Table des matières

Chapitre 1

Introduction générale

1.1 Historique

En statistique non-paramétrique, les outils naturels pour faire la prévision sont les modèles liés à la distribution conditionnelle. Parmi ces modèles les plus sollicités citons la régression classique.
Les premiers résultats sur l'estimation de ce modèle ont été obtenus par Tukey [21]. Trois ans plus tard Nadaraya [16] et Watson [24] ont utilisé la méthode du noyau pour estimer la fonction de régression et ils ont établi la convergence uniforme de l'estimateur construit. Stone [20] a considéré un estimateur de type robuste de la régression non-paramétrique et il a donné des conditions nécessaires et suffisantes à la convergence en norme L^p. D'autre auteurs ont étudié ce modèle non-paramétrique dans le cas multivarié, citons par exemple, Bosq et Lecoutre [2] pour la convergence en moyenne d'ordre p, Roussas [19] pour la normalité asymptotique.

Les premiers résultats asymptotiques sur l'estimation non-paramétrique de la fonction de régression sur les processus α-mélangeants ont été élaborés par Gyorfi et *al.* [13]. Dans ce cadre α-mélangeant, Vieu [23] a donné les termes asymptotiquement exacts de l'erreur quadratique de l'estimateur à noyau de la fonction de régression.

Ces derniers temps, ce modèle a pris un essor considérable, lorsque la variable explicative est de nature fonctionnelle. En effet, en 2002, Ferraty et *al.* [8] ont établi la convergence presque-complète d'un estimateur à noyau pour

ce modèle, en considérant des observations α-mélangantes. Dans la même année, Dobo-Niang et Rhomari [5] ont étudié la convergence en norme L^p pour le même estimateur et ils ont donné l'expression de la vitesse de convergence, ainsi que quelques applications à la prévision et à la discrimination des courbes. En supposant que les observations sont α-mélangeant Masry [15] a démontré la normalité asymptotique de cet estimateur. En collaboration avec Mas et Vieu, Ferraty [11] ont étudié la convergence en moyenne quadratique de l'estimateur à noyau de la fonction de régression à co-variable fonctionnelle. Ils ont obtenu l'expression asymptotiquement exacte de cette erreur quadratique, lorsque les observations sont indépendantes identiquement distribuées. Ce dernier résultat a été généralisé par Delsol [7] au cas où les observations sont dépendantes.
Dans un contexte assez général, Crambes et *al.* [4] ont étudié dans les deux cas (dépendant et indépendant) l'erreur d'ordre p d'une famille d'estimateurs robustes de la fonction de régression. Les expressions asymptotiquement exactes de ces erreurs sont étroitement liés à la normalité asymptotique de cette famille d'estimateurs initialement obtenu par Attouch et *al.* [1]. Il est à noter que ce type de résultats est important pour le choix du paramètre de lissage. A ce sujet, nous revoyons à Rachdi et Vieu [17] pour le choix local et globale du paramètre de lissage de ce modèle.

D'une manière générale, la modélisation statistique des variables fonctionnelles est devenue un sujet révélateur en statistique moderne comme en témoigne les nombreux numéros spéciaux accordés à cette thématique (citons par exemples, Ferraty et *al.* [10], Gonzalez-Manteiga and Vieu [12] ou Valderrama [22]).

Dans ce travail, on se propose de déterminer les constantes exactes intervenant dans l'erreur quadratique de l'estimateur à noyau de la régression non-paramétrique. L'expression de cette vitesse garde une forme usuelle du développement asymptotique de l'erreur quadratique obtenu par Vieu [23] ce qui est très pratique pour le choix du paramètre de lissage. Mentionnons, également que cette vitesse de convergence est liée à la régularité de notre modèle non-paramétrique et elle ne dépende pas de la corrélation des observations.

Ce manuscrit est présenté en quatre chapitres, et il est organisé comme suit : Après ce chapitre introductif destiné premièrement à un historique sur l'estimation de la fonction de régression par la méthode du noyau dans les deux cas multivarié et fonctionnel lorsque les observations sont indépendantes ou dépendantes ; on introduit dans la deuxième partie de ce chapitre un ensemble de définitions et notations permettant de clarifier le vocabulaire utilisé dans la suite du document, ainsi qu' un nombre important d'outils nécessaires pour l'élaboration de nos résultats.
Dans le deuxième chapitre on établit la convergence en moyenne quadratique d'un estimateur à noyau de la fonction de régression dans le cas où la variable explicative est scalaire (de dimension un) et en considérant le cas d'une suite d'observations indépendantes et identiquement distribuées. On étudiera également cette propriété asymptotique lorsque, les observations sont fortement mélangeantes.
Dans le troisième chapitre, on généralise les résultats du chapitre précédent au cas multivarié, dont l'objectif est d'étudier l'effet de la dimension sur l'erreur quadratique. On montre que la dimension a un effet inverse sur la vitesse de convergence. Autrement dit, on constate qu'il y a une dégradation de la vitesse de convergence par rapport à l'augmentation de la dimension.
Le quatrième chapitre est consacré au cas d'une variable explicative fonctionnelle. Dans ce contexte, nous établissons la vitesse de convergence en moyenne quadratique, de l'estimateur à noyau de la fonction de la régression dans les deux cas : cas i.i.d. et α-mélangeant. Nous terminons ce travail par une conclusion générale et quelques perspectives.

1.2 Préliminaires

1.2.1 Erreur quadratique moyenne

Pour évaluer la précision de notre estimateur non-paramétrique, nous utiliserons dans notre étude le critère de convergence dans L^2 appelé aussi l'erreur en moyenne quadratique. C'est un critère très répandu dans la littérature pour évaluer la précision d'une valeur estimée en un point x, et on a la définition suivante :

Définition 1.2.1.1. *si* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires définies sur l'espace probabilisé* $(\Omega, \mathcal{A}, \mathbb{P})$*, et* X *une variable aléatoire définie sur le même espace et telle que* $\mathbb{E}(|X|^2) < \infty$*, alors on dit que* X_n *converge dans* L^2 *vers* X *si :*

$$\lim_{n\longrightarrow\infty} \mathbb{E}(|X_n - X|^2) = 0$$

1.2.2 Mesures de dépendance

La manière la plus habituelle qui est utilisée pour modéliser la dépendance dans des problèmes de prévision sous hypothèse non-paramétrique est de supposer que le processus vérifi une condition de mélange. Nous donnons ici quelques définitions sur le coeffcient α de mélange fort.

Définition 1.2.2.1. *Soit* $\{\Delta_i, i \in \mathbb{Z}\}$ *une famille de variables aléatoires à valeurs dans un même espace probabilisable* $(E, \mathcal{E})$*. pour tout couple* (i, j) *dans* $\mathbb{Z} \cup \{-\infty, +\infty\}$*, on note* σ_i^j *la tribu engendré par* $\{\Delta_k, i < k < j\}$*. On appelle coefficients de mélange fort, les réels*

$$\alpha(n) = \sup_{\{k\in\mathbb{Z}, A\in\sigma_{-\infty}^k, B\in\sigma_{n+k}^{+\infty}\}} |P(A \cap B) - P(A)P(B)|$$

Définition 1.2.2.2. *On dit qu'une famille* $\{\Delta_i, i \in \mathbb{Z}\}$ *de variables aléatoires à valeurs dans un même espace probabilisable* $(E, \mathcal{E})$ *est fortement mélangeante, ou* α*-mélangeante , si l'on a*

$$\lim_{n\longrightarrow\infty} \alpha(n) = 0 \tag{1.1}$$

Définition 1.2.2.3. *On dit qu'une famille* $\{\Delta_i, i \in \mathbb{Z}\}$ *de variables aléatoires à valeurs dans un même espace probabilisable* $(E, \mathcal{E})$ *est algébriquement* α*-mélangente, s'il existe deux constantes* $c \in \mathbb{R}^{*+}$ *et* $a \in \mathbb{R}^{*+}$ *telles que les coefficients de mélange vérifient*

$$\alpha(n) \leq cn^{-a}$$

La littérature fait état de nombreux types de conditions de mélange. La définition suivante présente les deux autres types les plus utilsés.

Définition 1.2.2.4. *On dit que la famille* $\{\Delta_i, i \in \mathbb{Z}\}$ *est* φ*-mélangent (resp.* ρ*-mélangent) si la suite*

$$\varphi(n) = \sup_{k\in\mathbb{Z}} \sup_{A\in\sigma_{-\infty}^{k}, B\in\sigma_{n+k}^{+\infty}} |\mathbb{P}(B/A) - \mathbb{P}(A)|$$

(resp.

$$\rho(n) = \sup_{k\in\mathbb{Z}} \sup_{\{X\in L_2(\sigma_{-\infty}^{k}), Y\in L_2(\sigma_{n+k}^{+\infty})\}} \frac{\mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]}{[Var(X)\, Var(Y)]^2}$$

tend vers 0 quand n tend vers l'infinie[1].

La relation entre ces trois types des processus est donnée dans le lemme suivant :

Lemme 1.2.1. *Soit* $\{\Delta_i, i \in \mathbb{Z}\}$ *une famille de variables aléatoires définie sur* $(\Omega, \mathcal{A}, \mathbb{P})$ *à valeurs dans un espace probabélisable* $(E, \mathcal{E})$*. Pour tout* $n \in \mathbb{N}$*, on a,*

$$\alpha(n) \leq \rho(n) \leq 2\varphi(n)^{1/2}$$

En vertu de ce lemme on peut dire que le processus α-mélangeant est le plut fort. Ainsi, nous allons modélisé la notion de la dépendance par ce processus α-mélangeant, car, elle englobe les autres processus définies ci-dessus.

L'utilisation des inégalités exponentielles nécessite l'utilisation d'inégalités de covariance pour variables mélangeantes. Pour ce mémoire nous fonderons nos preuves sur l'inégalité donnée dans le lemme suivant

Lemme 1.2.2. *[3] Soit* $\{\Delta_i, i \in \mathbb{N}\}$ *une famille de variables aléatoires à valeurs dans* $\mathbb{R}$ *qui vérifient la condition de mélange fort (1.1), et telle que* $\|\Delta_i\|_\infty < \infty$, $\forall i$*. On a pour tout* $i \neq j$ *:*

$$|Cov(\Delta_i, \Delta_j)| \leq 4\|\Delta_i\|_\infty \|\Delta_j\|_\infty \alpha(|i-j|)$$

1. $L_2(\sigma_i^j)$ est l'espace des variables aléatoires σ_i^j-mesurables et de carrées sommable.

1.2.3 Le noyau d'ordre k

La définition suivante précise le noyau d'ordre k.

Définition 1.2.3.1. *Une fonction K de $\mathbb{R}^p$ est dit noyau d'ordre $k, k \in \mathbb{N}^*$, si :*

$$T_{(i_1,\ldots,i_p)}(K) = 0, \quad \textit{pour tout} \quad (i_1,\ldots,i_p) \in \mathbb{N}^{*p} \quad \textit{vérifiant } i_j < k \, (1 \leq j \leq p)$$

$$\textit{et} \quad T_j(K) \neq 0 \quad \textit{pour tout} \quad j \leq k$$

$$\textit{où} \quad T_{(i_1,\ldots,i_p)}(K) = \int_{\mathbb{R}^p} u_1^{i_1} \ldots u_p^{i_p} K(u_1,\ldots,u_p) du_1 \ldots du_p,$$

$$\textit{et} \quad T_j(K) = \int_{\mathbb{R}^p} u_j^k K(u_1,\ldots,u_p) du_1 \ldots du_p.$$

telle que $k > 0$, c'est à dire qu'il vérifie :

$$\int t^j K(t)dt = 0, \forall j = 1, ..., k-1 \qquad \textit{et} \qquad 0 < |\int t^k K(t)dt| < \infty$$

Chapitre 2

La covergence en moyenne quadratique : cas réel

Dans ce chapitre, nous étudions la convergence en moyenne quadratique de la fonction de régression lorsque la variable explicative est réelle. Ce chapitre est divisé en trois sections. On introduit notre modèle ainsi que son estimateur dans la première section. Nous établissons la convergence en moyenne quadratique de notre estimateur dans le cas où les variables sont indépendantes et identiquement distribuées dans la deuxième section. Le traitement du cas de mélange fort est explicitement donné dans la dernière section.

2.1 Le modèle et son estimateur

Soit (X, Y) un couple de variables aléatoires défini sur l'espace de probabilté $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R} \times \mathbb{R}$. On suppose que la variable explicative X à une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ notée f. Le modèle de régression est donné par :

$$Y = r(x) + \epsilon \tag{2.1}$$

Où ϵ est une variable aléatoire réelle centrée et indépendante de X et r une fonction de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ supposée de classe C^k. Notont que si ϵ est de loi normale centrée, la fonction r peut être exprimer par :

$$r(x) = \mathbb{E}(Y/X = x)$$

Ainsi, on peut estimer la fonction r par la méthode du noyau à partir des observations $(X_i, Y_i), (i = 1, \ldots, n)$ par :

$$\hat{r}(x) = \frac{\sum_{i=1}^{n} Y_i K\left(\frac{x - X_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)} \tag{2.2}$$

Où K est une fonction de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ appelée noyau et h est un paramètre réel strictement positif appelé paramètre de lissage.
Pour simplifier la notation on écrit :

$$\hat{r}(x) = \frac{\hat{g}(x)}{\hat{f}(x)}$$

Avec

$$\hat{g}(x) = \frac{1}{nh} \sum_{i=1}^{n} Y_i K\left(\frac{x - X_i}{h}\right)$$

Et

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

2.2 Cas des observations indépendantes et de même loi

Dans cette section, on considère n couples de variables aléatoires (X_i, Y_i) indépendantes et identiquementes distribuées dont la loi commune est celle du couple (X, Y). Dans ce qui suit, on fixe un point $x \in \mathbb{R}$ et on introduit les conditions suivantes :

2.2.1 Hypothèses

(H1) Les fonctions r et f sont 2-fois continûment défférentiables au voisinage de x.

(H2) La densité f de la variable explicative est strictement positive au point x.

(H3) Le paramètre de lissage est tel que :

$$\lim_{n \longrightarrow \infty} h = 0 \qquad et \qquad \lim_{n \longrightarrow \infty} nh = \infty$$

(H4) Le noyau K est d'ordre 2, borné et intégrable.

(H5) La variable réponse est telle que : $|Y| < M < \infty$

2.2.2 Propriété asymptotique

On établit le résultat suivant

Théorème 2.2.1. *Sous les hypothèses (H1)-(H5), on a :*

$$\mathbb{E}[\hat{r}(x) - r(x)]^2 = B^2(x)h^4 + V(x)\frac{1}{nh} + o(h^4 + \frac{1}{nh})$$

où

$$B(x) = \frac{(g^{(2)}(x) - r(x)f^{(2)}(x)) \displaystyle\int t^2 K(t)dt}{2f(x)}$$

et

$$V(x) = \frac{(\phi(x) - r^2(x))}{f(x)} \int K^2(t)dt$$

Avec

$$\phi(x) = \mathbb{E}(Y^2/X = x)$$

Démonstration du Théorème 2.2.1

La démonstration se fait par un calcul séparé des deux parties : Partie biais et partie variance. En effet, l'erreur quadratique peut être exprimé :

$$\mathbb{E}[\hat{r}(x) - r(x)]^2 = [\mathbb{E}(\hat{r}(x)) - r(x)]^2 + Var[\hat{r}(x)]$$

L'idée principale de cette preuve, dont la difficulté réside en la nécessité de calculer l'espérance de quotients aléatoires, est d'utiliser le développement usuel de $\frac{1}{z}$:

$$\frac{1}{z} = 1 - (z-1) + \ldots + (-1)^p (z-1)^p + (-1)^{p+1}\frac{(z-1)^{p+1}}{z}, \forall z \neq 0, \forall p \in \mathbb{N}^*$$

Une application de ce développement pour $z = \dfrac{\hat{f}(x)}{\mathbb{E}\hat{f}(x)}$ et $p = 1$ nous permet d'écrire

$$\begin{aligned}\hat{r}(x) - r(x) &= \left(\frac{\hat{g}(x)}{\mathbb{E}\hat{f}(x)} - r(x)\right) - \frac{(\hat{g}(x) - \mathbb{E}\hat{g}(x))}{(\mathbb{E}\hat{f}(x))^2}(\hat{f}(x) - \mathbb{E}\hat{f}(x)) \\ &\quad - \frac{\mathbb{E}\hat{g}(x)}{(\mathbb{E}\hat{f}(x))^2}\left(\hat{f}(x) - \mathbb{E}\hat{f}(x)\right) + \frac{\left(\hat{f}(x) - \mathbb{E}\hat{f}(x)\right)^2}{(\mathbb{E}\hat{f}(x))^2}\hat{r}(x)\end{aligned}$$

Ceci implique que :

$$\mathbb{E}(\hat{r}(x)) - r(x) = \left(\frac{\mathbb{E}\hat{g}(x)}{\mathbb{E}\hat{f}(x)} - r(x)\right) - \frac{Cov(\hat{g}(x), \hat{f}(x))}{(\mathbb{E}\hat{f}(x))^2} + \frac{\mathbb{E}[(\hat{f}(x) - \mathbb{E}\hat{f}(x))^2]\hat{r}(x)]}{(\mathbb{E}\hat{f}(x))^2}$$

Comme la variable Y est borné, on peut trouver une constante $C > 0$ telle que $\hat{r}(x) \leq C$. D'où,

$$\mathbb{E}(\hat{r}(x)) - r(x) = \left(\frac{\mathbb{E}\hat{g}(x)}{\mathbb{E}\hat{f}(x)} - r(x)\right) - \frac{Cov(\hat{g}(x), \hat{f}(x))}{(\mathbb{E}\hat{f}(x))^2} + \frac{Var(\hat{f}(x))}{(\mathbb{E}\hat{f}(x))^2}o(1)$$

La preuve de la partie variance, repose sur la même idée, c'est à dire une application du développement de $\dfrac{1}{z}$ pour $z = \frac{\hat{f}(x)}{\mathbb{E}\hat{f}(x)}$ et $p = 3$

$$Var(\hat{r}(x)) = \frac{Var(\hat{g}(x))}{(\mathbb{E}\hat{f}(x))^2} - 4\frac{\mathbb{E}\hat{g}(x)\, Cov(\hat{g}(x), \hat{f}(x))}{(\mathbb{E}\hat{f}(x))^3} + 3\, Var(\hat{f}(x))\frac{(\mathbb{E}\hat{g}(x))^2}{(\mathbb{E}\hat{f}(x))^4} + o(\frac{1}{nh})$$

Finalement, le Théorème 2.2.1 est une conséquence des lemmes suivants :

Lemme 2.2.1. *Sous les conditions du Théorème 2.2.1 on a :*

i) $\mathbb{E}[\hat{f}(x)] = f(x) + f^{(2)}(x)\dfrac{h^2}{2}\displaystyle\int t^2 K(t)dt + o(h^2),$

ii) $\mathbb{E}[\hat{g}(x)] = g(x) + g^{(2)}(x)\dfrac{h^2}{2}\displaystyle\int t^2 K(t)dt + o(h^2).$

Corollaire 2.2.1. *Sous les conditions du Théorème 2.2.1 on a :*

$$\frac{\mathbb{E}[\hat{g}(x)]}{\mathbb{E}[\hat{f}(x)]} - r(x) = B(x)h^2 + o(h^2)$$

Lemme 2.2.2. *Sous les hypothèses (H1)-(H2) et (H4) on a :*

$$Var[\hat{f}(x)] = \frac{1}{nh}f(x)\int K^2(t)dt + o\left(\frac{1}{nh}\right)$$

Lemme 2.2.3. *Sous les conditions du Théorème 2.2.1 on a :*

$$Var[\hat{g}(x)] = \frac{1}{nh} f(x)\phi(x) \int K^2(t)dt + o\left(\frac{1}{nh}\right)$$

avec

$$\phi(x) = \mathbb{E}[Y^2 | X = x]$$

Lemme 2.2.4. *Sous les conditions du Théorème 2.2.1 on a :*

$$Cov(\hat{g}(x), \hat{f}(x)) = \frac{1}{nh} g(x) \int K^2(t)dt + o\left(\frac{1}{nh}\right)$$

2.2.3 Démonstration des lemmes techniques

Démonstration du lemme 2.2.1 : Pour démontrer i), on a

$$\begin{aligned}
\mathbb{E}[\hat{f}(x)] &= \mathbb{E}\left[\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)\right] \\
&= \frac{1}{nh}\mathbb{E}\left[\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)\right] \\
&= \frac{1}{h}\mathbb{E}\left[K\left(\frac{x - X_1}{h}\right)\right] \\
&= \frac{1}{h}\int K\left(\frac{x - z}{h}\right) f(z)dz
\end{aligned}$$

Le changement de variable $t = \dfrac{x - z}{h}$ donne $\mathbb{E}[\hat{f}(x)] = \int K(t)f(x - ht)dt$
On utilise le développement de Taylor de f au voisinage de x :

$$f(x - ht) = f(x) + htf^{(1)}(x) + \frac{h^2}{2}t^2 f^{(2)}(x) + o(h^2)$$

D'où

$$\mathbb{E}[\hat{f}(x)] = f(x) + \frac{h^2}{2} f^{(2)}(x) \int t^2 K(t)dt + o(h^2)$$

Pour ii), nous avons

$$\begin{aligned}
\mathbb{E}[\hat{g}(x)] &= \mathbb{E}\Bigg[\frac{1}{nh}\sum_{i=1}^{n} Y_i K\Bigg(\frac{x-X_i}{h}\Bigg)\Bigg] \\
&= \frac{1}{nh}\mathbb{E}\Bigg[\sum_{i=1}^{n} Y_i K\Bigg(\frac{x-X_i}{h}\Bigg)\Bigg] \\
&= \frac{1}{h}\mathbb{E}\Bigg[Y_1 K\Bigg(\frac{x-X_1}{h}\Bigg)\Bigg] \\
&= \frac{1}{h}\mathbb{E}\Bigg[\mathbb{E}[Y_1|X=X_1] K\Bigg(\frac{x-X_1}{h}\Bigg)\Bigg] \\
&= \frac{1}{h}\int K\left(\frac{x-z}{h}\right) f(z)r(z)dz \\
&= \frac{1}{h}\int K\left(\frac{x-z}{h}\right) g(z)dz
\end{aligned}$$

Le changement de variable $t = \dfrac{x-z}{h}$ donne

$$\mathbb{E}[\hat{g}(x)] = \int K(t)g(x-ht)dt$$

On utilise le développement de Taylor de g au voisinage de x :

$$g(x-ht) = g(x) + htg^{(1)}(x) + \frac{h^2}{2}t^2 g^{(2)}(x) + o(h^2)$$

Ainsi,

$$\mathbb{E}[\hat{g}(x)] = g(x) + \frac{h^2}{2}g^{(2)}(x)\int t^2 K(t)dt + o(h^2) \qquad \blacksquare$$

Démonstration du Corollaire 2.2.1 : En vertu du Lemme2.2.1, on a

$$\begin{aligned}
\frac{\mathbb{E}[\hat{g}(x)]}{\mathbb{E}[\hat{f}(x)]} - r(x) &= \frac{g(x) + \frac{h^2}{2} g^{(2)}(x) \int t^2 K(t)dt + o(h^2)}{f(x) + \frac{h^2}{2} f^{(2)}(x) \int t^2 K(t)dt + o(h^2)} - r(x) \\
&= \frac{g(x) + \frac{h^2}{2} g^{(2)}(x) \int t^2 K(t)dt + o(h^2) - r(x)\left[f(x) + \frac{h^2}{2} f^{(2)}(x) \int t^2 K(t)dt + o(h^2)\right]}{f(x) + \frac{h^2}{2} f^{(2)}(x) \int t^2 K(t)dt + o(h^2)} \\
&= \frac{g(x) + \frac{h^2}{2} g^{(2)}(x) \int t^2 K(t)dt + o(h^2) - r(x)\left[f(x) + \frac{h^2}{2} f^{(2)}(x) \int t^2 K(t)dt + o(h^2)\right]}{f(x) + O(h^2)} \\
&= \frac{g(x) + \frac{h^2}{2} g^{(2)}(x) \int t^2 K(t)dt - r(x)\left[f(x) + \frac{h^2}{2} f^{(2)}(x) \int t^2 K(t)dt\right]}{f(x)} + O(h^2) \\
&= \frac{h^2 (g^{(2)}(x) - r(x) f^{(2)}(x)) \int t^2 K(t)dt}{2f(x)} + O(h^2) \qquad \blacksquare
\end{aligned}$$

Démonstration du lemme 2.2.2 On a

$$\begin{aligned}
Var[\hat{f}(x)] &= Var\left[\frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)\right] \\
&= \frac{1}{n^2 h^2} Var\left[\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)\right] \\
&= \frac{1}{nh^2}\left(\mathbb{E}\left[K^2\left(\frac{x - X_1}{h}\right)\right] - \left(\mathbb{E}\left[K\left(\frac{x - X_1}{h}\right)\right]\right)^2\right)
\end{aligned}$$

D'après le Lemme 2.2.1 partie i), on a

$$\frac{1}{h^2}\left(\mathbb{E}\left[K\left(\frac{x - X_1}{h}\right)\right]\right)^2 = (\mathbb{E}[\hat{f}(x)])^2 = O(1)$$

D'où

$$\begin{aligned} Var[\hat{f}(x)] &= \frac{1}{nh^2}\mathbb{E}\left[K^2\left(\frac{x-X_1}{h}\right)\right] + o\left(\frac{1}{nh}\right) \\ &= \frac{1}{nh^2}\int K^2\left(\frac{x-z}{h}\right) f(z)dz + o\left(\frac{1}{nh}\right) \end{aligned}$$

En utilisant le changement de variable $t = \dfrac{x-z}{h}$ et la continuité de la fonction $f(x)$, nous obtenons

$$Var[\hat{f}(x)] = \frac{1}{nh} f(x) \int K^2(t)dt + o\left(\frac{1}{nh}\right) \qquad \blacksquare$$

Démonstration du Lemme 2.2.3 La preuve est identique en remarquant que

$$\begin{aligned} Var[\hat{g}(x)] &= Var\left[\frac{1}{nh}\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h}\right)\right] \\ &= \frac{1}{n^2h^2} Var\left[\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h}\right)\right] \\ &= \frac{1}{nh^2}\left(\mathbb{E}\left[Y_1^2 K^2\left(\frac{x-X_1}{h}\right)\right] - \left(\mathbb{E}\left[Y_1 K\left(\frac{x-X_1}{h}\right)\right]\right)^2\right) \end{aligned}$$

D'après le Lemme 2.2.1 partie ii), on a

$$\frac{1}{h^2}\left(\mathbb{E}\left[Y_1 K\left(\frac{x-X_1}{h}\right)\right]\right)^2 = (\mathbb{E}[\hat{g}(x)])^2 = O(1)$$

D'où

$$\begin{aligned} Var[\hat{g}(x)] &= \frac{1}{nh^2}\mathbb{E}\left(K^2\left(\frac{x-X_1}{h}\right)\mathbb{E}[Y_1^2|X = X_1]\right) + o\left(\frac{1}{nh}\right) \\ &= \frac{1}{nh^2}\int K^2\left(\frac{x-z}{h}\right) f(z)\phi(z)dz + o\left(\frac{1}{nh}\right) \end{aligned}$$

Le changement de variable $t = \dfrac{x-z}{h}$ et la continuité de la fonction $f(x)\phi(x)$ permettent d'arriver à

$$Var[\hat{g}(x)] = \frac{1}{nh} f(x)\phi(x) \int K^2(z)dz + o\left(\frac{1}{nh}\right) \qquad \blacksquare$$

Démonstration du Lemme 2.2.4 on a

$$\begin{aligned}
Cov(\hat{g}(x),\hat{f}(x)) &= Cov\left(\frac{1}{nh}\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h}\right), \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{x-X_j}{h}\right)\right)\\
&= \frac{1}{n^2h^2} Cov\left(\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h}\right), \sum_{j=1}^{n} K\left(\frac{x-X_j}{h}\right)\right)\\
&= \frac{1}{nh^2}\left(\mathbb{E}\left[Y_1K^2\left(\frac{x-X_1}{h}\right)\right] - \left(\mathbb{E}\left[Y_1K\left(\frac{x-X_1}{h}\right)\right]\mathbb{E}\left[K\left(\frac{x-X_1}{h}\right)\right]\right)\right)
\end{aligned}$$

En vertu du calcul précédent, (voir le Lemme 2.2.1), on peut conclure.

$$\frac{1}{h^2}\left(\mathbb{E}\left[Y_1K\left(\frac{x-X_1}{h}\right)\right]\mathbb{E}\left[K\left(\frac{x-X_1}{h}\right)\right]\right) = O(1)$$

Ainsi,

$$\begin{aligned}
Cov(\hat{g}(x),\hat{f}(x)) &= \frac{1}{nh^2}\mathbb{E}\left[K^2\left(\frac{x-X_1}{h}\right)\mathbb{E}[Y_1|X=X_1]\right] + o\left(\frac{1}{nh}\right)\\
&= \frac{1}{nh^2}\int K^2\left(\frac{x-z}{h}\right) g(z)dz + o\left(\frac{1}{nh}\right)
\end{aligned}$$

En effectuant le changement de variable $t = \dfrac{x-z}{h}$ et la continuité de la fonction $g(x)$ ceci permet de conclure que

$$Cov(\hat{g}(x),\hat{f}(x)) = \frac{1}{nh}g(x)\int K^2(t)dt + o\left(\frac{1}{nh}\right) \qquad \blacksquare$$

En vue d'une application à des cas où les observations ne sont plus indépendantes, nous nous sommes intéressés dans la section suivante à l'extension du Théorème 2.2.1 au cas où les données sont $\alpha-$mélangeantes.

2.3 Cas dépendant

L'estimateur à noyau de la régression dans cette section, est construit à l'aide d'un échantillon de variables fortement mélangeantes. Pour cela, on considère le modèle (2.1) de variable réponse bornée, On garde les mêmes notations, ainsi, que les mêmes hypothèses de la section précédente et on ajoute les conditions suivantes

2.3.1 Hypothèses

(H6) La suite $(X_i, Y_i), i = 1, \ldots, n$ est α mélangeante.

(H7) Le couple de variables aléatoires (X_i, Y_j) admet une densité notée f_{ij} pour tout $i \neq j$.

(H8) Il existe deux constantes $C \in \mathbb{R}^{*+}$ et $a \in \mathbb{R}^{*+}$ telles que :

$$\alpha(n) \leq cn^{-a}$$

2.3.2 Propriété asymptotique

la vitesse de convergence est donnée par le théorème suivant :

Theorem 2.3.2.1. *Sous les hypothèses (H1)-(H7) et si la condition de décroissance algébrique (H8) est satisfaite pour une valeur de a vérifiant, relativement à un paramètre de lissage h, la condition suivante :*

$$h < c_1 n^{\frac{1}{1-a}}$$

Alors,

$$\mathbb{E}[\hat{r}(x) - r(x)]^2 = B^2(x)h^4 + V(x)\frac{1}{nh} + o(h^4 + \frac{1}{nh})$$

Démonstration du Théorème 2.3.2.1

La démonstration de ce théorème est de la même démarche que la démonstration du théorème précédent, c'est-à-dire on démontre les trois lemmes suivants, avec la partie biais reste la même que dans le cas i.i.d.

Lemme 2.3.1. *Sous les conditions du Théorème 2.3.2.1 on a :*

$$Var[\hat{f}(x)] = \frac{1}{nh} f(x) \int K^2(t)dt + o\left(\frac{1}{nh}\right)$$

Lemme 2.3.2. *Sous les conditions du Théorème 2.3.2.1 on a :*

$$Var[\hat{g}(x)] = \frac{1}{nh} f(x)\phi(x) \int K^2(t)dt + o\left(\frac{1}{nh}\right)$$

Lemme 2.3.3. *Sous les conditions du Théorème 2.3.2.1 on a :*

$$Cov(\hat{g}(x), \hat{f}(x)) = \frac{1}{nh} g(x) \int K^2(t)dt + o\left(\frac{1}{nh}\right)$$

2.3.3 Démonstrations des lemmes techniques

Démonstration du lemme 2.3.1 : On a :

$$Var\left[\hat{f}(x)\right] = \frac{1}{n^2h^2}\sum_{i,j} Cov(\Gamma_i, \Gamma_j)$$

Avec

$$\Gamma_i = \left[K\left(\frac{x - X_i}{h}\right) - \mathbb{E}K\left(\frac{x - X_i}{h}\right)\right]$$

Ce qui implique

$$Var\left[\hat{f}(x)\right] = \frac{1}{n^2h^2}\sum_{i\neq j} Cov(\Gamma_i, \Gamma_j) + \frac{1}{nh^2}Var[\Gamma_1] \tag{2.3}$$

Pour le deuxième terme de (2.3), nous nous sommes inspirés des idées du cas i.i.d, pour montrer que

$$\begin{aligned} \frac{1}{nh^2}Var[\Gamma_1] &= \frac{1}{nh^2}\mathbb{E}\left[K^2\left(\frac{x - X_1}{h}\right)\right] + o\left(\frac{1}{nh}\right) \\ &= \frac{1}{nh^2}\int K^2\left(\frac{x - z}{h}\right) f(z)dz + o\left(\frac{1}{nh}\right). \end{aligned}$$

On effectue Le changement de variable $t = \dfrac{x - z}{h}$ et on utilise la continuité de la fonction $f(x)$ pour obtenir

$$\frac{1}{nh^2}Var[\Gamma_1] = \frac{1}{nh}f(x)\int K^2(t)dt + o\left(\frac{1}{nh}\right)$$

Pour ce qui concerne le premier terme de (2.3), on a

$$S_n^{2*} = \frac{1}{n^2h^2}\sum_{i\neq j} Cov(\Gamma_i, \Gamma_j).$$

En effet, on a

$$Cov(\Gamma_i, \Gamma_j) = \mathbb{E}(\Gamma_i\Gamma_j)$$

Car les Γ_i sont centrées.

$$\begin{aligned}
\mathbb{E}(\Gamma_i\Gamma_j) &= \mathbb{E}\left[\left(K(\frac{x-X_i}{h}) - \mathbb{E}K(\frac{x-X_i}{h})\right)\left(K(\frac{x-X_j}{h}) - \mathbb{E}K(\frac{x-X_j}{h})\right)\right] \\
&= \mathbb{E}\left[K\left(\frac{x-X_i}{h})\right) K\left(\frac{x-X_j}{h}\right)\right] \\
&- \mathbb{E}\left[K\left(\frac{x-X_i}{h}\right)\right]\mathbb{E}\left[K\left(\frac{x-X_j}{h}\right)\right] \\
&= \int\int K\left(\frac{x-u}{h}\right) K\left(\frac{x-v}{h}\right) f_{ij}(u,v)dudv \\
&- \int\int K\left(\frac{x-u}{h}\right) K\left(\frac{x-v}{h}\right) f(u)f(v)dudv \\
&= \int\int K\left(\frac{x-u}{h}\right) K\left(\frac{x-v}{h}\right) [f_{ij}(u,v) - f(u)f(v)]dudv.
\end{aligned}$$

On prend les changements des variables $z = \dfrac{x-u}{h}, t = \dfrac{x-v}{h}$, pour avoir

$$|\mathbb{E}(\Gamma_i\Gamma_j)| \leq ch^2 \int\int |K(z)K(t)[f_{ij}(x-zh, x-th) - f(x-zh)f(x-th)]|dzdt.$$

Pour n assez grand, on peut trouver une constante C telle que

$$|\mathbb{E}(\Gamma_i\Gamma_j)| \leq ch^2[f_{ij}(x,x) - f^2(x)] \int\int |K(z)K(t)|dzdt.$$

Puisque les fonctions K, f et f_{ij} sont bornées, alors :

$$|\mathbb{E}(\Gamma_i\Gamma_j)| = O(h^2). \tag{2.4}$$

D'un autre côté, on peut aussi majorer cette covariance directement à partir de l'inégalité du Lemme 1.2.2, pour arriver à

$$|Cov(\Gamma_i\Gamma_j)| \leq C\alpha(|i-j|). \tag{2.5}$$

L'idée de la preuve consiste à introduire une suite u_n, et d'utiliser la borne (2.4) quand i est proche de j, et la borne (2.5) quand i et j sont éloignés ;

on arrive ainsi à

$$\begin{aligned} S_n^{2*} &\leq c\left[\sum_{0\leq|i-j|\leq u_n} h^2 + \sum_{|i-j|>u_n} \alpha(|i-j|)\right] \\ &= O(h^2 n u_n + n^2\alpha(u_n)) \end{aligned}$$

où u_n est une suite de nombres réels tendant vers ∞. Dans notre cas, on va prendre $u_n = \left[\frac{1}{h\log n}\right]$; ceci implique

$$\begin{aligned} S_n^{2*} &= O\left(nh^2\frac{1}{h\log n} + n^2\alpha\left(\frac{1}{h\log n}\right)\right) \\ &= O\left(\frac{nh}{\log n}\right) + O\left(n^2\alpha(\tfrac{1}{h\log n})\right). \end{aligned}$$

Il est évident que $\lim\limits_{n\longrightarrow\infty} = \dfrac{\frac{nh}{\log n}}{nh} = 0$ c'est-à-dire

$$S_n^{2*} = o(nh) + O(n^2\alpha(\frac{1}{h\log n}))$$

Pour achever la démonstration de ce lemme, il suffit de montrer que $O(n^2(h\log n)^a) = o(nh)$. En effet, on a

$$\begin{aligned} \frac{n^2(h\log n)^a}{nh} &= nh^{(a-1)}(\log n)^a \\ &= nn^{-1-\epsilon}(\log n)^a \\ &= \frac{(\log n)^a}{n^\epsilon} \longrightarrow 0 \ \text{ quand } \ n \longrightarrow \infty. \end{aligned}$$

Ainsi

$$Var[\hat{f}(x)] = \frac{1}{nh}f(x)\int K^2(t)dt + o\left(\frac{1}{nh}\right). \qquad \blacksquare$$

Démonstration du lemme 2.3.2 : La démonstration est similaire à celle du lemme précédent, il suffit d'écrire

$$Var\,[\hat{g}(x)] = \frac{1}{n^2h^2}\sum_{i,j} Cov(\Delta_i, \Delta_j)$$

Avec

$$\Delta_i = \left[Y_i K\left(\frac{x-X_i}{h}\right) - \mathbb{E}Y_i K\left(\frac{x-X_i}{h}\right)\right]$$

D'où

$$Var\left[\hat{g}(x)\right] = \frac{1}{n^2h^2}\sum_{i\neq j} Cov(\Delta_i,\Delta_j) + \frac{1}{nh^2}Var[\Delta_1]$$

Le calcul du cas i.i.d, montre que

$$\frac{1}{nh^2}Var[\Delta_1] = \frac{1}{nh}f(x)\phi(x)\int K^2(t)dt + o\left(\frac{1}{nh}\right).$$

Et par analogie du calcul précédent, on montre que

$$\frac{1}{n^2h^2}\sum_{i\neq j} Cov(\Delta_i,\Delta_j) = o\left(\frac{1}{nh}\right).$$

Ainsi

$$Var[\hat{g}(x)] = \frac{1}{nh}f(x)\phi(x)\int K^2(t)dt + o\left(\frac{1}{nh}\right). \qquad \blacksquare$$

Démonstration du Lemme 2.3.3 : De la même manière, on peut écrire

$$Cov\left(\hat{g}(x),\hat{f}(x)\right) = \frac{1}{n^2h^2}\sum_{i,j} Cov(\Delta_i,\Gamma_j)$$

Avec

$$\Delta_i = \left[Y_i K\left(\frac{x-X_i}{h}\right) - \mathbb{E}Y_i K\left(\frac{x-X_i}{h}\right)\right]$$

Et

$$\Gamma_j = \left[K\left(\frac{x-X_j}{h}\right) - \mathbb{E}K\left(\frac{x-X_j}{h}\right)\right]$$

Alors

$$Cov\left(\hat{g}(x),\hat{f}(x)\right) = \frac{1}{n^2h^2}\sum_{i\neq j} Cov(\Delta_i,\Gamma_j) + \frac{1}{nh^2}Cov(\Delta_1,\Gamma_1) \tag{2.6}$$

En reprenant les mêmes arguments du lemme précédent, on montre que le premier terme de (2.6) est de type $o\left(\frac{1}{nh}\right)$. En effet,

$$\begin{aligned}
\mathbb{E}(\Delta_i\Gamma_j) &= \mathbb{E}\left[\left(Y_iK(\frac{x-X_i}{h}) - \mathbb{E}Y_iK(\frac{x-X_i}{h})\right)\left(K(\frac{x-X_j}{h}) - \mathbb{E}K(\frac{x-X_j}{h})\right)\right] \\
&\leq C\mathbb{E}\left[K(\frac{x-X_i}{h})K(\frac{x-X_j}{h})\right] \\
&- C\mathbb{E}\left[K(\frac{x-X_i}{h})\right]\mathbb{E}\left[K(\frac{x-X_j}{h})\right] \\
&= C\int\int K\left(\frac{x-u}{h}\right)K\left(\frac{x-v}{h}\right)f_{ij}(u,v)dudv \\
&- C\int\int K\left(\frac{x-u}{h}\right)K\left(\frac{x-v}{h}\right)f(u)f(v)dudv \\
&= C\int\int K\left(\frac{x-u}{h}\right)K\left(\frac{x-v}{h}\right)[f_{ij}(u,v) - f(u)f(v)]dudv.
\end{aligned}$$

Ce qui nous permet d'écrire

$$|\mathbb{E}(\Delta_i\Gamma_j)| = o(h^2)$$

L'utilisation de l'inégalité de covariance, et en considérant la même suite du lemme précédent, permettent de conclure

$$S_n^{2*} = o(nh) + O\left(n^2\alpha(\frac{1}{h\log n})\right) = o(nh).$$

D'autre part, en utilisant le résultat du cas i.i.d pour le deuxième terme de (2.6), on montre que

$$\frac{1}{nh^2}Cov(\Delta_1,\Gamma_1) = \frac{1}{nh}g(x)\int K^2(t)dt + o\left(\frac{1}{nh}\right).$$

Finalement, en combinant ces deux résultats, on obtient

$$Cov\left(\hat{g}(x),\hat{f}(x)\right) = \frac{1}{nh}g(x)\int K^2(t)dt + o\left(\frac{1}{nh}\right). \qquad \blacksquare$$

Chapitre 3

La covergence en moyenne quadratique : cas vectoriel

Dans ce chapitre, nous généralisons les résultats déja obtenus dans le chapitre précédent à une variable explicative vectorielle. Ce chapitre est également divisé en trois sections. Nous présentons dans la première section un estimateur à noyau pour la régression vectorielle puis nous étudions l'erreur quadratique asymptotiquement exacte dans le cas où les observations sont indépendantes et identiquement distribuées. On abordera le cas dépendant dans la triosième section.

3.1 Le modèle et son estimateur

Soit (X, Y) un couple de variables aléatiores à valeurs dans $\mathbb{R}^p \times \mathbb{R}$, pour tout $(x_1, \ldots, x_p) \in \mathbb{R}^p$, on estime la régression vectorielle $r(x_1, \ldots, x_p) = \mathbb{E}[Y|X = (x_1, \ldots, x_p)]$ par :

$$\hat{r}(x) = \frac{\displaystyle\sum_{i=1}^{n} Y_i K\left(\frac{x_1 - X_i^1}{h}, \ldots, \frac{x_p - X_i^p}{h}\right)}{\displaystyle\sum_{i=1}^{n} K\left(\frac{x_1 - X_i^1}{h}, \ldots, \frac{x_p - X_i^p}{h}\right)} \tag{3.1}$$

Où K est une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$, h est un paramètre réel strictement positif et $((X_1, Y_1), \ldots, (X_n, Y_n))$ une suite des observations de (X, Y).

Cet estimateur est une généralisation à la dimension p de l'estimateur (2.2)
Où

$$\hat{g}(x) = \frac{1}{nh^p}\sum_{i=1}^{n} Y_i K\left(\frac{x_1 - X_i^1}{h}, \dots, \frac{x_p - X_i^p}{h}\right)$$

$$\hat{f}(x) = \frac{1}{nh^p}\sum_{i=1}^{n} K\left(\frac{x_1 - X_i^1}{h}, \dots, \frac{x_p - X_i^p}{h}\right)$$

3.2 Cas des observations indépendantes et de même loi

Dans cette section, nous supposons que les observations $(X_1, Y_1), \dots, (X_n, Y_n)$ sont indépendantes et nous allons étudier avec précision l'erreur quadratique de l'estimateur (3.1) au point fixé $(x_1, \dots, x_p) \in \mathbb{R}^p$.
On intrduit les conditions suivantes :

3.2.1 Hypothèses

(M1) Les fonctions r et f sont 2-fois continûment différentiables au voisinage de $(x_1, \dots, x_p)$.

(M2) La densité f de la variable explicative est strictement positive au point $(x_1, \dots, x_p)$.

(M3) Le paramètre de lissage est tel que :

$$\lim_{n \longrightarrow \infty} h = 0 \qquad et \qquad \lim_{n \longrightarrow \infty} nh^p = \infty$$

(M4) Le noyau K est d'ordre 2, borné et intégrable.

3.2.2 Propriété asymptotique

On établi le résultat suivant :

Théorème 3.2.1. *Sous les hypothèses (M1)-(M4) et (H5), on a :*

$$\mathbb{E}[\hat{r}(x_1, \dots, x_p) - r(x_1, \dots, x_p)]^2 = B^2(x_1, \dots, x_p)h^4 + V(x_1, \dots, x_p)\frac{1}{nh^p} + o(h^4 + \frac{1}{nh^p})$$

$$Où \quad B(x_1,\ldots,x_p) = \frac{\displaystyle\sum_{j=1}^{p} T_j(K)\left(\frac{\partial^2 g}{\partial x_j^2}(x_1,\ldots,x_p) - r(x_1,\ldots,x_p)\frac{\partial^2 f}{\partial x_j^2}(x_1,\ldots,x_p)\right)}{2f(x_1,\ldots,x_p)}$$

$$avec \qquad T_j(K) = \int_{\mathbb{R}^p} u_j^k K(u_1,\ldots,u_p)du_1,\ldots,du_p$$

$$et \qquad V(x) = \frac{\phi(x_1,\ldots,x_p) - r^2(x_1,\ldots,x_p)}{f(x_1,\ldots,x_p)} \int_{\mathbb{R}^p} K^2(t)dt$$

Démonstration du Théorème 3.2.1

En reprenant les mêmes arguments du cas réel, dont on calcul séparément la partie biais et la partie dispersion. Pour les deux parties, on considère le développement usuel de $\frac{1}{z}$ et on obtient

$$\begin{aligned}
\hat{r}(x_1,\ldots,x_p) - r(x_1,\ldots,x_p) &= \left(\frac{\hat{g}(x_1,\ldots,x_p)}{\mathbb{E}\hat{f}(x_1,\ldots,x_p)} - r(x_1,\ldots,x_p)\right) \\
&-\frac{(\hat{g}(x_1,\ldots,x_p) - \mathbb{E}\hat{g}(x_1,\ldots,x_p))}{(\mathbb{E}\hat{f}(x_1,\ldots,x_p))^2}\left(\hat{f}(x_1,\ldots,x_p) - \mathbb{E}\hat{f}(x_1,\ldots,x_p)\right) \\
&-\frac{\mathbb{E}\hat{g}(x_1,\ldots,x_p)}{(\mathbb{E}\hat{f}(x_1,\ldots,x_p))^2}\left(\hat{f}(x_1,\ldots,x_p) - \mathbb{E}\hat{f}(x_1,\ldots,x_p)\right) \\
&+\frac{\left(\hat{f}(x_1,\ldots,x_p) - \mathbb{E}\hat{f}(x_1,\ldots,x_p)\right)^2}{(\mathbb{E}\hat{f}(x_1,\ldots,x_p))^2}\hat{r}(x_1,\ldots,x_p)
\end{aligned}$$

Ceci implique

$$\begin{aligned}
\mathbb{E}[\hat{r}(x_1,\ldots,x_p)] - r(x_1,\ldots,x_p) &= \left(\frac{\hat{g}(x_1,\ldots,x_p)}{\mathbb{E}\hat{f}(x_1,\ldots,x_p)} - r(x_1,\ldots,x_p)\right) \\
&-\frac{Cov\left(\hat{g}(x_1,\ldots,x_p), \hat{f}(x_1,\ldots,x_p)\right)}{(\mathbb{E}\hat{f}(x_1,\ldots,x_p))^2} \\
&+\frac{\mathbb{E}\left(\hat{f}(x_1,\ldots,x_p) - \mathbb{E}\hat{f}(x_1,\ldots,x_p)\right)^2}{(\mathbb{E}\hat{f}(x_1,\ldots,x_p))^2}\hat{r}(x_1,\ldots,x_p)
\end{aligned}$$

Comme la variable Y est bornée, on peut trouver une constante $C > 0$ telle que $\hat{r}(x_1, \ldots, x_p) \leq C$. D'où,

$$\begin{aligned}
\mathbb{E}[\hat{r}(x_1, \ldots, x_p)] - r(x_1, \ldots, x_p) &= \left(\frac{\hat{g}(x_1, \ldots, x_p)}{\mathbb{E}\hat{f}(x_1, \ldots, x_p)} - r(x_1, \ldots, x_p) \right) \\
&- \frac{Cov\left(\hat{g}(x_1, \ldots, x_p), \hat{f}(x_1, \ldots, x_p)\right)}{(\mathbb{E}\hat{f}(x_1, \ldots, x_p))^2} \\
&+ \frac{Var\left(\hat{f}(x_1, \ldots, x_p)\right)}{(\mathbb{E}\hat{f}(x_1, \ldots, x_p))^2} O(1)
\end{aligned}$$

La preuve de la partie variance, repose sur la même idée, c'est à dire une application du développement de $\frac{1}{z}$ pour $z = \frac{\hat{f}(x)}{\mathbb{E}\hat{f}(x)}$ et $p = 3$, on aura :

$$\begin{aligned}
Var(\hat{r}(x_1, \ldots, x_p)) &= \frac{Var(\hat{g}(x_1, \ldots, x_p))}{(\mathbb{E}\hat{f}(x_1, \ldots, x_p))^2} \\
&-4 \frac{\mathbb{E}\hat{g}(x_1, \ldots, x_p)\, Cov\left(\hat{g}(x_1, \ldots, x_p), \hat{f}(x_1, \ldots, x_p)\right)}{(\mathbb{E}\hat{f}(x_1, \ldots, x_p))^3} \\
&+3\, Var(\hat{f}(x_1, \ldots, x_p)) \frac{(\mathbb{E}\hat{g}(x_1, \ldots, x_p))^2}{(\mathbb{E}\hat{f}(x_1, \ldots, x_p))^4} + o(\frac{1}{nh})
\end{aligned}$$

Finalement, le Théorème 3.2.1 est une conséquence des lemmes suivants :

Lemme 3.2.1. *Sous les conditions du Théorème 3.2.1 on a :*

i) $\mathbb{E}[\hat{f}(x_1, \ldots, x_p)] = f(x_1, \ldots, x_p) + \frac{h^2}{2} \left(\sum_{j=1}^{p} T_j(K) \frac{\partial^2 f}{\partial x_j^2}(x_1, \ldots, x_p) \right) + o(h^2)$

ii) $\mathbb{E}[\hat{g}(x_1, \ldots, x_p)] = g(x_1, \ldots, x_p) + \frac{h^2}{2} \left(\sum_{j=1}^{p} T_j(K) \frac{\partial^2 g}{\partial x_j^2}(x_1, \ldots, x_p) \right) + o(h^2)$

Corollaire 3.2.1. *Sous les conditions du Théorème 3.2.1 on a :*

$$\frac{\mathbb{E}[\hat{g}(x_1, \ldots, x_p)]}{\mathbb{E}[\hat{f}(x_1, \ldots, x_p)]} - r(x_1, \ldots, x_p) = B(x_1, \ldots, x_p) h^2 + o(h^2)$$

Lemme 3.2.2. *Sous les hypothèses (M1)-(M2) et (M4) on a :*

$$Var[\hat{f}(x_1,\ldots,x_p)] = \frac{1}{nh^p} f(x_1,\ldots,x_p) \int_{\mathbb{R}^p} K^2(t)dt + o\left(\frac{1}{nh^p}\right)$$

Lemme 3.2.3. *Sous les conditions du Théorème 3.2.1 on a :*

$$Var[\hat{g}(x_1,\ldots,x_p)] = \frac{1}{nh^p} f(x_1,\ldots,x_p)\phi(x_1,\ldots,x_p) \int_{\mathbb{R}^p} K^2(t)dt + o\left(\frac{1}{nh^p}\right)$$

$$\textit{avec} \qquad \phi(x_1,\ldots,x_p) = \mathbb{E}[Y^2|X=(x_1,\ldots,x_p)]$$

Lemme 3.2.4. *Sous les conditions du Théorème 3.2.1 on a :*

$$Cov(\hat{g}(x_1,\ldots,x_p),\hat{f}(x_1,\ldots,x_p)) = \frac{1}{nh^p} g(x_1,\ldots,x_p) \int_{\mathbb{R}^p} K^2(t)dt + o\left(\frac{1}{nh^p}\right)$$

3.2.3 Démonstration des lemmes techniques

Démonstration du Lemme 3.2.1 on a :

$$\begin{aligned}
\mathbb{E}[\hat{f}(x_1,\ldots,x_p)] &= \mathbb{E}\left[\frac{1}{nh^p}\sum_{i=1}^{n} K\left(\frac{x_1-X_i^1}{h},\ldots,\frac{x_p-X_i^p}{h}\right)\right] \\
&= \frac{1}{nh^p}\mathbb{E}\left[\sum_{i=1}^{n} K\left(\frac{x_1-X_i^1}{h},\ldots,\frac{x_p-X_i^p}{h}\right)\right] \\
&= \frac{1}{h^p}\mathbb{E}\left[K\left(\frac{x_1-X_1^1}{h},\ldots,\frac{x_p-X_1^p}{h}\right)\right] \\
&= \frac{1}{h^p}\int_{\mathbb{R}^p} K\left(\frac{x_1-z_1}{h},\ldots,\frac{x_p-z_p}{h}\right) f(z_1,\ldots,z_p)dz_1,\ldots,dz_p
\end{aligned}$$

Le changement de variable $t_i = \dfrac{x_i - z_i}{h}$ pour $i = 1,\ldots,p$ donne

$$\mathbb{E}[\hat{f}(x_1,\ldots,x_p)] = \int_{\mathbb{R}^p} K(t_1,\ldots,t_p) f(x_1-ht_1,\ldots,x_p-ht_p)dt_1,\ldots,dt_p$$

On utilise le développement du Taylor pour la fonction f au voisinage du point x à l'ordre $k=2$, on obtient :

$$f(x_1-hz_1,\ldots,x_p-hz_p)-f(x_1,\ldots,x_p) = \sum_{j=1}^{2}\frac{(-h)^j}{j!}\sum_{i_1+\ldots+i_p=j}\frac{\partial^j f(x_1,\ldots,x_p)}{\partial x_1^{i_1}\ldots\partial x_p^{i_p}}(z_1^{i_1},\ldots,z_p^{i_p})+o(h^2)$$

D'où

$$\mathbb{E}[\hat{f}(x_1,\dots,x_p)] - f(x_1,\dots,x_p) = \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 f}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)$$

De même, pour la deuxième équation

$$\begin{aligned}
\mathbb{E}[\hat{g}(x_1,\dots,x_p)] &= \mathbb{E}\left[\frac{1}{nh^p}\sum_{i=1}^{n} Y_i K\left(\frac{x_1-X_i^1}{h},\dots,\frac{x_p-X_i^p}{h}\right)\right] \\
&= \frac{1}{nh^p}\mathbb{E}\left[\sum_{i=1}^{n} Y_i K\left(\frac{x_1-X_i^1}{h},\dots,\frac{x_p-X_i^p}{h}\right)\right] \\
&= \frac{1}{h^p}\mathbb{E}\left[Y_1 K\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right] \\
&= \frac{1}{h^p}\mathbb{E}\left[\mathbb{E}[Y_1|X=X_1] K\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right] \\
&= \frac{1}{h^p}\int_{\mathbb{R}^p} K\left(\frac{x_1-z_1}{h},\dots,\frac{x_p-z_p}{h}\right) f(z_1,\dots,z_p)r(z_1,\dots,z_p)dz_1,\dots,dz_p \\
&= \frac{1}{h^p}\int_{\mathbb{R}^p} K\left(\frac{x_1-z_1}{h},\dots,\frac{x_p-z_p}{h}\right) g(z_1,\dots,z_p)dz_1,\dots,dz_p
\end{aligned}$$

après un changement de variable $t_i = \dfrac{x_i - z_i}{h}$ pour $i = 1,\dots,p$, puis un développement de Taylor de la fonction g, on obtient

$$\mathbb{E}[\hat{g}(x_1,\dots,x_p)] - g(x_1,\dots,x_p) = \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 g}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2) \qquad \blacksquare$$

Démonstration du Corollaire 3.2.1 : On a

$$\frac{\mathbb{E}[\hat{g}(x_1,\dots,x_p)]}{\mathbb{E}[\hat{f}(x_1,\dots,x_p)]} - r(x_1,\dots,x_p) = \frac{g(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 g}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)}{f(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 f}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)} - r(x_1,\dots,x_p)$$

$$= \frac{g(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 g}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)}{f(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 f}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)}$$

$$- \frac{r(x_1,\dots,x_p)\left(f(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 f}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)\right)}{f(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 f}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)}$$

$$= \frac{g(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 g}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)}{f(x_1,\dots,x_p) + O(h^2)}$$

$$- \frac{r(x_1,\dots,x_p)\left(f(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 f}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K) + o(h^2)\right)}{f(x_1,\dots,x_p) + O(h^2)}$$

$$= \frac{g(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 g}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K)}{f(x_1,\dots,x_p)}$$

$$- \frac{r(x_1,\dots,x_p)\left(f(x_1,\dots,x_p) + \frac{h^2}{2!}\sum_{j=1}^{p}\left[\frac{\partial^2 f}{\partial x_j^2}\right](x_1,\dots,x_p)T_j(K)\right)}{f(x_1,\dots,x_p)} + O(h^2)$$

$$= \frac{h^2}{2f(x_1,\dots,x_p)}\sum_{j=1}^{p}T_j(K)\left(\left[\frac{\partial^2 g}{\partial x_j^2}\right](x_1,\dots,x_p) - \left[r(x_1,\dots,x_p)\frac{\partial^2 f}{\partial x_j^2}(x_1,\dots,x_p)\right]\right) + O(h^2) \quad \blacksquare$$

Démonstration du Lemme 3.2.2 on a

$$\begin{aligned}
Var\left[\hat{f}(x_1,\dots,x_p)\right] &= Var\left[\frac{1}{nh^p}\sum_{i=1}^{n} K\left(\frac{x_1-X_i^1}{h},\dots,\frac{x_p-X_i^p}{h}\right)\right] \\
&= \frac{1}{n^2h^{2p}} Var\left[\sum_{i=1}^{n} K\left(\frac{x_1-X_i^1}{h},\dots,\frac{x_p-X_i^p}{h}\right)\right] \\
&= \frac{1}{nh^{2p}}\left(\mathbb{E}\left[K^2\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right]\right) \\
&- \frac{1}{nh^{2p}}\left(\left(\mathbb{E}\left[K\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right]\right)^2\right)
\end{aligned}$$

D'aprés le Lemme 3.2.1 la partie i), on a

$$\frac{1}{h^{2p}}\left(\mathbb{E}\left[K\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right]\right)^2 = \left(\mathbb{E}\left[\hat{f}(x_1,\dots,x_p)\right]\right)^2 = O(1)$$

D'où

$$Var\left[\hat{f}(x_1,\dots,x_p)\right] = \frac{1}{nh^{2p}}\mathbb{E}\left[K^2\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right] + o\left(\frac{1}{nh^p}\right)$$

$$= \frac{1}{nh^{2p}}\int_{\mathbb{R}^p} K^2\left(\frac{x_1-z_1}{h},\dots,\frac{x_p-z_p}{h}\right) f(z_1,\dots,z_p)dz_1\dots dz_p + o\left(\frac{1}{nh^p}\right)$$

Le changement de variable $t_i = \dfrac{x_i - z_i}{h}$ pour $i = 1,\dots,p$ et si on utilise la continuité de la fonction $f(x_1,\dots,x_p)$, on aura

$$Var\left[\hat{f}(x_1,\dots,x_p)\right] = \frac{1}{nh^p} f(x_1,\dots,x_p)\int_{\mathbb{R}^p} K^2(t_1,\dots,t_p)dt_1,\dots,dt_p + o\left(\frac{1}{nh^p}\right) \quad \blacksquare$$

Démonstration du Lemme 3.2.3

$$\begin{aligned}
Var[\hat{g}(x_1,\ldots,x_p)] &= Var\left[\frac{1}{nh^p}\sum_{i=1}^{n} Y_i K\left(\frac{x_1-X_i^1}{h},\ldots,\frac{x_p-X_i^p}{h}\right)\right] \\
&= \frac{1}{n^2h^{2p}} Var\left[\sum_{i=1}^{n} Y_i K\left(\frac{x_1-X_i^1}{h},\ldots,\frac{x_p-X_i^p}{h}\right)\right] \\
&= \frac{1}{nh^{2p}}\left(\mathbb{E}\left[Y_1^2K^2\left(\frac{x_1-X_1^1}{h},\ldots,\frac{x_p-X_1^p}{h}\right)\right]\right) \\
&- \frac{1}{nh^{2p}}\left(\left(\mathbb{E}\left[Y_1K\left(\frac{x_1-X_1^1}{h},\ldots,\frac{x_p-X_1^p}{h}\right)\right]\right)^2\right)
\end{aligned}$$

D'après le Lemme 3.2.1 la partie ii), on a

$$\frac{1}{h^{2p}}\left(\mathbb{E}\left[Y_1K\left(\frac{x_1-X_1^1}{h},\ldots,\frac{x_p-X_1^p}{h}\right)\right]\right)^2 = \left(\mathbb{E}\left[\hat{g}(x_1,\ldots,x_p)\right]\right)^2 = O(1)$$

D'où

$$Var[\hat{g}(x_1,\ldots,x_p)] = \frac{1}{nh^{2p}}\mathbb{E}\left[\mathbb{E}[Y_1^2|X=X_1]K^2\left(\frac{x_1-X_1^1}{h},\ldots,\frac{x_p-X_1^p}{h}\right)\right] + o\left(\frac{1}{nh^p}\right)$$

$$= \frac{1}{nh^{2p}}\int_{\mathbb{R}^p} K^2\left(\frac{x_1-z_1}{h},\ldots,\frac{x_p-z_p}{h}\right) f(z_1,\ldots,z_p)\phi(z_1,\ldots,z_p)dz_1\ldots dz_p + o\left(\frac{1}{nh^p}\right)$$

Le changement de variable $t_i = \dfrac{x_i-z_i}{h}$ pour $i=1,\ldots,p$ et si on utilise la continuité de la fonction $f(x_1,\ldots,x_p)\phi(x_1,\ldots,x_p)$ pour arriver à

$$Var[\hat{g}(x_1,\ldots,x_p)] = \frac{1}{nh^p}f(x_1,\ldots,x_p)\phi(x_1,\ldots,x_p)\int_{\mathbb{R}^p} K^2(t_1,\ldots,t_p)dt_1,\ldots,dt_p + o\left(\frac{1}{nh^p}\right) \quad \blacksquare$$

Démonstration du Lemme 3.2.4

$$\begin{aligned}
& Cov(\hat{g}(x_1,\dots,x_p),\hat{f}(x_1,\dots,x_p)) \\
= \;& Cov\Bigg(\frac{1}{nh^p}\sum_{i=1}^{n} Y_iK\left(\frac{x_1-X_i^1}{h},\dots,\frac{x_p-X_i^p}{h}\right),\frac{1}{nh^p}\sum_{i=1}^{n} K\left(\frac{x_1-X_i^1}{h},\dots,\frac{x_p-X_i^p}{h}\right)\Bigg) \\
= \;& \frac{1}{n^2h^{2p}}Cov\Bigg(\sum_{i=1}^{n} Y_iK\left(\frac{x_1-X_i^1}{h},\dots,\frac{x_p-X_i^p}{h}\right),\sum_{i=1}^{n} K\left(\frac{x_1-X_i^1}{h},\dots,\frac{x_p-X_i^p}{h}\right)\Bigg) \\
= \;& \frac{1}{nh^{2p}}\Bigg(\mathbb{E}\left[Y_1K^2\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right] \\
- \;& \left(\mathbb{E}\left[Y_1K\left(\tfrac{x_1-X_1^1}{h},\dots,\tfrac{x_p-X_1^p}{h}\right)\right]\mathbb{E}\left[K\left(\tfrac{x_1-X_1^1}{h},\dots,\tfrac{x_p-X_1^p}{h}\right)\right]\right)\Bigg)
\end{aligned}$$

Comme,

$$\frac{1}{h^{2p}}\left(\mathbb{E}\left[Y_1K\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right]\mathbb{E}\left[K\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right]\right)=O(1)$$

On peut conclure :

$$\begin{aligned}
& Cov(\hat{g}(x_1,\dots,x_p),\hat{f}(x_1,\dots,x_p)) \\
= \;& \frac{1}{nh^{2p}}\mathbb{E}\left[\mathbb{E}[Y_1|X=X_1]K^2\left(\frac{x_1-X_1^1}{h},\dots,\frac{x_p-X_1^p}{h}\right)\right]+o\left(\frac{1}{nh^p}\right) \\
= \;& \frac{1}{nh^{2p}}\int_{\mathbb{R}^p}K^2\left(\frac{x_1-z_1}{h},\dots,\frac{x_p-z_p}{h}\right)g(z_1,\dots,z_p)dz_1\dots dz_p+o\left(\frac{1}{nh^p}\right)
\end{aligned}$$

On considère le changement de variable $t_i=\dfrac{x_i-z_i}{h}$ pour $i=1,\dots,p$ et comme la fonction $g(x_1,\dots,x_p)$ est continue, il découle

$$Cov(\hat{g}(x_1,\dots,x_p),\hat{f}(x_1,\dots,x_p))=\frac{1}{nh^p}g(x_1,\dots,x_p)\int_{\mathbb{R}^p}K^2(t_1,\dots,t_p)dt_1,\dots,dt_p+O\left(\frac{1}{nh^p}\right)\blacksquare$$

3.3 Cas dépendant

On considère dans cette section le cas où les observations sont α-mélangeantes, on explicite la vitesse de convergence en moyenne quadratique de l'estimateur à noyau de la régression. Pour cela, on garde les mêmes notations de la section précédente et on ajoute les hypothèses suivantes :

3.3.1 Hypothèses

(M5) La suite $(X_i, Y_i), i = 1, \dots, n$ est α mélangeante.

(M6) Le couple des variables aléatoires (X_i, Y_j) admet une densité notée f_{ij} pour tout $i \neq j$.

(M7) Il existe deux constantes $C \in \mathbb{R}^{*+}$ et $a \in \mathbb{R}^{*+}$ telles que

$$\alpha(n) \leq cn^{-a}$$

3.3.2 Propriété asymptotique

La vitesse de convergence est donnée par le théorème suivant :

Théorème 3.3.1. *Sous les hypothèses (M1)-(M6) et (H5) et si la condition de décroissance algébrique (M7) est satisfaite pour une valeur de a vérifiant, relativement à un paramètre de lissage h, la condition suivante :*

$$\exists \epsilon > 0 \qquad h^p < c_1 n^{\frac{1}{1-a} - \epsilon}$$

Alors,

$$\mathbb{E}[\hat{r}(x_1, \dots, x_p) - r(x_1, \dots, x_p)]^2 = B^2(x_1, \dots, x_p)h^4 + V(x_1, \dots, x_p)\frac{1}{nh^p} + o(h^4 + \frac{1}{nh^p})$$

Démonstration du Théorème 3.3.1

Il suffit de suivre le même cheminement que pour la démonstration du théorème précédent, à l'exeption de la partie dispersion qui découle des les lemmes suivants

Lemme 3.3.1. *Sous les conditions du Théorème 3.3.1 on a :*

$$Var[\hat{f}(x_1, \dots, x_p)] = \frac{1}{nh^p} f(x_1, \dots, x_p) \int_{\mathbb{R}^p} K^2(t)dt + o\left(\frac{1}{nh^p}\right)$$

Lemme 3.3.2. *Sous les conditions du Théorème 3.3.1 on a :*

$$Var[\hat{g}(x_1, \dots, x_p)] = \frac{1}{nh^p} f(x_1, \dots, x_p)\phi(x) \int_{\mathbb{R}^p} K^2(t)dt + o\left(\frac{1}{nh^p}\right)$$

Lemme 3.3.3. *Sous les conditions du Théorème 3.3.1 on a :*

$$Cov(\hat{g}(x_1,\ldots,x_p),\hat{f}(x_1,\ldots,x_p)) = \frac{1}{nh^p}g(x_1,\ldots,x_p)\int_{\mathbb{R}^p}K^2(t)dt + o\left(\frac{1}{nh^p}\right)$$

3.3.3 Démonstrations des lemmes techniques

Démonstration du Lemme 3.3.1

Pour simplifier la notation, on pose

$$K\left(\frac{x-X_i}{h}\right) = K\left(\frac{x_1-X_i^1}{h},\ldots,\frac{x_p-X_i^p}{h}\right)$$

On a

$$Var\left[\hat{f}(x)\right] = \frac{1}{n^2h^{2p}}\sum_{i,j}Cov(\Gamma_i,\Gamma_j)$$

Avec

$$\Gamma_i = \left[K\left(\frac{x-X_i}{h}\right) - \mathbb{E}K\left(\frac{x-X_i}{h}\right)\right]$$

Ce qui implique

$$Var\left[\hat{f}(x)\right] = \frac{1}{n^2h^{2p}}\sum_{i\neq j}Cov(\Gamma_i,\Gamma_j) + \frac{1}{nh^{2p}}Var[\Gamma_1]$$

En utilisant le calcul du cas i.i.d., on montre que,

$$\frac{1}{nh^{2p}}Var[\Gamma_1] = \frac{1}{nh^p}f(x_1,\ldots,x_p)\int_{\mathbb{R}^p}K^2(t_1,\ldots,t_p)dt_1,\ldots,dt_p + o\left(\frac{1}{nh^p}\right)$$

Il suffit donc de calculer la quantité

$$S_n^{2*} = \frac{1}{n^2h^{2p}}\sum_{i\neq j}Cov(\Gamma_i,\Gamma_j).$$

En effet, on peut écrire

$$S_n^{2*} = \sum\sum_{0\leq|i-j|\leq u_n}Cov(\Gamma_i\Gamma_j) + \sum\sum_{|i-j|>u_n}Cov(\Gamma_i\Gamma_j) \qquad (3.2)$$

Avec $\qquad u_n = \left[\dfrac{1}{h^p\log n}\right]$

Pour la première somme de l'équation (3.2), on utilise l'approximation suivante :

$$\begin{aligned}\mathbb{E}(\Gamma_i\Gamma_j) &= \mathbb{E}\left[\left(K(\tfrac{x-X_i}{h}) - \mathbb{E}K(\tfrac{x-X_i}{h})\right)\left(K(\tfrac{x-X_j}{h}) - \mathbb{E}K(\tfrac{x-X_j}{h})\right)\right] \\ &= \mathbb{E}\left[K\left(\tfrac{x-X_i}{h}\right)\right) K\left(\tfrac{x-X_j}{h}\right)\Big] - \mathbb{E}\left[K\left(\tfrac{x-X_i}{h}\right)\right]\mathbb{E}\left[K\left(\tfrac{x-X_j}{h}\right)\right] \\ &= \int_{\mathbb{R}^p}\int_{\mathbb{R}^p} K\left(\frac{x-u}{h}\right) K\left(\frac{x-v}{h}\right) f_{ij}(u_1\dots u_p, v_1\dots v_p)du_1\dots du_p dv_1\dots dv_p \\ &- \int_{\mathbb{R}^p}\int_{\mathbb{R}^p} K\left(\frac{x-u}{h}\right) K\left(\frac{x-v}{h}\right) f(u_1\dots u_p)f(v_1\dots v_p)du_1\dots du_p dv_1\dots dv_p \\ &= \int_{\mathbb{R}^p}\int_{\mathbb{R}^p} K\left(\frac{x-u}{h}\right) K\left(\frac{x-v}{h}\right)[f_{ij}(u_1\dots u_p, v_1\dots v_p) \\ &-f(u_1\dots u_p)f(v_1\dots v_p)]du_1\dots du_p dv_1\dots dv_p.\end{aligned}$$

On prend le changement des variables $z = \dfrac{x-u}{h}, t = \dfrac{x-v}{h}$, d'où,

$$|\mathbb{E}(\Gamma_i\Gamma_j)| \leq ch^{2p}\int_{\mathbb{R}^p}\int_{\mathbb{R}^p} |K(z)K(t)[f_{ij}(x-zh, x-th) - f(x-zh)f(x-th)]|dzdt.$$

Puisque les fonctions K, f et f_{ij} sont bornées, alors :

$$|\mathbb{E}(\Gamma_i\Gamma_j)| = O(h^{2p}).$$

D'autre part, on peut aussi majorer cette covariance pour la deuxième somme de (3.2) en utilisant l'inégalité de covariance

$$|Cov(\Gamma_i, \Gamma_j)| \leq 4\|\Delta_i\|_\infty\|\Delta_j\|_\infty\alpha(|i-j|)$$

$$\begin{aligned}S_n^{2*} &= O\left(nh^{2p}\frac{1}{h^p\log n} + n^2\alpha\left(\frac{1}{h^p\log n}\right)\right) \\ &= O\left(\frac{nh^p}{\log n}\right) + O\left(n^2\alpha(\tfrac{1}{h^p\log n})\right).\end{aligned}$$

Il est évident que $\lim\limits_{n\longrightarrow\infty} = \dfrac{\dfrac{nh^p}{\log n}}{nh^p} = 0$ c'est-à-dire

$$S_n^{2*} = o(nh^p) + O(n^2\alpha(\frac{1}{h^p \log n}))$$

Pour achever la démonstration de ce lemme il suffit de montrer que $O(n^2(h\log n)^a) = o(nh)$. En effet, on a

$$\begin{aligned}
\frac{n^2(h^p\log n)^a}{nh^p} &= nh^{p(a-1)}(\log n)^a \\
&= nn^{-1-\epsilon}(\log n)^a \\
&= \frac{(\log n)^a}{n^\epsilon} \longrightarrow 0 \text{ quand } n \longrightarrow \infty.
\end{aligned}$$

Ainsi

$$Var[\hat{f}(x_1,\ldots,x_p)] = \frac{1}{nh^p} f(x_1,\ldots,x_p)\int_{\mathbb{R}^p} K^2(t)dt + o\left(\frac{1}{nh^p}\right) \quad \blacksquare$$

Démonstration du Lemme 3.3.2

On utilise la même démarche que lors de la preuve du lemme précédent, il suffit d'écrire

$$Var\left[\hat{g}(x_1,\ldots,x_p)\right] = \frac{1}{n^2h^{2p}}\sum_{i,j} Cov(\Delta_i,\Delta_j)$$

Avec

$$\Delta_i = \left[Y_iK\left(\frac{x-X_i}{h}\right) - \mathbb{E}Y_iK\left(\frac{x-X_i}{h}\right)\right]$$

D'où

$$Var\left[\hat{g}(x_1,\ldots,x_p)\right] = \frac{1}{n^2h^{2p}}\sum_{i\neq j} Cov(\Delta_i,\Delta_j) + \frac{1}{nh^{2p}} Var[\Delta_1]$$

Le calcul du cas i.i.d., montre que

$$\frac{1}{nh^{2p}} Var[\Delta_1] = \frac{1}{nh^p} f(x_1,\ldots,x_p)\phi(x_1,\ldots,x_p)\int_{\mathbb{R}^p} K^2(t_1\ldots t_p)dt_1\ldots dt_p + o\left(\frac{1}{nh^p}\right).$$

Et par analogie du calcul précédent, on montre que

$$\frac{1}{n^2h^2}\sum_{i\neq j} Cov(\Delta_i,\Delta_j)=o\left(\frac{1}{nh^p}\right).$$

Ainsi

$$Var[\hat{g}(x_1,\ldots,x_p)]=\frac{1}{nh^p}f(x_1,\ldots,x_p)\phi(x_1,\ldots,x_p)\int_{\mathbb{R}^p}K^2(t_1\ldots t_p)dt_1\ldots dt_p+o\left(\frac{1}{nh^p}\right). \quad \blacksquare$$

Démonstration du Lemme 3.3.3
De la même manière, on peut écrire

$$Cov\left(\hat{g}(x_1,\ldots,x_p),\hat{f}(x_1,\ldots,x_p)\right)=\frac{1}{n^2h^{2p}}\sum_{i,j}Cov(\Delta_i,\Gamma_j)$$

Avec

$$\Delta_i=\left[Y_iK\left(\frac{x-X_i}{h}\right)-\mathbb{E}Y_iK\left(\frac{x-X_i}{h}\right)\right]$$

Et

$$\Gamma_j=\left[K\left(\frac{x-X_j}{h}\right)-\mathbb{E}K\left(\frac{x-X_j}{h}\right)\right]$$

Alors

$$Cov\left(\hat{g}(x_1,\ldots,x_p),\hat{f}(x_1,\ldots,x_p)\right)=\frac{1}{n^2h^{2p}}\sum_{i\neq j}Cov(\Delta_i,\Gamma_j)+\frac{1}{nh^{2p}}Cov(\Delta_1,\Gamma_1)$$

En reprenant les même arguments du lemme précédent, on montre que le premier terme est de type $o\left(\frac{1}{nh^p}\right)$. En effet,

$$\begin{aligned}
\mathbb{E}(\Delta_i\Gamma_j) &= \mathbb{E}\left[\left(Y_iK(\frac{x-X_i}{h}) - \mathbb{E}Y_iK(\frac{x-X_i}{h})\right)\left(K(\frac{x-X_j}{h}) - \mathbb{E}K(\frac{x-X_j}{h})\right)\right] \\
&\leq c\mathbb{E}\left[K(\frac{x-X_i}{h})K(\frac{x-X_j}{h})\right] - c\mathbb{E}\left[K(\frac{x-X_i}{h})\right]\mathbb{E}\left[K(\frac{x-X_j}{h})\right] \\
&= c\int_{\mathbb{R}^p}\int_{\mathbb{R}^p} K\left(\frac{x-u}{h}\right)K\left(\frac{x-v}{h}\right) f_{ij}(u_1\dots u_p, v_1\dots v_p)du_1\dots du_p dv_1\dots dv_p \\
&- c\int_{\mathbb{R}^p}\int_{\mathbb{R}^p} K\left(\frac{x-u}{h}\right)K\left(\frac{x-v}{h}\right) f(u_1\dots u_p)f(v_1\dots v_p)du_1\dots du_p dv_1\dots dv_p \\
&= c\int\int K\left(\frac{x-u}{h}\right)K\left(\frac{x-v}{h}\right)[f_{ij}(u_1\dots u_p, v_1\dots v_p) \\
&- f(u_1\dots u_p)f(v_1\dots v_p)]du_1\dots du_p dv_1\dots dv_p
\end{aligned}$$

Ce qui nous permet d'écrire

$$|\mathbb{E}(\Delta_i\Gamma_j)| = o(h^{2p})$$

L'utilisation de l'inégalité de covariance, et en considérant la même suite du lemme précédent nous permet de conclure

$$S_n^{2*} = o(nh^p) + O\left(n^2\alpha(\frac{1}{h^p\log n})\right) = o(nh^p).$$

D'autre part, en utilisant le résultat du cas i.i.d. pour le deuxième terme, on montre que

$$\frac{1}{nh^{2p}}Cov(\Delta_1, \Gamma_1) = \frac{1}{nh^p}g(x_1,\dots,x_p)\int_{\mathbb{R}^p} K^2(t)dt + o\left(\frac{1}{nh^p}\right).$$

Finalement, en combinant ces deux résultats, on obtient

$$Cov\left(\hat{g}(x), \hat{f}(x)\right) = \frac{1}{nh^p}g(x_1,\dots,x_p)\int_{\mathbb{R}^p} K^2(t)dt + o\left(\frac{1}{nh^p}\right). \qquad \blacksquare$$

Chapitre 4

La covergence en moyenne quadratique : cas fonctionnel

Ce chapitre est consacré à l'étude de modèle du régression lorsque la variable explicative est à valeurs dans un espace vectoriel semi-normé. Nous partagons ce chapitre en trois sections. Nous présentons dans la première section le modèle ainsi que son estimateur. On s'intérèsse à un échantillon indépendant et identiquement distribué dans la deuxième section. Le cas dépendant est traité dans la dernière section.

4.1 Le modèle et son estimateur

Soit (X, Y) un couple de variables aléatoires à valeurs dans $\mathcal{F} \times \mathbb{R}$, où $\mathcal{F}$ est un espace vectoriel semi-normé, on fixe un point x dans $\mathcal{F} \times \mathbb{R}$ et on note $\mathcal{V}_x$ un voisinage de ce point. Le modèle de régression entre (X, Y) est écrit sous la forme :

$$Y = r(X) + \epsilon$$

Où ϵ est une variable aléatoire réelle centrée et indépendante de X, et r est supposée dans un espace fonctionnel approprié. Étant donné $(X_i, Y_i)_{i=1,\dots,n}$ des observations i.i.d. de (X, Y), l'estimateur à noyau de la fonction de régression r est défini par :

$$\hat{r}(x) = \frac{\sum_{i=1}^{n} Y_i K\left(\frac{\|x - X_i\|}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{\|x - X_i\|}{h}\right)}$$

Avec la convention $\frac{0}{0} = 0$. K est un noyau et h est une suite de nombres réels positifs.

4.2 Cas des observations indépendantes et de même loi

On suppose que les observations $(X_1, Y_1), \ldots, (X_n, Y_n)$ sont indépendantes et identiquement distribuées selon la loi du couple (X, Y). Afin d'étudier la convergence en moyenne quadratique de l'estimateur $\hat{r}(x)$ vers $r(x)$, on introduit les hypothèses suivantes.

4.2.1 Hypothèses

(H_1) On suppose que l'espace fonctionnel $\mathcal{F}$ est muni d'une mesure positive σ-finie notée μ, et que pour tout $r > 0$ la variable aléatoire $Z = \frac{(x - X)}{r}$ est absolument continue par rapport à cette mesure et que sa densité $g(r, x, v)$ est strictement positive sur $B(0, 1)$ et peut être écrite sous la forme :

$$g(r, x, v) = \phi(r)h(x, v) + o(\phi(r)) \text{ pour tout } v \in B(0, 1)$$

Où ϕ est une fonction différentiable, croissante à valeurs dans $\mathbb{R}^+$ et h définie sur $\mathcal{F} \times \mathcal{F}$ à valeurs dans $\mathbb{R}^+$, est telle que $0 < \int_{B(0,1)} h(x, v) d\mu(v) < \infty$.

(H_2) K est un noyau de support $[0, 1]$ tel que $0 < C_1 < K(t) < C_2 < \infty$.

(H_3) $\lim_{n \longrightarrow \infty} h = 0$, et $\lim_{n \longrightarrow \infty} n\phi_x(h) = \infty$.

4.2.2 Propriété asymptotique

Pour simplifier les notations, on pose $K_i = K\left(\frac{\|x - X_i\|}{h}\right)$, pour tout $i = 1, \ldots, n$ et on pose

$$\begin{aligned}
\hat{g}(x) &= \frac{1}{n\phi_x(h)} \sum_{i=1}^{n} Y_i K_i \\
\hat{f}(x) &= \frac{1}{n\phi_x(h)} \sum_{i=1}^{n} K_i
\end{aligned}$$

On établit le résultat suivant

Théorème 4.2.1. *Sous les hypothèses (H_1)-(H_3), on a :*

$$\mathbb{E}[\hat{r}(x) - r(x)]^2 = B^2(x)h^2 + V(x)\frac{1}{n\phi_x(h)} + o\left(h^2 + \frac{1}{n\phi_x(h)}\right)$$

Où

$$B(x) = \frac{\displaystyle\int_{B(0,1)} K(\|v\|) r'(x)[v] h(x,v) du(v)}{\displaystyle\int_{B(0,1)} K(\|v\|) h(x,v) du(v)}$$

Et

$$V(x) = \frac{(\Phi(x) - r^2(x)) \displaystyle\int_{B(0,1)} K^2(\|v\|) h(x,v) du(v)}{\left(\displaystyle\int_{B(0,1)} K(\|v\|) h(x,v) du(v)\right)^2}$$

Démonstration du Théorème 4.2.1

La démonstration est basée sur l'évaluation séparée des deux parties : partie biais et partie variance. Pour la partie déterministe, on utilise la décomposition suivante :

$$\mathbb{E}(\hat{r}(x)) = \frac{\mathbb{E}\hat{g}(x)}{\mathbb{E}\hat{f}(x)} + \frac{A_1}{(\mathbb{E}\hat{f}(x))^2} + \frac{A_2}{(\mathbb{E}\hat{f}(x))^2}$$

avec

$$A_1 = \mathbb{E}[\hat{g}(x)(\hat{f}(x) - \mathbb{E}\hat{f}(x))]$$

et

$$A_2 = \mathbb{E}[(\hat{f}(x) - \mathbb{E}\hat{f}(x))^2 \hat{r}(x)]$$

De même pour la partie dispersion nous avons,

$$Var(\hat{r}(x)) = \frac{Var(\hat{g}(x))}{(\mathbb{E}\hat{f}(x))^2} - 4\frac{\mathbb{E}\hat{g}(x)\,Cov(\hat{g}(x),\hat{f}(x))}{(\mathbb{E}\hat{f}(x))^3} + 3\,Var(\hat{f}(x))\frac{(\mathbb{E}\hat{g}(x))^2}{(\mathbb{E}\hat{f}(x))^4} + o(\frac{1}{nh})$$

Ainsi, la démonstration du Théorème 4.2.1 repose sur les lemmes suivants :

Lemme 4.2.1. *Sous les conditions du Théorème 4.2.1 on a :*

i) $\mathbb{E}\left[\hat{f}(x)\right] = \int_{B(0,1)} K(\|v\|)h(x,v)du(v) + o(1)$

ii) $\mathbb{E}\left[\hat{g}(x)\right] = r(x)\int_{B(0,1)} K(\|v\|)h(x,v)du(v) - h\int_{B(0,1)} K(\|v\|)r'(x)[v]h(x,v)du[v] + o(h)$

Corollaire 4.2.1. *Sous les conditions du Théorème 4.2.1 on a :*

$$\frac{\mathbb{E}[\hat{g}(x)]}{\mathbb{E}[\hat{f}(x)]} - r(x) = B(x)h + o(h)$$

Lemme 4.2.2. *Sous les conditions* (H_1),(H_2) *et* (H_3) *on a :*

$$Var\left[\hat{f}(x)\right] = \frac{1}{n\phi_x(h)}\int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right)$$

Lemme 4.2.3. *Sous les conditions du Théorème 4.2.1 on a :*

$$Var\left[\hat{g}(x)\right] = \frac{\Phi(x)}{n\phi_x(h)}\int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right)$$

Lemme 4.2.4. *Sous les conditions du Théorème 4.2.1 on a :*

$$Cov(\hat{g}(x),\hat{f}(x)) = \frac{r(x)}{n\phi_x(h)}\int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right)$$

4.2.3 Démonstration des lemmes techniques

Démonstration du Lemme 4.2.1 on a :

$$
\begin{aligned}
\mathbb{E}\left[\hat{f}(x)\right] &= \mathbb{E}\left[\frac{1}{n\phi_x(h)}\sum_{i=1}^{n} K_i\right] \\
&= \frac{1}{n\phi_x(h)}\mathbb{E}\left[\sum_{i=1}^{n} K_i\right] \\
&= \frac{1}{\phi_x(h)}\int_{B(0,1)} K\left(\frac{\|x-z\|}{h}\right) dP^X(z) \\
&= \frac{1}{\phi_x(h)}\int_{B(0,1)} K(\|v\|) g(h,x,v) du(v)
\end{aligned}
$$

Comme $\quad h(x,v) = \dfrac{g(h,x,v)}{\phi(h)} + o(1)$

Donc $\quad \mathbb{E}\left[\hat{f}(x)\right] = \displaystyle\int_{B(0,1)} K(\|v\|) h(x,v) du(v) + o(1)$

Pour la deuxième équation :

$$
\begin{aligned}
\mathbb{E}\left[\hat{g}(x)\right] &= \mathbb{E}\left[\frac{1}{n\phi_x(h)}\sum_{i=1}^{n} Y_i K_i\right] \\
&= \frac{1}{n\phi_x(h)}\mathbb{E}\left[\sum_{i=1}^{n} Y_i K_i\right] \\
&= \frac{1}{\phi_x(h)}\mathbb{E}\left[Y_1 K_1\right] \\
&= \frac{1}{\phi_x(h)}\mathbb{E}\left[\mathbb{E}[Y_1|X=X_1] K_1\right] \\
&= \frac{1}{\phi_x(h)}\mathbb{E}\left[r(X_1) K_1\right] \\
&= \frac{1}{\phi_x(h)}\int_{B(0,1)} r(z) K\left(\frac{\|x-z\|}{h}\right) dP^X(z) \\
&= \frac{1}{\phi_x(h)}\int_{B(0,1)} K\left(\|v\|\right) r(x-hv) g(h,x,v) du(v)
\end{aligned}
$$

En utilisant le fait que dans un voisinage du point x, on a

$$r(x-hv) = r(x) - hr'(x)[v] + o(h)$$

Ainsi $\mathbb{E}\left[\hat{g}(x)\right] = \dfrac{1}{\phi_x(h)} \displaystyle\int_{B(0,1)} K\left(\|v\|\right)\left[r(x) - hvr'(x)\right] g(h,x,v)du[v] + o(h)$

Alors

$$\begin{aligned}
\mathbb{E}\left[\hat{g}(x)\right] &= \frac{1}{\phi_x(h)} \int_{B(0,1)} K\left(\|v\|\right) r(x) g(h,x,v)du[v] \\
&- \frac{h}{\phi_x(h)} \int_{B(0,1)} K\left(\|v\|\right) r'(x)[v] g(h,x,v)du[v] + o(h) \\
&= r(x) \int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v] - h \int_{B(0,1)} K\left(\|v\|\right) r'(x)[v] h(x,v)du[v] + o(h) \quad \blacksquare
\end{aligned}$$

Démonstration du Corollaire 4.2.1 D'après le calcul précédent :

$$\begin{aligned}
\frac{\mathbb{E}[\hat{g}(x)]}{\mathbb{E}[\hat{f}(x)]} - r(x) &= \frac{r(x) \int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v]}{\displaystyle\int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v]} \\
&- \frac{h \displaystyle\int_{B(0,1)} K\left(\|v\|\right) r'(x)[v] h(x,v)du[v] + o(h)}{\int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v]} - r(x) \\
&= \frac{r(x) \displaystyle\int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v] - h \int_{B(0,1)} K\left(\|v\|\right) r'(x)[v] h(x,v)du[v]}{\displaystyle\int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v]} \\
&- \frac{r(x) \left[\displaystyle\int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v]\right]}{\displaystyle\int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v]} + o(h) \\
&= \frac{-h \displaystyle\int_{B(0,1)} K\left(\|v\|\right) r'(x)[v] h(x,v)du[v]}{\displaystyle\int_{B(0,1)} K\left(\|v\|\right) h(x,v)du[v]} + o(h)
\end{aligned}$$

Alors :

$$\frac{\mathbb{E}[\hat{g}(x)]}{\mathbb{E}[\hat{f}(x)]} - r(x) = B(x)h + o(h) \qquad \blacksquare$$

Démonstration du Lemme 4.2.2 On a,

$$\begin{aligned} Var\left[\hat{f}(x)\right] &= Var\left[\frac{1}{n\phi_x(h)}\sum_{i=1}^{n} K_i\right] \\ &= \frac{1}{n^2\phi_x(h)^2} Var\left[\sum_{i=1}^{n} K_i\right] \\ &= \frac{1}{n\phi_x(h)^2}\left(\mathbb{E}[K_1^2] - (\mathbb{E}K_1)^2\right) \end{aligned}$$

Du fait que

$$\frac{1}{\phi_x(h)^2}\left(\mathbb{E}[K_1]\right)^2 = \left(\mathbb{E}\left[\hat{f}(x)\right]\right)^2 = O(1)$$

Alors :

$$\begin{aligned} Var\left[\hat{f}(x)\right] &= \frac{1}{n\phi_x(h)^2}\mathbb{E}[K_1^2] + o\left(\frac{1}{n\phi_x(h)}\right) \\ &= \frac{1}{n\phi_x(h)^2}\int_{B(0,1)} K^2\left(\frac{\|x-z\|}{h}\right) dP^X(z) + o\left(\frac{1}{n\phi_x(h)}\right) \\ &= \frac{1}{n\phi_x(h)^2}\int_{B(0,1)} K^2\left(\|v\|\right) g(h,x,v)du[v] + o\left(\frac{1}{n\phi_x(h)}\right) \\ &= \frac{1}{n\phi_x(h)}\int_{B(0,1)} K^2\left(\|v\|\right) h(x,v)du[v] + o\left(\frac{1}{n\phi_x(h)}\right) \end{aligned}$$

Ainsi :

$$Var\left[\hat{f}(x)\right] = \frac{1}{n\phi_x(h)}\int_{B(0,1)} K^2\left(\|v\|\right) h(x,v)du[v] + o\left(\frac{1}{n\phi_x(h)}\right) \qquad \blacksquare$$

Démonstration du Lemme 4.2.3 On a

$$\begin{aligned}
Var\left[\hat{g}(x)\right] &= Var\left[\frac{1}{n\phi_x(h)}\sum_{i=1}^{n} Y_iK_i\right] \\
&= \frac{1}{n^2\phi_x(h)^2} Var\left[\sum_{i=1}^{n} Y_iK_i\right] \\
&= \frac{1}{n\phi_x(h)^2}\left(\mathbb{E}[Y_1^2K_1^2] - (\mathbb{E}Y_1K_1)^2\right)
\end{aligned}$$

Du fait que

$$\frac{1}{\phi_x(h)^2}\left(\mathbb{E}[Y_1K_1]\right)^2 = \left(\mathbb{E}\left[\hat{g}(x)\right]\right)^2 = O(1)$$

Alors

$$\begin{aligned}
Var\left[\hat{g}(x)\right] &= \frac{1}{n\phi_x(h)^2}\mathbb{E}\left[\mathbb{E}[Y_1^2|X = X_1]K_1^2\right] + o\left(\frac{1}{n\phi_x(h)}\right) \\
&= \frac{1}{n\phi_x(h)^2}\int_{B(0,1)} \Phi(z)K^2\left(\frac{\|x-z\|}{h}\right) dP^X(z) + o\left(\frac{1}{n\phi_x(h)}\right) \\
&= \frac{1}{n\phi_x(h)^2}\int_{B(0,1)} \Phi(x-hv)K^2\left(\|v\|\right) g(h,x,v)du[v] + o\left(\frac{1}{n\phi_x(h)}\right)
\end{aligned}$$

On utilise le développement de Taylor pour l'opérateur Φ au voisinage du point x.

$$\Phi(x-hv) = \Phi(x) - h\Phi'(x)[v] + o(h)$$

Ainsi,

$$Var\left[\hat{g}(x)\right] = \frac{1}{n\phi_x(h)^2}\int_{B(0,1)} K^2\left(\|v\|\right)\left[\Phi(x) - h\Phi'(x)\right]g(h,x,v)du[v] + o\left(\frac{1}{n\phi_x(h)}\right)$$

Finalement, on a :

$$Var\left[\hat{g}(x)\right] = \frac{1}{n\phi_x(h)}\Phi(x)\int_{B(0,1)} K^2\left(\|v\|\right) h(x,v)du[v] + o\left(\frac{1}{n\phi_x(h)}\right) \qquad \blacksquare$$

Démonstration du Lemme 4.2.4

$$\begin{aligned}
Cov(\hat{g}(x),\hat{f}(x)) &= Cov\left(\frac{1}{n\phi_x(h)}\sum_{i=1}^{n}Y_iK_i,\frac{1}{n\phi_x(h)}\sum_{i=1}^{n}K_i\right)\\
&= \frac{1}{n\phi_x(h)^2}\left(\mathbb{E}\left[Y_1K_1^2\right]-(\mathbb{E}K_1\mathbb{E}Y_1K_1)\right)
\end{aligned}$$

En vertu du Lemme 4.2.1, on a :

$$\frac{\mathbb{E}(Y_1K_1)}{\phi_x(h)}=\mathbb{E}\left[\hat{g}(x)\right]=O(1)$$

$$\frac{\mathbb{E}(K_1)}{\phi_x(h)}=\mathbb{E}\left[\hat{f}(x)\right]=O(1)$$

D'où,

$$\begin{aligned}
Cov(\hat{g}(x),\hat{f}(x)) &= \frac{1}{n\phi_x(h)^2}\mathbb{E}\left[Y_1K_1^2\right]+o\left(\frac{1}{n\phi_x(h)}\right)\\
&= \frac{1}{n\phi_x(h)^2}\mathbb{E}\left[\mathbb{E}[Y_1|X=X_1]K_1^2\right]+o\left(\frac{1}{n\phi_x(h)}\right)\\
&= \frac{1}{n\phi_x(h)^2}\int_{B(0,1)}r(z)K^2\left(\frac{\|x-z\|}{h}\right)dP^X(z)+o\left(\frac{1}{n\phi_x(h)}\right)\\
&= \frac{1}{n\phi_x(h)^2}\int_{B(0,1)}K^2\left(\|v\|\right)r(x)g(h,x,v)du[v]\\
&- \frac{h}{n\phi_x(h)^2}\int_{B(0,1)}K^2\left(\|v\|\right)r'(x)[v]g(h,x,v)du[v]+o\left(\frac{1}{n\phi_x(h)}\right)\\
&= \frac{r(x)}{n\phi_x(h)}\int_{B(0,1)}K^2\left(\|v\|\right)h(x,v)du[v]+o\left(\frac{1}{n\phi_x(h)}\right)
\end{aligned}$$

Ainsi,

$$Cov(\hat{g}(x),\hat{f}(x))=\frac{r(x)}{n\phi_x(h)}\int_{B(0,1)}K^2\left(\|v\|\right)h(x,v)du[v]+o\left(\frac{1}{n\phi_x(h)}\right)\qquad\blacksquare$$

4.3 Cas dépendant

Dans cette section, l'étude sera orientée vers l'estimation de la fonction de régression dans le cas où les observations sont à la fois fortement mélangeantes et la co-variable est fonctionnelle. Pour cela, on considère les hypothèses suivantes :

4.3.1 Hypothèses

(H_5) La suite des observations $\{(X_i, X_j), i, j \in \mathbb{N}\}$ est α-mélangeante dont le coefficient de mélange satisfait :

$$\exists a > (5 + \sqrt{17}/2), \exists \nu > 0 : \forall n \in \mathbb{N}, \alpha(n) \leq \nu n^{-a}$$

(H_6) $0 < \sup_{i \neq j} P((X_i, X_j) \in B(x,h) \times B(x,h)) = O\left(\frac{(\phi_x(h))^{(a+1)/a}}{n^{1/a}}\right)$, où $\phi_x(h) = \phi(r) \int_{B(0,1)} h(x,v) d\mu(v)$

4.3.2 Propriété asymptotique

La vitesse de convergence est donnée par le théorème suivant :

Théorème 4.3.1. *Sous les hypothèses(H_1)-(H_6) et si (H_5) est telle que*

$$\phi(h) < c_1 n^{\frac{1}{1-a} - \epsilon} \qquad \epsilon > 0$$

Alors,

$$\mathbb{E}[\hat{r}(x) - r(x)]^2 = B^2(x)h^2 + V(x)\frac{1}{n\phi_x(h)} + o\left(h^2 + \frac{1}{n\phi_x(h)}\right)$$

Démonstration du Théorème 4.3.1

Comme la partie biais est indépendante à la corrélation des observations, il suffit donc d'évaluer la partie variance lorsque les observations sont α-mélangeantes. Alors, nous présentons ci-dessous les lemmes qui servent à établir le résultat de cette section.

Lemme 4.3.1. *Sous les conditions du Théorème 4.3.1 on a :*

$$Var\left[\hat{g}(x)\right] = \frac{\Phi(x)}{n\phi_x(h)} \int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right)$$

Lemme 4.3.2. *Sous les conditions du Théorème 4.3.1 on a :*

$$Var\left[\hat{f}(x)\right] = \frac{1}{n\phi_x(h)} \int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right)$$

Lemme 4.3.3. *Sous les conditions du Théorème 4.3.1 on a :*

$$Cov(\hat{g}(x), \hat{f}(x)) = \frac{r(x)}{n\phi_x(h)} \int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right)$$

4.3.3 Démonstration des lemmes techniques

Démonstration du Lemme 4.3.1 On peut écrire

$$Var\left[\hat{g}(x)\right] = \frac{1}{n^2\phi_x(h)^2} \sum_{i \neq j} Cov(\Delta_i, \Delta_j) + \frac{1}{n\phi_x(h)^2} Var[\Delta_1]$$

Avec $\Delta_i = [Y_i K_i - \mathbb{E}(Y_i K_i)]$
En suivant la même démarche que pour le cas i.i.d. et du fait que

$$\frac{1}{\phi_x(h)^2} (\mathbb{E}[K_1 Y_1])^2 = (\mathbb{E}[\hat{g}(x)])^2 = O(1).$$

On montre que

$$\frac{1}{n\phi_x(h)^2} Var[\Delta_1] = \frac{1}{n\phi_x(h)^2} (\mathbb{E}[K_1^2 Y_1^2]) + o\left(\frac{1}{n\phi_x(h)}\right)$$

Ainsi,

$$\begin{aligned}
\frac{1}{n\phi_x(h)^2} Var\left[\Delta_1\right] &= \frac{1}{n\phi_x(h)^2} \mathbb{E}\left[\mathbb{E}[Y_1^2 | X = X_1] K_1^2\right] + o\left(\frac{1}{n\phi_x(h)}\right) \\
&= \frac{1}{n\phi_x(h)^2} \int_{B(0,1)} \Phi(z) K^2\left(\frac{\|x - z\|}{h}\right) dP^X(z) + o\left(\frac{1}{n\phi_x(h)}\right) \\
&= \frac{1}{n\phi_x(h)^2} \int_{B(0,1)} \Phi(x - hv) K^2\left(\|v\|\right) g(h, x, v) du[v] + o\left(\frac{1}{n\phi_x(h)}\right)
\end{aligned} \tag{4.1}$$

On a

$$\Phi(x-hv)=\Phi(x)-h\Phi'(x)[v]+o(h)$$

en remplaçant ce développement dans l'équation (4.1), on arrive à

$$\frac{1}{n\phi_x(h)^2}Var[\Delta_1]=\frac{\Phi(x)}{n\phi_x(h)}\int_{B(0,1)}K^2(\|v\|)h(x,v)du[v]+o\left(\frac{1}{n\phi_x(h)}\right)$$

Il nous reste maintenant d'évaluer la quantité

$$S_n^{2*}(x)=\sum_{i\neq j}Cov(\Delta_i(x),\Delta_j(x)).$$

Pour calculer cette quantité, on défini les ensembles

$$E_1=\{(i,j)\text{ tel que }1\leq|i-j|\leq m_n\}$$

Et

$$E_2=\{(i,j)\text{ tel que }m_n+1\leq|i-j|\leq n-1\}$$

Avec $\lim_{n\to+\infty} m_n=+\infty$

On note par $j_{1,n}$ et $j_{2,n}$ la somme sur les ensembles E_1 et E_2 respectivement. Alors,

$$j_{1,n}=\sum_{E_1}|Cov(\Delta_i(x),\Delta_j(x))|\leq\sum_{E_1}|\mathbb{E}[K_i(x)K_j(x)]-\mathbb{E}^2[K_1(x)]|.$$

Sous les hypothèses $(H_1),(H_5)$ et (H_6) on obtient,

$$j_{1,n}\leq Cnm_n\phi_x(h)\left(\left(\frac{\phi_x(h)}{n}\right)^{\frac{1}{a}}+\phi_x(h)\right)$$

Sur l'ensemble E_2, on utilise l'inégalité de Covariance laquelle nous permet d'écrire, pour tout $i\neq j$, que

$$|Cov(K_i(x),K_j(x))|\leq C\alpha(|i-j|).$$

Donc,

$$j_{2,n}=\sum_{E_2}|Cov(K_i(x),K_j(x))|\leq n^2m_n^{-a}$$

En choisissant $m_n = \left(\dfrac{\phi_x(h)}{n\log n}\right)^{\frac{-1}{a}}$ pour que

$$j_{1,n} = o(n\phi_x(h)) \quad \text{et} \quad j_{2,n} = o(n\phi_x(h))$$

Ainsi,

$$\frac{1}{n^2\phi_x(h)^2}\sum_{i\neq j} Cov(\Delta_i,\Delta_j) = o\left(\frac{1}{n\phi_x(h)}\right)$$

par conséquent

$$Var\left[\hat{g}(x)\right] = \frac{1}{n\phi_x(h)^2} Var[\Delta_1] + o\left(\frac{1}{n\phi_x(h)}\right)$$

Finalement,

$$Var\left[\hat{g}(x)\right] = \frac{\Phi(x)}{n\phi_x(h)}\int_{B(0,1)} K^2(\|v\|)h(x,v)du[v] + o\left(\frac{1}{n\phi_x(h)}\right)$$

Ce qui achève la preuve de Lemme 4.3.1 ■

Démonstration du Lemme 4.3.2 La démonstration est de la même démarche que le lemme précédent, il suffit d'écrire

$$Var\left[\hat{f}(x)\right] = \frac{1}{n^2\phi_x(h)^2}\sum_{i\neq j} Cov(\Gamma_i,\Gamma_j) + \frac{1}{n\phi_x(h)^2} Var[\Gamma_1]$$

Avec $\Gamma_i = K_i - \mathbb{E}K_i$

Le calcul du cas i.i.d., montre que

$$\frac{1}{n\phi_x(h)^2} Var[\Gamma_1] = \frac{1}{n\phi_x(h)}\int_{B(0,1)} K^2\left(\|v\|\right) h(x,v)du[v] + o\left(\frac{1}{n\phi_x(h)}\right)$$

Et par analogie on montre que

$$\frac{1}{n^2\phi_x(h)^2}\sum_{i\neq j} Cov(\Gamma_i,\Gamma_j) = o\left(\frac{1}{n\phi_x(h)}\right)$$

Ainsi,

$$Var\left[\hat{f}(x)\right] = \frac{1}{n\phi_x(h)}\int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right) \quad \blacksquare$$

Démonstration du Lemme 4.3.3 De la même manière, on peut écrire

$$Cov(\hat{g}(x),\hat{f}(x)) = \frac{1}{n^2\phi_x(h)^2}\sum_{i,j} Cov(\Delta_i,\Gamma_j)$$

Il est claire que

$$Cov(\hat{g}(x),\hat{f}(x)) = \frac{1}{n^2\phi_x(h)^2}\sum_{i\neq j} Cov(\Delta_i,\Gamma_j) + \frac{1}{n\phi_x(h)^2} Cov(\Delta_1,\Gamma_1)$$

Le premier terme est évaluée en considérant les deux ensembles E_1 et E_2 (en reprenant les même arguments du Lemme 4.3.2) :

$$\sum_{E_1} |Cov(\Delta_i(x),\Gamma_j(x))| \leq \sum_{E_1} |\mathbb{E}[K_i(x)K_j(x)] - \mathbb{E}^2[K_1(x)]|.$$

Sous les hypothèses $(H_1), (H_5)$ et (H_6) on obtient,

$$\sum_{E_1} |Cov(\Delta_i(x),\Gamma_j(x))| \leq Cnm_n\phi_x(h)\left(\left(\frac{\phi_x(h)}{n}\right)^{\frac{1}{a}} + \phi_x(h)\right)$$

Sur l'ensemble E_2, on a :

$$\sum_{E_2} |Cov(K_i(x),K_j(x))| \leq n^2 m_n^{-a}$$

Choisissant $m_n = \left(\frac{\phi_x(h)}{n\log n}\right)^{\frac{-1}{a}}$ pour que

$$\frac{1}{n^2\phi_x(h)^2}\sum_{i\neq j} Cov(\Delta_i,\Gamma_j) = o\left(\frac{1}{n\phi_x(h)}\right)$$

Pour le deuxième terme, en utilisant le résultat du cas i.i.d., pour arriver à

$$\frac{1}{n\phi_x(h)^2} Cov(\Delta_1,\Gamma_1) = \frac{r(x)}{n\phi_x(h)}\int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right)$$

Finalement, on obtient

$$Cov(\hat{g}(x),\hat{f}(x)) = \frac{r(x)}{n\phi_x(h)}\int_{B(0,1)} K^2(\|v\|)h(x,v)du(v) + o\left(\frac{1}{n\phi_x(h)}\right) \qquad \blacksquare$$

Conclusion et perspectives

Conclusion

A travers ce mémoire, nous avons étudié l'erreur moyenne quadratique d'un estimateur à noyau de la régression lorsque la variable explicative prenne ses valeurs dans un espace unidimensionnel, multidimensionnel et finalement fonctionnel, et la variable réponse est toujours un réel.
Nous avons obtenu l'expression asymptotiquement exacte de l'erreur quadratique, lorsque les observations sont indépendantes identiquement distribuées, et ces résultats sont généralisés au cas où les observations sont α-mélangeantes
Tous les résultats des vitesses de convergence, c'est-à-dire les théorèmes dans les trois cas mettent en évidence le rôle du paramètre de lissage h, en regardant que le terme de biais est proportionnel à h tandis que le terme de variance est inversement proportionnel à h, ce qui signifie que tous les résultats établis nous indiquent que dans le but de minimiser l'erreur quadratique, la largeur de fenêtre ne doit pas être choisie trop grande puisque dans ce cas elle augmentera la composante proportionnelle à h^4, ni trop petite car elle augmentera la composante proportionnelle à h, et alors l'optimisation de l'erreur moyenne quadratique nous permettent de déterminer le paramètre de lissage optimal.
On peut aussi remarquer que grâce aux hypothèses supplémentaires du cas dépendant nous avons réussi à établir des vitesses de convergence semblables au cas i.i.d.
Finalement, il faut mentionner que la dimensionalité des observations (resp. du modèle) est bien exploitée dans l'expression de la vitesse de convergence.

On trouvera la dimensionalité du modèle dans la partie biais, tandis que, la dimensionalité de la variable fonctionnelle est dans la partie dispersion à travers la propriété de concentration de la mesure de probabilité de la variable fonctionnelle qui est étroitement liée à l'espace fonctionnel de la variable explicative. Ceci met en évidence le rôle de la semi-métrique dans la qualité de notre estimation, lorsqu'on considère des observations fonctionnelles. Un choix convenable de ce paramètre nous permet d'établir une solution originale pour le problème de fléau de la dimension.

perspectives

Le sujet que nous avons abordé dans ce mémoire offre de nombreuses perspectives. Parmi lesquelles :

Le paramètre de lissage est un paramètre crucial en estimation non-paramétrique car ce paramètre intervient dans toutes les propriétés asymptotiques qu'on a étudié. La sélection des paramètres de lissage est un sujet de recherche. Nos résultats asymptotiques constituent une étape préliminaire indispensable permettant d'envisager cette perspective de recherche.

La deuxième perspective à traiter à cours terme est la question liée à la normalité asymptotique de notre estimateur. Notons que cette propriété asymptotique permettra d'optimiser les intervalles de confiance et de faire des tests statistiques.

Il est bien possible de généraliser les résultats de ce mémoire en considérant la variable réponse fonctionnelle.

Bibliographie

[1] Attouch, M., Laksaci, A., Ould-Saïd, E. (2009). Asymptotic Distribution of Robust Estimator for Functional Nonparametric Models. *Communications in Statistics : Theory and Methods.* **38**, 1317-1335.

[2] Bosq, D., Lecoutre, J.P., (1987). *Théorie de l'estimation fonctionnelle*, Edition Economica, Paris.

[3] Bosq, D. (1996). Nonparametric statistics for stochastic processes : estimation and prediction, *Lecture Notes in statistics,* **110**, Spriger-Verlag.

[4] Carmbes, C., Laurent, D., Laksaci, A., (2007). Robust nonparametric estimation for functional data. *J. Nonparametric Stat.* **20**, No. 7, 573-598.

[5] Dabo-Niang. S. , Rhomari, N. (2004). Estimation non-paramétrique de la régression avec variable explicative dans un espace métrique. *C. R. Math. Acad. Sci. Paris* **336** (1) 75-80.

[6] Delsol, L. (2007). CTL and $\mathbb{L}^q$ errors in nonparametric functional regression C.R. Math. Acad. Sci. **345**, (7) 411-414.

[7] Delsol, L. (2009). Advances on asymptotic normality in nonparametric functional Time Series Analysis. *Statistics*, **43**, 13-33.

[8] Ferraty, F. et Vieu, P. (2002) Statistique fonctionnelle : Modèles Non-Paramétriques de Régression, Notes de cours de DEA.

[9] Ferraty, F. (2003). Modélisation statistique pour variables aléatoires fonctionnelles : Théorie et Application, Habiliation à diriger des recherches, *Université Paul Sabatier Toulouse III.*

[10] Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis theory and practice.* Springer-Verlag.

[11] Ferraty, F., Mas, A. and Vieu, P. (2007). Advances on nonparamtric regression for fonctional variables. *Austr. New-Zeal. J. Stat.*, **49**, 1-20.

[12] González-Manteiga, W. and Vieu, P. (2007). Statistics for functional data. *Computational Statistics and Data Analysis*, **51**, 4788-4792.

[13] Gyorfi, L., Hardle, W., Sarda P. and Vieu, P. (1989) Nonparametric Curve Estimation for Time Series. *Lecture Notes in statistics*, **60** Springer-Verlag, Berlin.

[14] Laksaci, A. (2005). Contribution aux modèles non-paramétrique conditionnels pour variables explicatives fonctionnelles, Thèse de Doctorat, *Université Paul Sabatier Toulouse III.*

[15] Masry, E. (2005). Nonparametric regression estimation for dependent functional data : Asymptotic normality. *Stoch. Proc. and their Appl.*, **115**, 155-177.

[16] Nadaraya, E. (1964). On estimating regression. *Teor. Prob.Appl.*,**10**, 186-196.

[17] Rachdi, M., Vieu, P. (2007). Nonparametric regression for functional data : automatic smoothing parameter selection. *Journal of Statistical Planning and Inference*, **137**, 2784-2801.

[18] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U. S. A.*, **42**, 43-47.

[19] Roussas, G. (1989) Consistent regression estimation with fixed design points under dependence conditions. *Stat. Probab. Lett.* **8**, No.1, 41-50.

[20] Stone, C., J. (1977) Consistent nonparametric regression. *Discussion. Ann. Stat.* **5**, 595-645.

[21] Tukey, J. W. (1961) Curves as parameters, and touch estimation *Proc. 4th Berkeley Symp. Math. Stat. Probab.* **1**, 681-694.

[22] Valderrama, M.J. (2007). An overview to modelling functional data. *Computational Statistics*, **22**, 331-334.

[23] Vieu, P. (1991) Quadratic errors for nonparametric estimates under dependence. *J. Multivariate Anal.* **39**, No.2, 324-347.

[24] Watson, G.S. (1964) Smooth regression analysis. *Sankhya, Ser. A*, **26**, 359-372.