université de Saïda, Monsieur Madani Fethi Maître de conférence à l'université de Saïda, Monsieur Kandouci Abdeljabbar, Professeur à l'université de Saïda.

Je remercie, monsieur Larbi Ait Henneni, enseignant á université de Lille Roubaix, France pour sa collaboration.

Mes remeriements aussi tous les enseignants du laboratoire de statistiques-probabilités de l'université de Sidi Bel-Abbés ainsi que mes collègues de la faculté des Siences Exactes de l'université de Sidi Bel-Abbés.

Finalement, mes plus vifs remerciement s'adressent à tout personne qui a contribué de prés ou de loin à l'élaboration de ce travail.

Notation et Lexique

$\mathbb{R}$ L'ensemble des nombres réels.
$N(0,1)$ La loi normale centré réduite.
$\mathcal{P}$ La probabilité.
$(\Omega, A, \mathcal{P})$ l'espace probabilité où les variables aléatoires sont considérées.
$\mathcal{K}, \mathcal{H}$ Les fonctions du noyau.
$\mathcal{B}, \mathcal{F}$ Les espaces fonctionnelles.
χ La variable explicative.
$\mathcal{Y}$ La variable de réponse.
$d(.,.)$ La distance entre les vecteurs de l'espace semi-métrique.
$\mathcal{P}(./\chi = x)$ Mésuré la probabilité conditionnel.
$f^{\chi}(.)$ La densité conditionnelle de Y sachant X.
$\widehat{f}_n$ l'estimateur de densité.
θ Le mode conditionnel.
$\widehat{\theta}$ l'estimateur de mode.
h Le paramêtre de lissage.
a_n, b_n resp $(h_{k,n}, h_{H,n})$ Des suites de nobmres réels positifs.
$r(.)$ Fonction de regression.
$B(x,h)$ la boule de centre x et de rayon h.
$\phi(x,r)$ Mesuré la boule $B(x,h)$.
i.i.d indépendantes et identiquement distribuées.
p.c.o La convergence presque complète.

Table des matières

Chapitre 1

Introduction

1.1 Statistique des données fonctionnelles

1.1.1 Motivation

Les problèmes statistiques en rapport avec la modalisation et avec l'étude des variables aléatoires fonctionnelles connaissent depuis fort longtemps un important intérêt en statistique. Pour arriver à adapter des techniques statistiques multivariés classiques, il faudrait que la descritisation de ces observations fonctionnelles soit la base des premiers travaux conséquemment à l'avancée et le progrés du monde informatique, qui recupère des données de plus en plus volumineuses, une alternative fût prise pour traiter convenablement ce type de données dans sa propre dimensions c'est-à-dire conserver le caractére fonctionnelles. Nombreux sont les domaines qui ont bénéficient de la statistiques fonctionnelles à savoir la climatologie (Besse (2000)), l'économie (Kneip (2005)), Bujaen en (2003) s'est occupé du linquistique , et l'environnement a été notament étudie par Cardot et Vieu (2004)...ect. La forte liaison de ces travaux avec les méthodes de classification résulte quand le variable explicative est une courbe. Par conséquent , l'estimation de mode conditionnel occupe une grande place dans la statistique fonctionnelle.

1.1.2 Statistique non paramétrique fonctionnelle : Note bibliographique

Ces derniéres années il existe une vaste littérature consacré à l'étude des modéles statistiques, en effet, Roussas en (1967) a fourni les premiers résultats pour la convergence en norme L^2 d'un estimateur à noyau de la densité de transition. Deux aux plutard, Rosenblat (1969) s'est occupé du cas de la densité de la loi stationnaire sous la condition G_2. Des résultats sur des modéles de régularité un peu différents pour la fonction de régression a été établis aussi par Mamman (1999a) et (1991b). la littérature est trés présente dans le cas fonctionnel et ce sont justement Farraty et Vieu (2000) qui ont permis d'obtenir les premiers travaux. Ils ont étudié la convergence presque complète d'un estimateur à noyau de la fonction de régression pour des données indépendantes identiquement distribuées. Ferraty et al. (2002), ont prolongé ces travaux, aux problémes non-standards de la régression telle la prévision des séries chronologiques. En reprenant l'estimateur de Ferraty et Vieu (2004) et en utilisant la propriété de concentration de la mesure de probabilité de la variable explicative fonctionnelle, Niang et Rhomari (2004) ont étudié la convergence en norme L^p de l'estimateur de la régression. Ils ont appliqué leur résultats à la discrimination et la classification des courbes. Masry (2005) a obtenu la normalité asymptotique de l'estimateur de la regression dans le cas dépendant. Nous revoyons à Ferraty et al. (2004) pour la convergence presque compléte pour le cas i.i.d. La monographie de Ferraty et al. (2006) présente une variété importante d'outils statistique pour la prévision par des variables fonctionnelles. Dans ce travail pionnier, les auteurs sont donnés la convergence presque compléte de l'estimateur du noyau de la fonction de distribution conditionnelle et ses dérivées. La normalité asymptotiques de cet outils a été étudiée par Ezzaharioui et Oul-Saïd (2005). Autres sont intéressés à l'estimation de quantiles conditionnels , à titre d'exemple Roussas (1991) pour un processus de Markov. Il a utilisé un estimateur de la fonction de répartition conditionnelle et il a établi la vitesse de la convergence presque sûre d'un estimateur à noyau. Laksaci et Maref (2009) donnent la version uniforme du taux de convergence de l'estimateur du noyau proposé par Ferraty et al. (2006). Dabo-Niang et Laksaci (2006) ont généralisés leur résultats de la convergence en norme L_p de l'estimateur. Dans le cas dépendant Lemdani et al. (2011) ont étudié la convergence presque compléte et la normalité asymptotique de l'estimateur L^1 des quantiles conditionnels. Dabo-Niang en association avec Laksaci (2012) ont établi la convergence en norme

L^p pour un estimateur à double noyau. L'approuche robuste est une allure d'estimer les quantiles conditionelles et d'avoir des résultats asymptotiques. Citons à titre d'exemple Gooijer et Gannoun (2000) pour la détermination des intervalles prédictives, nous renvoyons à Zhang en (2001) pour la convergence en moyenne quadratique, dans les différent cas i.i.d, α-méleangeant et β-méleangeant. Cependant, tous les résultats impliques ci-dessus sont dans le cas où les échantillions sont observés complatement. Certaines paires d'observations peuvent être incomplète, comme dans le cas de données manquantes. On trouvra dans la littérature, plusieurs référence pour les données manquantes. Citons, par exemple Graciela Boentea et al. (2009) ont introduit deux familles d'estimateurs, pour estimer la fonction de régressionn par la méthode rébustes l'orsque les données sont manquantes. Ces derniers ont montrés, la normalité asymptotique de ces estimateurs, Boente .al (1990) Pour la convergence presque sûre d'un estimateur d'autogréssion pour la distribution dans le cas dépendant.

1.2 Problématique de la thèse

L'apport principale de cette thèse est d'étudier l'estimation du mode conditionnel dans le cas oú la variable explicative est fonctionnelle , qui prend ces valeurs dans un espace semi-métrique et une variable réponse réelle sous la condition de l'ergodicité. Plus précisément , noun nous proposons d'étudier les propriétés asymptotiques d'un estimateur pour le mode conditionnel d'une variable aléatoire réelle sera notée Y sachant une autre variable de dimention eventiellement infini sera notée X. Cette autre estimation est basée sur une méthode récursive du noyau. Dans cette partie, nous mettons en évidence un avantage important de l'estimateur récursif par rapport à l'estimateur classique qui utiliser par Ferraty et Vieu (2006). Ce concerne le gain de temps de calcul pour un phénoméne lorsque de nouvelle valeurs de la variable explicative sont ajouté séquentiellement à la base de données.

Ainsi, dans un premier temps, nous construisons un estimateur par la méthode récursive, pour le mode conditionnel adapté on notre cas fonctionnel, ensuite nous obtenuons sa vitesse de convergence presque complète de l'estimateur proposé.

La deuxième étape consiste à traiter le cas oú les observations sont ergodique, dans cette partie nous allons ajouté une condition de dépendante faible, pour obtenir deux version de convergence, la convergence presque complète et la normalité asymptotique aves une vitesse de convergence. Notre approche est illustré par une application de données réelles.

Dans ce paragraphe, nous ouvrons une parenthèse sur les hypothèses, utilisons dans notre travail. Nos conditions sont assez légères dans ce contexte de statistique non paramétrique fonctionnelle, dans le cas ergodique, nous avons utilisé une condition qui est trés similaire à ce utilisé par Gheriballah et al (2013), cette condition est trés importante dans le cas fonctionnelle aussi bien que dans le cas de dimension fini et qui est mois restrictive aux condition imposée par Laib et Louani (2011), car n'est pas nécessaire de utilisé les propriétés sur les fonctions de probabilités de petites boules et de changer par une autre propriété sur la fonction de concentration fonctionnelle qui définies par :

$$\mathcal{P}(X_i \in B(x,r)/\mathcal{F}_{i-1})$$

Tandis que des conditions sont les mêmes utilisé que dans Ferraty et al. (2006). Ces hypothéses peut être vu comme des hypothéses de régularités qui caractérise l'espéce fonctionnel de notre modéle telque des conditions sont necessaire pour la continuité du modéle et d'obtenir la convergence. Ainsi des autres sont basée sur la condition de type lipschitz et qui est utilisé pour préciser le taux de convergence de l'estimation.

Mots clés :

Données fonctionnelles, mode conditionnel, distribution conditionnel récursif, estimation, condition ergodique.

1.3 Estimation de mode conditionnel en statistique fonctionnelle

Le mode conditionnel, joue un rôle crucial dans la prévision non paramétrique, et a stimulé quelque nombre de chercheurs dans l'étude du mode estimateurs, et constitue une méthode alternative pour estimer les régression. La littérature est trés abondante sur l'estimation du mode conditionnel lorsque la variable explicative est de dimenstion fini. Le cas indépendant, fut traité par Samanta et Thaaneswason (1990). Ces derniers ont obtenu la normalité asymptotiquede de l'estimateur par la méthode de noyau du mode conditionnel. De nombreux auteurs ont ajouté des résultats sur ce sujet dans Le cas dépendant, citons a titre d'exemple Collomb en (1987) sur le processus φ-mélangeant, Quanttela et Vieu (1997) sur le processus α-mélangeant, Ould-Saïd (1997) pour le cas ergodique, et (1999) pour le cas α-mélangeant sur la normalité asymptotique.

Le premier article sur l'estimation de la densité conditionnelle a été obtenu par Roussas (1968). Il a montré la convergence en moyenne quadratique, la convergence en probabilité et la normalité asymptotique.Quanttela et Vieu (1997) ont construit un estimateur pour estimer mode conditionnel comme étant une solution de la dérivée d'ordre un de l'estimateur de la densité conditionnelle et ils one démontré la vitesse de convergence presque complète dans le cas d'un processus de mélange. Dans le cas d'une co-variable de dimension infini, les premiers résultats ont été obtenu en (2004) par Farraty et al. Ils ont étudiés la convergence presque complète d'un estimateur de la densité conditionnelle et aborder une application de la prévision via le mode conditionnel dans le cas i.i.d. Ces derniers ont généralisé ce résultat au cas α-mélangeant en (2005). Niang et Laksaci en (2007) se sont intéressé à la convergence en norme L_p d'un estimateur à noyau du mode de la densité d'une variable aléatoire réelle conditionnée par une variable fonctionnelle. Ezzahrioui et Ould Saïd (2008) ont étudié la normalité asymptotique de l'estimateur non paramétrique du mode conditionel dans le cas i.i.d. Deux ans plus tard ils ont démontré le cas dépendant. Récemment, Dabo-Niang et al. (2012) ont traité du probléme de prédiction spatial de l'estimation du mode conditionel. Ils ont établi la convergence presque complète avec un taux de l'estimateur adaptatif. Ling et Xu (2012) ont étudié les propriétés asymptotiques de l'estimation de la densité conditionnelle pour des séries chronologiques fonctionnelles. Se sont intéréssés à la normalité asymptotique de l'estimateur pour la densité conditionnelle et le mode conditionnel à la fois dans le cas α-mélangeant. En autre, Rachdi et al. (2014), ont obtenu également des données théoriques et propriétées pratiques de l'estimateur de la densité conditionnelle et le mode via l'approche linéaire locale. Deux des méthodes les plus courantes dans la regression non paramétrique la méthode de noyau et la la méthode des K-plus proches voisins cette derniére méthode a été utilisé par Attouch et Bouabca (2013). Ils ont construit un estimateur à noyau de mode conditionnel d'une fenêtre locale et ils ont établi la convergence presque complète de cet estimateur. On peut se reporter à la monographie de Ferraty et Vieu (2006) pour le cas oú les variables explicatives fonctionnelle.Ils ont établis le taux de convergence presque complète de l'estimateur du noyau du mode conditionel. La convergence L_p du même estimateur a été obtenu par Dabo-Niang et Laksaci (2007). Ezzahrioui et Ould-Saïd (2008) ont donné la normalité asymptotique du noyau d'estimateur du mode conditionnel dans le cas i.i.d.

1.4 Estimation récursive

L'estimation fonctionnelle est un domaine central des statistiques, l'accent étant généralement mis sur les estimateurs de densité et de régression. Dans cette partie nous mettons en évidence un aventage important de l'estimateur récursive de la régression et la densité par rapport á l'estimateur á noyau classique. Ce concerne le gain de temps de calculs lorsque de nouvelle valeurs de la variable explicative sont ajoutés séquenciellement á la base de données.

L'estimation de la fonction de régression est un probléme d'intérêt considérable dans de nombreux domaines appliqués telsque la théorie de la décision et la prévision. La littérature sur ce sujet est toujour en croissance. Les premiers estimateurs récursive de la régressions ont été introduits par Ahmed et Lin (1976) et par Devroye et Wagner (1980). Ensuite Amiri (2012) propose une nouvelle famille générale d'estimateurs récursives, définis par :

$$\widehat{r}_n(x) = \frac{1}{\sum_{i=1}^{n} h_i^{(1-l)d}} \frac{\sum_{i=1}^{n} \frac{m(y_i)}{h_i^{dl}} K\left(\frac{x-X_i}{h_i}\right)}{\sum_{i=1}^{n} \frac{1}{h_i^{dl}} K\left(\frac{x-X_i}{h_i}\right)}. \tag{1.1}$$

Déffirents résultats pour cet estimateur dans le cas i.i.d ont été établis par de nombreux auteurs. Nous citons, ente autre Devroye (1981), Krzyzak (1992) et Walk (2001). Griblicki et Palwlak (1987) ont donné des conditions sur le paramétre lisse h_n et ils ont obtenu la convergence ponctuelle et la convergence presque sûre. Nous renvoyons á Amiri (2012) a étudié la normalité asymptotique, la convergence en moyenne quadratique et la vitesse de convergence presque sûre pour des observations i.i.d. Dans le cas dépendant, Roussas (1990) a donné la convergence uniforme. Deux ans plus tard, Tranetal a étudié la normalité asymptotique sous des conditions de φ-mélangeant pour cet estimateur.

En ce qui concerne l'estimateur récursive de la densité , le premier estimateur a utilisé par Wagner et al. (1969) qui définie par

$$\widehat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} h_d^{-1} K\left(\frac{x-X_i}{h_i}\right), \forall x \in \mathrm{I\!R}^d. \tag{1.2}$$

Liang et Back (2004) ont étudié la normalité asymptotique pour des variables négatives associés à cet estimateurs. Notons que cet estimateur a été utilisé par Wagman et Davies en (1979) dans le cas indépendant. Dans le cas dépendant, la normalité asymptotique a été démontré par Masry (1986) sous des divers hypothèses. Györfi (1981) a prouvé la cohérence forte de l'estimateur de la densité. Nous renvoyons á Masry (1987) pour la convergence presque sûre de cet estimateur dans le cas φ-mélangeant. Tran (1989) a étudié la convergence uniforme et la normalité asymptotique sous une hypothèse dépendante. D'autre autre se sont intéressé á une sous famille des estimateurs récursifs initiés par cet estimateur. Citons á titre d'exemple Amiri en (2009), ce dernier a établi l'erreur quadratique asymptotique, ensuite il a introduit des critères pour comparer ces estimateurs. En renvoyons à Bouadjemia (2014) a considéré le cas où les observations sont i.i.d et il a élaboré un résultat sur la normalité asymptotique d'un estimateur récursive de la fonction de répartition conditionnelle.

1.5 La théorie Ergodique

La théorie ergodique est une condition fondamentental de la physique statistique. Plus précisement, il assemble les propriétés thermodynamiques des gaz, des atomes, des électrons ou des plasmes. Cette condition est également utilisée dans le traitement du signal, pour l'étude de l'évolution d'un signal aléatoire. Ergodique est construit à partir des termes grecs, c'est a dire (action,travail), telque Ludwing Boltzmann en 1871 semble être le premier qui pose cette théorie pour les besoins de sa théorie cinitique des gaz. Cés ce moment la théorie ergodique s'est développé au sein de la mathématique, analyse fonctionnelle...ect

La littérateure sur les données fonctionnelles ergodiques est encore limité. En effet, ce problème a été initié par Liab et Louani (2010). Ils ont considéré le problème de l'estimation fonctionnelle pour la fonction de régression non paramétrique dans le cas dépendant mois faible qui est l'état d'ergodicité. Ces derniers ont introduit la même version de l'estimateur que l'estimateur de Nadarya-Watson de type noyau de la fonction de régression classique construite par la méthode du noyau et ils ont établi la convergence uniforme et la vitesse presque complète lorsque les données fonctionnels sont ergodiques stationnaires. En 2011, ils ont étudié la convergence presque sûre et la normalité asymptotique de cet estimateur. Récemment Benziadi et Laksaci (2012) ont considéré deux estimateurs récursifs du quantiles conditionel lorsque la variable explicative est de type ergodique telque le premier, est donnée en enversant de l'estimation à double noyau de la fonction distribution conditionnelle, et le second est obtenu de utiliser l'approche robuste. Ils ont établi la convergence presque complète de ces deux estimateurs. Nous renvoyons á Sakal et Laksaci (2013) ont obtenu la convergence presque complète avec un taux de l'estimation de la regression robuste ensuite la normalité asymptotique d'une famille d'estimateur robuste dans le cas où les observations sont échantillonnés à partir d'un processus ergodique fonctionnel. Des résultats a été obtenu sur l'estimateur à noyau de la densité conditionnelle par Ould-Saïd (1997), qui a étudié la convergence uniforme presque sûre d'un processus stationnaire ergodique par un modèle de régression pour des séries temporelles (modèle autorégression), et un résultat similaire sur la médiane conditionnelle a été donné par Rosa (1992). Une famille d'estimateurs d'autorégression non paramétrique ont été étudié dans le papier de Laib et Ould Saïd (2000). Ils ont démontré la convergence uniforme de cette famille sous une condition d'ergodicité dans le cas où les observations ne sont pas bornées. Cependant

Collomb et Härdle (1986) ont étudié cette propriété asymptotique lorsque la fonction être bornée pour ces estimateurs. Ainsi, l'originalité de notre résultat à la structure de dépendance faible imposée au processus d'observation en supposant simplement son ergodicité. De plus comme l'ont souligné Laib et Louani (2011), les variables aléaitoires générées avec les modèles fonctionnels autorégressifs c'est un exemple particulier des variables aléatoires ergodiques fonctionnelles.

1.6 Contribution de la thèse

S*oit* (X,Y) un couple de variables aléatoires à valeurs dans $\mathcal{F}\times \mathbb{R}$, où $\mathcal{F}$ est un espace semi-métrique. On note d la semi métrique sur $\mathcal{F}$. Pour tout $\chi \in \mathcal{F}$, on désigne par f^χ la densité conditionnelle de Y sachant $X=\chi$. Par la suite, on fixe un point χ dans $\mathcal{F}$, on note N_χ un voisinage de ce point et on suppose qu'il existe un sous ensemble compact $S=[\theta-\xi,\theta+\xi]$, $\xi>0$ tel que le mode θ de la densité conditionnelle f^χ soit unique dans ce compact. l'estimateur à noyau de la densité conditionnelle f^χ noté $\widehat{f}^\chi$, est défini par

$$\forall \chi \in \mathcal{F} \text{ et } \forall y \in \mathbb{R} \qquad \widehat{f}^\chi(y)=\frac{b_n^{-1}\sum_{i=1}^n K(a_n^{-1}d(\chi,X_i))H(b_n^{-1}(y-Y_i))}{\sum_{i=1}^n K(a_n^{-1}d(\chi,X_i))},$$

avec la convention $\frac{0}{0}=0$. K et H sont des noyaux et $a_n=h_{K,n}$ (resp. $b_n=h_{H,n}$) est une suite de nombres réels positifs. Notons que cet estimateur a été utilisé par Ferraty *et al.* (2006) pour le cas fonctionnel.

La version recursive de cette estimateur est définie par :

$$\forall \chi \in \mathcal{F} \text{ et } \forall y \in \mathbb{R} \qquad \widehat{f}^\chi(y)=\frac{\displaystyle\sum_{i=1}^n K\left(a_i^{-1}d(\chi,X_i)\right) b_i^{-1}H\left(b_i^{-1}(y-Y_i)\right)}{\displaystyle\sum_{i=1}^n K\left(a_i^{-1}d(\chi,X_i)\right)}.$$

Un estimateur de θ est la variable aléatoire $\widehat{\theta}$ solution du problème doptimisation suivant :

$$\widehat{f}^\chi(\widehat{\theta})=\sup_{y\in(\alpha_\chi,\beta_\chi)}\widehat{f}^\chi(y). \tag{1.3}$$

En général, cet estimateur n'est pas unique. Ainsi, dans toute la suite de cette Note, $\widehat{\theta}$ désigne n'importe quelle variable aléatoire vérifiant (1.3).

Dans ce paragraphe, nous donnons une brève présentation des différents résultats obtenu dans cette thèse.

Résultats : l'estimation du mode conditionnel avec les variables fonctionnelles

Dans ce paragraphe, nous traitons le cas où le variable réponse est banachique et la variable explicative prend ces valeurs dans un espace semi métrique pour un modèle de regression fonctionnelle de mode conditionnel. Nous étudions par suite la version uniforme de la convergence presque complète.

Théorème 1.1 *Sous les hypothèses (H1)- (H7), on a pour la convergence presque complète uniforme*

$$\sup_{x\in S_{\mathcal{F}}}\left[||\widehat{\theta(x)}-\theta(x)||^2\right]=O(h_K^{b_1})+O(h_H^{b_2})+O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)+\psi_{S_{\mathcal{B}}}(h_2\epsilon)}{ng(h_2)\phi(h_1)}}\right).$$

Les épreuves de cette résultats se trouve au chapitre 2.

Résultats : l'estimateur récursive de mode conditonnel

Dans ce paragraphe, nous construisons un estimateur récursive lisse de noyau et nous traitons la convergence presque complète, de l'estimateur récursive de mode conditionnel.

Théorème 1.2 *Sous certains hypothèses détaillées dans le chapitre 3, on a*

$$\hat{\theta}-\theta=O\left(\frac{1}{n\varphi_n(x)}\sum_{i=1}^{n}\left(a_i^{\beta_1}\phi(x,a_i)\right)\right)^{1/2}$$

$$+O\left(\frac{1}{n\varphi_n(x)}\sum_{i=1}^{n}\left(b_i^{\beta_2}\phi(x,a_i)\right)\right)^{1/2}+O\left(\sqrt{\frac{\log n}{nb_n\varphi_n(x)}}\right)^{1/2}\quad p.co.$$

Les épreuves des résultats auxiliaires seront présentées au chapitre 3.

Résultat : Cas des données ergodiques fonctionnelles

Dans ce paragraphes, on considéres le cas où les données sont ergodiques fonctionnelles à la fois, on présente la convergence presque complète, ainsi la normalité asymptotique de cet estimateur

Théorème 1.3 *Sous la condition de ergodicité et les hypothèses de concentraction fonctionnelle, on a*

$$\hat{\theta}-\theta=O\left(a_n^{\beta_1/2}\right)+O\left(b_n^{\beta_2/2}\right)+\left(\sqrt{\frac{\log n}{nb_n\psi_n(x)}}\right)^{1/2}\quad p.co.$$

Théorème 1.4 *Sous certaines hypothèses portant sur l'estimateur, nous aurons pour la normalité asymptotique*

$$\left(\frac{nb_n^3\varphi_n(x)}{\sigma^2(x)}\right)^{1/2}(\hat{\theta}-\theta)\longrightarrow \mathcal{N}(0,1) \quad as \quad n\to\infty.$$

où

$$\sigma(x)=\frac{f^x(\theta)\left(K^2(1)\delta(1)-\int_0^1(K^2(s))'\delta(s)\beta_x(s)ds\right)}{\left(f^{x(2)}(\theta)\left(K(1)-\int_0^1 K'(s)\beta_x(s)ds\right)\right)^2}\int H^{'2}(t)dt.$$

La démonstration et les hypothèses nécessaires de cette résultat, sont détaillées au chapitre 4 .

1.7 Plan de la thèse

Ce travail est présenté essentiellement en quatre chapitres, qu'ils ont organisées comme suit :
ce premier chapitre est consacré à la présentation des résultats asymptotiques obtenus. Plus précisément, l'objectif de ce chapitre introductif est de donner un contexte bibliographique sur la densité conditionnelle et le mode conditionnel, pareillement nous avons ouvrir une parenthèse sur l'estimation récursive une part et d'autre part d'introduire le principe de la thèorie ergodique.

Dans le chapitre 2 nous construisons un estimateur par la méthode de noyau de mode conditionnel, l'orsque la variable réponse est banachique (appartenant à un espace de Banach) et la variable explicative est fonctionnelle . Nons étudions par suite la convergence uniforme presque complète de cet estimateur avec un taux de convergence.

Dans le chapitre 3, nous présentons une version de convergence, la convergence presque complète d'un nouvel estimateur à noyau récursif .

La cas fonctionnelle ergodique est étudié dans le chapitre quatre. Ce chapitre est divisé en cinq Sections. Dans la deuxième Section nous construisons un estimateur du mode conditionnel par la méthode récursive à noyau sous l'hypothèse ergodique. Ensuite, nous donnons le taux de convergence presque complète dans la section 3, tandis que la normalité asymptotique est présenté dans la quatrième section. L'efficacité de cet méthodes est également démontrée par des simulation numériques mettant en évidence leur temps de calcul considérablement réduit se trouve dans la cinquième section. Enfin, les épreuves des résultats auxiliaires sont données en annexe.

Enfin, on donne une conclusion du travail effectué et nous présentons quelques persepectives présumables de nos tavaux de recherche dans la fin de cette thése.

Chapitre 2

Estimation of the conditional mode for the two functional variables

Ce chapitre fait l'objet d'une soumission dans le Journal of Applicationes Mathematicae.

ESTIMATION OF THE CONDITIONAL MODE FOR THE TWO FuNCTIONAL VARIABLES

Ardjoun Fatime-Zohra, Laksaci Ali and Amel Tadj
Laboratoire de Statistique et Processus Stochastiques
Univ. Djillali Liabès
BP 89, 22000 Sidi Bel Abbès, Algeria
e-mail : (alilak or bchouaf or maila.nabila)@yahoo.fr

Abstract

In this paper, we consider a kernel estimates of the modal regression when the response variable is in a Banach space and the explanatory variable takes values in a semi-metric space. We quantify the asymptotic properties of these estimates, by giving the uniform almost complete convergence rate.

Subject Classification : Primary, 62G20 ; secondary 62G08, 62G35, 62E20.

Key Words : Uniform almost complete convergence, kernel estimators, functional data, entropy, semi-metric space.

2.1 Introduction

It is well known that in statistics, regression model is main tool to examine the relation of a dependent variable (response variable) to specified independent variables (explanatory variables). However, the latter may not be sufficiently informative, when the conditional distribution possesses multimodality or a highly skewed profile with heteroscedastic noise. In such cases, a pertinent predictor can be obtained by estimating the conditional mode or the conditional median. In the last few years these models was widely studied when the explicative variable is functional and the response variable is real (see, for instance, Ferraty *et al.* (2005,2006) and Dabo-Niang and Laksaci (2007)). In this paper, we are interested in the nonparametric estimation of the conditional mode when the both variables (response and explanatory) are curves. We note that, the study of statistical models adapted to such kind of data has been an increasing interest in recent literature (see, for example, Ramsay and Silverman (1997, 2005), Bosq (2000), for parametric

models and Ferraty and Vieu (2006) in the nonparametric context). The first results on conditional nonparametric analysis for functional responses dates back to Bosq and Delecroix (1985). They consider the problem of forecasting by the nonparametric estimation of the Markovian transition operator valued in a Hilbert space. Lecoutre (1990) studied the regression function estimation when the response variable taking values in a Banach space and the explicative variable is real. This author gives the uniform almost completely convergence of the regressogram estimate of this model, in the case where the observations are i.i.d. Recently, under some differentiation condition Niang and Rhomari (2003) stated the convergence in L^p norm of a kernel estimator of the regression operator with the response and the explanatory variable are functionals. Recently, Ferraty and Laksaci *et al.* (2011) obtained a rate of the uniform almost complete convergence of the kernel estimator of this nonparametric model in the i.i.d. case. Among the recent lot of papers on the conditional mode estimation in functional statistics, we refer to the papers by Dabo-Niang and Laksaci (2010), Ezzahrioui and Ould-Said (2008) and Ferraty and Laksaci *et al.* (2010).
The principal aim of the present work is to generalize the results obtained by Ferraty and Laksaci *et al.* (2010) of the functional conditional mode with scalar response variable, to the case where the response variable is functional. More precisely, we study, the almost completely convergence rate of the kernel estimate of this models uniformly over some set of the functional space.
The paper is organized as follows. We present our model in Section 2. In Section 3 we introduce notations, Assumptions and state the main results. Section 4 is devoted to the application of our study in the prediction problem. The proofs of the results are relegated to the last section.

2.2 Conditional mode estimation

Consider a sample of independent pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$ identically distributed as (X, Y) which is a random vector valued in $\mathcal{F} \times \mathcal{B}$, where $\mathcal{F}$ is a semi-metric space, d denoting the semi-metric and $\mathcal{B}$ in Banach space. We suppose that, the conditional mode is uniquely defined over some fixed subset $S_{\mathcal{B}}$ of $\mathcal{B}$:

$$\theta(x) = \arg\sup_{y \in S_{\mathcal{B}}} f^x(y)$$

It is well know that, the estimation of the conditional mode is strongly linked to kernel estimate of the conditional density of Y given X. For this, we assume that there exists a regular version of the conditional probability of Y given X, which is absolutely continuous with respect an measure μ on $\mathcal{B}$ and we will denote by f^x the conditional density of Y given $X = x$. We define the kernel estimator $\widehat{f^x}$ of f^x as follows : $\forall y \in \mathcal{B},\ \forall x \in \mathcal{F}$,

$$\widehat{f^x}(y) = \frac{(E[K(h_2^{-1}||y-Y||)])^{-1}\sum_{i=1}^{n} K(h_1^{-1}d(x,X_i))K(h_2^{-1}||y-Y_i||)}{\sum_{i=1}^{n} K(h_1^{-1}d(x,X_i))},$$

with the functions K and H are kernels and $h_1 = h_{K,n}$ (resp. $h_2 = h_{H,n}$) is a sequence of positive real numbers.
A natural extension of the kernel estimator $\widehat{\theta(x)}$ of the conditional mode $\theta(x)$ to the functional framework is given by :

$$\widehat{\theta(x)} \;=\; \arg\sup_{y \in S_{\mathcal{B}}} \widehat{f}^x(y). \tag{2.1}$$

2.3 Notations and hypotheses

In order to establish the uniform almost complete convergence of the estimator $\theta(x)$, uniformly on fixed compact subset $S_{\mathcal{F}}$ of $\mathcal{F}$, we consider the following additional assumptions :

(H1) $\forall x \in S_{\mathcal{F}}$,

$$0 < C\phi(h) \leq P\left(X \in B(x,h)\right) \leq C'\phi(h) < \infty$$

and $\forall y \in S_{\mathcal{B}}$,

$$0 < Cg(h) \leq P\left(Y \in B(y,h)\right) \leq C'g(h) < \infty.$$

(H2) f^x is uniformly continuous and twice differentiable on $S_{\mathcal{B}}$ such that

$$\inf \frac{|D^2 f^x(\theta(x))(u,u)|}{\|u\|^2} > C > 0,$$

and $\forall (y_1, y_2) \in S_{\mathcal{B}} \times S_{\mathcal{B}}$

$$\forall (x_1,x_2) \in S_{\mathcal{F}} \times S_{\mathcal{F}}, ||f^{x_1}(y_1) - f^{x_2}(y_2)|| \leq C\left(d^{b_1}(x_1,x_2) + ||y_1 - y_2||^{b_2}\right),$$

(H3) $\forall \epsilon_0 > 0\ \exists \eta > 0,\ \forall r : S \to S_{\mathbb{R}}$,

$$\sup_{x \in S_{\mathcal{F}}} |\theta(x) - r(x)| \geq \epsilon_0 \Rightarrow \sup_{x \in S_{\mathcal{F}}} |f^x(r(x)) - f^x(\theta(x))| \geq \eta.$$

(H4) K is a bounded and Lipschitz kernel on its support $[0,1)$, and if $K(1) = 0$, the kernel K has to fulfill the additional condition $-\infty < C < K'(t) < C' < 0$.

(H5) The functions g, $\psi_{S_{\mathcal{B}}}$ and $\psi_{S_{\mathcal{F}}}$ are such that : $\exists C > 0,\ \exists \eta_0 > 0,\ \forall \eta < \eta_0,\ g'(\eta) < C$, and if $K(1) = 0$, the function $g(.)$ has to fulfill the additional condition :

$$\exists C > 0,\ \exists \eta_0 > 0,\ \forall 0 < \eta < \eta_0,\ \int_0^\eta g(u)\, du\ >\ C\, \eta\, g(\eta).$$

(H6) It exists $v_n = O\left(\dfrac{\log n}{n g(h_2)\, \phi(h_1)}\right)$ such that, $\forall \epsilon = O\left(\dfrac{\log n}{n}\right)$, $v_n \log n < \psi_{S_{\mathcal{F}}}(\epsilon) + \psi_{S_{\mathcal{B}}}(h_2 \epsilon) < \dfrac{1}{v_n}$,

(H7) The Kolmogorov's ϵ-entropy of $S_{\mathcal{F}}$ and $S_{\mathcal{B}}$ satisfies

$$\sum_{n=1}^{\infty} \exp\left\{(1-\beta)\left(\psi_{S_{\mathcal{F}}}(\epsilon) + \psi_{S_{\mathcal{B}}}(h_2 \epsilon)\right)\right\} < \infty, \text{ for some } \beta > 1,$$

2.4 Main Results

2.4.1 the uniform almost complete convergence

Proposition 2.1 *Under the hypotheses (H1)-(H3) and (H5)-(H7), we have*

$$\sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathcal{B}}} ||\widehat{f^x}(y) - f^x(y)|| = O(h_1^{b_1}) + O(h_2^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon) + \psi_{S_{\mathcal{B}}}(h_2 \epsilon)}{n g(h_2) \phi(h_1)}}\right).$$

Proof. The proof is based on the following decomposition

$$\widehat{f^x}(y) - f^x(y) = \frac{1}{\widehat{f}(x)}\Big[(\widehat{f^x_N}(y) - E\widehat{f^x_N}(y)) - (f^x(y) - E\widehat{f^x_N}(y))\Big] + \frac{f^x(y)}{\widehat{f}(x)}\Big[E\widehat{f}(x) - \widehat{f}(x)\Big] \tag{2.2}$$

where

$$\widehat{f}_N^x(y) = \frac{1}{nE[K(h_2^{-1}||y-Y||)]E[K(h_1^{-1}d(x,X_1))]} \sum_{i=1}^{n} K\left(h_1^{-1}d(x,X_i)\right) H\left(h_2^{-1}||y-Y_i||\right).$$

Theorem 2.1 can be deduced from the following intermediate results, together with Lemma 2.1 and Corollary 2.1.

Lemma 2.1 *Under the hypotheses (H1) and (H4)-(H7), we have*

$$\sup_{x\in S_{\mathcal{F}}} ||\widehat{f}(x)-1|| = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right).$$

Corollary 2.1 *Under the hypotheses of Lemma 2.1, we have*

$$\sum_{i=1}^{\infty} P\left(\inf_{x\in S_{\mathcal{F}}} \widehat{f}(x) < \frac{1}{2}\right) < \infty.$$

Lemma 2.2 *Under the hypotheses (H1)-(H4), we have*

$$\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathcal{B}}} ||f^x(y) - E\widehat{f}_N^x(y)|| = O(h_1^{b_1}) + O(h_2^{b_2}).$$

Lemma 2.3 *Under the assumptions of Theorem 2.1, we have*

$$\sup_{x\in S_{\mathcal{F}}}\sup_{y\in S_{\mathcal{B}}} ||\widehat{f}_N^x(y) - E\widehat{f}_N^x(y)|| = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)+\psi_{S_{\mathcal{B}}}(h_2\epsilon)}{ng(h_2)\phi(h_1)}}\right).$$

□

From Theorem 2.1 we derive the following result.

Theorem 2.1 *Under the hypotheses (H1)- (H7), we have*

$$\sup_{x\in S_{\mathcal{F}}} \left[||\widehat{\theta(x)}-\theta(x)||^2\right] = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)+\psi_{S_{\mathcal{B}}}(h_2\epsilon)}{ng(h_2)\phi(h_1)}}\right).$$

□

2.5 Appendix : Proofs

In the following, we will denote, for all $i = 1, \ldots, n$, by

$$K_i(x) = K(h_1^{-1} d(x, X_i)) \qquad \text{and} \qquad H_i(y) = K(h_2^{-1}||y - Y_i||).$$

First of all, according to (H1) and (H4), it is clear that if $K(1) > C > 0$,

$$\forall x \in S_{\mathcal{F}},\ \exists 0 < C < C' < \infty,\ C\phi(h_1) < E[K_1(x)] < C'\phi(h_1). \qquad (2.3)$$

and

$$\forall y \in S_{\mathcal{B}},\ \exists 0 < C < C' < \infty,\ Cg(h_2) < E[H_1(y)] < C'g(h_2).$$

In the situation when $K(1) = 0$, the combination of (H1) and (H5) allows to get the same result (see p44, Lemma 4.4. in Ferraty and Vieu, 2006).

Proof of Lemma 2.1.

Let $x_1, \ldots x_{N_\epsilon(S_{\mathcal{F}})}$ be an ϵ-net for $S_{\mathcal{F}}$ and for all $x \in S_{\mathcal{F}}$, one sets $k(x) = \arg \min\limits_{k \in \{1,2,\ldots,N_\epsilon(S_{\mathcal{F}})\}} d(x, x_k)$.

One considers the following decomposition :

$$\sup_{x \in S_{\mathcal{F}}} ||\widehat{f}(x) - E\widehat{f}(x)|| \leq \underbrace{\sup_{x \in S_{\mathcal{F}}} ||\widehat{f}(x) - \widehat{f}(x_{k(x)})||}_{F_1} + \underbrace{\sup_{x \in S_{\mathcal{F}}} ||\widehat{f}(x_{k(x)}) - E\widehat{f}(x_{k(x)})||}_{F_2}$$
$$+ \underbrace{\sup_{x \in S_{\mathcal{F}}} ||E\widehat{f}(x_{k(x)}) - E\widehat{f}(x)||}_{F_3}.$$

• Let us study F_1. By using (2.3) and the boundness of K, one can write :

$$\begin{aligned} F_1 &\leq \sup_{x \in S_{\mathcal{F}}} \frac{1}{n} \sum_{i=1}^{n} \left|\left| \frac{1}{E[K_1(x)]} K_i(x) - \frac{1}{E[K_1(x_{k(x)})]} K_i(x_{k(x)}) \right|\right| \\ &\leq \frac{C}{\phi(h_1)} \sup_{x \in S_{\mathcal{F}}} \frac{1}{n} \sum_{i=1}^{n} \left|\left| K_i(x) - K_i(x_{k(x)}) \right|\right| 1\!\!1_{B(x,h_1) \cup B(x_{k(x)},h_1)}(X_i). \end{aligned}$$

Let us first consider the case $K(1) = 0$. Because K is Lipschitz on $[0, 1]$ in this case, it comes

$$F_1 \ \leq \ \frac{C}{n} \sum_{i=1}^{n} Z_i \text{ with } Z_i = \frac{\epsilon}{h_1\, \phi(h_1)} 1\!\!1_{B(x,h_1) \cup B(x_{k(x)},h_1)}(X_i),$$

with $Z_1 = O\left(\dfrac{\epsilon}{h_1\phi(h_1)}\right)$, $EZ_1 = O\left(\dfrac{\epsilon}{h_1}\right)$ and $Var(Z_1) = O\left(\dfrac{\epsilon^2}{h_K^{12}\phi(h_1)}\right)$. A standard inequality for sums of bounded random variables (see Corollary A.9 in Ferraty and Vieu, 2006) with (H5b) allows to get

$$F_1 \;=\; O\left(\frac{\epsilon}{h_1}\right) + O_{a.co.}\left(\frac{\epsilon}{h_1}\sqrt{\frac{\log n}{n\,\phi(h_1)}}\right).$$

Now, it suffices to combine (H5) and (H6) to get

$$F_1 \;=\; O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\,\phi(h_1)}}\right).$$

Now, let $K(1) > C > 0$. In this situation K is Lipschitz on $[0,1)$. One has to decompose F_1 into three terms as follows :

$$F_1 \le C \sup_{x\in S_{\mathcal{F}}} \left(F_{11} + F_{12} + F_{13}\right),$$

with

$$\begin{aligned}
F_{11} &= \frac{1}{n\phi(h_1)}\sum_{i=1}^{n}\left|K_i(x) - K_i(x_{k(x)})\right|1\!\!1_{B(x,h_1)\cap B(x_{k(x)},h_1)}(X_i),\\
F_{12} &= \frac{1}{n\phi(h_1)}\sum_{i=1}^{n}K_i(x)1\!\!1_{B(x,h_1)\cap\overline{B(x_{k(x)},h_1)}}(X_i),\\
F_{13} &= \frac{1}{n\phi(h_1)}\sum_{i=1}^{n}K_i(x_{k(x)})1\!\!1_{\overline{B(x,h_1)}\cap B(x_{k(x)},h_1)}(X_i).
\end{aligned}$$

One can follow the same steps (i.e. case $K(1) = 0$) for studying F_{11} and one gets the same result :

$$F_{11} \;=\; O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\,\phi(h_1)}}\right).$$

Following same ideas for studying F_{12}, one can write :

$$F_{12} \;\le\; \frac{C}{n}\sum_{i=1}^{n}W_i \text{ with } W_i = \frac{1}{\phi(h_1)}1\!\!1_{B(x,h_1)\cap\overline{B(x_{k(x)},h_1)}}(X_i),$$

and by using again the same inequality for sums of bounded random variables, one has :

$$F_{12} \;=\; O\left(\frac{\epsilon}{\phi(h_1)}\right) + O_{a.co.}\left(\sqrt{\frac{\epsilon \log n}{n\,\phi(h_1)^2}}\right).$$

Similarly, one can state the same rate of convergence for F_{13}. To end the study of F_1, it suffices to put together all the intermediate results and to use again (H5b) for getting :

$$F_1 \;=\; O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right).$$

• Now, concerning F_2, we have, for all $\eta > 0$,

$$\begin{aligned} & P\left(F_2 > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right) \\ & \qquad = P\left(\max_{k\in\{1,\cdots,N_\epsilon(S_{\mathcal{F}})\}} ||\widehat{f}(x_{k(x)}) - E\widehat{f}(x_{k(x)})|| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right) \\ & \qquad \le N_\epsilon(S_{\mathcal{F}}) \max_{k\in\{1,\cdots,N_\epsilon(S_{\mathcal{F}})\}} P\left(||\widehat{f}(x_k) - E\widehat{f}(x_k)|| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right). \end{aligned}$$

Let $\Delta_{ki} = \dfrac{1}{E[K_1(x_k)]}\left(K_i(x_k) - E[K_i(x_k)]\right)$. We show, under (H1) and (H4), that $\forall k = 1,\ldots,N_\epsilon(S_{\mathcal{F}})$, $\forall i = 1,\ldots,n$,

$$\Delta_{ki} = O\left(\phi(h_1)^{-1}\right) \quad \text{ and also } \quad Var(\Delta_{ki}) = O\left(\phi(h_1)^{-1}\right).$$

So, one can apply the Bernstein-type inequality (see, Corollary A.9 in Ferraty and Vieu, 2006) which gives directly :

$$\begin{aligned} P\left(|\widehat{f}(x_k) - E\widehat{f}(x_k)| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right) \;&=\; P\left(\frac{1}{n}\left|\sum_{i=1}^{n}\Delta_{ki}\right| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right) \\ &\le\; 2\exp\{-C\eta^2\psi_{S_{\mathcal{F}}}(\epsilon)\}. \end{aligned}$$

Thus, by using the fact that $\psi_{S_{\mathcal{F}}}(\epsilon) = \log N_\epsilon(S_{\mathcal{F}})$ and by choosing η such that $C\eta^2 = \beta$, we have

$$N_\epsilon(S_{\mathcal{F}}) \max_{k\in\{1,\cdots,N_\epsilon(S_{\mathcal{F}})\}} P\left(|\widehat{f}(x_k) - E\widehat{f}(x_k)| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right) \le C' N_\epsilon(S_{\mathcal{F}})^{1-\beta}. \tag{2.4}$$

Because $\sum_{n=1}^{\infty} N_\epsilon(S_{\mathcal{F}})^{1-\beta} < \infty$, we obtain that

$$F_2 = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right).$$

• For F_3, it is clear that $F_3 \le E(\sup_{x\in S_{\mathcal{F}}} ||\widehat{f}(x) - \widehat{f}(x_{k(x)})||)$ and by following a similar proof to the one used for studying F_1, it comes $F_3 \ = \ O\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\phi(h_1)}}\right)$.
$\square$

Proof of Corollary 2.1. It is easy to see that,

$$\inf_{x\in S_{\mathcal{F}}} \widehat{f}(x) \le \frac{1}{2} \Rightarrow \exists x \in S_{\mathcal{F}}, \quad \text{such that} \quad 1-\widehat{f}(x) > \frac{1}{2} \Rightarrow \sup_{x\in S_{\mathcal{F}}} |1-\widehat{f}(x)| > \frac{1}{2}.$$

We deduce from Lemma 2.1 that

$$P\left(\inf_{x\in S_{\mathcal{F}}} \widehat{f}(x) \le \frac{1}{2}\right) \le P\left(\sup_{x\in S_{\mathcal{F}}} |1 - \widehat{f}(x)| > \frac{1}{2}\right).$$

Consequently ,

$$\sum_{i=1}^{\infty} P\left(\inf_{x\in S_{\mathcal{F}}} |\widehat{f}(x)| < \frac{1}{2}\right) < \infty.$$

$\square$

Proof of Lemma 2.2. One has

$$\begin{aligned} E\widehat{f_N^x}(y) - f^x(y) &= \frac{1}{nEH_1(x)EK_1(x)} E\left[\sum_{i=1}^{n} K_i(x)H_i(y)\right] - f^x(y) \\ &= \frac{1}{EH_1(x)EK_1(x)} E\left(K_1(x)\left[E(H_1(y)|X_1) - f^x(y)\right]\right). \end{aligned}$$

Moreover,

$$\begin{aligned} \tfrac{1}{EH_1(x)} E(H_1(y)|X_1) - f^x(y) &= \frac{1}{EH_1(x)} \int K(||\frac{y-z}{h_2}||) |f^{X_1}(z) d\mu - f^x(y) \\ &= \frac{1}{EH_1(x)} \int K(||\frac{y-z}{h_2}||) f^{X_1}(y - h_1(\frac{y-z}{h_1})) d\mu - f^x(y) \\ &= \frac{1}{EH_1(x)} \int K(||v||) f^{X_1}(y - h_1 v) d\mu^{\frac{y-Y}{h_1}}(v) - f^x(y) \\ &= \frac{1}{EH_1(x)} \int K(||v||)[f^{X_1}(y - h_1 v) - f^x(y)] d\mu^{\frac{y-Y}{h_1}}. \end{aligned}$$

Finally, the use of (H2) and (H4) implies that

$$||\frac{1}{EH_1(x)}E(H_1(y)|X_1) - f^x(y)|| \leq C(h_1^{b_1} + h_2^{b_2}).$$

This inequality is uniform on (x, y) in $S_{\mathcal{F}} \times S_{\mathcal{B}}$ and the use of (H4) states Lemma 2.2.

□

Proof of Lemma 2.3. Let us keep the definition of $k(x)$ as in Lemma 2.1 and let $t_1, \ldots t_{N_{h_2\epsilon}(S_{\mathcal{B}})}$ be an $h_2\epsilon$ -net for $S_{\mathcal{B}}$
for all $y \in S_{\mathcal{B}}$ we put

$$j(y) = \arg \min_{j \in \{1,2,\ldots,N_{h_2\epsilon}(S_{\mathcal{B}})\}} ||y - t_j||,$$

we have the following decomposition :

$$\begin{aligned} \left|\left|\widehat{f}_N^x(y) - E\widehat{f}_N^x(y)\right|\right| \leq\ & \underbrace{\sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathcal{B}}} \left|\left|\widehat{f}_N^x(y) - \widehat{f}_N^{x_{k(x)}}(y)\right|\right|}_{T_1} + \underbrace{\sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathcal{B}}} \left|\left|\widehat{f}_N^{x_{k(x)}}(y) - \widehat{f}_N^{x_{k(x)}}(t_{j(y)})\right|\right|}_{T_2} \\ & + \underbrace{\sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathcal{B}}} \left|\left|\widehat{f}_N^{x_{k(x)}}(t_{j(y)}) - E\widehat{f}_N^{x_{k(x)}}(t_{j(y)})\right|\right|}_{T_3} \\ & + \underbrace{\sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathcal{B}}} \left|\left|E\widehat{f}_N^{x_{k(x)}}(t_{j(y)}) - E\widehat{f}_N^{x_{k(x)}}(y)\right|\right|}_{T_4} \\ & + \underbrace{\sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathcal{B}}} \left|\left|E\widehat{f}_N^{x_{k(x)}}(y) - E\widehat{f}_N^x(y)\right|\right|}_{T_5}. \end{aligned}$$

Similarly to the study of the term F_1 and by replacing (H5b) with (H5b′), it comes

$$T_1 = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon) + \psi_{S_{\mathcal{B}}}(h_2\epsilon)}{ng(h_2)\phi(h_1)}}\right) \text{ and } T_5 = O\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon) + \psi_{S_{\mathcal{B}}}(h_2\epsilon)}{ng(h_2)\phi(h_1)}}\right). \tag{2.5}$$

Concerning the term T_2, by using the Lipschitz's condition on the kernel K, one can write

$$\begin{aligned}\left|\left|\widehat{f}_N^{x_{k(x)}}(y)-\widehat{f}_N^{x_{k(x)}}(t_{j(y)})\right|\right| &\leq C\frac{1}{n\,g(h_2)\,\phi(h_1)}\sum_{i=1}^n K_i(x_{k(x)})\,\left|\left|H_i(y)-H_i(t_{j(y)})\right|\right| \\ &\leq \frac{C}{n}\sum_{i=1}^n Z_i,\end{aligned}$$

where $Z_i = \dfrac{\epsilon\, K_i(x_{k(x)})}{g(h_2)\,\phi(h_1)}$. Once again, a standard exponential inequality for a sum of bounded variables allow us to write

$$\widehat{f}_N^{x_{k(x)}}(y)-\widehat{f}_N^{x_{k(x)}}(t_{j(y)}) = O\left(\frac{\epsilon}{g(h_2)}\right)+O_{a.co.}\left(\frac{\epsilon}{g(h_2)}\sqrt{\frac{\log n}{n\,\phi(h_1)}}\right).$$

In the same way that (2.5) we have

$$T_2 = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)+\psi_{S_{\mathcal{B}}}(h_2\epsilon)}{ng(h_2)\phi(h_1)}}\right) \text{ and } T_4 = O\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)+\psi_{S_{\mathcal{B}}}(h_2\epsilon)}{ng(h_2)\phi(h_1)}}\right). \tag{2.6}$$

Now, by using analogous arguments as for Lemma 2.1, we can show for all $\eta > 0$

$$\begin{aligned}&P\left(T_3 > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)+\psi_{S_{\mathcal{B}}}(h_2\epsilon)}{n\,g(h_2)\,\phi(h_1)}}\right) \\ &= P\left(\max_{j\in\{1,2,\dots,N_{h_2\epsilon}(S_{\mathcal{B}})\}}\max_{k\in\{1,\dots,N_\epsilon(S_{\mathcal{F}})\}}\left|\widehat{f}_N^{x_k}(t_j)-E\widehat{f}_N^{x_k}(t_j)\right| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)+\psi_{S_{\mathcal{B}}}(h_2\epsilon)}{n\,g(h_2)\,\phi(h_1)}}\right) \\ &\leq N_{h_2\epsilon}(S_{\mathcal{B}})N_\epsilon(S_{\mathcal{F}})\max_{j\in\{1,2,\dots,N_{h_2\epsilon}(S_{\mathcal{B}})\}}\max_{k\in\{1,\dots,N_\epsilon(S_{\mathcal{F}})\}} \\ &\qquad P\left(\left|\widehat{f}_N^{x_k}(t_j)-E\widehat{f}_N^{x_k}(t_j)\right| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)+\psi_{S_{\mathcal{B}}}(h_2\epsilon)}{n\,g(h_2)\,\phi(h_1)}}\right).\end{aligned}$$

Let

$$\Delta_i = \frac{1}{g(h_2)\,\phi(h_1)}\left[K_i(x_k)H_i(t_j) - E\left(K_i(x_k)H_i(t_j)\right)\right].$$

To apply the Bernstein exponential inequality (see, Corollary A.9 in Ferraty and Vieu, 2006), we must calculate the asymptotic behavior of $E|\Delta_i|$ and

$E\Delta_i^2$. Indeed, it follows from the fact that the kernel K is bounded that $E|\Delta_i| \leq C\,(g(h_2)\,\phi(h_1))^{-1}$. Secondly, the use of the same analytic arguments as of Lemma 2.2, allows us to get :

$$\lim_{n\to\infty} \frac{1}{g(h_2)} E[H_1^2(y)|X_1] = O(1),$$

which implies that

$$E|\Delta_i|^2 \leq \frac{C}{g(h_2)\,\phi(h_1)}.$$

Thus, we are now in position to apply the Bernstein exponential inequality and we get :

$$\forall j \leq N_{h_2\epsilon}(S_{\mathcal{B}}),\ P\left(\left|\widehat{f}_N^{x_k}(t_j) - E\widehat{f}_N^{x_k}(t_j)\right| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon) + \psi_{S_{\mathcal{B}}}(h_2\epsilon)}{n\,g(h_2)\,\phi(h_1)}}\right)$$

$$\leq 2\exp\{-C\eta^2\psi_{S_{\mathcal{F}}}(\epsilon) + \psi_{S_{\mathcal{B}}}(h_2\epsilon)\}.$$

Therefore, by choosing $C\eta^2 = \beta$ one has

$$N_{h_2\epsilon}(S_{\mathcal{B}})N_\epsilon(S_{\mathcal{F}}) \max_{j\in\{1,2,\dots,N_{h_2\epsilon}(S_{\mathcal{B}})\}} \max_{k\in\{1,\dots,N_\epsilon(S_{\mathcal{F}})\}} P\left(\left|\widehat{f}_N^{x_k}(t_j) - E\widehat{f}_N^{x_k}(t_j)\right| > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon)}{n\,h_2\,\phi(h_1)}}\right)$$

$$\leq C'\left(N_{h_2\epsilon}(S_{\mathcal{B}}N_\epsilon(S_{\mathcal{F}})\right)^{1-C\eta^2}.$$

Finally one obtains

$$T_3 = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}(\epsilon) + \psi_{S_{\mathcal{B}}}(h_2\epsilon)}{ng(h_2)\phi(h_1)}}\right) \tag{2.7}$$

So, Lemma 2.3 can be easily deduced from (2.5)-(2.7).

□

Proof of Theorem 2.1. By a simple manipulation, we show that

$$||f^x(\widehat{\theta(x)}) - f^x(\theta(x))|| \leq 2\sup_{y\in S}|\widehat{f^x}(y) - f^x(y)|. \tag{2.8}$$

We use the following Taylor expansion of the function f^x :

$$f^x(\widehat{\theta(x)}) = f^x(\theta(x)) + \frac{1}{2}D^2 f^x(\theta'(x))(\widehat{\theta(x)} - \theta(x), \widehat{\theta(x)} - \theta(x)) + o\left(\|\widehat{\theta(x)} - \theta(x)\|^2\right),$$

Clearly, it follows by (H7)

$$\sup_{x \in S_{\mathcal{F}}} ||\widehat{\theta(x)} - \theta(x)||^2 \leq C \sup_{x \in S_{\mathcal{F}}} \sup_{y \in S_{\mathcal{B}}} ||\widehat{f}^x(y) - f^x(y)||, \qquad a.co.$$

By combining this result with Proposition 2.1 we obtain the claimed result.

$\square$

References

BOSQ, D. (2000). *Linear processs in function spaces. Theory and Application. Lectures Notes in Statistics*, Vol 149. Springer-Verlag.

BOSQ, D. and DELECROIX, M. (1985). Nonparametric prediction of a Hilbert-space valued random variable. *Stochastic Process. Appl.* **19** 271-280.

DABO-NIANG, S., RHOMARI, N. (2009). Kernel regression estimation in a Banach space. *J. Statist. Plann. Inference* **139**, 1421-1434.

DABO-NIANG, S., and LAKSACI, A. (2010). Note on conditional mode estimation for functional dependent data, *Statistica*, **70**, 83-94.

DABO-NIANG, S., and LAKSACI, A. (2007). *Estimation non paramïétrique du mode conditionnel pour variable explicative fonctionnelle, C. R., Math., Acad. Sci. Paris*, **344**, 49-52.

DABO-NIANG, S., and LAKSACI, A. (2010). Note on conditional mode estimation for functional dependent data, *Statistica*, **70**, 83-94.

EZZAHRIOUI, M. and OULD-SAID, E. (2008). Asymptotic normality of nonparametric estimators of the conditional mode function for functional data *J. Nonparametric Statist.* , **20**, 3-18.

FERRATY, F., LAKSACI, A. AND VIEU, P.(2005). Functional times series prediction via conditional mode. *C. R., Math., Acad. Sci. Paris*, **340**, No.5, 389-392.

FERRATY, F., LAKSACI, A., VIEU, P. (2006). Estimating some chacteristics of the conditional distribution in nonparametric functional models. *Statist. Inf. for Stoch. Proc.*, **9**, 47-76.

FERRATY, F., LAKSACI, A., TADJ, A. and VIEU, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. Inference.* **140** 335-352.

FERRATY, F., LAKSACI, A., TADJ, A. and VIEU, P. (2011). Kernel regression with functional response. *Electronic Journal of Statistics.* **5** 159-171.

FERRATY, F., VIEU, P. (2006). *Nonparametric functional data analysis. Theory and Practice.* Springer-Verlag.

DABO-NIANG, S., RHOMARI, N. (2003). Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *C. R., Math., Acad. Sci. Paris.* **336**, 75-80.

LECOUTRE, J. P. (1990). Uniform consistency of a class of regression function estimators for Banach-space valued random variable. *Statist. Probab. Lett.* **10** 145-149.

RAMSAY, J. O. and SILVERMAN, B. W. (1997). *Functional data analysis.*

Springer, New York.
RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd ed. Springer,New-York.

Chapitre 3

Recursive estimation of the conditional mode for functional explanatory variable

Ce chapitre fait l'objet d'une soumission dans le Journal of Applicationes Mathematicae.

RECURSIVE ESTIMATION OF THE CONDITIONAL MODE FOR FUNCTIONAL EXPLANTORY VARIABLE

Ali Laksaci and Ardjoun Fatima Zohra
Laboratoire de Statistique et Processus Stochastiques
Univ. Djillali Liabès
BP 89, 22000 Sidi Bel Abbès, Algeria
e-mail : (alilak or maila.nabila)@yahoo.fr

Abstract

In this paper, we consider an alternative estimator of the conditional mode when the explanatory variable takes values in a semi-metric space. This alternative estimate is based in a recursive kernel method. We quantify the asymptotic properties of this estimate, by giving the almost complete convergence rate.

Mathematics Subject Classification : *62G05, 62G08, 62G20*

Key Words : The almost complete convergence, kernel estimators, functional data, recursive estimate, semi-metric space.

3.1 Introduction

It is well known that in statistics, regression model is main tool to examine the relation of a dependent variable (response variable) to specified independent variables (explanatory variables). However, the latter may not be sufficiently informative, when the conditional distribution possesses multimodality or a highly skewed profile with heteroscedastic noise. In such cases, a pertinent predictor can be obtained by estimating the conditional mode or the conditional median.
The literature on conditional mode estimation is quite numerous when the explanatory variable X is real (see for instance Collomb *et al.*, (1987), for previous works and Quintela and Vieu, (1997), Berlinet *et al.*, (1998), or Louani and Ould-Saïd, (1999), for recent advances and references). In the last few years these models were widely studied when the explanatory variable is functional and the response variable is real. The first results was obtained

in the i.i.d. case by Ferraty *et al.* (2006). They established the almost complete convergence rate of the kernel estimate of the conditional mode. The L_p consistency of the same estimator was obtained by Dabo-Niang and Laksaci (2007, 2010). Ezzahrioui and Ould-Saïd (2008) stated the asymptotic normality of the kernel estimate of the conditional mode. Recently, Dabo-Niang *et al.* (2012) consider the spatial version of this kernel estimator. They obtained the almost complete convergence rate of the adaptive estimator. Unlike in these cited works, where the authors consider the kernel estimator of the conditional mode, we are interested, here, in the nonparametric estimation of the conditional mode by the recursive method. Clearly the interest in this method comes from the fact that this method permits to update our estimator for each additional information. This feature is very adequate in this context of functional sadistics when the date are very bulky and are available by real time monitoring. Moreover, the kernel method can be viewed as particular case of our study. Specifically, as asymptotic result we prove the almost complete convergence and rate of this estimator. Among the lot of papers concerning the nonparametric recursive estimation or modelization of variable taking values in infinite dimensional spaces, we only refer to the papers by Roussas (1989), Thiam (2006) Ferraty *et al.* (2010), Bosq (2000), Ramsay and Silverman (2002-2005), Ferraty and Vieu (2006).

The paper is organized as follows. We present our model in Section 2. In Section 3 we introduce notations, Assumptions and we state the main results. The proofs of the results are relegated to the last section.

3.2 Recursive conditional mode estimation

Consider n pairs of random variables (X_i, Y_i) for $i = 1, \ldots, n$ that we assume drawn from the pair (X, Y) which is valued in $\Im \times \mathbb{R}$ where $\Im$ is a semi-metric vector space, $d(.,.)$ 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 an unique mode θ on S. In the remainder of the paper x is fixed in $\Im$ and N_x denotes a neighborhood of x. Let f^x be the conditional density of the variable Y given $X = x$. We define the kernel estimator $\widehat{f^x}$ of f^x as follows :

$$\widehat{f^x}(y) = \frac{\sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right) b_i^{-1} H\left(b_i^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right)}, \tag{3.1}$$

where K and H are kernels and a_n (resp. b_n) is a sequence of positive real numbers. Note that we can treat the classical kernel method studied by Ferraty *et al.* (2006) as particular case of the present study by taking $a_i = a_n$ and $b_i = b_n$.

Now, a natural extension of the kernel estimator $\hat{\theta}$ of the conditional mode θ to the functional framework is given by :

$$\hat{\theta} \;=\; \arg\sup_{y \in S} \widehat{f^x}(y). \tag{3.2}$$

The estimate $\hat{\theta}$ is not necessarily unique, and if this is the case, all the remaining of our paper concerns any value $\hat{\theta}$ satisfying (3.2). This work establishes the almost complete convergence of the kernel estimate $\hat{\theta}$ of θ under some standard conditions.

3.3 Hypotheses and main result

Throughout the paper, when no confusion is possible, we will denote by C or C' some strictly positive generic constants, and we will use the notation $B(x, h) = \{x' \in \mathcal{F} \,:\, d(x', x) < h\}$. Our nonparametric model will be quite general in the sense that we will just need the following assumptions :

(H1) $P\left(X \in B(x, r)\right) = \phi(x, r) > 0$, for all $r > 0$

(H2) The conditional density f^x satisfies

$$\begin{cases} \text{i) } f^x \text{ is } 2- \text{ times continuously differentiable with respect to } y \text{ on } (\theta - \xi, \theta + \xi), \\ \qquad\qquad \text{such that } 0 < \left| f^{x(2)}(\theta) \right| < \infty \\ \text{ii) } \forall (y_1, y_2) \in S \times S,\ \forall (x_1, x_2) \in N_x \times N_x \\ \qquad\qquad |f^{x_1}(y_1) - f^{x_2}(y_2)| \leq C_x \left(d(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2} \right), \\ \text{iii) } f^x \text{ is strictly increasing on } (\theta - \xi, \theta) \text{ and strictly decreasing on } (\theta, \theta + \xi). \end{cases}$$

(H3) K is a function with support $(0, 1)$ such that

$$0 < C < K(t) < C' < \infty,$$

(H4) The kernel H is bounded such that

$$\forall (y_1, y_2) \in \mathbb{R}^2, \ \ |H(y_1) - H(y_2)| \leq C|y_1 - y_2| \ \text{ and } \ \int |t|^{b_2} H(t)dt < \infty.$$

(H5) The sequences a_n and b_n such that

$$\begin{cases} \text{i) There existe } C > 0 \quad \sup_{i \leq n} \left(\dfrac{b_n}{b_i} \right) \leq C < \infty \quad \text{and} \quad \exists \gamma > 0 \qquad \lim_{n\to\infty} n^{\gamma}\, b_n = \infty \\ \text{ii) } \lim_{n\to\infty} \dfrac{\log n}{n\varphi_n(x)} = 0 \end{cases}$$

where $\varphi_n(x) = n^{-1} \sum_{i=1}^{n} \phi(x, a_i)$.

Observe that, these conditions are very standard in this context. Indeed, condition (H1)-(H4) are the same as that used by Ferraty and Vieu (2006). Assumption (H5) is a similar technical condition considered by Thiam (2006) for the vectorial case.
The following Theorem gives the almost-complete convergence [1] (a.co.) of $\widehat{\theta}$.

Theorem 3.1 *Under the hypotheses (H1)-(H5) we have*

$$\hat{\theta} - \theta = O\left(\frac{1}{n\varphi_n(x)} \sum_{i=1}^{n} \left(a_i^{\beta_1} \phi(x, a_i) \right) \right)^{1/2} + O\left(\frac{1}{n\varphi_n(x)} \sum_{i=1}^{n} \left(b_i^{\beta_2} \phi(x, a_i) \right) \right)^{1/2} + O\left(\sqrt{\frac{\log n}{n b_n \varphi_n(x)}} \right)^{1/2}$$

Remarque 3.1

Here, we point out that, if $(a_i, b_i) = (a_n, b_n)$ for all $i = 1, \ldots n$, we obtain the same rate given in Ferraty *et al.* (2006) for the classical kernel case.

1. Let $(z_n)_{n\in\mathbb{N}}$ be a sequence of real r.v.'s ; we say that z_n converges almost completely (a.co.) to zero if, and only if, $\forall \epsilon > 0$, $\sum_{n=1}^{\infty} P(|z_n| > \epsilon) < \infty$. Moreover, we say that the rate of almost complete convergence of z_n to zero is of order u_n (with $u_n \to 0$) and we write $z_n = O_{a.co.}(u_n)$ if, and only if, $\exists \epsilon > 0$, $\sum_{n=1}^{\infty} P(|z_n| > \epsilon u_n) < \infty$. This kind of convergence implies both almost sure convergence and convergence in probability.

Proof of Theorem 3.1 Firstly, observe that by a Taylor expansion of the function f^x we have

$$|(\theta - \hat{\theta})|^2 \leq \frac{2}{f^{x(2)}(\theta^*)} \sup_{y \in S} |\hat{f}^x(y) - f^x(y)|,$$

for some θ^* between θ and $\hat{\theta}$. Moreover, by simple analytic arguments, we can show that

$$|f^x(\hat{\theta}) - f^x(\theta)| \leq 2 \sup_{y \in S} |\hat{f}^x(y) - f^x(y)|. \tag{3.3}$$

Therefore, all what is left to do, is to study the convergence rate of

$$\sup_{y \in S} |\hat{f}^x(y) - f^x(y)|$$

and to show that

$$\exists c > 0, \ \sum_{n=1}^{\infty} P\left(f^{x(2)}(\theta^*) < c\right) \ < \ \infty.$$

The latter is based on the following decomposition

$$\widehat{f}(y) - f^x(y) = \widehat{B}_n(x,y) + \frac{\widehat{R}_n(x,y)}{\widehat{f}_D^x} + \frac{\widehat{Q}_n(x,y)}{\widehat{f}_D^x}$$

where

$$\begin{aligned}
\widehat{Q}_n(x,y) \ &:= (\widehat{f}_N^x(y) - \mathbb{E}[\widehat{f}_N^x(y)]) - f^x(y)(\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x]) \\
\widehat{B}_n(x,y) \ &:= \frac{\mathbb{E}[\widehat{f}_N^x(y)]}{\mathbb{E}[\widehat{f}_D^x]} - f^x(y), \quad \text{and} \quad \widehat{R}_n(x,y) := -\widehat{B}_n(x,y)(\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x])
\end{aligned}$$

with

$$\widehat{f}_N^x(y) = \frac{1}{n\,\varphi_n(x)} \sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right) b_i^{-1} H\left(b_i^{-1}(y - Y_i)\right)$$

and

$$\widehat{f}_D^x \ = \ \frac{1}{n\varphi_n(x)} \sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right).$$

Finally, Theorem **3.1** can be deduced from the Lemmas ((3.1), (3.2), (3.3), (3.4)) and Corollary (3.1) :

Lemma 3.1
Under the hypotheses (H1), (H3) and (H5) we have

$$\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x] = O\left(\sqrt{\frac{\log n}{n\varphi_n(x)}}\right), \quad a.co.$$

Corollary 3.1
Under the hypotheses of Lemma 3.1, we have,

$$\exists\, C > 0 \quad \sum_{n=1}^{\infty} \mathbb{P}\left(\widehat{f}_D^x < C\right) < \infty.$$

Lemma 3.2
Under the hypotheses of Theorem 3.1 we have

$$\sup_{y\in S} |\widehat{B}_n(x,y)| = O\left(\frac{1}{n\varphi_n(x)} \sum_{i=1}^{n} \left(a_i^{\beta_1}\phi(x,a_i)\right)\right) + O\left(\frac{1}{n\varphi_n(x)} \sum_{i=1}^{n} \left(b_i^{\beta_2}\phi(x,a_i)\right)\right), \quad a.co.$$

Lemma 3.3
Under the hypotheses of Theorem 3.1 we have

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

Lemma 3.4
Under the hypotheses of Theorem 3.1 we have

$$\exists C > 0, \quad \sum_{n=1}^{\infty} P\left(f^{x(j)}(\theta^*) < C\right) \; < \; \infty.$$

3.4 Appendix

Proof of Lemma 3.1
For all $i = 1, \ldots, n$ we put $K_i(x) = K(a_i^{-1}d(x, X_i))$ and $\Delta_i(x) = K_i(x) - \mathbb{E}\left[K_i(x)\right]$. Then, it can be seen that

$$\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x] = \frac{1}{\varphi_n(x)} \sum_{i=1}^{n} \Delta_i(x).$$

Now, we apply the Bernstein's inequality on the random variable $\Delta_i(x)$. Clearly, under (H3)

$$|\Delta_i(x)| \leq C \text{ and } \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} Var\left[\Delta_i(x)\right] \leq \frac{C'}{n}\sum_{i=1}^{n}\phi(x,a_i) = C'\varphi_n(x).$$

It follows that for all $\varepsilon > 0$, we have

$$\begin{aligned}
\mathbb{P}\left\{\left|\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x]\right| > \varepsilon\right\} &= \mathbb{P}\left\{\left|\frac{1}{n\varphi_n(x)}\sum_{i=1}^{n}\Delta_i(x)\right| > \varepsilon\right\} \\
&= \mathbb{P}\left\{\frac{1}{n}\left|\sum_{i=1}^{n}\Delta_i(x)\right| > 2\varepsilon\varphi_n(x)\right\} \\
&\leq 2\exp\left\{-\frac{n\varepsilon^2\varphi_n^2(x)}{2\sigma^2 + 2C\varepsilon\varphi_n(x)/3}\right\} \\
&\leq 2\exp\left\{-\varepsilon^2 n\varphi_n(x)\left(\frac{1}{2C' + 2C\varepsilon/3}\right)\right\}.
\end{aligned}$$

Finally, taking $\varepsilon = \epsilon_0\sqrt{\dfrac{\log n}{n\varphi_n(x)}}$ to show that

$$\mathbb{P}\left\{\left|\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x]\right| > \epsilon_0\sqrt{\frac{\log n}{n\varphi_n(x)}}\right\} \leq n^{-C\epsilon_0^2}.$$

Consequently an appropriate choice of ϵ_0 complete the proof of this lemma.□
Proof of Corollary 3.1 We have, under (H3), there exists $0 < C < C' < \infty$

$$0 < C\frac{1}{n\varphi_n(x)}\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i)) < \mathbb{E}[\widehat{f}_D^x] < \left|\widehat{F}_D^x - \mathbb{E}[\widehat{f}_D^x]\right| + \widehat{f}_D^x.$$

Since $\varphi_n(x) = n^{-1}\sum_{i=1}^{n}\phi(x,a_i)$, we have

$$C - \left|\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x]\right| < \widehat{f}_D^x.$$

It follows that

$$\mathbb{P}\left(\widehat{f}_D^x \leq \frac{C}{2}\right) \leq \mathbb{P}\left(\left|\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x]\right| > \frac{C}{2}\right).$$

It is obvious that the previous Lemma allows to get the desired result. $\square$

Proof of Lemma 3.2 For all $i = 1,\ldots,n$ we put $H_i(y) = H(b_i^{-1}(y - Y_i))$. Conditioning by X_i we gives

$$\begin{aligned}
\widehat{B}_n(x,y) &= \frac{\mathbb{E}[\widehat{f}_N^x(y)] - f^x(y)\mathbb{E}[\widehat{f}_D^x]}{\mathbb{E}[\widehat{f}_D^x]} \\
&= \frac{1}{n\varphi_n(x)\mathbb{E}[\widehat{f}_D^x]}\sum_{i=1}^n \left(\mathbb{E}\left[K_i(x)b_i^{-1}\mathbb{E}[H_i(y)|X_i]\right] - f^x(y)\mathbb{E}\left[K_i(x)\right]\right) \\
&\leq \frac{1}{n\varphi_n(x)\mathbb{E}[\widehat{f}_D^x]}\sum_{i=1}^n \mathbb{E}\left[K_i(x)\left|\mathbb{E}\left[b_i^{-1}H_i(y)|X\right] - f^x(y)\right|\right]
\end{aligned}$$

Now, (H2ii) allows to write

$$\left|E\left(b_i^{-1}H_i(y)|X\right) - f^x(y)\right| \;\leq\; C_x\int_{\mathbb{R}} H(t)\left(a_i^{b_1} + |t|^{b_2}b_i^{b_2}\right)dt \qquad (3.4)$$

By combining this result with the fact that,

$$\mathbb{E}\left[K_i(x)\right] \leq C'\phi(x,a_i) \;\text{ and }\; \mathbb{E}[\widehat{f}_D^x] = O(1)$$

we obtain the desired result. $\square$

Proof of Lemma 3.3 Using the compactness of S, we can write that $S \subset \bigcup_{k=1}^{s_n} S_k$ where $S_k = (t_k - l_n,\ t_k + l_n)$. Taking $t_y = \arg\min_{t\in\{t_1,\ldots,t_{s_n}\}}|y-t|$. Furthermore, we consider the following decomposition :

$$\begin{aligned}
\frac{1}{\widehat{f}_D^x}\sup_{y\in S}\left|\widehat{f}_N^x(y) - \mathbb{E}[\widehat{f}_N^x(y)]\right| \;\leq\;& \underbrace{\frac{1}{\widehat{f}_D^x}\sup_{y\in S}\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_y))\right|}_{T_1}\underbrace{\frac{1}{\widehat{f}_D^x}\sup_{y\in S}\left|\widehat{f}_N^x(t_y) - \mathbb{E}[\widehat{f}_N^x](t_y)\right|}_{T_2} \\
&+ \underbrace{\frac{1}{\widehat{f}_D^x}\sup_{y\in S}\left|\mathbb{E}[\widehat{f}_N^x](t_y) - \mathbb{E}[\widehat{f}_N^x(y)]\right|}_{T_3} \qquad (3.5)
\end{aligned}$$

where $t_y = \arg\min_{t\in\{y_1,\ldots,y_{d_n}\}}|y-t|$.

- Firstly, concerning (T_1) : By the first part of (H5ii), we have

$$\begin{aligned}
\frac{1}{\widehat{f}_D^x}\sup_{y\in S}\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_y)\right| &\leq \frac{1}{n\,\varphi_n(x)\,\widehat{f}_D^x}\sup_{y\in S}\sum_{i=1}^n\left|H_i(y) - H_i(t_y)\right|K_i, \\
&\leq \frac{C}{n\,\varphi_n(x)\,\widehat{f}_D^x}\sup_{y\in S}|y - t_y|\left(\sum_{i=1}^n\frac{K_i}{b_i^2}\right), \\
&\leq C\frac{l_n}{b_n^2}.
\end{aligned}$$

By taking $l_n = n^{-3\gamma/2-1/2}$ we obtain that

$$\frac{l_n}{b_n^2} = o\left(\sqrt{\frac{\log n}{nb_n\varphi_n(x)}}\right).$$

So, for n large enough, we can found a $\eta > 0$ such that

$$\mathbb{P}\left(\frac{1}{\widehat{f}_D^x}\sup_{y\in S}\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_y)\right| > \eta\sqrt{\frac{\log n}{nb_n\varphi_n(x)}}\right) \;=\; 0. \tag{3.6}$$

and

$$\frac{1}{\widehat{f}_D^x}\sup_{y\in S}\left|\mathbb{E}[\widehat{f}_N^x(y)] - \mathbb{E}[\widehat{f}_N^x(t_y))]\right| \leq C\frac{l_n}{b_n^2} = o\left(\sqrt{\frac{\log n}{nb_n\varphi_n(x)}}\right) \tag{3.7}$$

- Secondly, for (T_2) we use the same arguments as for the proof of Lemma 3.1, under the fact that

$$\forall\, i = 1, \ldots, n \quad \mathbb{E}\left|K_i b_i^{-1} H_i\right| \leq C b_n^{-1}$$

and

$$\mathbb{E}\left[K_i b_i^{-1} H_i - \mathbb{E}\left[K_i b_i^{-1} H_i\right]\right]^2 \leq C b_i^{-1}\phi(x, a_i) \leq C b_n^{-1}\phi(x, a_i).$$

Therefore, for all $\eta > 0$, we have

$$\mathbb{P}\left(\sup_{y\in S}\left|\widehat{f}_N^x(t_y) - \mathbb{E}[\widehat{f}_N^x](t_y))\right| > \eta\sqrt{\frac{\log n}{nb_n\varphi_n(x)}}\right) \leq 2s_n\exp\{-C\eta^2\log n\}.$$

Consequently, an appropriate choice of η done

$$\sum_n \mathbb{P}\left(\sup_{y\in S}\left|\widehat{f}_N^x(y) - \mathbb{E}[\widehat{f}_N^x](t_y))\right| > \eta\sqrt{\frac{\log n}{nb_n\varphi(x)}}\right) < \infty. \tag{3.8}$$

Now, Lemma 3.3 can be easily deduced from (3.6), (3.7), (3.8) and (3.5). □
Proof of Lemma 3.4 The proof of this Lemma is very similar to Lemma 3.1.4 in Ferraty *et al.* (2005). Indeed, it based in the almost complete convergence of

$$\hat{\theta} - \theta \to 0,\ a.co. \tag{3.9}$$

Because of (H2ii), we have :

$$\forall \epsilon > 0,\ \exists \delta(\epsilon) > 0, \forall y \in (\theta - \xi, \theta + \xi), |f^x(y) - f^x(\theta)| \leq \delta(\epsilon) \Rightarrow |y - \theta| \leq \epsilon.$$

Thus,

$$\forall \epsilon > 0,\ \exists \delta(\epsilon) > 0,\ \ P\left(|\hat{\theta} - \theta| > \epsilon\right) \leq P\left(|f^x(\hat{\theta}) - f^x(\theta)| \geq \delta(\epsilon)\right). \quad (3.10)$$

Then, $\hat{\theta} - \theta$ tend to 0 by means of (3.3). Finally, Lemma 3.4 follows directly by using the continuity of $f^{(2)}$ together with (H2i) and (3.9). $\square$

Bibliographie

Berlinet, A., Gannoun, A. and Matzner-Lober, E. (1998). Normalité asymptotique d'estimateurs convergents du mode conditionnel (french). *La Rev. Canad. de Statist.*, **26**, 365-380.

Bosq, D. (2000). *Linear Processes in Function Spaces : Theory and applications.* Lecture Notes in Statistics, 149, Springer.

Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf.*, **15**, 227-236.

Dabo-Niang, S., and Laksaci, A. (2007). *Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle, C. R., Math., Acad. Sci. Paris*, **344**, 49-52.

Dabo-Niang, S. and Laksaci, A. (2010). Note on conditional mode estimation for functional dependent data, *Statistica*, **70**, 83-94.

Dabo-Niang, S. and Laksaci, A. (2012). On spatial conditional mode estimation for a functional regressor, *Statistics Probability Letters, Elsevier*, vol. **82(7)**, pages 1413-1421.

Ezzahrioui, M. and Ould-Said, E. (2008). Asymptotic normality of nonparametric estimators of the conditional mode function for functional data *J. Nonparametric. Statist*, **20**, 3-18

Ferraty, F., Laksaci, A. and Vieu, P.(2005). Functional times series prediction via conditional mode. *C. R., Math., Acad. Sci. Paris*, **340**, No. 5, 389-392.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis.* Springer Series in Statistics, Springer New York.

Ferraty, F. and Laksaci, A. and Vieu, P. (2006). Estimating some chacteristics of the conditional distribution in nonparametric functional models. *Statist. Inf. for Stoch. Proc.* **9**, 47-76.

Ferraty, F. Romain, Y (2010) Handbook of functional data anlysis rela-

ted fields OXford university press.OXford.
Quintela del Rio, A. and Vieu, P. (1997). A nonparametric conditionnal mode estimate. *Nonparametric Statistics*, **8**, 253-266.
Louani, D. and Ould-Saïd, E. (1999). Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametric Stat.*, **11** (4), 413-442.
Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis.* Second Edition. Springer, New York.
Roussas, G.R. and Tran, L.T. (1989). Asymptotic normality of the recursive kernel regression estimate under dependence conditions. Ann. Statist., **20**, 98à120.
Thiam, B. (2006). Recursive estimation of functional. Thesis Doctarat, University Versailies.

Chapitre 4

A recursive kernel estimates of the functional modal regression under ergodic dependence condition

Ce chapitre fait l'objet d'une publication dans le Journal of Statistical Theory and Practice.

A RECURSIVE KERNEL ESTIMATES OF THE FUNCTIONAL MODEL REGRESSION UNDER ERGODIC DEPENDENCE CONDITION

Fatima Zohra Ardjoun[†], **Larbi Ait Hennani** [††] and **Ali Laksaci,**[†, *]

[†] Laboratoire de Statistique et Processus Stochastiques,
Université de Sidi Bel Abbès,
BP 89 Sidi Bel Abbès 22000. Algeria.
[*] Corresponding author : alilak@yahoo.fr
[††] Institut universitaire de technologie , Lille 2,
Rond-point de l'Europe, BP 557 59060 Roubaix, France,

Abstract
In this paper, we consider an alternative estimator of the conditional mode when the explanatory variable takes values in a semi-metric space. This alternative estimate is based in a recursive kernel method. Under the ergodicity hypothesis, we quantify the asymptotic properties of this estimate, by giving the almost complete convergence rate. The asymptotic normality of this estimate is also given. Our approach is illustrated by a real data application.

AMS subject classifications : 60G42, 62F12, 62G20, 62G05.

Key words : Functional data, Conditional mode Conditional distribution, recursive estimation, ergodic condition.

4.1 Introduction

Usually, the time series data is modeled by using the mixing conditions, our main aim in this paper is to analyze the functional time series data under an alternative dependency. More precisely, we consider the problem of the nonparametric functional analysis under the ergodicity assumption. This dependence structure covers several case does not satisfy the usual mixing structures. On the other hand, the ergodicity condition is one of a principal

postulate of statistical physics. It models the thermodynamic properties of gases, atoms, electrons or plasmas. This hypothesis is also used in signal processing, for studying the evolution of a random signal.
The litterateur in egrodic functional data is still limited. This problem has been initiated by Laib and Louani (2010, 2011). They consider the problem of functional estimation for nonparametric regression operators under the ergodicity condition. However, the latter may not be sufficiently informative, when the conditional distribution possesses multi-modality or a highly skewed profile with heteroscedastic noise. In such cases, a pertinent predictor can be obtained by estimating the conditional mode or the conditional median.
The literature on conditional mode estimation is quite numerous when the explanatory variable is real (see for instance Collomb *et al.*, (1986), Quintela and Vieu, (1997), Louani and Ould-Saïd, (1999) and references therein). In the last few years these models were widely studied when the explanatory variable is functional and the response variable is real. The first results was obtained in the i.i.d. case by Ferraty *et al.*, (2006). They established the almost complete convergence rate of the kernel estimate of the conditional mode. The L_p consistency of the same estimator was obtained by Dabo-Niang and Laksaci (2007). Ezzahrioui and Ould-Saïd (2008) stated the asymptotic normality of the kernel estimate of the conditional mode in the i.i.d case. These results are extended to the dependent case by Ezzahrioui and Ould-Saïd (2010). Recently, Dabo-Niang *et al.* (2012) consider the spatial version of this kernel estimator. They obtained the almost complete convergence rate of the adaptive estimator. Unlike in these cited works, where the authors consider the kernel estimator of the conditional mode, we are interested, here, in the nonparametric estimation of the conditional mode by the recursive method. Clearly the interest in this method comes from the fact that this method permits to update our estimator for each additional information. This feature is very adequate in this context of functional statistics when the date are very bulky and are available by real time monitoring. Moreover, the kernel method can be viewed as particular case of our study. Specifically, as asymptotic result we prove the almost complete convergence and the asymptotic normality of this estimator. Among the lot of papers concerning the nonparametric recursive estimation or modelization of variable taking values in infinite dimensional spaces, we only refer to the papers by Roussas *et al.* (1989), Thiam (2006), Amiri et al. (2014), Ling and Xu (2012) Ferraty *et al.* (2010), Bosq (2000), Ramsay and Silverman (2005), Ferraty and Vieu (2006). The paper is organized as follows. We present our model in Section 2. In

Section 3 we treat the almost complete convergence of our estimate. The asymptotic normality of the proposed estimator is given in Section 4. A simulation study is given in Section 5. Finally, the proofs of the auxiliary results are given in the Appendix.

4.2 Recursive conditional mode estimation

Let $Z_i = (X_i, Y_i)_{i=1,\dots n}$ be a $\mathcal{F} \times \mathbb{R}$-valued measurable strictly stationary process, defined on a probability space $(\Omega,\ \mathcal{A}, \mathbb{P})$, where $\mathcal{F}$ is a semi-metric space, d 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 a compact subset $S = [\theta - \xi, \theta + \xi]$, $\xi > 0$, such that the conditional density of Y given $X = x$ has an unique mode θ on S. In the remainder of the paper x is fixed in $\mathcal{F}$ and N_x denotes a neighborhood of x. Let f^x be the conditional density of the variable Y given $X = x$. We define the kernel estimator $\widehat{f^x}$ of f^x as follows :

$$\widehat{f^x}(y) = \frac{\sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right) b_i^{-1} H\left(b_i^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right)}, \tag{4.1}$$

where K and H are kernels and a_n (resp. b_n) is a sequence of positive real numbers. Note that we can treat the classical kernel method studied by Ferraty *et al.* (2006) as particular case of the present study by taking $a_i = a_n$ and $b_i = b_n$; for all $i = 1, \dots, n$. Now, a natural extension of the kernel estimator $\hat{\theta}$ of the conditional mode θ to the functional framework is given by :

$$\hat{\theta} \ = \ \arg\sup_{y \in S} \widehat{f^x}(y). \tag{4.2}$$

The estimate $\hat{\theta}$ is not necessarily unique, and if this is the case, all the remaining of our paper concerns any value $\hat{\theta}$ satisfying (4.2). This work establishes the almost complete convergence and the asymptotic normality of the kernel estimate $\hat{\theta}$ of θ under the functional ergodicity condition (see Laib and Louani (2011) for the definition and some examples). Note that this estimate is a generalization of the classical kernel method studied by Ferraty et al. (2006) and the k-NN method studied by Attouch and Boubaca (2013). Indeed, in both cases the bandwidth parameter is the same as of all observations, while in our recursive case the smoothing parameter is linked to X_i which permits to update our estimator for each additional observation. Moreover, the latter is computed by a recursive formula (see, Amiri et al. (2014)) which makes this estimator very fast in practice.

4.3 Consistency

Throughout the paper, when no confusion is possible, we will denote by C or C' some strictly positive generic constants, and we will use the notation $B(x,h) = \{x' \in \Im\ : d(x',x) < h\}$. Moreover, for $i = 1,\dots,n$, we put $\mathfrak{F}_k$ is the σ-field generated by $((X_1,Y_1),\dots(X_k,Y_k)))$, we pose $\mathfrak{G}_k$ is the σ-field generated by $((X_1,Y_1),\dots$
$(X_k,Y_k),X_{k+1})$ and we assume that the conditional distribution of Y given $\mathfrak{G}_k$ depends only on X_{k+1}.

In order to establish our asymptotic results we need the following hypotheses :
(H1)

$$\begin{cases}\text{(i) The function } \phi(x,r) := \mathbb{P}(X \in B(x,r)) \text{ is such that } \phi(x,r) > 0, \ \ \forall\, r > 0 \\ \text{(ii) For all } i = 1,\dots,n \text{ there exists a deterministic function } \phi_i(x,\cdot) \text{ such that} \\ \quad \text{almost surely } 0 < \mathbb{P}\left(X_i \in B(x,r)|\mathfrak{F}_{i-1}\right) \leq \phi_i(x,r),\ \forall\, r > 0, \text{ and } \phi_i(x,r) \to 0 \\ \quad \text{as } r \to 0 \\ \\ \text{(iii) For all sequence } (r_i)_{i=1,\dots n} > 0,\ \dfrac{\sum_{i=1}^n \mathbb{P}\left(X_i \in B(x,r_i)|\mathfrak{F}_{i-1}\right)}{\sum_{i=1}^n \phi(x,r_i)} \to 1 \qquad a.co.\end{cases}$$

(H2) The conditional density f^x satisfies

$$\begin{cases}\text{i) } f^x \text{ is } 2- \text{ times continuously differentiable with respect to } y \text{ on } [\theta-\xi,\theta+\xi], \\ \qquad\qquad \text{such that } 0 < \left|f^{x(2)}(\theta)\right| < \infty \\ \text{ii) } \forall (y_1,y_2) \in S \times S,\ \forall (x_1,x_2) \in N_x \times N_x \\ \qquad\qquad |f^{x_1}(y_1) - f^{x_2}(y_2)| \leq C_x \left(d(x_1,x_2)^{\beta_1} + |y_1-y_2|^{\beta_2}\right), \qquad \beta_1,\beta_2 > 0 \\ \text{iii) } f^x \text{ is strictly increasing on } (\theta-\xi,\theta) \text{ and strictly decreasing on } (\theta,\theta+\xi).\end{cases}$$

(H3) K is a function with support $(0,1)$ such that

$$0 < C < K(t) < C' < \infty$$

,

(H4) The kernel H is bounded such that

$$\forall (y_1, y_2) \in \mathbb{R}^2, \;\; |H(y_1) - H(y_2)| \leq C|y_1 - y_2| \;\text{ and }\; \int |t|^{\beta_2} H(t)dt < \infty.$$

(H5) The sequences a_n and b_n such that

$$\begin{cases} \text{i) There exists } \gamma > 0 \text{ such that } \lim\limits_{n\to\infty} n^{\gamma} b_n = \infty \quad \text{ and } \sup\limits_i \left(\dfrac{b_n}{b_i}\right) < C, \\ \text{ii) } \lim\limits_{n\to\infty} \dfrac{\log n}{n b_n \psi_n(x)} = 0 \text{ where } \psi_n(x) = n^{-1} \dfrac{\left(\sum\limits_{i=1}^{n} \phi(x, a_i)\right)^2}{\sum\limits_{i=1}^{n} \phi_i(x, a_i)} \\ \text{iii)} \dfrac{\sum_{i=1}^{n} \left(h_i^{\beta} \phi_i(x, a_i)\right)}{h_n^{\beta} \sum_{i=1}^{n} \phi(x, a_i)} = O(h_n^{\beta}) \text{ with } h_i \in \{a_i, b_i\} \text{ and } \beta \in \{\beta_1, \beta_2\} \end{cases}$$

We note that, these conditions are very standard in this context. Indeed, condition (H1) is the same as used by Gheriballah et al (2013) and (H2)-(H4) are the same as that used by Ferraty and Vieu (2006). Assumption (H5) is a similar technical condition considered by Thiam (2006) for the vectorial case.
The following Theorem gives the almost-complete convergence [1] (a.co.) of $\widehat{\theta}$.

Theorem 4.1 *Under the hypotheses (H1)-(H5) we have*

$$\hat{\theta} - \theta = O\left(a_n^{\beta_1/2}\right) + O\left(b_n^{\beta_2/2}\right) + O\left(\sqrt{\frac{\log n}{n b_n \psi_n(x)}}\right)^{1/2} \qquad a.co.$$

Proof of Theorem 4.1
Firstly, observe that by a Taylor expansion of the function f^x we have

$$|(\theta - \hat{\theta})|^2 \leq \frac{2}{f^{x(2)}(\theta^*)} |f^x(\hat{\theta}) - f^x(\theta)|$$

1. Let $(z_n)_{n\in\mathbb{N}}$ be a sequence of real r.v.'s ; we say that z_n converges almost completely (a.co.) to zero if, and only if, $\forall\epsilon > 0$, $\sum_{n=1}^{\infty} P(|z_n| > \epsilon) < \infty$. Moreover, we say that the rate of almost complete convergence of z_n to zero is of order u_n (with $u_n \to 0$) and we write $z_n = O_{a.co.}(u_n)$ if, and only if, $\exists\epsilon > 0$, $\sum_{n=1}^{\infty} P(|z_n| > \epsilon u_n) < \infty$. This kind of convergence implies both almost sure convergence and convergence in probability.

for some θ^* between θ and $\hat{\theta}$. Moreover, by simple analytic arguments, we can show that

$$|f^x(\hat{\theta}) - f^x(\theta)| \leq 2 \sup_{y \in S} |\widehat{f^x}(y) - f^x(y)|. \tag{4.3}$$

Therefore, all what is left to do, is to study the convergence rate of

$$\sup_{y \in S} |\widehat{f^x}(y) - f^x(y)|$$

and to show that

$$\exists c > 0, \ \sum_{n=1}^{\infty} P\left(f^{x(2)}(\theta^*) < c\right) \ < \ \infty.$$

The latter is based on the following decomposition

$$\widehat{f^x}(y) - f^x(y) = \widehat{B}_n(x,y) + \frac{\widehat{R}_n(x,y)}{\widehat{f}_D(x)} + \frac{\widehat{Q}_n(x,y)}{\widehat{f}_D(x)}$$

where

$$\begin{aligned}
\widehat{Q}_n(x,y) \ &:= (\widehat{f}_N^x(y) - \bar{f}_N^x(y)) - f^x(y)(\widehat{f}_D(x) - \bar{f}_D(x)),\\
\widehat{B}_n(x,y) \ &:= \frac{\bar{f}_N^x(y)}{\bar{f}_D(x)} - f^x(y) \quad \text{and} \quad \widehat{R}_n(x,y) := -\widehat{B}_n(x,y)(\widehat{f}_D(x) - \bar{f}_D(x))
\end{aligned}$$

with

$$\begin{aligned}
\widehat{f}_N^x(y) \ &:= \frac{1}{n\varphi_n(x)} \sum_{i=1}^n K(a_i^{-1} d(x, X_i)) b_i^{-1} H\left(b_i^{-1}(y - Y_i)\right),\\
\bar{f}_N^x(y) \ &:= \frac{1}{n\varphi_n(x)} \sum_{i=1}^n \mathbb{E}\left[K(a_i^{-1} d(x, X_i))) b_i^{-1} H\left(b_i^{-1}(y - Y_i)\right) |\mathfrak{F}_{i-1}\right],\\
\widehat{f}_D(x) \ &:= \frac{1}{n\varphi_n(x)} \sum_{i=1}^n K(a_i^{-1} d(x, X_i)),\\
\bar{f}_D(x) \ &:= \frac{1}{n\varphi_n(x)} \sum_{i=1}^n \mathbb{E}\left[K(a_i^{-1} d(x, X_i)) |\mathfrak{F}_{i-1}\right].
\end{aligned}$$

where $\varphi_n(x) = n^{-1} \sum_{i=1}^n \phi(x, a_i)$

Finally, Theorem 4.1 can be deduced from the following Lemmas and Corollary.

Lemma 4.1
Under the hypotheses (H1), (H3) and (H5) we have

$$\widehat{f}_D(x) - \bar{f}_D(x) = O\left(\sqrt{\frac{\log n}{n\psi_n(x)}}\right), \quad a.co.$$

Corollary 4.1
Under the hypotheses of Lemma 4.1, we have,

$$\exists\, C > 0 \quad \sum_{n=1}^{\infty} \mathbb{P}\left(\widehat{f}_D(x) < C\right) < \infty.$$

Lemma 4.2
Under the hypotheses of Theorem 4.1 we have

$$\sup_{y\in S} |\widehat{B}_n(x,y)| = O\left(a_n^{\beta_1}\right) + O\left(b_n^{\beta_2}\right), \quad a.co.$$

Lemma 4.3
Under the hypotheses of Theorem 4.1 we have

$$\frac{1}{\widehat{f}_D(x)} \sup_{y\in S} |\widehat{f}_N^x(y) - \bar{f}_N^x(y))| = O\left(\sqrt{\frac{\log n}{nb_n\psi_n(x)}}\right), \quad a.co.$$

Lemma 4.4
Under the hypotheses of Theorem 4.1 we have

$$\exists C > 0, \quad \sum_{n=1}^{\infty} P\left(f^{x(2)}(\theta^*) < C\right) \; < \; \infty.$$

4.4 Asymptotic normality

Now, we study the asymptotic normality of $\hat{\theta}$. To, do that we replace the conditions (H1)-(H5) the following hypotheses,

(N1) The property (H1) holds. Moreover there exists functions $\beta_x(\cdot)$ and $\delta(\cdot)$ such that

$$\forall s \in [0,1], \quad \lim_{n\to\infty} \frac{\sum_{i=1}^{n} \phi(x, sa_i)}{\sum_{i=1}^{n} \phi(x, a_i)} = \beta_x(s)$$

and

$$\lim_{n\to\infty} b_n^3 \frac{\sum_{i=1}^n b_i^{-3}\pi_i(x, sa_i)}{\sum_{i=1}^n \phi(x, sa_i)} = \delta(s) \qquad \text{In probability.}$$

where $\pi_i(x,r) = \mathbb{P}\left(X_i \in B(x,r)|\mathfrak{F}_{i-1}\right)$

(N2) The conditional density f^x satisfies (H2). Moreover $\forall (y_1, y_2) \in S \times S$, $\forall (x_1, x_2) \in N_x \times N_x$

$$|f^{x_1^{(j)}}(y_1) - f^{x_2^{(j)}}(y_2)| \leq C_x \left(d(x_1, x_2)^{\beta_1^j} + |y_1 - y_2|^{\beta_2^j} \right)$$

where $\beta_1^j, \beta_2^j > 0$ for $j = 1, 2$.

(N3) The Kernel K is a differentiable function supported on $[0,1]$ satisfies (H3). Its derivative K' such that

$$K^2(1)\delta(1) - \int_0^1 (K^2(u))'\delta(u)\beta_x(u)du > 0 \quad \text{and} \qquad K(1) - \int_0^1 K'(u)\beta_x(u)du \neq 0.$$

(N4) The kernel H is of class C^2, of compact support, satisfies

$$\int H'^2(t)dt \quad \text{and} \qquad \int |t|^{\beta_i^j} H(t)dt < \infty, \qquad i, j = 1, 2.$$

(N5) The sequences a_n and b_n satisfy (H5i) and such that :

$$\frac{nb_n^3 \varphi_n(x)}{\log n} \to \infty,$$

and

$$\sqrt{\frac{b_n^3}{n\varphi_n(x)}} \sum_{i=1}^n \left(a_i^{\beta_1^1} + b_i^{\beta_2^1} \right) \pi_i(x, a_i) \to 0 \;\text{ as }\; n \to \infty \quad \text{In probability.}$$

The asymptotic normality of $\hat{\theta}$ is given in the following Theorem.

Theorem 4.2 *Under the hypotheses (N1)-(N5) we have*

$$\left(\frac{nb_n^3 \varphi_n(x)}{\sigma^2(x)} \right)^{1/2} (\hat{\theta} - \theta) \longrightarrow \mathcal{N}(0,1) \quad \text{as} \quad n \to \infty \tag{4.4}$$

where

$$\sigma^2(x) = \frac{f^x(\theta)\left(K^2(1)\delta(1) - \int_0^1 (K^2(s))'\delta(s)\beta_x(s)ds \right)}{\left(f^{x(2)}(\theta) \left(K(1) - \int_0^1 K'(s)\beta_x(s)ds \right) \right)^2} \int H'^2(t)dt.$$

Proof of Theorem 4.2 Consider the Taylor expansion of function $\widehat{f'}^x$ around $\widehat{\theta}$, we get :

$$\widehat{f'^x}(\widehat{\theta}) = \widehat{f'^x}(\theta) + (\widehat{\theta} - \theta)\widehat{f''^x}(\theta^*),$$

where θ^* is between θ and $\widehat{\theta}$. Therefore,

$$V_n(\widehat{\theta} - \theta) = -\frac{V_n \widehat{f'^x}(\theta)}{\widehat{f''^x}(\theta^*)}$$

where

$$V_n = \left(\frac{n b_n^3 \varphi_n(x)}{f^x(\theta)\left(K^2(1)\delta(1) - \int_0^1 (K^2(s))'\delta(s)\beta_x(s)ds\right)\int H'^2(t)dt} \right)^{(1/2)}$$

observe that

$$\widehat{f'^x}(y) = \widehat{B}'_n(x,y) + \frac{\widehat{R}'_n(x,y)}{\widehat{f}_D(x)} + \frac{\widehat{Q}'_n(x,y)}{\widehat{f}_D(x)}$$

where

$$\begin{aligned}
\widehat{Q}'_n(x,y) &:= (\widehat{f'}_N(x,y) - \bar{f'}_N(x,y)) - f'^x(y)(\widehat{f}_D(x) - \bar{f}_D(x)) \\
\widehat{B}'_n(x,y) &:= \frac{\bar{f'}_N(x,y)}{\bar{f}_D(x)} \quad \text{and} \quad \widehat{R}'_n(x,y) := -\widehat{B}'_n(x,y)(\widehat{f'}^x_N(y) - \bar{f'}^x_N(y))
\end{aligned}$$

with

$$\begin{aligned}
\widehat{f'}^x_N(y) &:= \frac{1}{n\varphi_n(x)} \sum_{i=1}^n K(h^{-1}d(x,X_i)) b_i^{-2} H'\left(b_i^{-1}(y - Y_i)\right), \\
\bar{f'}^x_N(y) &:= \frac{1}{n\varphi_n(x)} \sum_{i=1}^n \mathbb{E}\left[K(h^{-1}d(x,X_i)) b_i^{-2} H'\left(b_i^{-1}(y - Y_i)\right) |\mathfrak{F}_{i-1}\right].
\end{aligned}$$

Thus, Theorem 4.2 is a consequence of the following intermediate results.

Lemma 4.5 *Under the hypotheses of Theorem 4.2, we have*

$$V_n \widehat{Q}'_n(x,\theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \text{as } n \to \infty.$$

Lemma 4.6 *Under hypotheses (N2), (N4) and (N5) we have*

$$V_n(\widehat{B}'_n(x,y)) = o\,(1)\,, \;\; as \quad n \to +\infty \quad In\ probability.$$

Lemma 4.7 *Under hypotheses (N2), (N4) and (N5) we have*

$$\widehat{f}_D(x) \to K(1) - \int_0^1 (K(u))'\beta_x(u)du, \;\; in\ probability\ , \;\; as \quad n \to +\infty.$$

Lemma 4.8 *Under the conditions of the Theorem (4.2), we have*

$$\widehat{f''}^{x}_{N}(\theta^*) \longrightarrow f''^{x}(\theta) \;\; in\ probability\ .$$

4.5 A simulation study

This section illustrates the finite sample performance of the proposed procedure. For this aim, we compare the efficiency of the recursive estimation method to the classical kernel method studied by Ferraty et al. (2005). To do that, we consider the following functional nonparametric model :

$$Y_i = r(X_i) + \epsilon_i \text{ for } i = 1, \ldots, n \tag{4.5}$$

where the ϵ_i's are generated independently according to a normal distribution with mean 0 and standard deviation $\sigma = 0.05$. The operator r is defined by $r(x) = \int_0^1 \frac{1}{1+x^2(t)}dt$. Concerning the functional regressors, we precise that we have used a functional autoregressive process with Wiener noise. The latter is generated by the R-routine *simul.far.wiener* of the *far* package in **R**. This routine use the Karhunen-Loève expansion of the noise. The linear operator ρ is expressed in the Karhunen-Loève basis, by using a diagonal matrix d_ρ and a perturbation coefficient $c = 0.05$. We note that this process satisfies the ergodicity condition (see, Laib and Louani (2011) for more details). The curves X_i's are discretized in the same grid composed by 100 points and are plotted in Figure 4.1.

It is clear that the conditional distribution of Y given $X = x$ is explicitly given by the distribution of ϵ_i shifted by $r(x)$, which permits to determine

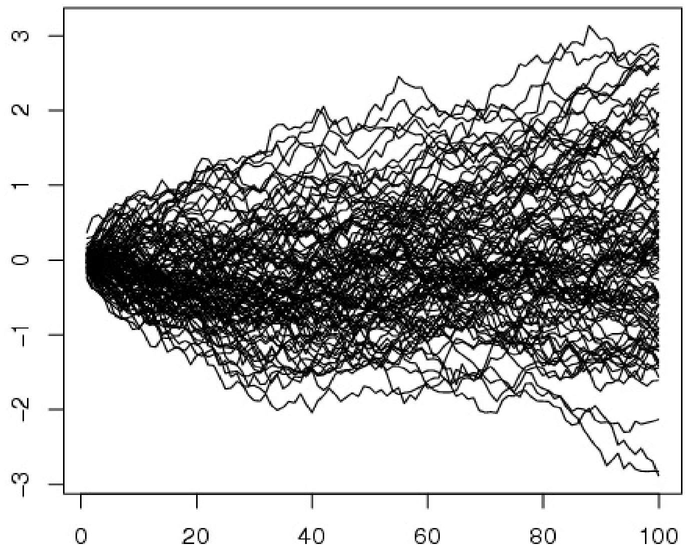

FIGURE 4.1 – A sample of 100 curves, for $d_\rho = (0.45, 0.90, 0.34, 0.45)$

the theoretical conditional mode $\theta(x)$. The efficiency of the predictors is evaluated by the empirical mean square error :

$$MSE(KM) = n^{-1}\sum_{i=1}^{n}(\theta(X_i) - \tilde{\theta}(X_i))^2 \text{ and } MSE(RKM) = n^{-1}\sum_{i=1}^{n}(\theta(X_i) - \hat{\theta}(X_i))^2$$

where

$$\hat{\theta}(x) = \arg\sup_y \frac{\sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right) b_i^{-1} H\left(b_i^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right)},$$

and

$$\tilde{\theta}(x) = \arg\sup_y \frac{\sum_{i=1}^{n} K\left(\tilde{a}_n^{-1} d(x, X_i)\right) \tilde{b}_n^{-1} H\left(\tilde{b}_n^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} K\left(\tilde{a}_n^{-1} d(x, X_i)\right)}$$

For this comparison study, we treat both estimators in the same conditions. In particular, we take the same sequence of bandwidths used by Amiri et al (2014) $(a_i, b_i) = (C\max_j(d(X_j, X_i)i^{-v}, C\max_j(|Y_j - Y_i)i^{-v})$, for the recursive case, and for the standard case $(\tilde{a}_n, \tilde{b}_n) = (C\max_{i,j}(d(X_j, X_i)n^{-v}, C\max_{i,j}(|Y_j - Y_i|n^{-v})$ where C and v are chosen among the subset $\{0.5, 1, 5\}$ and $\{1, 1/2, 1/5, 1/10\}$ (respectively) by the following cross-validation

$$CV(C, v) = n^{-1}\sum_{i=1}^{n}(Y_i - \ddot{\theta}(X_i))^2, \quad \text{where } \ddot{\theta}(X_i) = \hat{\theta}(X_i) \quad \text{or} \quad \tilde{\theta}(X_i) \quad .$$

Finally, we point out that we considered a quadratic kernel on $(0, 1)$ and we used the L_2 metric.

Now, in order to explore the two structural axes of our study such as the correlation of data and the recursivity of the estimate, we compare the performance of both estimates with various values of n and various parameters of the functional autoregressive X_i. Typically, we consider three values of n ($n = 100$, $n = 200$ and $n = 500$) and three matrix $d_\rho = (0.45, 0.90, 0.34, 0.45)$, $d_\rho = (0.22, 0.45, 0.17, 0.22)$, $d_\rho = (0.90, 1.80, 0.68, 0.90)$. We point out that, the results of our simulation study are evaluated over 50 independent replications. The obtained results are presented in Table 1.

d_ρ	n	MSE(KM)	MSE(RKM)	CT(KM)	CT(RKM)
$(0.22, 0.45, 0.17, 0.22)$	n=100	0.21	0.34	0.17	0.08
$\|\rho\| = 0.19$	n=200	0.18	0.19	1.63	0.17
	n=500	0.16	0.08	5.47	1.93
$(0.45, 0.90, 0.34, 0.45)$	n=100	0.24	0.54	0.81	0.11
$\|\rho\| = 0.31$	n=200	0.20	0.16	1.54	0.16
	n=500	0.17	0.11	6.04	2.07
$(0.90, 1.80, 0.68, 0.90)$	n=100	0.39	0.71	0.89	0.14
$\|\rho\| = 0.59$	n=200	0.36	0.21	1.72	0.27
	n=500	0.31	0.17	6.53	2.93

Table 1 MSE and Time Results.

It appears clearly that the performance of both estimates is closely linked to the degree of correlation expressed by $\|\rho\|$. In sense that the values of MSE increases substantially with respect to the value of $\|\rho\|$. Moreover, we observe that the accuracy of the recursive estimate increases significantly with respect to n, unlike to the standard case which is not substantially affected by the change of the sample size. The last column presents the computation time of the estimates (CT), in seconds. These results are obtained by the R-routine *system.time* . Although the computation time is closely related to the machine characteristics, we observe that the RKM estimate is faster to compute than the KM. For instance in the case where $n = 500$ and $d_\rho = (0.90, 1.80, 0.68, 0.90)$ we win 56% in computation time for the recursive case compared to the standard case. Overall, both methods give a satisfactory level of accuracy with slight superiority for the kernel method.

4.6 Real data example

The goal of this section is to show through some real data example the applicability of our procedure in practice. The main purpose is to give a comparison between the functional conditional mode recursive estimate and the recursive estimate of the functional nonparametric regression studied by Amiri et al (2014) defined by

$$\hat{r}(x) = \frac{\displaystyle\sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right) Y_i}{\displaystyle\sum_{i=1}^{n} K\left(a_i^{-1} d(x, X_i)\right)}.$$

For this aim we consider the road traffic data available on the web site *"http ://traffic.berkeley.edu"*. We precise that, we have used the loop detector data, obtained from the Freeway Performance Measurement System (PeMS) database. The considered time series date concerns the traffic flow characterized by the number of vehicle detected per 30 seconds, between 08/02/2008 at 18 :00 :00 and 09/02/2008 at 02 :00 :00. The studied time series are plotted in the figure (4.2)

Noting that, there are several metrics to study the traffic flow such as the traffic density, the vehicle speeds, headway times or the connectivity between the vehicles. In this application study we analyze the traffic flow by predicting the traffic density a ten minutes ahead via both recursive estimates (mode and regression). We recall that this functional data set is usually employed as

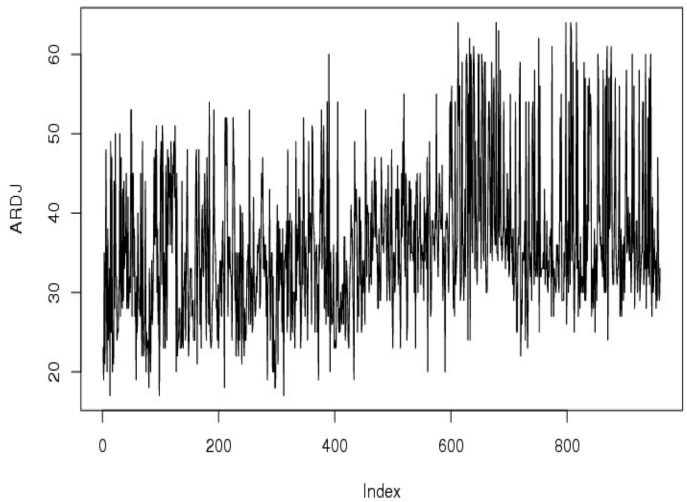

FIGURE 4.2 – The functional time series data

example of functional autoregressive of order 1 (see, Besse and Crdot 1996) which is an important example of functional ergodic data (see, Laib and Louani (2011)).
Now, for our prediction aim we denote by $X(\cdot)$ the curves of traffic, livered by a period of 30 seconds, observed during ten minutes. Then, between 08/02/2008 at 18 :00 :00 and 09/02/2008 at 02 :00 :00 we obtain 296 curves $(X_i)_{i=1,\ldots,296}$. In order to forecast the traffic density of the last ten minutes we keep the 294 first ones as training sample and we build our estimates for any j^{th}-30 second $(j = 1, \ldots, 20)$ by the sample $(Y_i^j, X_i)_{i=1\ldots294}$ with $Y_i^j = X_{i+1}(j)$. For both models we use the same sequence of bandwidths of the previous section defined by $(a_i, b_i) = (a_i, b_i) = (C \max_j(d(X_j, X_i)i^{-v}, (C \max_j(|Y_j - Y_i)i^{-v})$. We simulate by using quadratic kernel

$$K(x) = \frac{3}{2}(1 - x^2)1\!\!1_{[0,1]},$$

and we consider, a semi-metric based on the D first eigenfunctions of the empirical covariance operator associated to the D greatest eigenvalues. The optimal parameters D, C v are chosen by the same cross-validation method of the previous section.

The results are given in Figure (4.3) where we draw three curves : the first one (solid curve) corresponds to the predicted values by the conditional mode ($MAE_{mode} = \frac{1}{30}\sum_{j=1}^{30}(\hat{\theta}(X_{295}) - Y_{295}) \approx 4$), the second one corresponds to the observed values (dashed curve) and the last one (pointed curve), corresponds to the predicted values by the regression model ($MAE_{reg} = \frac{1}{30}\sum_{j=1}^{30}(\hat{r}(X_{295}) - Y_{295}) \approx 9$).

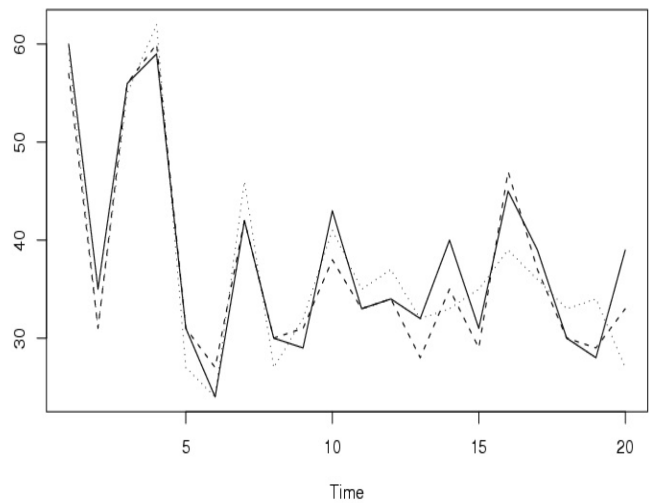

FIGURE 4.3 – The prediction results

4.7 Appendix

We first state the following exponential inequality for partial sums of martingale differences.

Lemma 4.9 *(see, Laib and Louani (2011, P.365)) Let $(Z_n)_{n\geq 1}$ be a sequence of real martingale differences with respect to the sequence of σ- fields $\mathfrak{F}_n = \sigma(Z_1, Z_2, \ldots, Z_n)_{n\geq 1}$, generated by the random variables $Z_1, Z_2, \ldots, Z_n$. Set $S_n = \sum_{i=1}^{n} Z_i$. For any $p \geq 2$ and for any $n \geq 1$, assume that there exist some nonnegative constants C and d_n such that*

$$\mathbb{E}\left[Z_n^p | \mathfrak{F}_{n-1}\right] \leq C^{p-2} p! d_n^2 \qquad \textit{almost surely.}$$

Then, for any $\epsilon > 0$, we have

$$\mathbb{P}\left(|S_n| > \epsilon\right) \leq 2\exp\left\{\frac{-\epsilon^2}{2(D_n + C\epsilon)}\right\}$$

where $D_n = \sum_{i=1}^{n} d_i^2$.

Proof of Lemma 4.1
We put, for all $i = 1, \ldots, n$, $K_i(x) = K(a_i^{-1}d(x, X_i))$ and $\Delta_i(x) = (K_i(x) - \mathbb{E}\left[K_i(x)|\mathfrak{F}_{i-1}\right])$. Hence, we can write

$$\widehat{f}_D(x) - \bar{f}_D(x) = \frac{1}{n\varphi_n(x)} \sum_{i=1}^{n} \Delta_i(x).$$

Using the fact that $\Delta_i(x)$ is a triangular array of martingale differences according to the σ-fields $(\mathfrak{F}_{i-1})_i$ to apply the inequality of Lemma 4.9. To do that, we must evaluate the quantity $\mathbb{E}\left[\Delta_i(x)^p|\mathfrak{F}_{i-1}\right]$. Indeed, under (H1) and (H3) we have, for $p \geq 2$

$$\begin{aligned}
\mathbb{E}\left[\Delta_i(x)^p|\mathfrak{F}_{i-1}\right] &\leq C\mathbb{E}\left[\Delta_i(x)^2|\mathfrak{F}_{i-1}\right] \\
&\leq C\mathbb{E}\left[K_i^2(x)|\mathfrak{F}_{i-1}\right] \\
&< C\mathbb{P}(X_i \in B(x, a_i)|\mathfrak{F}_{i-1}) \\
&\leq C\phi_i(x, a_i) \qquad .
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{P}\left\{\left|\widehat{f}_D(x)-\bar{f}_D(x)\right|>\varepsilon\right\} &= \mathbb{P}\left\{\left|\frac{1}{n\varphi_n(x)}\sum_{i=1}^{n}\Delta_i(x)\right|>\varepsilon\right\}\\
&\leq 2\exp\left\{-\frac{\varepsilon^2 n^2\varphi_n^2(x)}{2\left(\sum_{i=1}^{n}\phi_i(x,a_i)+C\varepsilon n\varphi_n(x)\right)}\right\}\\
&= 2\exp\left\{\frac{-\varepsilon^2 n^2\varphi_n^2(x)}{2\displaystyle\sum_{i=1}^{n}\phi_i(x,a_i)}\left(\frac{1}{1+\frac{C\varepsilon n\varphi_n(x)}{\displaystyle\sum_{i=1}^{n}\phi_i(x,a_i)}}\right)\right\}.\\
&= 2\exp\left\{\frac{-\varepsilon^2 n\psi_n(x)}{2}\left(\frac{1}{1+\frac{C\varepsilon n\varphi_n(x)}{\displaystyle\sum_{i=1}^{n}\phi_i(x,a_i)}}\right)\right\}.
\end{aligned}$$

Taking $\varepsilon=\epsilon_0\sqrt{\dfrac{\log n}{n\psi_n(x)}}$, then,

$$\mathbb{P}\left\{\left|\widehat{f}_D(x)-\bar{f}_D(x)\right|>\epsilon_0\sqrt{\frac{\log n}{n\psi_n(x)}}\right\}\leq 2\exp\left\{\frac{-\epsilon_0^2\log n}{2}\left(\frac{1}{1+\frac{C\epsilon_0\sqrt{\log n}}{\sqrt{\sum_{i=1}^{n}\phi_i(x,a_i)}}}\right)\right\}.$$

Clearly, under (H5), we can write

$$\lim_{n\mapsto\infty}\frac{\log n}{\sum_{i=1}^{n}\phi_i(x,a_i)}=0.$$

Therefore,

$$\mathbb{P}\left\{\left|\widehat{f}_D(x)-\bar{f}_D(x)\right|>\epsilon_0\sqrt{\frac{\log n}{n\psi_n(x)}}\right\}\leq 2\exp\left\{\frac{-\epsilon_0^2\log n}{2C}\right\}\leq 2n^{-C'\epsilon_0^2}.$$

Consequently an appropriate choice of ϵ_0 completes the proof of this lemma.

□

Proof of Corollary 4.1 Conditions (H1)(iii) and (H3) allow to write , there exists $C > 0$ such that

$$0 < C\frac{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i)|\mathfrak{F}_{i-1})}{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i))} \le \bar{f}_D(x) \le \left|\bar{f}_D(x) - \widehat{f}_D(x)\right| + \widehat{f}_D(x).$$

Thus,

$$C\frac{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i)|\mathfrak{F}_{i-1})}{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i))} - \left|\widehat{f}_D(x) - \bar{f}_D(x)\right| < \widehat{f}_D(x).$$

So,

$$\begin{aligned}
\mathbb{P}\left(\widehat{f}_D(x) \le \tfrac{C}{2}\right) \quad & \le \mathbb{P}\left(C\frac{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i)|\mathfrak{F}_{i-1})}{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i))} - \left|\widehat{f}_D(x) - \bar{f}_D(x)\right| < \frac{C}{2}\right) \\
& \le \mathbb{P}\left(\left|C\frac{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i)|\mathfrak{F}_{i-1})}{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i))} - \left|\widehat{f}_D(x) - \bar{f}_D(x)\right| - C\right| > \frac{C}{2}\right).
\end{aligned}$$

Finally Lemma 4.1 and (H1)(iii) give

$$\sum_n \mathbb{P}\left(\left|C\frac{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i)|\mathfrak{F}_{i-1})}{\sum_{i=1}^{n}\mathbb{P}(X_i \in B(x,a_i))} - \left|\widehat{f}_D(x) - \bar{f}_D(x)\right| - C\right| > \frac{C}{2}\right) < \infty$$

which gives the result.

□

Proof of Lemma 4.2 Letting $H_i(y) = b_i^{-1}H(b_i^{-1}(y - Y_i))$, for $i = 1, \dots, n$. It is easy to see that

$$\begin{aligned}
|\widehat{B}_n(x,y)| \quad & = \left|\frac{1}{n\varphi_n(x)\bar{f}_D(x)}\sum_{i=1}^{n}\left\{\mathbb{E}\left(K_i(x)\mathbb{E}[H_i(y)|\mathfrak{G}_{i-1}]|\mathfrak{F}_{i-1}\right) - f^x(y)\mathbb{E}\left[K_i(x)|\mathfrak{F}_{i-1}\right]\right\}\right| \\
& = \left|\frac{1}{n\varphi_n(x)\bar{f}_D(x)}\sum_{i=1}^{n}\left\{\mathbb{E}\left(K_i(x)\mathbb{E}[H_i(y)|X_i]|\mathfrak{F}_{i-1}\right) - f^x(y)\mathbb{E}\left[K_i(x)|\mathfrak{F}_{i-1}\right]\right\}\right| \\
& \le \frac{1}{n\varphi_n(x)\bar{f}_D(x)}\sum_{i=1}^{n}\left\{\mathbb{E}\left(K_i(x)\left|\mathbb{E}[H_i(y)|X_i] - f^x(y)\right|\mathfrak{F}_{i-1}\right)\right\}.
\end{aligned}$$

Next, by a change of variable we get

$$\mathbb{E}\left[H_i(y)|X\right] \;=\; \int_{\mathbb{R}} H(t) f^{X_i}(y-b_i t)dt.$$

Thus, we have

$$\left|\mathbb{E}\left[H_i(y)|X_i\right] - f^x(y)\right| \;\leq\; \int_{\mathbb{R}} H(t)\left|f^{X_i}(y-b_i t) - f^x(y)\right| dt.$$

Moreover, it follows by (H2) that

$$1\!\!1_{\{B(x,a_i)\}}(X_i)\left|\mathbb{E}\left[H_i(y)|X_i\right] - f^x(y)\right| \;\leq\; C\int_{\mathbb{R}}\left(a_i^{\beta_1} + |t|^{\beta_2} b_i^{\beta_2}\right) H(t)dt.$$

Condition (H5) achieves the proof of lemma. □

Proof of Lemma 4.3 The compactness of S, allow to write, $S \subset \bigcup_{k=1}^{s_n} S_k$ where $S_k = (t_k - l_n,\ t_k + l_n)$ and the sequences (l_n, s_n) such that $l_n s_n \leq C$. Taking $t_y = \arg\min_{t\in\{t_1,\dots,t_{s_n}\}} |y-t|$. Furthermore, we consider the following decomposition

$$\begin{aligned}
\frac{1}{\widehat{f}_D(x)}\sup_{y\in S}\left|\widehat{f}_N^x(y) - \bar{f}_N^x(y)\right| \;&\leq\; \underbrace{\frac{1}{\widehat{f}_D(x)}\sup_{y\in S}\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_y)\right|}_{T_1} + \underbrace{\frac{1}{\widehat{f}_D(x)}\sup_{y\in S}\left|\widehat{f}_N^x(t_y) - \bar{f}_N^x(t_y)]\right|}_{T_2} \\
&\quad + \underbrace{\frac{1}{\widehat{f}_D(x)}\sup_{y\in S}\left|\bar{f}_N^x(t_y) - \bar{f}_N^x(y)\right|}_{T_3} \qquad (4.6)
\end{aligned}$$

- To evaluate (T_1) we use (H4) and the first part of (H5) to write

$$\begin{aligned}
\frac{1}{\widehat{f}_D(x)}\sup_{y\in S}\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_y)\right| \;&\leq\; \frac{1}{n\,\varphi_n(x)\,\widehat{f}_D(x)}\sup_{y\in S}\sum_{i=1}^{n}\left|H_i(y) - H_i(t_y)\right| K_i(x), \\
&\leq\; \frac{C}{n\,\varphi_n(x)\,\widehat{f}_D(x)}\sup_{y\in S}|y-t_y|\left(\sum_{i=1}^{n}\frac{K_i(x)}{b_i^2}\right), \\
&\leq\; C\frac{l_n}{b_n^2}.
\end{aligned}$$

By taking $l_n = n^{-3\gamma/2-1/2}$ we obtain that

$$\frac{l_n}{b_n^2} = o\left(\sqrt{\frac{\log n}{nb_n\psi_n(x)}}\right).$$

So, for n large enough, we can found a $\eta > 0$ such that

$$\mathbb{P}\left(\frac{1}{\widehat{f}_D(x)}\sup_{y\in S}\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_y)\right| > \eta\sqrt{\frac{\log n}{nb_n\psi_n(x)}}\right) \;=\; 0. \qquad (4.7)$$

and

$$\frac{1}{\widehat{f}_D(x)}\sup_{y\in S}\left|\bar{f}_N^x(t_y) - \bar{f}_N^x(y)\right| \leq C\frac{l_n}{b_n^2} = o\left(\sqrt{\frac{\log n}{nb_n\varphi_n(x)}}\right) \qquad (4.8)$$

- Now, we treat (T_2). To do that we use the same arguments as for the proof of Lemma 4.1, where $K_i(x)$ is replaced by $K_i(x)H_i(t_y)$

$$\mathbb{E}\left[K_i(x)H_i(t_y)|\mathfrak{F}_{i-1}\right]^p \leq Cb_i^{-p+1}\phi_i(x,a_i) \leq Cb_n^{-(p-1)}\phi_i(x,a_i).$$

Therefore, we can apply Lemma 4.9 with $C = b_n^{-1}$ and $d_i = b_n\phi_i(x,a_i)$. So, for all $\varepsilon > 0$, we have

$$\mathbb{P}\left\{\left|\widehat{f}_N^x(t_y) - \bar{f}_N^x(t_y)]\right| > \varepsilon\right\} \;\leq\; 2\exp\left\{-\frac{\varepsilon^2 n^2\varphi_n^2(x)}{2\left(b_n\sum_{i=1}^n\phi_i(x,a_i) + Cb_n^{-1}\varepsilon n\varphi_n(x)\right)}\right\}.$$

Thus, for $\varepsilon = \eta\sqrt{\dfrac{\log n}{nb_n\psi_n(x)}}$

$$\mathbb{P}\left(\sup_{y\in S}\left|\widehat{f}_N^x(t_y) - \bar{f}_N^x(t_y)]\right| > \eta\sqrt{\frac{\log n}{nb_n\varphi_n(x)}}\right) \leq 2s_n\exp\{-C\eta^2\log n\}.$$

Consequently, an appropriate choice of η done

$$\sum_n\mathbb{P}\left(\sup_{y\in S}\left|\widehat{f}_N^x(t_y) - \bar{f}_N^x(t_y)\right| > \eta\sqrt{\frac{\log n}{nb_n\psi_n(x)}}\right) < \infty. \qquad (4.9)$$

Finally, Lemma 4.3 can be easily deduced from (4.6), (4.7), (4.9) and (4.8). □

Proof of Lemma 4.4 The proof of this Lemma is very similar to Lemma 3.1.4 in Ferraty *et al.* (2005). As $f^{(2)}$ is continuous function, the proof of this lemma is a consequence of the almost complete convergence of

$$\hat{\theta} - \theta \to 0, \ a.co.$$

For this we use (H2iii) to get

$$\forall \epsilon > 0, \ \exists \delta(\epsilon) > 0, \forall y \in (\theta - \xi, \theta + \xi), |f^x(y) - f^x(\theta)| \leq \delta(\epsilon) \Rightarrow |y - \theta| \leq \epsilon.$$

Thus,

$$\forall \epsilon > 0, \ \exists \delta(\epsilon) > 0, \ \ P\left(|\hat{\theta} - \theta| > \epsilon\right) \leq P\left(|f^x(\hat{\theta}) - f^x(\theta)| \geq \delta(\epsilon)\right).$$

Then, $\hat{\theta} - \theta$ results from the previous lemmas □

Proof of Lemma 4.5 For all $i = 1, \ldots, n$ we put

$$\eta_{ni} = V_n H_i'(\theta) \frac{K_i(x)}{\sqrt{n}\varphi_n(x)}$$

and define $\zeta_{ni} := \eta_{ni} - \mathbb{E}[\eta_{ni} \mid |\mathfrak{F}_{i-1}]$. Thus,

$$V_n \widehat{Q}_n'(x, \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \zeta_{ni}$$

As ζ_{ni} is a triangular array of martingale differences according the σ-fields $(\mathfrak{F}_{i-1})_i$, we are in position to apply the central limit theorem based on unconditional Lindeberg condition (see, Gaenssler et al. (1978)). More precisely, the asymptotic normality of $V_n \widehat{Q}_n'(x, \theta)$ is consequence of the following preliminaries results

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[\zeta_{ni}^2 \mid \mathfrak{F}_{i-1}] \to 1 \quad \text{in probability.} \tag{4.10}$$

and

$$\text{for every} \quad \epsilon > 0 \qquad \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[\zeta_{ni}^2 \mathbb{1}_{\zeta_{ni}^2 > \epsilon n}] \to 0. \tag{4.11}$$

Firstly, we prove (4.10). To do that,we write

$$\mathbb{E}[\zeta_{ni}^2|\mathfrak{F}_{i-1}] = \mathbb{E}[\eta_{ni}^2|\mathfrak{F}_{i-1}] - \mathbb{E}^2[\eta_{ni}|\mathfrak{F}_{i-1}].$$

Therefore, it suffices to show the

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}^2[\eta_{ni}|\mathfrak{F}_{i-1}] \to^{\mathbb{P}} 0$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[\eta_{ni}^2|\mathfrak{F}_{i-1}] \to^{\mathbb{P}} 1. \tag{4.12}$$

$$\mathbb{E}\left[K_i H_i^{'}(y)|\mathfrak{F}_{i-1}\right] = \left(f^{'x}(y)\int H'(t)dt + O\left(a_i^{b_1} + b_i^{b_2}\right)\right)\mathbb{E}\left[K_i|\mathfrak{F}_{i-1}\right], \quad j = 1, 2.$$

On other hand, under (N1) and (N3) we have

$$C\pi_i(x, a_i) \leq \mathbb{E}\left[K_i(x)|\mathfrak{F}_{i-1}\right] \leq C'\pi_i(x, a_i).$$

It follows that

$$\begin{aligned}(|\ \mathbb{E}[\eta_{ni}|\mathfrak{F}_{i-1}])^2 \ &\leq\ C\frac{b_n^3}{\varphi_n(x)}\pi_i^2(x, a_i)\\ &\leq\ C\frac{b_n^3}{\varphi_n(x)}\pi_i(x, a_i).\end{aligned}$$

Finally, under the fact that (see, (H1iii)) we have

$$\frac{1}{n\varphi_n(x)}\sum_{i=1}^{n}\pi_i(x, a_i) \to 1 \quad \text{In probability}$$

which implies that

$$\frac{1}{n}\sum_{i=1}^{n}(\mathbb{E}[\eta_{ni}|\mathfrak{F}_{i-1}])^2 = o(1) \quad \text{In probability.}$$

Now, we treat the convergence (4.12), Indeed, we write

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[\eta_{ni}^2|\mathfrak{F}_{i-1}] \ = \tfrac{V_n^2}{n^2\varphi_n^2(x)}\sum_{i=1}^{n}\mathbb{E}[(H_i'(\theta))^2 K_i^2(x)|\mathfrak{F}_{i-1}].$$

By a simpl algebra we show that

$$\mathbb{E}\left[K_i^2 b_i^3 H_i'^2(\theta)|\mathfrak{F}_{i-1}\right] = \left(f^x(\theta)\int (H'^2(t))dt + O\left(a_i^{b_1} + b_i^{b_2}\right)\right)\mathbb{E}\left[K_i^2|\mathfrak{F}_{i-1}\right], \quad j = 1, 2. \tag{4.13}$$

Thus,

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[\eta_{ni}^2|\mathfrak{F}_{i-1}] = \frac{V_n^2 f^x(\theta)}{n^2\varphi_n^2(x)}\int (H'^2(t))dt \sum_{i=1}^{n} b_i^{-3}\mathbb{E}\left[K_i^2(x)|\mathfrak{F}_{i-1}\right] + o(1).$$

Next, we use the same arguments as those used in Ferraty et al. (2009) to get

$$\mathbb{E}\left[K_i^2(x)|\mathfrak{F}_{i-1}\right] = K^2(1)\pi_i(x,h) - \int_0^1 (K^2(u))'\pi_i(x,uh)du.$$

Thus, from (N1) and (N5) we have

$$\begin{aligned}
&\frac{b_n^3}{n\varphi_n(x)}\sum_{i=1}^{n} b_i^{-3}\mathbb{E}\left[(K_i^2(x)|\mathfrak{F}_{i-1}\right] \ \frac{b_n^3}{n\varphi_n(x)}\sum_{i=1}^{n} b_i^{-3}\mathbb{E}\left[(K_i^2(x)|\mathfrak{F}_{i-1}\right] \\
&= \frac{b_n^3 K^2(1)}{n\varphi_n(x)}\sum_{i=1}^{n} b_i^{-3}\pi_i(x,a_i) - \int_0^1 (K^2(u))'\frac{b_n^3}{n\varphi_n(x)}\sum_{i=1}^{n} b_i^{-3}\pi_i(x,uh)du \\
&= \frac{b_n^3 K^2(1)}{\sum_{i=1}^n \phi_i(x,a_i)}\sum_{i=1}^{n} b_i^{-3}\pi_i(x,a_i) \\
&- \int_0^1 (K^2(u))'\frac{b_n^3\sum_{i=1}^n \phi(x,ua_i)}{\sum_{i=1}^n \phi(x,ua_i)\sum_{i=1}^n \phi(x,a_i)}\sum_{i=1}^{n} b_i^{-3}\pi_i(x,ua_i)du + o_p(1) \\
&= \delta(1)K^2(1) - \int_0^1 (K^2(u))'\delta(u)\frac{\sum_{i=1}^n \phi(x,ua_i)}{\sum_{i=1}^n \phi_i(x,a_i)}du + o_p(1). \\
&= \delta(1)K^2(1) - \int_0^1 (K^2(u))'\delta(u)\beta_x(u)du + o_p(1).
\end{aligned}$$

We deduce that

$$\lim_{n\mapsto\infty}\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[\eta_{ni}^2|\mathfrak{F}_{i-1}] = 1.$$

Concerning (4.11), we write

$$\zeta_{ni}^2 1\!\!1_{\zeta_{ni}^2>\epsilon n} \le \frac{|\zeta_{ni}|^{2+\nu}}{\sqrt{(\epsilon n)^\nu}} \quad \text{for every} \quad \nu > 0.$$

Observe that

$$\begin{aligned}
\mathbb{E}[\zeta_{ni}^{2+\nu}] &= \mathbb{E}\left[\left|\eta_{ni}(x) - \mathbb{E}[\eta_{ni}|\mathfrak{F}_{i-1}]\right|^{2+\nu}\right] \\
&\leq 2^{1+\nu}\mathbb{E}\left[|\eta_{ni}(x)|^{2+\nu}\right] + 2^{1+\nu}|\mathbb{E}\left[\mathbb{E}[\eta_{ni}|\mathfrak{F}_{i-1}]\right]^{2+\nu}\right]|.
\end{aligned}$$

By the Jensen inequality we obtain

$$\mathbb{E}[\zeta_{ni}^{2+\nu}] \leq C\mathbb{E}\left[|\eta_{ni}(x)|^{2+\nu}\right].$$

So, it remains to evaluate $\mathbb{E}\left[|\eta_{ni}(x)|^{2+\nu}\right]$. To do so, we write

$$\mathbb{E}\left[|\eta_{ni}(x)|^{2+\nu}\right] \leq C\left(\frac{b_n^3}{\varphi_n(x)}\right)^{1+\nu/2}\mathbb{E}\left[K_i^{2+\nu}(x)H_i^{'(2+\nu)}(\theta)\right].$$

Simleraly to (4.13) we have

$$\mathbb{E}\left[K_i^{2+\nu}(x)H_i^{'(2+\nu)}(\theta)\right] = O(b_i^{-(3+2\nu)}\mathbb{E}[K_i^{2+\nu}(x)]).$$

It follows that

$$\begin{aligned}
\mathbb{E}\left[|\eta_{ni}(x)|^{2+\nu}\right] &\leq C\left(\frac{b_n^3}{\varphi_n(x)}\right)^{1+\nu/2} b_i^{-(3+2\nu)}\mathbb{E}([K_i(x)]^{2+\nu}) \\
&\leq C\left(\frac{b_n^3}{\varphi_n(x)}\right)^{1+\nu/2} b_i^{-(3+2\nu)}\phi(x,a_i).
\end{aligned}$$

Consequently

$$\begin{aligned}
\tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbb{E}[\zeta_{ni}^2 1\!\!1_{\zeta_{ni}^2>\epsilon n}] &\leq C\epsilon^{-\nu/2}n^{-\nu/2}\left(\tfrac{b_n^3}{\varphi_n(x)}\right)^{1+\nu/2}\tfrac{1}{n}\textstyle\sum_{i=1}^n b_i^{-(3+2\nu)}\phi(x,a_i) \\
&\leq C\epsilon^{-\nu/2}n^{-\nu/2}b_n^{-\nu/2}\left(\tfrac{1}{\varphi_n(x)}\right)^{1+\nu/2}\varphi_n(x) \\
&\leq C\epsilon^{-\nu/2}(nb_n\varphi_n(x))^{-\nu/2} \to 0.
\end{aligned}$$

which completes the proof of the Lemma .

□

Proof of Lemma 4.6 Similarly to Lemma (4.2), we write

$$|\widehat{B}'_n(x,y)| \;\; \leq \frac{1}{n\varphi_n(x)\bar{f}_D(x)} \sum_{i=1}^{n} \left\{ \mathbb{E}\left(K_i(x) \left| \mathbb{E}[H'_i(y)|X_i] - f'^{x}(y) \right| \mathfrak{F}_{i-1} \right) \right\}.$$

Let us calculate the quantity $\mathbb{E}[H'_i(y)|X]$. Clearly,

$$\mathbb{E}[H'_i(y)|X] = \frac{1}{b_i^2} \int H'\Big(\frac{y-z}{b_i}\Big) f^x(z)dz.$$

Simple algebra gives,

$$\mathbb{E}[H'_i(y)|X] = \int H(t) f'^x(y - b_i t)dt.$$

Now, (N2) permits to write

$$\left| E\left(H'_i(y)|X\right) - f^{x'}(y) \right| \;\leq\; C \int_{\mathbb{R}} H(t) \left(a_i^{\beta_1^1} + |t|^{\beta_2^1} b_i^{b_2^1} \right) dt$$

Therefore,

$$|\widehat{B}'_n(x,y)| = O_p\left(\frac{1}{n\varphi_n(x)} \sum_{i=1}^{n} \left(a_i^{\beta_1^1} \pi_i(x,a_i) \right) \right) + O_p\left(\frac{1}{n\varphi_n(x)} \sum_{i=1}^{n} \left(b_i^{\beta_2^1} \pi_i(x,a_i) \right) \right),$$

Hence, the claimed result of this lemma is then a direct consequence of Assumption (N5) □

Proof of Lemma 4.7 We use a similar ideas of Ferraty et al. (2009) to write

$$\mathbb{E}\left[K_i(x)\right] = K(1)\phi(x,a_i) - \int_0^1 (K(u))'\phi(x,ua_i)du$$

and

$$\mathbb{E}\left[K_i^2(x)\right] = K^2(1)\phi(x,a_i) - \int_0^1 (K^2(u))'\phi(x,ua_i)du$$

It follows that

$$\mathbb{E}[\widehat{f}_D(x)] = K(1) - \int_0^1 (K(u))' \frac{\sum_{i=1}^n \phi_i(x,ua_i)}{\sum_{i=1}^n \phi_i(x,a_i)} du$$

and

$$Var[\widehat{f}_D(x)] = \frac{1}{n\varphi_n(x)} \left(K^2(1) - \int_0^1 (K^2(u))' \frac{\sum_{i=1}^n \phi_i(x, ua_i)}{\sum_{i=1}^n \phi_i(x, a_i)} du \right)$$

Therefore, under (N1) we get

$$\mathbb{E}[\widehat{f}_D(x)] \to K(1) - \int_0^1 (K(u))\beta_x(u)du.$$

and

$$Var[\widehat{f}_D(x)] \to 0.$$

which complete the proof of this lemma □

Proof of Lemma 4.8 We write

$$\left|\widehat{f''}^x_N(\theta^*) - f''^x(\theta)\right| \leq \left|\widehat{f''}^x_N(\theta^*) - f''^x(\theta^*)\right| + \left|f''^x_N(\theta^*) - f''^x(\theta)\right|$$

Now, using the same arguments as those used in the proof of Lemma 4.3, we get the almost complete convergence of $sup_{y\in S}\left|\widehat{f''}^x_N(y) - f''^x(y)\right|$ to 0 which implies its convergence in probability. While the consistency of the second term is consequence of results (4.7) and the fact that the conditional density is of class C^2

□

Acknowledgements. The authors would like to thank the Associate-Editor and the anonymous reviewer for their valuable comments and suggestions which improved substantially the quality of an earlier version of this paper.

Bibliographie

Amiri, A., Crambes, Ch. and Thiam, B. (2014). Recursive estimation of nonparametric regression with functional covariate. *Comput. Statist. Data Anal.*, **69** , 154–172.

Attouch, M. K. and Bouabça, W. (2013). The k-nearest neighbors estimation of the conditional mode for functional data. *Rev. Roumaine Math. Pures Appl.*, **58**, 393–415.

Bosq, D. *Linear Processes in Function Spaces : Theory and applications.* Lecture Notes in Statistics, 149, Springer, 2000.

Besse, P. and Cardot, H. (1996). Approximation spline de la prévision d'un processus fonctionnel autorégressif d'ordre 1. *Canad. J. Statist.*, **24**, 467–487.

Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf.*, **15**, 227-236.

Dabo-Niang, S. and Laksaci, A. (2007). Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. *C. R., Math., Acad. Sci. Paris*, **344**, 49-52.

Dabo-Niang, S. and Laksaci, A. (2012). On spatial conditional mode estimation for a functional regressor, *Stat. Probab. Lett.*, **82**, 1413–1421.

Ezzahrioui, M. and Ould-Said, E. (2008). Asymptotic normality of nonparametric estimators of the conditional mode function for functional data *J. Nonparametric. Statist*, **20**, 3-18.

Ezzahrioui, M. and Ould-Said, E. (2010). Some asymptotic results of a non-parametric conditional mode estimator for functional time-series data. *Stat. Neerl.* **64**, 171–201

Ferraty, F. and Vieu, P. *Nonparametric Functional Data Analysis.* Springer Series in Statistics, Springer New York, 2006.

Ferraty, F., Laksaci, A. and Vieu, P. (2006). Estimating some chacteristics of the conditional distribution in nonparametric functional models. *Statist. Inf. for Stoch. Proc.* **9**, 47-76.

Ferraty F., Laksaci A., Tadj A. and Vieu P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *J. Statist. Plann. and Inf.* **140**, 335-352.

Gheriballah, A., Laksaci, A. and Sekkal, S. (2013) Nonparametric M-regression for functional ergodic data. *Stat. Probab. Lett.* **83** , 902–908

Quintela del Rio, A. and Vieu, P. (1997). A nonparametric conditionnal mode estimate. *J. Nonparametric. Statist*, **8**, 253-266.

Laïb, N. and Louani, D. (2011). Strong consistency of the regression function estimator for functional stationary ergodic data. *J. Statist. Plann. and Inf.* **141**, 359-372.

Ling, N. and Xu, Q. (2012). Asymptotic normality of conditional density estimation in the single index model for functional time series data. *Stat. Probab. Lett.*, **82**, 2235–2243.

Louani, D. and Ould-Saïd, E. (1999)Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametric. Statist*, **11**, 413-442.

Ramsay, J.O. and Silverman, B.W. *Functional Data Analysis.* Second Edition. Springer, New York, 2005.

Roussas, G.R. and Tran, L.T. (1989). Asymptotic normality of the recursive kernel regression estimate under dependence conditions. *Ann. Statist.*, **20**, 98–120.

Thiam, B. (2006). Recursive estimation of functional. Thesis Doctarat, University Versailies.

Conclusion et Perspectives

Conclusion

Dans ce travail, nous nous concentrons sur l'estimation récursive qui est une nouvelle méthode pour estimer le mode conditionnel avec des données fonctionnelles. Comme résultats asymptotique nous donnons la convergence presque complète de cet estimateur.

Nous rappelons que l'importance de ce travail est motivée pour le fait que la méthode d'estimation récursive présente des divers avantages par rapport à la méthode classique du noyau. Eneffet il est bien connue que la méthode récursive est plus rapide que la méthode du noyau classique.

En ce qui concerne la dépendance faible des observations nous avons étudiés ce modèle non paramétrique dans le cas ergodique, nous obtenons la vitesse de convergence presque complète. Sous cette hypothèse d'ergodicité on donne également la normalité asymptotique.

Perspectives

A fini, de mettre en évidence l'originalité de cette thèse nous allons laisser quelques questions ouvertes á traiter pour des futurs recherche.

• Estimation récursive de mode conditionnel pour des données ergodiques au cas spatial,

Une question naturelle de ce travail est d'étendre notre approche sur la méthode de récursive du cas des données spatiales.

• Estimation récursive de mode conditionnel pour des données ergodiques au cas des données incomplètes,

Le cas des données incomplètes sous l'hypothèse d'ergodicité n'est pas encore traité, ce qui permettre de étudiée nos résultats président pour des données censurées ou bien tronquées. Ainsi, on peut envisager de généraliser notre l'étude à ce cas trés intéressant en pratique.

Bibliographique générale

AMIRI, A. (2009). Sur une famille paramétrique d'estimateurs séquentiels de la densité pour un processus fortement mélangeant, C. R. Acad. Sci. Paris, Ser. I, **347**, 309-314.

AMIRI, A. (2012). Recursive regression estimators with application to nonparametric prediction. J. Nonparametr. Stat. **24**(1),169à186.

AMIRI, A. (2012). Estimateurs fonctionnels récursifs et leurs applications à la prévision,thèse de doctorat, académie d'aix-marseille université dŠavignon et des pays de vaucluse. Ecole Doctorale ED 536 Sciences et Agrosciences Laboratoire d'Analyse Non Linéaire et Géométrie (EA 2151).

ANOUK BARBEROUSSE Ludwig Boltzmann, les théories physiques et les atomes Institut d'Histoire et de Philosophie des Sciences et des Techniques, CNRS, Université Paris 1, ENS, Paris. Une thèse.

ATTOUCH, M. K. ET BOUABÇA, W. (2013). The k -nearest neighbors estimation of the conditional mode for functional data. Rev. Roumaine Math. Pures Appl., **58**, 393à415.

AHMAD, I. (1976). Uniform strong convergence of a generalized failure rate estimate. *Bull. Math. Statist.* **17**, 77-84.

ANEIROS-PEREZ, G., CARDOT, H., ESTEVEZ-PEREZ, G. ET VIEU, P. (2004). Maximum ozone concentration forecasting by functional nonparametric approaches. Environmetrics, **15**, 675-685.

BOUADJEMI, A., (2004) Asymptotic Normality of the Recursive Kernel Estimate of Conditional Cumulative Distribution Function. Journal of Probability and Statistical Science **12**(2), 117-126.

BESSE, P., CARDOT, H. ET STEPHENSON, D. (2000). Autoregressive forecasting of some functional climatic variations. Scandinavian Journal of Statistics, **27**, 673-687.

BENKO, M., HÄRDLE, W. ET KNEIP, A. (2005). Common functional principal components. SFB 649 Economic Risk Discussion Paper, 2006-010.

BOENTE, G., FRAIMAN,R., (1990).Asymptotic distribution of robust estimates for nonparametric models from mixing observations.*Ann.Statist.***18**,891à906.

BOENTE, G., GONZÄLEZ W.,GONZÄLAZ A. (2009). Robust nonparametric estimation with missing data. *Journal of statistical planning and inference* **139**, 571-592.

BENZIADI, F., LAKSACI, A., TEBBOUNE F. (2012) Recursive kernel estimate of the conditional quantile for functional ergodic data Communications in Statistics - Theory and Methods Pages 3097-3113.

COLLOMB, G., HÄRDLE, W., AND HASSANI, S. (1987). A note on prediction via conditional mode estimation. J. Statist. Plann. Inference,**15**, 227-236.

COLLOMB, G., HÄRDLE, W. (1986). *Strong uniform convergence rates in robust nonparametric time series analysis and prediction : Kernel regression estimation from dependent observations.* Stochastic Processes and Their Applications, **23**, 77-89.

CARDOT, H., CRAMBES, C., SARDA, P. (2004), *Spline estimation of conditional quantities for functional covariates.* C. R., Math., Acad. Sci. Paris **339**, No.2, 141-144.

DABO-NIANG, S., RHOMARI, N., (2002), Nonparametric regression estimation when the regressor takes values in a metric space, Technical Repport 2002-9, LSTA Univ. paris 6, 2001.

DABO-NIANG, S., RHOMARI, N., (2003), Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *C. R., Math., Acad. Sci. Paris* **336**, No.1, 75-80.

DABO-NIANG, S., LAKSACI, A. (2007). Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. C. R., Math., Acad. Sci. Paris, **344**, 49-52.

DABO-NIANG, S. AND LAKSACI, A. (2012). On spatial conditional mode estimation for a functional regressor, Stat. Probab. Lett., **82**, 1413à1421.

DEVROYE, L., WAGNER, T.J., (1980). Distribution-free consistency results in nonparametric discrimination and regression function estimation. Ann. Statist. **8**, 231à239.

DEVROYE, L. (1981). On the almost everywhere convergence of nonparamertic regression fuction estimates, Ann. Statist. **9**, 1310-1319.

DABO-NIANG, S., RHOMARI, N. (2003). Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *C. R., Math., Acad. Sci. Paris.* **336**, 75-80.

Dabo-Niang, S. and Laksaci, A. (2012). On spatial conditional mode estimation for a functional regressor, *Statistics Probability Letters, Elsevier*, vol. **82(7)**, pages 1413-1421.

EZZAHRIOUI, M. et OULD-SAÏD, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. Nonparametr. Stat.* **20**, 3-18.

FERRATY, F. et VIEU, P. (2006). *Nonparametric functional data analysis. Theory and Practice.* Springer Series in Statistics. Springer. New York.

FERRATY, F., VIEU, P. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *C. R. Acad. Sci., Paris.* **330**, No.2, 139-142.

FERRATY, F., VIEU, P. (2002). The functional nonparametric model and application to spectrometric data. *Comput. Statist. and Data Anal.* **17**, 545-564.

FERRATY, F., GOIA, A., VIEU, PH., (2002), Functional nonparametric model for time series : a fractal approach for dimension reduction. Test **11**, No.2, 317-344.

FERRATY, F., GOIA, A., VIEU, PH., (2002), Régression non-paramétrique pour des variables aléatoires fonctionnelles mélangeantes. *C. R., Math., Acad. Sci. Paris* **334**, No. 3, 217-220.

FERRATY, F., LAKSACI, A. AND VIEU, P. (2004), Estimating some chacteristics of the conditional distribution in nonparametric functional models. *Statist. Inf. for Stoch. Proc.*. Sous presee.

FERRATY, F., LAKSACI, A. AND VIEU, P. (2005), Functional times series prediction via conditional mode. *C. R., Math., Acad. Sci. Paris* **340**, No. 5, 389-392.

FERRATY, F., VIEU, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. *J. Nonparametric Statist.* **16**, 111-125.

Ferraty, F. Laksaci, A. et Vieu, P. (2005). *Functional times series prediction via conditional mode. C. R., Math., Acad. Sci. Paris.* **340**, No. 5, 389-392.

Ferraty, F. and Vieu, P. Nonparametric Functional Data Analysis. Springer Series in Statistics, Springer New York, 2006.

Ferraty, F. Romain, Y (2010) Handbook of functional data anlysis related fields OXford university press.OXford.
Gheriballah, A., Laksaci, A. and Sekkal, S. (2013) Nonparametric M -regression for functional ergodic data. Stat. Probab. Lett. **83** , 902à908.

Greblecki, W. and Pawlak, M, (1987). Necessary and sufficient consistency conditions for a recursive kernel regression estimate, J. Multivariate Anal. **23**, 67-76.

Gooijer, J.G., Gannoun, A. (2000). *Nonparamertric Predictive Region for Stochastic Processes. Computational Statistics and Data Analysis,***33**, 259-275.

Györfi, L. (1981) Strong consistent density estimate from ergodic sample. J Multivariate Anal **11**, 81à84.

Hastie, T., Buja, A. and Tibshirani, R. (1995). Penalized discriminant analysis. Annals of Statistics, **23**, 73-102.

Krzyzak, A. (1992). Global convergence of the recursive kernel regression estimates with applications in classification and nonlinear system estimation, IEEE Trans. In- form. Theory **38**, 1323-1338.

Laksaci, M. and Maref, F. (2009). Conditional cumulative distribution estimation and its applications Journal of probability and statistical sciences, **13**, Pages 47-56.

Ling,N.X. and Xu, Q. (2012). Asymptotic normality of conditional density estimation in the single index model for functional time series data, Statistics and Probability Letters. **82** 2235-2243.

Laïb, C. Louani, D. Strong consistency of the regression function estimator for functional stationary ergodic data, J. Statist. Plann. Inference (2010), in press.

LAÏB, N. AND LOUANI, D. (2011). Strong consistency of the regression function estimator for functional stationary ergodic data. J. Statist. Plann. and Inf. **141**, 359-372.

LAKSACI, A., LEMDANI, M. ET OULD-SAÏD, E. (2009) A generalized L^1-approach for a kernal estimator for conditional quantile with functional regressors : Consistency and asymptotic normality Stat. Probab . **79** 1065-1073.

LAKSACI, A. LEMDANI, M. OULD-SAÏD, E. (2011). Asymptotic results for an L1-norm kernel estimator of the conditional quantile for functional dependent data with application to climatology. Sankhya, A **73** 125à14.

LOUANI, D. et OULD-SAÏD, E. (1999). *Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. J. Nonparametric Stat.* **11**, 413-442.

LIANG, H.Y,. AND BAEK, J. (2004). Asymptotic normality of recursive density estimates under some dependence assumptions, Metrika **60**, 155-166.

MASRY, E. (1986). Recursive probability density estimation for weakly dependent stationary processes, IEEE Trans. Inform. Theory 32, no **2**, 254-267.

MASRY, E (1987). Almost sure convergence of recursive density estimators for stationary mixing processes, Statist. Probab. Lett. **5**, 249-254.

MAMMAN, E. (1991a). Estiming a smooth monotone regression function. Ann. Of Statist, **19,** 724-740.

MAMMAN, E. (1991b). Nonparametric regression under qualitative smoothness assumptions. Ann. Of Statist, **19**, 741-759.

MASRY, E.A (2005). Nonparametric regression estimation for dependent distribution, *Theo. Prob. Appl.* **19**, 794-799.

OULD-SAÏD, E., (1997), A note on ergodic processes prediction via estimation of the conditional mode function. *Scand. J. Stat.* **24**, No.2, 231-239.

QUINTELA DEL RIO, A. AND VIEU, PH., (1997), A nonparametric conditionnal mode estimate. *Nonparametric Statistics*, **8**, 253-266.

RACHDI,M., LAKSACI,A., DEMONGEOT,J., ABDALI,A. AND MADANI,F., (2014) Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data, Computational Statistics and Data Analysis,**73** : 53-68.

ROUSSAS, G.G. (1968). On some properties of nonparametric estimates of probability density functions. *Bull. Soc. Math Gréce.,* (N.S.) **9** fasc. 1,29-43.

ROUSSAS, G.G. (1969). Nonparametric estimation of the transition distribution function of a Markov process, *Ann. Math. Statist.,* **40**, 1386-1400.

ROUSSAS, G.R. AND TRAN, L.T. (1989). Asymptotic normality of the recursive kernel regression estimate under dependence conditions. Ann. Statist., **20**, 98à120.

ROUSSAS, G.G. (1990), Nonparametric Regression Estimation Under Mixing Conditions, Stochastic Processes and their Applications, **36**(1), 107-116.

ROSENBLATT, M., (1969), Conditional probability density and regression estimators. In Multivariate Analysis II, Ed. P.R. Krishnaiah. Academic Press, New York and London.

ROSA, A.C.M. (1992), Estimation non-paramétrique de la médiane conditionnelle sous hypothèse ergodique. Application à la prévision des processus. XXIV e Journees de statistique, Bruxelles, pp. 325-328.

SAMANTA, M., THAVANESWARAN, A. (1990). *Non-parametric estimation of conditional mode. Comm. Statist. Theory and Meth.* **16**, 4515-4524.

WOLVERTON, C. AND WAGNER, T. J. (1969). Recursive estimates of probability densities, IEEE Trans. Syst. Cybern. **5**, 307-308.

WALK, H. (2001). Strong universal pointwise consistency of recursive regression estimates, Ann. Inst. Statist. Math. **53** (4), 691-707.

WEGMAN,E. J. AND DAVIES, H. I. (1979). Remarks on some recursive estimators of a probability density, Ann. Statist. **7** (2), 316-327.

ZHANG, J., (2001), Estimation of conditional quantiles using artificial neural networks Estimation de quantiles conditionnels par réseaux de neurones, Vol **332**, P 569-574.