DEA de Statistique et Modèles Aléatoires en Economie et en Finance
Université Paris VII

Mémoire de DEA

Année 2002-2003

Déconvolution de densité : vitesses de convergence et estimation adaptative

Claire Lacour

sous la direction de F. Comte

Déconvolution de densité : vitesses de convergence et estimation adaptative

1 Introduction

On suppose qu'on observe les variables aléatoires suivantes :

$$Y_i = X_i + \varepsilon_i \quad i = 1, \dots, n \quad (1)$$

où X_i et ε_i sont des variables aléatoires indépendantes de densités respectives f et q . On cherche à estimer f en supposant que la densité du bruit ε_i est connue. Ce problème est appelé problème de déconvolution puisque la densité des observations est égale au produit de convolution $f * q$.

Ce problème a été largement étudié et on connaît les taux de convergence optimaux pour le risque point par point (Fan (1991)) comme pour le risque L_p (Fan (1993)). L'utilisation des techniques d'ondelettes a permis de construire des estimateurs optimaux (voir Pensky et Vidakovic (1999), Pensky (2002)).

Ici, nous utiliserons des estimateurs à noyaux. Ils sont inspirés des estimateurs à noyaux de l'estimation de densité et utilisent l'analyse de Fourier (voir Zhang (1990)). Plusieurs travaux ont montré la consistance de ces estimateurs (voir Devroye (1989), Liu et Taylor (1989), Stefanski et Carroll (1990), Taylor et Zhang (1990)). Carroll et Hall (1988), Stefanski (1990), Piterbarg et Penskaya (1993) ont donné des taux de convergence pour le risque de ces estimateurs dans les cas i.i.d. tandis que Masry a étudié le cas des variables mélangeantes (voir Masry (1993a), Masry (1993b)). Pour la densité q de l'erreur, on considère deux cas : q ordinary smooth ou q supersmooth. Les mauvais résultats obtenus lorsque q est supersmooth (vitesse de convergence logarithmique) nous amènent à considérer des fonctions f également supersmooth et non plus ordinary smooth. Ainsi on aura quatre cas à envisager selon les régularités de q et f . Cependant les vitesses trouvées ne sont atteignables que si la fenêtre de l'estimateur est convenablement choisie et ce choix dépend de la densité inconnue f . On va donc introduire un autre estimateur défini par minimum de contraste (Birgé et Massart (1998)). Pour le rendre adaptatif, on utilise la sélection de modèles par pénalisation introduite par Birgé et Massart (voir Birgé et Massart (1997), Barron, Birgé et Massart (1999)).

Dans la section 2, on présente les notations utilisées, notamment en ce qui concerne la régularité de f et q . Puis, on définit l'estimateur à noyaux et on calcule ses vitesses de convergence pour le risque ponctuel et pour le risque L_2 . La section 4 contient la définition de l'estimateur adaptatif et le calcul de son risque à l'aide de l'inégalité de Talagrand.

2 Présentation du problème

On considère le problème suivant :

$$Y_i = X_i + \varepsilon_i$$

- où (H1) les X_i sont i.i.d. de densité f inconnue
 (H2) les ε_i sont i.i.d. de densité q connue
 (H3) (X_i) est indépendante de (ε_i)

On note, pour toute fonction g considérée, $\tilde{g}$ ou $\mathcal{F}(g)$ la transformée de Fourier de g :

$$\tilde{g}(u) = \int e^{-ixu} g(u) du = \mathcal{F}(g)(u)$$

et $\bar{\mathcal{F}}(g)$ la transformée de Fourier inverse :

$$\bar{\mathcal{F}}(g)(u) = \frac{1}{2\pi} \int e^{ixu} g(u) du$$

On note p la densité de Y_1 :

$$p(x) = f * q(x) = \int f(x-y)q(y)dy \quad \text{donc} \quad \tilde{p} = \tilde{f}\tilde{q}$$

On va considérer deux cas :

- soit f est *ordinary smooth* i.e. $f \in OS(\delta)$ où

$$OS(\delta) = \{f \text{ densité, } \exists C > 0 \int |\tilde{f}(u)|^2 (1+u^2)^\delta du \leq C\}$$

- soit f est *super smooth* i.e. $f \in SS(A, r)$ où

$$SS(A, r) = \{f \text{ densité, } \exists C > 0 \int |\tilde{f}(u)|^2 \exp(2A|u|^r) du \leq C\}$$

De même pour q , on aura :

- (H4) soit q est *ordinary smooth* i.e. $q \in OS(\gamma)$ où

$$OS(\gamma) = \{q \text{ densité, } \exists Q_1, Q_2 > 0, Q_1 \leq |\tilde{q}(u)|^2 (1+u^2)^\gamma \leq Q_2\}$$

- (H5) soit q est *super smooth* i.e. $q \in SS(B, s)$ où

$$SS(B, s) = \{q \text{ densité, } \exists Q_1, Q_2 > 0, Q_1 \leq |\tilde{q}(u)|^2 \exp(2B|u|^s) \leq Q_2\}$$

On supposera aussi que q est paire.

3 Vitesse de convergence d'un estimateur à noyau

3.1 Définition de l'estimateur

On cherche un estimateur $\hat{f}$ de f sachant qu'on n'observe que les Y_i . On peut estimer p par un estimateur à noyau :

$$\hat{p}(x) = \frac{1}{nh} \sum_{k=1}^n G\left(\frac{Y_k - x}{h}\right) \quad \text{où} \quad \tilde{G}(u) = 1_{[-1,1]}(u)$$

On définit $\hat{f}$ par $\hat{p} = \hat{f} * q$. Si on pose $G_k(x) = \frac{1}{h} G\left(\frac{Y_k - x}{h}\right)$, on a

$$\hat{f} * q(x) = \frac{1}{n} \sum_{k=1}^n G_k(x) \quad \text{et donc}$$

$$\tilde{f}(u) \tilde{q}(u) = \frac{1}{n} \sum_{k=1}^n \tilde{G}_k(u)$$

Or $\tilde{G}_k(u) = e^{-iuY_k} \tilde{G}(-uh)$ donc :

$$\tilde{f}(u) = \frac{1}{n} \sum_{k=1}^n e^{-iuY_k} \frac{\tilde{G}(-uh)}{\tilde{q}(u)}$$

On définit K par $\tilde{K}(y) = \frac{\tilde{G}(y)}{\tilde{q}(-y/h)}$ et ainsi

$$\tilde{f}(u) = \frac{1}{n} \sum_{k=1}^n e^{-iuY_k} \tilde{K}(-uh) = \frac{1}{n} \sum_{i=1}^n \tilde{K}_i(u)$$

On a donc l'estimateur suivant :

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{Y_i - x}{h}\right) \quad (2)$$

On suppose que $\tilde{q}$ est paire donc K est pair et

$$\tilde{K}(u) = \frac{\tilde{G}(u)}{\tilde{q}(u/h)} = \frac{1_{[-1,1]}(u)}{\tilde{q}(u/h)} \quad (3)$$

On considère que $h > 0$ et $h \rightarrow 0$ (quand $n \rightarrow \infty$).

On s'intéresse à deux types de risques :

- le risque point par point : $MSE(x) = \mathbb{E}|\hat{f}_n(x) - f(x)|^2$
- et le risque L_2 : $MISE = \mathbb{E}\|\hat{f}_n - f\|_2^2$

Si on note

$$b(x) = |f(x) - \mathbb{E}\hat{f}_n(x)|$$

$$\sigma^2(x) = \mathbb{E}|\hat{f}_n(x) - \mathbb{E}\hat{f}_n(x)|^2,$$

on a $MSE(x) = b^2(x) + \sigma^2(x)$
et $MISE = \int (b^2(x) + \sigma^2(x)) dx$

3.2 Etude du biais

Proposition 1 Soit $\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{Y_i - x}{h}\right)$ et $b(x) = |f(x) - \mathbb{E}\hat{f}_n(x)|$.
Alors

- $b(x) \leq \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} |\tilde{f}(u)| du$
- $\int b^2(x) dx = \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} |\tilde{f}(u)|^2 du$

Démonstration :

i)

$$\begin{aligned}\mathbb{E}\hat{f}_n(x) &= \frac{1}{nh} \sum_{i=1}^n \mathbb{E} \left[K \left(\frac{Y_i - x}{h} \right) \right] = \frac{1}{h} \int K \left(\frac{y - x}{h} \right) p(y) dy \\ &= K_h * p(x)\end{aligned}$$

$$\text{où } K_h(z) = \frac{1}{h} K\left(\frac{z}{h}\right)$$

$$\begin{aligned}\tilde{K}_h(u) &= \int \frac{1}{h} K\left(\frac{z}{h}\right) e^{-iuz} dz = \int K(y) e^{-iuhy} dy \\ &= \tilde{K}(uh) = 1_{[-1,1]}(uh) \frac{1}{|u|}\end{aligned}$$

$$\mathcal{F}(\mathbb{E}\hat{f}_n)(u) = \tilde{K}_h(u) \tilde{p}(u) = \tilde{K}_h(u) \tilde{f}(u) \tilde{q}(u) = 1_{[-1,1]}(uh) \tilde{f}(u)$$

$$\text{Donc } \mathcal{F}(f - \mathbb{E}\hat{f}_n)(u) = 1_{\{|u| > \frac{1}{h}\}} \tilde{f}(u)$$

$$f(x) - \mathbb{E}\hat{f}_n(x) = \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} e^{ixu} \tilde{f}(u) du$$

$$b(x) \leq \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} |\tilde{f}(u)| du$$

ii) La formule de Parseval donne

$$\int b^2(x) dx = \|f - \mathbb{E}\hat{f}_n\|_2^2 = \frac{1}{2\pi} \int |\mathcal{F}(f - \mathbb{E}\hat{f}_n)|^2(u) du = \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} |\tilde{f}(u)|^2 du$$

□

3.3 Etude de la variance

$$\begin{aligned}\sigma^2(x) &= \text{Var}(\hat{f}_n(x)) = \frac{1}{nh^2} \text{Var}(K(\frac{Y_1 - x}{h})) \\ &\leq \frac{1}{nh^2} \int K^2\left(\frac{y-x}{h}\right) p(y) dy\end{aligned}$$

Lemme 1 Si $\int |\tilde{p}(u)| du < \infty$ alors

$$\frac{1}{h} \int K^2\left(\frac{y-x}{h}\right) p(y) dy = O\left(\int K^2(u) du\right)$$

et donc $\sigma^2(x) = O\left(\frac{1}{nh} \|K\|_2^2\right)$.

Remarque 1 Si, par exemple, $f \in SS(A, r)$,

$$\begin{aligned}\left(\int |\tilde{p}(u)| du\right)^2 &\leq \left(\int |\tilde{f}(u)| e^{A|u|^r} |\tilde{q}(u)| e^{-A|u|^r}\right)^2 \\ &\leq C \int e^{-2A|u|^r} |\tilde{q}(u)|^2 du\end{aligned}$$

On voit que la condition $\int |\tilde{p}(u)| du < \infty$ n'est pas trop restrictive pour q .

Démonstration : Pour tout x , $|p(x)| \leq \frac{1}{2\pi} \int |\tilde{p}(u)| du$. Donc, sous l'hypothèse du lemme, $\|p\|_\infty < \infty$. Ainsi

$$\begin{aligned} \left| \frac{1}{h} \int K^2\left(\frac{y-x}{h}\right) p(y) dy \right| &\leq \|p\|_\infty \frac{1}{h} \int K^2\left(\frac{y-x}{h}\right) dy \\ &\leq \|p\|_\infty \int K^2(u) du \end{aligned}$$

□

Calculons maintenant $\int \sigma^2(x) dx$.

$$\begin{aligned} \int \sigma^2(x) dx &= \mathbb{E} \|\hat{f}_n - \mathbb{E} \hat{f}_n\|_2^2 \\ &\leq \frac{1}{nh^2} \int p(y) \int K^2\left(\frac{y-x}{h}\right) dx dy \\ &\leq \frac{1}{nh} \int p(y) \int K^2(z) dz dy \\ &\leq \frac{1}{nh} \|K\|_2^2 \end{aligned}$$

La variance s'exprime donc en fonction de $\|K\|_2^2$.

$$\|K\|_2^2 = \frac{1}{2\pi} \int |\tilde{K}(u)|^2 = \frac{1}{2\pi} \int_{-1}^1 \frac{du}{|\tilde{q}(u/h)|^2} = \frac{h}{2\pi} \int_{-1/h}^{1/h} \frac{dv}{|\tilde{q}(v)|^2}$$

Finalement, on a la proposition suivante :

Proposition 2 Soit $\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{Y_i-x}{h}\right)$ et $\sigma^2(x) = \mathbb{E} |\hat{f}_n(x) - \mathbb{E} \hat{f}_n(x)|^2$. Alors

i) Si $\int |\tilde{p}(u)| du < \infty$ alors $\sigma^2(x) = O\left(\frac{1}{nh} \|K\|_2^2\right)$

ii) $\int \sigma^2(x) dx \leq \frac{1}{nh} \|K\|_2^2$

iii) $\|K\|_2^2 = \frac{h}{2\pi} \int_{-1/h}^{1/h} \frac{dv}{|\tilde{q}(v)|^2}$

3.4 Vitesses de convergence

3.4.1 cas 1 : $f \in OS(\delta)$ et $q \in OS(\gamma)$

On utilise la proposition 1 i). D'après l'inégalité de Cauchy-Schwartz :

$$\begin{aligned} b^2(x) &\leq \frac{1}{4\pi^2} \left(\int_{|u| > \frac{1}{h}} |\tilde{f}(u)| (1+u^2)^{\delta/2} (1+u^2)^{-\delta/2} du \right)^2 \\ &\leq \frac{C}{2\pi^2} \int_{\frac{1}{h}}^{\infty} (1+u^2)^{-\delta} du \\ &\leq \frac{C}{2\pi^2} \int_{\frac{1}{h}}^{\infty} (uh)(1+u^2)^{-\delta} du \\ &\leq \frac{Ch}{4\pi^2(1-\delta)} [(1+u^2)^{1-\delta}]_{\frac{1}{h}}^{\infty} \quad \text{si } \delta \neq 1 \\ &= O(h^{2\delta-1}) \end{aligned}$$

Si $\delta = 1$, on a directement

$$b^2(x) \leq \frac{C}{2\pi^2} [\text{Arctan}(u)]_{\frac{1}{h}}^{\infty} = O(h)$$

En utilisant la proposition 1 ii) :

$$\begin{aligned} \int b^2(x) dx &= \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} |\tilde{f}(u)|^2 du \\ &= \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} |\tilde{f}(u)|^2 (1+u^2)^\delta (1+u^2)^{-\delta} du \\ &\leq \frac{C}{2\pi} (1 + \frac{1}{h^2})^{-\delta} \end{aligned}$$

Donc $\int b^2(x) dx = O(h^{2\delta})$.

Calculons $\|K\|_2^2$

$$\begin{aligned} \|K\|_2^2 &= \frac{h}{2\pi} \int_{-1/h}^{1/h} \frac{dv}{|\tilde{q}(v)|^2} \\ &\leq \frac{h}{2\pi Q_1} \int_{-1/h}^{1/h} (1+v^2)^\gamma dv \\ &\leq \frac{h}{2\pi Q_1} \frac{2}{h} (1 + \frac{1}{h^2})^\gamma \\ &\leq \frac{2}{\pi Q_1} h^{-2\gamma} \end{aligned}$$

Donc, d'après la proposition 2 ii)

$$\int \sigma^2(x) dx = O\left(\frac{1}{nh^{2\gamma+1}}\right)$$

$$MISE = O\left(h^{2\delta} + \frac{1}{nh^{2\gamma+1}}\right)$$

On note pour deux suites u_n et v_n , $u_n \asymp v_n$ si $u_n = O(v_n)$ et $v_n = O(u_n)$. On obtient la meilleure majoration du risque L^2 si $h^{2\delta} \asymp \frac{1}{nh^{2\gamma+1}}$ c'est-à-dire si $h \asymp n^{-\frac{1}{2\gamma+2\delta+1}}$. On a alors

$$MISE = O\left(n^{-\frac{2\delta}{2\gamma+2\delta+1}}\right)$$

Pour $MSE(x)$, il faut calculer $\int |\tilde{p}(u)| du$ pour pouvoir majorer la variance.

$$\begin{aligned} \left(\int |\tilde{p}(u)| du\right)^2 &\leq \left(\int |\tilde{f}(u)| (1+u^2)^{\delta/2} |\tilde{q}(u)| (1+u^2)^{-\delta/2} du\right)^2 \\ &\leq C Q_2 \int (1+u^2)^{-\gamma} (1+u^2)^{-\delta} du \end{aligned}$$

Si $2\delta + 2\gamma > 1$, cette quantité est finie et la proposition 2 i) s'applique.

$$\sigma^2(x) = O\left(\frac{1}{nh^{2\gamma+1}}\right)$$

et donc

$$MSE(x) = O\left(h^{2\delta-1} + \frac{1}{nh^{2\gamma+1}}\right)$$

On obtient la meilleure majoration si $h^{2\delta-1} \asymp \frac{1}{nh^{2\gamma+1}}$ c'est-à-dire si $h \asymp n^{-\frac{1}{2\gamma+2\delta}}$. On a alors

$$MISE = O\left(n^{\frac{1-2\delta}{2\gamma+2\delta}}\right)$$

Proposition 3 Si $f \in OS(\delta)$ et $q \in OS(\gamma)$, l'estimateur à noyau défini par (2) et (3) vérifie

1. si $h = n^{-\frac{1}{2\gamma+2\delta+1}}$ alors $MISE = O(n^{-\frac{2\delta}{2\gamma+2\delta+1}})$
2. si $2\delta + 2\gamma > 1$ et si $h = n^{-\frac{1}{2\gamma+2\delta}}$ alors $MSE(x) = O(n^{\frac{1-2\delta}{2\gamma+2\delta}})$

Remarque 2 Pour le risque L_2 , Fan (1993) montre la borne inférieure correspondante pour $f \in H(\delta)$ où

$$H(\delta) = \{f \text{ densité, } |f^{(\delta-1)}(x) - f^{(\delta-1)}(y)| \leq C|x - y|\}.$$

On a donc cette même borne inférieure pour f appartenant à $OS(\delta)$ en remarquant que $H(\delta) \subset Sobolev(\delta) \subset OS(\delta)$. Ainsi, la vitesse $n^{-\frac{2\delta}{2\gamma+2\delta+1}}$ est optimale. Lorsqu'on s'intéresse aux classes de Hölder, le risque ponctuel est le même que le risque L_2 (voir Fan (1991)). Ce n'est pas le cas ici et on peut comprendre ce décalage en observant que $OS(\delta) \subset H(\delta - 1/2)$. Ainsi, le biais est d'ordre $h^{2\delta}$ si $f \in H(\delta)$ alors qu'il n'est plus que d'ordre $h^{2\delta-1}$ si $f \in OS(\delta)$.

3.4.2 cas 2 : $f \in OS(\delta)$ et $q \in SS(B, s)$

Le terme de biais n'a pas changé puisque f appartient toujours à $OS(\delta)$. Calculons $\|K\|_2^2$ et $\int |\tilde{p}(u)| du$.

Lemme 2 Pour $B > 0$, $s > 0$, on a

$$\int_0^{1/h} e^{2Bv^s} dv = O(h^{s-1} e^{2B/h^s})$$

Démonstration : Il faut distinguer 2 cas.

si $s \geq 1$

$$\begin{aligned} \int_0^{1/h} e^{2Bv^s} dv &= \int_0^{1/h} e^{2Bv^{s-1}v} dv \leq \int_0^{1/h} e^{\frac{2B}{h^{s-1}}v} dv \\ &\leq \left[\frac{h^{s-1}}{2B} e^{\frac{2B}{h^{s-1}}v} \right]_0^{1/h} \leq \frac{h^{s-1}}{2B} (e^{2B/h^s} - 1) \\ &\leq O(h^{s-1} e^{2B/h^s}) \end{aligned}$$

si $s < 1$

$$\begin{aligned} v \leq 1/h &\Rightarrow (vh)^{s-1} \geq 1 \\ \int_0^{1/h} e^{2Bv^s} dv &\leq \int_0^{1/h} (vh)^{s-1} e^{2Bv^s} dv \\ &\leq \frac{h^{s-1}}{2Bs} \int_0^{1/h} 2Bsv^{s-1} e^{2Bv^s} dv \\ &\leq \frac{h^{s-1}}{2Bs} [e^{2Bv^s}]_0^{1/h} \leq \frac{h^{s-1}}{2Bs} (e^{2B/h^s} - 1) \\ &\leq O(h^{s-1} e^{2B/h^s}) \end{aligned}$$

□

Remarque 3 Par le même raisonnement, on obtient

$$\exists K > 0 \int_0^{1/h} e^{2Bv^s} dv \geq Kh^{s-1} e^{2B/h^s}$$

$$\begin{aligned}
\|K\|_2^2 &= \frac{h}{2\pi} \int_{-1/h}^{1/h} \frac{dv}{|\tilde{q}(v)|^2} \\
&\leq \frac{h}{2\pi Q_1} \int_{-1/h}^{1/h} e^{2B|v|^s} dv \\
&\leq \frac{h}{\pi Q_1} \int_0^{1/h} e^{2Bv^s} dv
\end{aligned}$$

D'après le lemme 2

$$\|K\|_2^2 = O(e^{2B/h^s} h^s)$$

$$\begin{aligned}
\left(\int |\tilde{p}(u)| du\right)^2 &\leq \left(\int |\tilde{f}(u)|(1+u^2)^{\delta/2} |\tilde{q}(u)|(1+u^2)^{-\delta/2}\right)^2 \\
&\leq CQ_2 \int e^{-2B|u|^s} (1+u^2)^{-\delta} du < \infty
\end{aligned}$$

Donc on peut appliquer le i) de la proposition 2

$$\sigma^2(x) = O\left(\frac{h^{s-1}}{n} e^{2B/h^s}\right)$$

On a de même :

$$\int \sigma^2(x) dx = O\left(\frac{1}{nh} \|K\|_2^2\right) = O\left(\frac{h^{s-1}}{n} e^{2B/h^s}\right)$$

Ainsi

$$MSE(x) = O\left(h^{2\delta-1} + \frac{h^{s-1}}{n} e^{2B/h^s}\right)$$

$$MISE = O\left(h^{2\delta} + \frac{h^{s-1}}{n} e^{2B/h^s}\right)$$

$$\text{Soit } G(h) = h^{2\delta} + \frac{h^{s-1}}{n} e^{2B/h^s}$$

$$G'(h) = 2\delta h^{2\delta-1} + \frac{1}{nh^2} e^{2B/h^s} ((s-1)h^s - 2Bs)$$

La meilleure majoration du risque s'obtient en minimisant la fonction G et donc en trouvant h qui annule sa dérivée. On cherche donc $h = h_n$ tel que

$$h^{-2\delta-1} e^{2B/h^s} \asymp n$$

On vérifie que

$$h = \left(\frac{\ln n - \frac{2\delta+1}{s} \ln \ln n}{2B}\right)^{-1/s}$$

convient. On a alors :

$$MISE = O((\ln n)^{-\frac{2\delta}{s}})$$

$$\text{Soit maintenant } L(h) = h^{2\delta-1} + \frac{h^{s-1}}{n} e^{2B/h^s}$$

$$L'(h) = (2\delta-1)h^{2\delta-2} + \frac{1}{nh^2} e^{2B/h^s} ((s-1)h^s - 2Bs)$$

La meilleure majoration du risque ponctuel s'obtient donc en trouvant $h = h_n$ tel que

$$h^{-2\delta} e^{2B/h^s} \asymp n$$

On vérifie que

$$h = \left(\frac{\ln n - \frac{2\delta}{s} \ln \ln n}{2B} \right)^{-1/s}$$

convient. On a alors :

$$MSE(x) = O((\ln n)^{\frac{1-2\delta}{s}})$$

Proposition 4 Si $f \in OS(\delta)$ et $q \in SS(B, s)$, l'estimateur à noyau défini par (2) et (3) vérifie

1. si $h = \left(\frac{\ln n - \frac{2\delta}{s} \ln \ln n}{2B} \right)^{-1/s}$ alors $MSE(x) = O((\ln n)^{\frac{1-2\delta}{s}})$
2. si $h = \left(\frac{\ln n - \frac{2\delta+1}{s} \ln \ln n}{2B} \right)^{-1/s}$ alors $MISE = O((\ln n)^{-\frac{2\delta}{s}})$

Remarque 4 Pour le risque L_2 , Fan (1993) a montré la borne inférieure si $f \in H(\delta)$. pour les mêmes raisons que précédemment (voir remarque 2), cela prouve que la vitesse obtenue est optimale.

En ce qui concerne le risque ponctuel, Goldenschluger (1999) a prouvé qu'on ne pouvait atteindre une meilleure vitesse que $(\ln n)^{\frac{1-2\delta}{s}}$.

3.4.3 cas 3 : $f \in SS(A, r)$ et $q \in OS(\gamma)$

Lemme 3 Pour $A > 0$, $r > 0$, on a

$$\int_{1/h}^{\infty} e^{-2Au^r} du = O(h^{r-1} e^{-2A/h^r})$$

Démonstration :

Soit $I_p = \int_{1/h}^{\infty} e^{-2Au^r} u^{-pr} du$. On cherche à majorer I_0 .

$$\begin{aligned} I_p &= \int_{1/h}^{\infty} e^{-2Au^r} (-2Ar u^{r-1}) \frac{u^{-(p+1)r+1}}{-2Ar} du \\ &= \left[e^{-2Au^r} \frac{u^{-(p+1)r+1}}{-2Ar} \right]_{1/h}^{\infty} - \int_{1/h}^{\infty} e^{-2Au^r} \frac{-(p+1)r+1}{-2Ar} u^{-(p+1)r} du \\ &= \frac{1}{2Ar} e^{-2A/h^r} h^{(p+1)r-1} + \frac{1-(p+1)r}{2Ar} I_{p+1} \end{aligned}$$

Soit $k = \lceil \frac{1}{r} \rceil$ de telle sorte que $r \geq \frac{1}{k}$. Par récurrence,

$$I_0 \leq K e^{-2A/h^r} (h^{r-1} + h^{2r-1} + \dots + h^{(k-1)r-1}) + K' I_{k-1}$$

Or $u \geq 1/h \Rightarrow (uh)^{kr-1} \geq 1$ donc

$$\begin{aligned} I_{k-1} &\leq \int_{1/h}^{\infty} e^{-2Au^r} u^{-(k-1)r} (uh)^{kr-1} du \\ &\leq h^{kr-1} \int_{1/h}^{\infty} e^{-2Au^r} u^{r-1} du \\ &\leq \frac{1}{2Ar} h^{kr-1} e^{-2A/h^r} \end{aligned}$$

$$I_0 \leq K'' e^{-2A/h^r} (h^{r-1} + h^{2r-1} + \dots + h^{(k-1)r-1} + h^{kr-1})$$

Donc $I_0 = O(e^{-2A/h^r} h^{r-1})$ □

D'après l'inégalité de Cauchy-Schwartz :

$$\begin{aligned} b^2(x) &\leq \frac{1}{4\pi^2} \left(\int_{|u| > \frac{1}{h}} |\tilde{f}(u)| e^{A|u|^r} e^{-A|u|^r} du \right)^2 \\ &\leq \frac{C}{2\pi^2} \int_{1/h}^{\infty} e^{-2Au^r} du \end{aligned}$$

D'après le lemme 3,

$$b^2(x) = O(h^{r-1} e^{-2A/h^r})$$

De plus (proposition 1 ii) :

$$\begin{aligned} \int b^2(x) dx &= \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} |\tilde{f}(u)|^2 du \\ &= \frac{1}{2\pi} \int_{|u| > \frac{1}{h}} |\tilde{f}(u)|^2 e^{2A|u|^r} e^{-2A|u|^r} du \\ &\leq \frac{C}{2\pi} e^{-2A/h^r} \end{aligned}$$

Calculons maintenant $\|K\|_2^2$ et $\int |\tilde{p}(u)| du$ pour l'étude du terme de variance. Comme au cas 1,

$$\begin{aligned} \|K\|_2^2 &\leq \frac{2}{\pi Q_1} h^{-2\gamma} \\ \left(\int |\tilde{p}(u)| du \right)^2 &\leq \left(\int |\tilde{f}(u)| e^{A|u|^r} |\tilde{q}(u)| e^{-A|u|^r} \right)^2 \\ &\leq C Q_2 \int e^{-2A|u|^r} (1+u^2)^{-\gamma} du < \infty \end{aligned}$$

Donc on peut appliquer la proposition 2

$$\sigma^2(x) = O\left(\frac{1}{nh^{2\gamma+1}}\right) \quad \text{et}$$

$$\int \sigma^2(x) dx = O\left(\frac{1}{nh} \|K\|_2^2\right) = O\left(\frac{1}{nh^{2\gamma+1}}\right)$$

Ainsi

$$\begin{aligned} MSE(x) &= O(h^{r-1} e^{-2A/h^r} + \frac{1}{nh^{2\gamma+1}}) \\ MISE &= O(e^{-2A/h^r} + \frac{1}{nh^{2\gamma+1}}) \end{aligned}$$

Soit $L(h) = h^{r-1} e^{-2A/h^r} + \frac{1}{nh^{2\gamma+1}}$

$$L'(h) = [(r-1)h^{r-2} + h^{r-1} 2Arh^{-r-1}] e^{-2A/h^r} - \frac{2\gamma+1}{nh^{2\gamma+2}}$$

La meilleure majoration du risque ponctuel s'obtient pour $h = h_n$ tel que

$$h^{-2\gamma} e^{2A/h^r} \asymp n$$

On vérifie que

$$h = \left(\frac{\ln n - \frac{2\gamma}{r} \ln \ln n}{2A} \right)^{-1/r}$$

convient. On a alors :

$$MSE(x) = O\left(\frac{(\ln n)^{\frac{2\gamma+1}{r}}}{n}\right)$$

$$\text{Soit maintenant } G(h) = e^{-2A/h^r} + \frac{1}{nh^{2\gamma+1}}$$

$$G'(h) = 2Arh^{-r-1}e^{-2A/h^r} - \frac{2\gamma+1}{nh^{2\gamma+2}}$$

La meilleure majoration du risque L_2 s'obtient pour $h = h_n$ tel que

$$h^{r-2\gamma-1}e^{2A/h^r} \asymp n$$

$$h = \left(\frac{\ln n + \frac{r-2\gamma-1}{r} \ln \ln n}{2A}\right)^{-1/r}$$

convient. On a alors :

$$MISE = O\left(\frac{(\ln n)^{\frac{2\gamma+1}{r}}}{n}\right)$$

Proposition 5 *Si $f \in SS(A, r)$ et $q \in OS(\delta)$, l'estimateur à noyau défini par (2) et (3) vérifie*

$$1. \text{ si } h = \left(\frac{\ln n - \frac{2\gamma}{r} \ln \ln n}{2A}\right)^{-1/r} \text{ alors } MSE(x) = O\left(\frac{(\ln n)^{\frac{2\gamma+1}{r}}}{n}\right)$$

$$2. \text{ si } h = \left(\frac{\ln n + \frac{r-2\gamma-1}{r} \ln \ln n}{2A}\right)^{-1/r} \text{ alors } MISE = O\left(\frac{(\ln n)^{\frac{2\gamma+1}{r}}}{n}\right)$$

Remarque 5 *Butucea (2003) trouve la même vitesse et montre qu'il s'agit de la vitesse optimale, pour le risque ponctuel comme pour le risque intégré*

3.4.4 cas 4 : $f \in SS(A, r)$ et $q \in SS(B, r)$

Le biais n'a pas changé, on a toujours :

$$b^2(x) = O(h^{r-1}e^{-2A/h^r})$$

Par ailleurs

$$\int b^2(x)dx \leq \frac{C}{2\pi}e^{-2A/h^r}$$

Pour la variance, on a, comme au cas 2 :

$$\|K\|_2^2 = O(h^s e^{2B/h^s})$$

De plus,

$$\begin{aligned} \left(\int |\tilde{p}(u)|du\right)^2 &\leq \left(\int |\tilde{f}(u)e^{A|u|^r} \tilde{q}(u)|e^{-A|u|^r}\right)^2 \\ &\leq CQ_2 \int e^{-2B|u|^s} e^{-2A|u|^r} du < \infty \end{aligned}$$

Donc

$$\sigma^2(x) = O\left(\frac{h^{s-1}}{n}e^{2B/h^s}\right) \quad \text{et}$$

$$\int \sigma^2(x)dx = O\left(\frac{1}{nh} \|K\|_2^2\right) = O\left(\frac{h^{s-1}}{n}e^{2B/h^s}\right)$$

Donc

$$MSE(x) = O(h^{r-1}e^{-2A/h^r} + \frac{h^{s-1}}{n}e^{2B/h^s})$$

$$MISE = O(e^{-2A/h^r} + \frac{h^{s-1}}{n}e^{2B/h^s})$$

$$\text{Soit } G(h) = e^{-2A/h^r} + \frac{h^{s-1}}{n}e^{2B/h^s}$$

$$G'(h) = 2Arh^{-r-1}e^{-2A/h^r} + \frac{1}{nh^2}[(s-1)h^s - 2Bs]e^{2B/h^s}$$

Pour minimiser $G(h)$ et donc le risque, on va donc chercher $h = h_n$ vérifiant

$$h^{r-1}e^{2A/h^r+2B/h^s} \asymp n$$

Plus généralement, on va chercher $h = h_n$ vérifiant

$$h^\alpha e^{2A/h^r+2B/h^s} \asymp n$$

En effet, $MSE(x) = O(L(h))$ où $L(h) = h^{r-1}e^{-2A/h^r} + \frac{h^{s-1}}{n}e^{2B/h^s}$. Dans ce cas, le risque est minimum si h vérifie

$$e^{2A/h^r+2B/h^s} \asymp n$$

Soit $W(h) = h^\alpha e^{2A/h^r+2B/h^s}$ où $\alpha = 0$ si on s'intéresse au risque ponctuel et $\alpha = r - 1$ sinon. On cherche $h = h_n$ tel que $W(h_n) \asymp n$.

Il faut considérer 3 cas :

cas $r = s$

$$\text{Posons } h = \left(\frac{\ln n}{2A + 2B}\right)^{-1/s}$$

$$\begin{aligned} W(h) &= h^\alpha e^{(2A+2B)/h^s} \\ &= \left(\frac{\ln n}{2A + 2B}\right)^{-\alpha/s} e^{\ln n} \\ &= Kn(\ln n)^{-\alpha/s} \end{aligned}$$

Ce h ne convient que si $\alpha = 0$.

Si $\alpha \neq 0$, on prend $h = \left(\frac{\ln n + \frac{\alpha}{s} \ln \ln n}{2A+2B}\right)^{-1/s}$. On a alors

$$\begin{aligned} W(h) &= h^\alpha e^{(2A+2B)/h^s} \\ &= \left(\frac{\ln n + \frac{\alpha}{s} \ln \ln n}{2A + 2B}\right)^{-\alpha/s} e^{\ln n + \frac{\alpha}{s} \ln \ln n} \\ &= Kn(\ln n + \frac{\alpha}{s} \ln \ln n)^{-\alpha/s} (\ln n)^{\alpha/s} \\ &= Kn\left(1 + \frac{\frac{\alpha}{s} \ln \ln n}{\ln n}\right)^{-\alpha/s} \\ &= Kn(1 + o(1)) \asymp n \end{aligned}$$

Ce $h = h_{opt}$ convient donc quel que soit α .

On peut calculer le risque correspondant :

$$\begin{aligned} G(h_{opt}) &= e^{-\frac{A}{A+B}(\ln n + \frac{\alpha}{s} \ln \ln n)} + \left(\frac{\ln n + \frac{\alpha}{s} \ln \ln n}{2A + 2B}\right)^{\frac{1-s}{s}} \frac{1}{n} e^{\frac{B}{A+B}(\ln n + \frac{\alpha}{s} \ln \ln n)} \\ &= n^{-\frac{A}{A+B}} [(\ln n)^{-\frac{A}{A+B} \frac{\alpha}{s}} + \left(\frac{\ln n + \frac{\alpha}{s} \ln \ln n}{2A + 2B}\right)^{\frac{1-s}{s}} (\ln n)^{\frac{B}{A+B} \frac{\alpha}{s}}] \\ &= O(n^{-\frac{A}{A+B}} [(\ln n)^{-\frac{A}{A+B} \frac{s-1}{s}} + (\ln n)^{\frac{s-1}{s}(\frac{B}{A+B}-1)]) \quad \text{car } \alpha = s - 1 \\ &= O(n^{-\frac{A}{A+B}} (\ln n)^{-\frac{A}{A+B} \frac{s-1}{s}}) \end{aligned}$$

Donc

$$MISE = O(n^{-\frac{A}{A+B}} (\ln n)^{-\frac{A}{A+B} \frac{s-1}{s}})$$

Pour le risque ponctuel, on prend $\alpha = 0$ dans $L(h_{opt})$:

$$MSE(x) = O(n^{-\frac{A}{A+B}} (\ln n)^{\frac{1-s}{s}})$$

Remarque 6 Si $r = s = 1$, on obtient une vitesse linéaire ce qui est bien meilleur que la vitesse logarithmique du cas 2. Le fait que f soit supersmooth compense ici la trop grande régularité de q .

cas $r < s$

Notons $\lambda = \frac{r}{s}$ ($\lambda < 1$) et $c = \frac{2A}{(2B)^\lambda}$.

Le h optimal va dépendre de la valeur de λ .

Soit $k \in \mathbb{N}$ et

$$h = (2B)^{1/s} [\ln n + \frac{\alpha}{s} \ln \ln n + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}]^{-1/s}$$

où les b_i sont des réels qui seront précisés ultérieurement.

Calculons $W(h)$:

$$\begin{aligned} W(h) &= \exp\left[\frac{2B}{h^s} + \frac{2A}{h^r} + \alpha \ln h\right] \\ &= \exp\left[\ln n + \frac{\alpha}{s} \ln \ln n + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right. \\ &\quad \left. + \frac{2A}{(2B)^{r/s}} (\ln n)^{r/s} \left(1 + \frac{\alpha \ln \ln n}{s \ln n} + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-(i+1)}\right)^{r/s}\right. \\ &\quad \left. + \frac{\alpha}{s} \ln(2B) - \frac{\alpha}{s} \ln \ln n - \frac{\alpha}{s} \ln\left(1 + \frac{\alpha \ln \ln n}{s \ln n} + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-(i+1)}\right)\right] \\ &= Kn \exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right. \\ &\quad \left. + c(\ln n)^\lambda (1 + \lambda u_n + \dots + \frac{\lambda(\lambda-1)\dots(\lambda-k)}{(k+1)!} u_n^{k+1} + o(u_n^{k+1}))\right. \\ &\quad \left. - \frac{\alpha}{s} \ln(1 + o(1))\right] \end{aligned}$$

car $u_n \triangleq \frac{\alpha \ln \ln n}{s \ln n} + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-(i+1)} = o(1)$

On remarque que pour $j \in \{0, \dots, k-1\}$

$$u_n = \sum_{i=0}^j b_i (\ln n)^{(i+1)\lambda-(i+1)} + o((\ln n)^{(j+1)\lambda-(j+1)})$$

$$\begin{aligned} W(h) &= Kn \exp\left[o(1) + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i} + c(\ln n)^\lambda + c\lambda \sum_{i=0}^k b_i (\ln n)^{(i+2)\lambda-(i+1)}\right. \\ &\quad \left. + c \frac{\lambda(\lambda-1)}{2} (\ln n)^\lambda \left[\sum_{i=0}^{k-1} b_i (\ln n)^{(i+1)\lambda-(i+1)} + o((\ln n)^{k\lambda-k})\right]^2 + A_3 + \dots + A_k\right. \\ &\quad \left. + c \frac{\lambda(\lambda-1)\dots(\lambda-k)}{(k+1)!} (\ln n)^\lambda [b_0 (\ln n)^{\lambda-1} + o((\ln n)^{\lambda-1})]^{k+1}\right. \\ &\quad \left. + o((\ln n)^\lambda (\ln n)^{(k+1)(\lambda-1)})\right] \end{aligned}$$

avec, pour $2 \leq j \leq k+1$,

$$A_j = c \frac{\lambda \cdot (\lambda - j + 1)}{j!} (\ln n)^\lambda \left[\sum_{i=0}^{k-j+1} b_i (\ln n)^{(i+1)\lambda - (i+1)} + o((\ln n)^{(k+2-j)\lambda - (k+2-j)}) \right]^j$$

En fait

$$\begin{aligned} A_j &= (\ln n)^\lambda \left[\sum_{i=j}^{k+1} d_{i,j} (\ln n)^{i(\lambda-1)} + o((\ln n)^{(k+1)(\lambda-1)}) \right] \\ &= \sum_{i=j}^{k+1} d_{i,j} (\ln n)^{(i+1)\lambda - i} + o((\ln n)^{(k+2)\lambda - (k+1)}) \end{aligned}$$

où $\forall i \geq j$

$$d_{i,j} = c \frac{\lambda \cdot (\lambda - j + 1)}{j!} \sum_{p_1 + \dots + p_j = i} b_{p_1-1} \dots b_{p_j-1} \quad (4)$$

En effet, si on note $z = (\ln n)^{\lambda-1}$ (donc $z = o(1)$) et $l = k - j + 2$,

$$\begin{aligned} A_j &= c \frac{\lambda \cdot (\lambda - j + 1)}{j!} (\ln n)^\lambda \left[\sum_{p=1}^l b_{p-1} z^p + o(z^l) \right]^j \\ A_j &= c \frac{\lambda \cdot (\lambda - j + 1)}{j!} (\ln n)^\lambda \left[\sum_{i=j}^{k+1} \left(\sum_{p_1 + \dots + p_j = i} b_{p_1-1} \dots b_{p_j-1} \right) z^i + o(z^{k+1}) \right] \end{aligned}$$

Ainsi

$$\begin{aligned} W(h) &= Kn \exp[o(1) + (b_0 + c)(\ln n)^\lambda + \sum_{i=1}^k (b_i + \lambda c b_{i-1})(\ln n)^{(i+1)\lambda - i} \\ &\quad + \lambda c b_k (\ln n)^{(k+2)\lambda - (k+1)} + \sum_{i=2}^{k+1} d_{i,2} (\ln n)^{(i+1)\lambda - i} \\ &\quad + \dots + d_{k+1,k+1} (\ln n)^{(k+2)\lambda - (k+1)} + o((\ln n)^{(k+2)\lambda - (k+1)})] \\ &= Kn \exp[o(1) + \sum_{i=0}^{k+1} M_i (\ln n)^{(i+1)\lambda - i} + o((\ln n)^{(k+2)\lambda - (k+1)})] \end{aligned}$$

où

$$\begin{aligned} M_0 &= b_0 + c \\ M_1 &= b_1 + \lambda c b_0 \\ M_j &= b_j + \lambda c b_{j-1} + d_{j,2} + \dots + d_{j,j} \quad \forall 2 \leq j \leq k \\ M_{k+1} &= \lambda c b_k + d_{k+1,2} + \dots + d_{k+1,k+1} \end{aligned}$$

On peut observer sur la formule (4) que pour tout j et tout $i \geq j$, $d_{i,j}$ ne dépend que de $b_0, \dots, b_{i-1}$. On peut donc déterminer les b_i de manière unique en imposant les conditions $M_0 = 0, M_1 = 0, \dots, M_k = 0$. Dans ce cas, on a

$$W(h) = Kn \exp[o(1) + M_{k+1} (\ln n)^{(k+2)\lambda - (k+1)} + o((\ln n)^{(k+2)\lambda - (k+1)})]$$

Si $\lambda < \frac{k+1}{k+2}$ alors $(\ln n)^{(k+2)\lambda - (k+1)} \xrightarrow[n \rightarrow \infty]{} 0$.

Si $\lambda = \frac{k+1}{k+2}$ alors $\exp[M_{k+1} (\ln n)^{(k+2)\lambda - (k+1)}] = cste$.

Ainsi dans ces deux cas, $W(h) = Kn(1 + o(1)) \asymp n$. Si $\lambda > \frac{k+1}{k+2}$, ce n'est plus vrai et il faut modifier l'expression de h :

$$h = (2B)^{1/s} [\ln n + \frac{\alpha}{s} \ln \ln n + \sum_{i=0}^{k+1} b_i \lambda^i (\ln n)^{(i+1)\lambda-i}]^{-1/s}$$

Le même raisonnement avec un b_{k+1} adéquat montre que ce h convient si et seulement si $\lambda \leq \frac{k+2}{k+3}$.

Conclusion 1 h_{opt} dépend de la valeur de $\lambda : \forall k \in \mathbb{N}$

si $\frac{k}{k+1} < \lambda \leq \frac{k+1}{k+2}$ alors

$$h_{opt} = (2B)^{1/s} [\ln n + \frac{\alpha}{s} \ln \ln n + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}]^{-1/s}$$

avec

$$\begin{aligned} b_0 + c &= 0 \\ b_1 + \lambda c b_0 &= 0 \\ b_j + \lambda c b_{j-1} + d_{j,2} + \dots + d_{j,j} &= 0 \quad \forall 2 \leq j \leq k \end{aligned}$$

et $\forall i \geq j$, $d_{i,j} = c \frac{\lambda \cdot (\lambda - j + 1)}{j!} \sum_{p_1 + \dots + p_j = i} b_{p_1-1} \dots b_{p_j-1}$

On peut calculer le risque correspondant pour $\frac{k}{k+1} < \lambda \leq \frac{k+1}{k+2}$.

$$\begin{aligned} G(h_{opt}) &= \exp\left[-\frac{2A}{(2B)^{r/s}} (\ln n)^{r/s} (1 + u_n)^{r/s}\right] \\ &\quad + \frac{1}{n} \exp\left[\ln n + \frac{\alpha}{s} \ln \ln n + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right] \left(\frac{\ln n}{2B}\right)^{\frac{1-s}{s}} (1 + u_n)^{\frac{1-s}{s}} \\ &= \exp[-c(\ln n)^\lambda (1 + u_n)^\lambda] \\ &\quad + \exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right] K (\ln n)^{\frac{\alpha}{s} + \frac{1-s}{s}} (1 + o(1)) \end{aligned}$$

Or

$$\begin{aligned} c(\ln n)^\lambda (1 + u_n)^\lambda &= c(\ln n)^\lambda (1 + \lambda u_n) + A_2 + \dots + A_{k+1} + o((\ln n)^{(k+2)\lambda - (k+1)}) \\ &= c(\ln n)^\lambda + c\lambda \frac{\alpha}{s} \ln \ln n (\ln n)^{\lambda-1} + c\lambda \sum_{i=0}^k b_i (\ln n)^{(i+2)\lambda - (i+1)} \\ &\quad + \sum_{j=2}^{k+1} \sum_{i=j}^{k+1} d_{i,j} (\ln n)^{(i+1)\lambda - i} + o((\ln n)^{(k+2)\lambda - (k+1)}) \\ &= \sum_{i=0}^k (M_i - b_i) (\ln n)^{(i+1)\lambda - i} + o(1) \\ &= -\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda - i} + o(1) \end{aligned}$$

$$\begin{aligned}
G(h_{opt}) &= \exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i} + o(1)\right] \\
&\quad + \exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right] K (\ln n)^{\frac{\alpha}{s} + \frac{1-s}{s}} (1 + o(1)) \\
&= O\left(\exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right] (1 + (\ln n)^{\frac{\alpha}{s} + \frac{1-s}{s}})\right)
\end{aligned}$$

Pour $\alpha = r - 1$,

$$\frac{\alpha}{s} + \frac{1-s}{s} = \frac{r-s}{s} < 0$$

Donc

$$MISE = O\left(\exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right]\right)$$

Pour le risque ponctuel :

$$\begin{aligned}
L(h_{opt}) &= \exp\left[-\frac{2A}{(2B)^{r/s}} (\ln n)^{r/s} (1 + u_n)^{r/s}\right] \left(\frac{\ln n}{2B} (1 + u_n)\right)^{\frac{1-r}{s}} \\
&\quad + \frac{1}{n} \exp\left[\ln n + \frac{\alpha}{s} \ln \ln n + \sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right] \left(\frac{\ln n}{2B} (1 + u_n)\right)^{\frac{1-s}{s}} \\
&= O\left(\exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right] ((\ln n)^{\frac{1-r}{s}} + (\ln n)^{\frac{\alpha}{s} + \frac{1-s}{s}})\right) \\
&= O\left((\ln n)^{\frac{1-r}{s}} \exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right]\right)
\end{aligned}$$

puisque ici $\alpha = 0$ et $\frac{1-s}{s} < \frac{1-r}{s} < 0$. Ainsi

$$MSE(x) = O\left((\ln n)^{\frac{1-r}{s}} \exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}\right]\right)$$

cas $r > s$

Notons $\mu = \frac{s}{r}$ ($\mu < 1$) et $d = \frac{2B}{(2A)^\mu}$.

Le raisonnement est le même que dans le cas $r < s$ avec cette fois pour

$$\frac{k}{k+1} < \mu \leq \frac{k+1}{k+2}$$

$$h_{opt} = (2A)^{1/r} \left[\ln n + \frac{\alpha}{r} \ln \ln n + \sum_{i=0}^k a_i (\ln n)^{(i+1)\mu-i} \right]^{-1/r}$$

avec

$$\begin{aligned}
a_0 + d &= 0 \\
a_1 + \mu d a_0 &= 0 \\
a_j + \mu d a_{j-1} + d_{j,2} + \dots + d_{j,j} &= 0 \quad \forall 2 \leq j \leq k
\end{aligned}$$

$$\text{et } \forall i \geq j, \quad d_{i,j} = d^{\frac{\mu \cdot (\mu - j + 1)}{j!}} \sum_{p_1 + \dots + p_j = i} a_{p_1-1} \dots a_{p_j-1}$$

Ce h vérifie bien

$$h^\alpha e^{2A/h^r + 2B/h^s} \asymp n.$$

On note $v_n = \frac{\alpha}{r} \frac{\ln \ln n}{\ln n} + \sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - (i+1)}$.

Calculons le risque pour $\frac{k}{k+1} < \mu \leq \frac{k+1}{k+2}$.

$$\begin{aligned} G(h_{opt}) &= \exp\left[-\ln n + \frac{\alpha}{r} \ln \ln n + \sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i}\right] \\ &\quad + \frac{1}{n} \exp\left[\frac{2B}{(2A)^{s/r}} (\ln n)^{s/r} (1+v_n)^{s/r}\right] \left(\frac{\ln n}{2A} (1+v_n)\right)^{\frac{1-s}{r}} \\ &= \frac{1}{n} (\ln n)^{-\frac{\alpha}{r}} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i}\right] \\ &\quad + \frac{1}{n} \exp[d(\ln n)^\mu (1+v_n)^\mu] K(\ln n)^{\frac{1-s}{r}} (1+o(1)) \end{aligned}$$

Or, comme précédemment,

$$d(\ln n)^\mu (1+v_n)^\mu = -\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i} + o(1)$$

Donc

$$\begin{aligned} G(h_{opt}) &= \frac{1}{n} (\ln n)^{-\frac{\alpha}{r}} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i}\right] \\ &\quad + \frac{1}{n} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i} + o(1)\right] K(\ln n)^{\frac{1-s}{r}} (1+o(1)) \\ &= O\left(\frac{1}{n} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i}\right] \left((\ln n)^{-\frac{\alpha}{r}} + (\ln n)^{\frac{1-s}{r}}\right)\right) \end{aligned}$$

Or $\alpha = r - 1$ et $\frac{1-r}{r} < \frac{1-s}{r}$ donc

$$MISE = O\left(\frac{(\ln n)^{\frac{1-s}{r}}}{n} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i}\right]\right)$$

Calculons maintenant $L(h_{opt})$ (cas $\alpha = 0$)

$$\begin{aligned} L(h_{opt}) &= \exp\left[-\ln n + \frac{\alpha}{r} \ln \ln n + \sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i}\right] \left(\frac{\ln n}{2A} (1+v_n)\right)^{\frac{1-r}{r}} \\ &\quad + \frac{1}{n} \exp\left[\frac{2B}{(2A)^{s/r}} (\ln n)^{s/r} (1+v_n)^{s/r}\right] \left(\frac{\ln n}{2A} (1+v_n)\right)^{\frac{1-s}{r}} \\ &= O\left(\frac{1}{n} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i}\right] \left((\ln n)^{-\frac{\alpha}{r} + \frac{1-r}{r}} + (\ln n)^{\frac{1-s}{r}}\right)\right) \end{aligned}$$

Le résultat est le même,

$$MSE(x) = O\left(\frac{(\ln n)^{\frac{1-s}{r}}}{n} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu - i}\right]\right)$$

3.4.5 Récapitulatif

On peut récapituler les vitesses de convergence pour le risque L_2 dans un tableau.

	$q \in OS(\gamma)$	$q \in SS(B, s)$
$f \in OS(\delta)$	$n^{-\frac{2\delta}{2\delta+2\gamma+1}}$	$(\ln n)^{-\frac{2\delta}{s}}$
$f \in SS(A, r)$	$\frac{(\ln n)^{\frac{2\gamma+1}{r}}}{n}$	$n^{-\frac{A}{A+B}} (\ln n)^{-\frac{A}{A+B} \frac{s-1}{s}} \text{ si } r = s$ $\exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\frac{r}{s}-i}\right] \text{ si } \frac{k}{k+1} < \frac{r}{s} \leq \frac{k+1}{k+2}$ $\frac{(\ln n)^{\frac{1-s}{r}}}{n} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\frac{s}{r}-i}\right] \text{ si } \frac{k}{k+1} < \frac{s}{r} \leq \frac{k+1}{k+2}$

Pour le risque ponctuel, le tableau est légèrement différent :

	$q \in OS(\gamma)$	$q \in SS(B, s)$
$f \in OS(\delta)$	$n^{\frac{1-2\delta}{2\delta+2\gamma}}$	$(\ln n)^{\frac{1-2\delta}{s}}$
$f \in SS(A, r)$	$\frac{(\ln n)^{\frac{2\gamma+1}{r}}}{n}$	$n^{-\frac{A}{A+B}} (\ln n)^{\frac{1-s}{s}} \text{ si } r = s$ $(\ln n)^{\frac{1-r}{s}} \exp\left[\sum_{i=0}^k b_i (\ln n)^{(i+1)\frac{r}{s}-i}\right] \text{ si } \frac{k}{k+1} < \frac{r}{s} \leq \frac{k+1}{k+2}$ $\frac{(\ln n)^{\frac{1-s}{r}}}{n} \exp\left[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\frac{s}{r}-i}\right] \text{ si } \frac{k}{k+1} < \frac{s}{r} \leq \frac{k+1}{k+2}$

Dans tous les cas, les meilleures vitesses de convergence s'obtiennent avec une certaine fenêtre h_{opt} qui dépend de façon complexe de la régularité de f . On aimerait donc trouver un estimateur adaptatif c'est-à-dire un estimateur calculable sans connaître la régularité de f mais convergeant à la même vitesse vers f .

4 Mise en place d'un estimateur adaptatif

4.1 Contraste et estimateur

On suppose que $f \in L^2$. On note $\| \cdot \|$ la norme de L_2 . Soit t un estimateur appartenant à S_m où, quel que soit m «modèle», S_m est un sous-espace de L^2 . Par exemple,

$$S_m = \text{Vect}(\varphi(2^m x - j), j \in \mathbb{Z})$$

où $\varphi(x) = \frac{\sin \pi x}{\pi x}$. On a

$$\begin{aligned} \arg \min_{t \in L^2} \|t - f\|^2 &= \arg \min_{t \in L^2} (\|t\|^2 - 2 \int t(x)f(x)dx) \\ &= \arg \min_{t \in L^2} (\|t\|^2 - 2E[t(X_1)]) \end{aligned}$$

On cherche ψ_t telle que $E[\psi_t(Y_1)] = E[t(X_1)]$. Ainsi,

$$\gamma_n(t) = \|t\|^2 - \frac{2}{n} \sum_{i=1}^n \psi_t(Y_i)$$

sera un contraste.

En utilisant la formule de Plancherel, on a

$$\begin{aligned} E[t(X_1)] &= \int t(x)f(x)dx \\ &= \frac{1}{2\pi} \int \overline{\tilde{t}(u)} \tilde{f}(u) du \\ &= \frac{1}{2\pi} \int \frac{\overline{\tilde{t}(u)}}{\tilde{q}(u)} \tilde{p}(u) du \\ &= \frac{1}{2\pi} \int \frac{\overline{\tilde{t}(u)}}{\tilde{q}(u)} \int p(y) e^{-iyu} dy du \\ &= \frac{1}{2\pi} \int \int \frac{\overline{\tilde{t}(u)}}{\tilde{q}(u)} e^{-iyu} p(y) du dy \end{aligned}$$

Pour justifier la dernière égalité, il faut montrer que $\int \int |\frac{\tilde{t}(u)}{\tilde{q}(u)} p(y)| du dy < \infty$.

Or $\|p\|_1 = 1$. Et si on suppose que $\tilde{t}$ est bornée à support compact et que $q \in OS(\gamma)$ ou $q \in SS(B, s)$:

$$\int |\frac{\tilde{t}(u)}{\tilde{q}(u)}| du < \infty$$

$$\text{Donc } \int \int |\frac{\tilde{t}(u)}{\tilde{q}(u)} p(y)| du dy = \|p\|_1 \int |\frac{\tilde{t}(u)}{\tilde{q}(u)}| du < \infty.$$

De plus $\overline{\tilde{t}(u)} = \int e^{iux} t(x) dx = \tilde{t}(-u)$.

Donc en supposant $\tilde{t}$ bornée à support compact

$$E[t(X_1)] = \frac{1}{2\pi} \int \int \frac{\tilde{t}(-u)}{\tilde{q}(u)} e^{-iyu} p(y) du dy = \int \bar{\mathcal{F}}\left(\frac{\tilde{t}}{\tilde{q}}\right)(y) p(y) dy$$

car $\tilde{q}$ est paire.

Et finalement

$$\psi_t = \bar{\mathcal{F}}\left(\frac{\tilde{t}}{\tilde{q}}\right).$$

On remarque que ψ_t est linéaire en t et qu'elle est bien définie pour $t \in S_m$. En effet $\tilde{\varphi} = 1_{[-\pi, \pi]}$ est bornée à support compact. On a même $\psi_\varphi(x) = V_0(x)$ où

$$\tilde{V}_0 = \frac{1_{[-\pi, \pi]}}{\tilde{q}}.$$

Si on note $\varphi_{m,j}(x) = 2^{m/2} \varphi(2^m x - j)$, on a $\tilde{\varphi}_{m,j}(u) = 2^{-m/2} e^{-ij \frac{u}{2^m}} \tilde{\varphi}\left(\frac{u}{2^m}\right)$.

$$\begin{aligned} \psi_{\varphi_{m,j}}(x) &= \bar{\mathcal{F}}\left[2^{-m/2} e^{-ij \frac{u}{2^m}} \tilde{V}_m\left(\frac{u}{2^m}\right)\right] \\ &= 2^{m/2} V_m(2^m x - j) \\ &= V_{m,j}(x) \end{aligned}$$

où $\tilde{V}_m(u) = \frac{1_{[-\pi, \pi]}(u)}{\tilde{q}(2^m u)}$ et $V_{m,j}(x) = 2^{m/2} V_m(2^m x - j)$.

Si $t \in S_m$, $t = \sum_{j \in \mathbb{Z}} a_j \varphi_{m,j}$ et donc

$$\psi_t(x) = \sum_{j \in \mathbb{Z}} a_j V_{m,j}(x)$$

Puisque γ_n est un contraste, on définit l'estimateur suivant

$$\hat{f}_m = \arg \min_{t \in S_m} \gamma_n(t)$$

Or si $t \in S_m$, $t = \sum_{j \in \mathbb{Z}} a_j \varphi_{m,j}$ et

$$\gamma_n(t) = \sum_{j \in \mathbb{Z}} a_j^2 - \frac{2}{n} \sum_{i=1}^n \sum_{j \in \mathbb{Z}} a_j V_{m,j}(Y_i)$$

car $(\varphi_{m,j})_{j \in \mathbb{Z}}$ est une base orthonormale.

$$\frac{\partial \gamma_n(t)}{\partial a_k} = 2a_k - \frac{2}{n} \sum_{i=1}^n V_{m,k}(Y_i)$$

$$\Rightarrow \hat{a}_k = \frac{1}{n} \sum_{i=1}^n V_{m,k}(Y_i)$$

$$\hat{f}_m = \sum_{j \in \mathbb{Z}} \hat{a}_j \varphi_{m,j}$$

Remarquons que :

$$\begin{aligned} \mathbb{E}[\hat{a}_k] &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\psi_{\varphi_{m,k}}(Y_i)] \\ &= \mathbb{E}[\psi_{\varphi_{m,k}}(Y_1)] \\ &= \mathbb{E}[\varphi_{m,k}(X_1)] \\ &= \langle \varphi_{m,k}, f \rangle \end{aligned}$$

4.2 Evaluation du risque

On va calculer $\mathbb{E}\|\hat{f}_m - f\|^2$. Soit

$$f_m = \mathbb{E}[\hat{f}_m] = \sum_{j \in \mathbb{Z}} \mathbb{E}[\hat{a}_j] \varphi_{m,j} = \sum_{j \in \mathbb{Z}} \langle \varphi_{m,j}, f \rangle \varphi_{m,j}$$

f_m est la projection orthogonale de f sur S_m .

$$\|\hat{f}_m - f\|^2 = \|\hat{f}_m - f_m\|^2 + \|f_m - f\|^2$$

par le théorème de Pythagore.

Or $\mathbb{E}\|\hat{f}_m - f_m\|^2 = \sum_{j \in \mathbb{Z}} \mathbb{E}(\hat{a}_j - a_j)^2 = \sum_{j \in \mathbb{Z}} \text{Var}(\hat{a}_j)$

$$\begin{aligned} \sum_{j \in \mathbb{Z}} \text{Var}(\hat{a}_j) &\leq \sum_{j \in \mathbb{Z}} \frac{1}{n^2} \sum_{i=1}^n \mathbb{E}[V_{m,j}(Y_i)^2] \\ &\leq \sum_{j \in \mathbb{Z}} \frac{1}{n} \int V_{m,j}^2(x) p(x) dx \\ &\leq \frac{1}{n} \int \sum_{j \in \mathbb{Z}} V_{m,j}^2(x) p(x) dx \quad \text{par convergence monotone} \\ &\leq \frac{1}{n} \left\| \sum_{j \in \mathbb{Z}} V_{m,j}^2 \right\|_{\infty} \end{aligned}$$

On va maintenant majorer $\|\sum_{j \in \mathbb{Z}} V_{m,j}^2\|_\infty$

$$\begin{aligned} V_{m,j}(x) &= 2^{m/2} V_m(2^m x - j) = 2^{m/2} \bar{\mathcal{F}}(\tilde{V}_m)(2^m x - j) \\ &= \frac{1}{2\pi} \int 2^{m/2} e^{iu(2^m x - j)} \tilde{V}_m(u) du \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iuj} W_m(u) du = c_j(W_m) \end{aligned}$$

où W_m est la fonction 2π -périodique qui vaut $2^{m/2} e^{iu2^m x} \tilde{V}_m(u)$ sur l'intervalle $[-\pi, \pi]$ et $c_j(W_m)$ son j -ième coefficient de Fourier.

D'après le théorème de Parseval, on a

$$\sum_{j \in \mathbb{Z}} |c_j(W_m)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |W_m(u)|^2 du$$

$$\begin{aligned} \text{Donc } \sum_{j \in \mathbb{Z}} |V_{m,j}(x)|^2 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} 2^m |\tilde{V}_m(u)|^2 du \\ &= \frac{1}{2\pi} 2^m \int_{-\pi}^{\pi} |\tilde{q}(2^m u)|^{-2} du \end{aligned}$$

On note $D_m = 2^m$ et $\Delta(m) = \frac{1}{2\pi^2} \int_{-\pi}^{\pi} |\tilde{q}(2^m u)|^{-2} du$.

$$\begin{aligned} \Delta(m) &= \frac{1}{2\pi^2 D_m} \int_{-\pi D_m}^{\pi D_m} |\tilde{q}(u)|^{-2} du \\ &= \frac{h}{2\pi} \int_{-1/h}^{1/h} |\tilde{q}(u)|^{-2} du \end{aligned}$$

en posant $h = \frac{1}{\pi D_m}$. Ainsi, on retrouve exactement le terme $\|K\|_2^2$ de l'estimateur à noyau. Finalement

$$\|\sum_{j \in \mathbb{Z}} V_{m,j}^2\|_\infty = \pi D_m \Delta(m)$$

ce qui implique

$$\mathbb{E}\|\hat{f}_m - f\|^2 \leq \mathbb{E}\|f_m - f\|^2 + \frac{\pi D_m}{n} \Delta(m)$$

La variance est donc la même que pour l'estimateur à noyau (toujours avec $h = \frac{1}{\pi D_m}$). Calculons maintenant le biais.

$$\begin{aligned} \|f - f_m\|^2 &= \inf_{t \in S_m} \|f - t\|^2 \\ &\leq \|f - f * 2^{m/2} \varphi_{m,0}\|^2 \\ &\leq \frac{1}{2\pi} \int |\tilde{f}(u) - 2^{m/2} \tilde{f}(u) \tilde{\varphi}_{m,0}(u)|^2 du \\ &\leq \frac{1}{2\pi} \int |\tilde{f}(u)|^2 |1 - \tilde{\varphi}(u/2^m)|^2 du \\ &\leq \frac{1}{2\pi} \int_{|u| > 2^m \pi} |\tilde{f}(u)|^2 du \end{aligned}$$

Le biais est donc majoré de la même façon que pour l'estimateur à noyau. On obtient ainsi les mêmes vitesses de convergence que précédemment en choisissant m convenablement. Cependant, ce choix dépend de la régularité de f qui est inconnue. On cherche donc maintenant à déterminer à partir des seules observations $Y_1, \dots, Y_n$ le meilleur m possible, c'est-à-dire le modèle qui permet de se rapprocher le plus près possible des vitesses de convergence déterminées précédemment.

4.3 Sélection de modèles

Soit $\mathcal{M}_n$ l'ensemble des modèles m considérés.

$$\mathcal{M}_n = \{m \in \mathbb{N}, \frac{\pi D_m}{n} \Delta(m) \leq 1\}$$

(de telle sorte que la variance soit toujours bornée).

On pose

$$\hat{m} = \arg \min_{m \in \mathcal{M}_n} \{\gamma_n(\hat{f}_m) + pen(m)\}$$

où $pen(m)$ est une fonction croissante de m qui sera déterminée par la suite.

On s'intéresse à l'estimateur $\hat{f}_{\hat{m}}$. Par définition de $\hat{m}$,

$$\gamma_n(\hat{f}_{\hat{m}}) + pen(\hat{m}) \leq \gamma_n(f_m) + pen(m)$$

Or,

$$\begin{aligned} \gamma_n(\hat{f}_{\hat{m}}) - \gamma_n(f_m) &= \|\hat{f}_{\hat{m}}\|^2 - \|f_m\|^2 - \frac{2}{n} \sum_{i=1}^n \psi_{\hat{f}_{\hat{m}} - f_m}(Y_i) \\ &= \|\hat{f}_{\hat{m}} - f\|^2 - \|f_m - f\|^2 + 2 \langle \hat{f}_{\hat{m}} - f_m, f \rangle - \frac{2}{n} \sum_{i=1}^n \psi_{\hat{f}_{\hat{m}} - f_m}(Y_i) \end{aligned}$$

En posant $\nu_n(t) = \frac{1}{n} \sum_{i=1}^n \{\psi_t(Y_i) - \langle t, f \rangle\}$, on a

$$\gamma_n(\hat{f}_{\hat{m}}) - \gamma_n(f_m) = \|\hat{f}_{\hat{m}} - f\|^2 - \|f_m - f\|^2 - 2\nu_n(\hat{f}_{\hat{m}} - f_m)$$

Ainsi

$$\|\hat{f}_{\hat{m}} - f\|^2 - \|f_m - f\|^2 - 2\nu_n(\hat{f}_{\hat{m}} - f_m) \leq pen(m) - pen(\hat{m})$$

On peut donc calculer le risque de l'estimateur $\hat{f}_{\hat{m}}$:

$$\begin{aligned} \|\hat{f}_{\hat{m}} - f\|^2 &\leq \|f_m - f\|^2 + 2\nu_n(\hat{f}_{\hat{m}} - f_m) + pen(m) - pen(\hat{m}) \\ &\leq \|f_m - f\|^2 + 2\|\hat{f}_{\hat{m}} - f_m\| \nu_n\left(\frac{\hat{f}_{\hat{m}} - f_m}{\|\hat{f}_{\hat{m}} - f_m\|}\right) + pen(m) - pen(\hat{m}) \\ &\leq \|f_m - f\|^2 + \frac{1}{6}\|\hat{f}_{\hat{m}} - f_m\|^2 + 6\nu_n^2\left(\frac{\hat{f}_{\hat{m}} - f_m}{\|\hat{f}_{\hat{m}} - f_m\|}\right) + pen(m) - pen(\hat{m}) \\ &\leq \|f_m - f\|^2 + \frac{1}{3}\|\hat{f}_{\hat{m}} - f\|^2 + \frac{1}{3}\|f_m - f\|^2 + 6 \sup_{t \in \mathcal{T}(m, \hat{m})} \nu_n^2(t) \\ &\quad + pen(m) - pen(\hat{m}) \end{aligned}$$

où $\mathcal{T}(m, m') = \{t \in S_m + S_{m'}, \|t\| = 1\} = \{t \in S_{m \vee m'}, \|t\| = 1\}$

Soit $p(m, m')$ une fonction telle que

$$\forall m, m' \quad 6p(m, m') \leq pen(m) + pen(m') \quad (5)$$

$$\begin{aligned} \frac{2}{3}\|\hat{f}_{\hat{m}} - f\|^2 &\leq \frac{4}{3}\|f_m - f\|^2 + 6\left(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - p(m, \hat{m})\right)_+ \\ &\quad + 6p(m, \hat{m}) + pen(m) - pen(\hat{m}) \\ \|\hat{f}_{\hat{m}} - f\|^2 &\leq 2\|f_m - f\|^2 + 9 \sum_{m' \in \mathcal{M}_n} \left(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - p(m, m')\right)_+ + 3pen(m) \end{aligned}$$

Pour majorer le terme $\sum_{m' \in \mathcal{M}_n} \left(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - p(m, m')\right)_+$, on utilise le lemme suivant.

Lemme 4 (Talagrand(1996))

On suppose que

1. $\sup_{t \in \mathcal{T}(m, m')} \|\psi_t\|_\infty \leq M_1$
2. $\mathbb{E}(\sup_{t \in \mathcal{T}(m, m')} |\nu_n(t)|) \leq H$
3. $\sup_{t \in \mathcal{T}(m, m')} \text{Var}[\psi_t(Y_1)] \leq v.$

Alors, il existe $K > 0$, $K_1 > 0$, $K_2 > 0$ telles que pour tout réel ξ

$$\mathbb{E}(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - 2(4 + \xi^2)H^2)_+ \leq K \left[\frac{v}{n} e^{-K_1 \frac{nH^2\xi^2}{v}} + \frac{M_1^2}{n^2} e^{-K_2 \frac{nH\xi}{M_1}} \right]$$

On note pour la suite $m'' = m \vee m'$

1. Calcul de M_1

$$\begin{aligned} t \in \mathcal{T}(m, m') &\Rightarrow t = \sum_{j \in \mathbb{Z}} a_j \varphi_{m'', j} \Rightarrow \psi_t = \sum_{j \in \mathbb{Z}} a_j V_{m'', j} \\ \psi_t^2(x) &\leq \sum_{j \in \mathbb{Z}} a_j^2 \sum_{j \in \mathbb{Z}} V_{m'', j}^2(x) \leq \|t\|^2 \sum_{j \in \mathbb{Z}} V_{m'', j}^2(x) \\ \sup_{t \in \mathcal{T}(m, m')} \|\psi_t^2\|_\infty &\leq \left\| \sum_{j \in \mathbb{Z}} V_{m'', j}^2 \right\|_\infty \leq \pi D_m'' \Delta(m'') \\ M_1 &= \sqrt{\pi D_m'' \Delta(m'')} \end{aligned}$$

2. Calcul de H

Remarquons d'abord que

$$[\mathbb{E}(\sup_{t \in \mathcal{T}(m, m')} |\nu_n(t)|)]^2 \leq \mathbb{E}(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t))$$

Si $t \in \mathcal{T}(m, m')$ alors $\nu_n(t) = \sum_{j \in \mathbb{Z}} a_j \nu_n(\varphi_{m'', j})$

Donc

$$\begin{aligned} \nu_n^2(t) &\leq \sum_{j \in \mathbb{Z}} a_j^2 \sum_{j \in \mathbb{Z}} \nu_n^2(\varphi_{m'', j}) \leq \|t\|^2 \sum_{j \in \mathbb{Z}} \nu_n^2(\varphi_{m'', j}) \\ \sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) &\leq \sum_{j \in \mathbb{Z}} \nu_n^2(\varphi_{m'', j}) \end{aligned}$$

Or $\nu_n(\varphi_{m'', j}) = \frac{1}{n} \sum_{i=1}^n \{V_{m'', j}(Y_i) - \langle \varphi_{m'', j}, f \rangle\}$

$$\begin{aligned} \mathbb{E} \sum_{j \in \mathbb{Z}} [\nu_n(\varphi_{m'', j})]^2 &\leq \sum_{j \in \mathbb{Z}} \frac{1}{n^2} \sum_{i=1}^n \mathbb{E}[V_{m'', j}(Y_i)^2] \\ &\leq \sum_{j \in \mathbb{Z}} \frac{1}{n} \int V_{m'', j}^2(x) p(x) dx \\ &\leq \frac{1}{n} \int \sum_{j \in \mathbb{Z}} V_{m'', j}^2(x) p(x) dx \\ &\leq \frac{1}{n} \left\| \sum_{j \in \mathbb{Z}} V_{m'', j}^2 \right\|_\infty \\ &\leq \frac{\pi D_m''}{n} \Delta(m'') \end{aligned}$$

Ainsi $\mathbb{E}(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t)) \leq \frac{\pi D_{m''}}{n} \Delta(m'')$ et

$$H = \sqrt{\frac{\pi D_{m''}}{n} \Delta(m'')}$$

3. Calcul de v

Soit $t \in \mathcal{T}(m, m')$

$$\begin{aligned} \text{Var}[\psi_t(Y_1)] &\leq \int \psi_t^2(x) p(x) dx \leq \int \left(\sum_{j \in \mathbb{Z}} a_j V_{m'', j}(x) \right)^2 p(x) dx \\ &\leq \sum_{j, k \in \mathbb{Z}} a_j a_k \int V_{m'', j}(x) V_{m'', k}(x) p(x) dx \end{aligned}$$

Soit $\rho_{j, k} = \int V_{m'', j}(x) V_{m'', k}(x) p(x) dx$

$$\text{Var}[\psi_t(Y_1)] \leq \sum_{j \in \mathbb{Z}} a_j^2 \sqrt{\sum_{j, k \in \mathbb{Z}} |\rho_{j, k}|^2}$$

$$\sup_{t \in \mathcal{T}(m, m')} \text{Var}[\psi_t(Y_1)] \leq \sqrt{\sum_{j, k \in \mathbb{Z}} |\rho_{j, k}|^2}$$

$$\begin{aligned} \rho_{j, k} &= D_{m''} \int V_{m''}(D_{m''}x - j) V_{m''}(D_{m''}x - k) p(x) dx \\ &= \frac{D_{m''}}{4\pi^2} \int \int e^{-iu j} \tilde{V}_{m''}(u) e^{-iv k} \tilde{V}_{m''}(v) \int e^{iD_{m''}x(u+v)} p(x) dx dudv \\ &= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{-iu j} e^{-iv k} Z_{m''}(u, v) dudv = c_{j, k}(Z_{m''}) \end{aligned}$$

où $Z_{m''}$ est la fonction 2π -périodique qui vaut

$D_{m''} \tilde{V}_{m''}(u) \tilde{V}_{m''}(v) \tilde{p}(-D_{m''}(u+v))$ sur $[-\pi, \pi] \times [-\pi, \pi]$ et $c_{j, k}(Z_{m''})$ son j, k -ième coefficient de Fourier.

D'après le théorème de Parseval, on a

$$\sum_{j, k \in \mathbb{Z}} |c_{j, k}(Z_{m''})|^2 = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |Z_{m''}(u, v)|^2 dudv$$

$$\sum_{j, k \in \mathbb{Z}} |\rho_{j, k}|^2 = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} D_{m''}^2 |\tilde{V}_{m''}(u) \tilde{V}_{m''}(v) \tilde{p}(-D_{m''}(u+v))|^2 dudv \quad (6)$$

$$\begin{aligned} &\leq \frac{D_{m''}^2}{4\pi^2} \sqrt{\int \int |\tilde{V}_{m''}(u)|^4 |\tilde{p}(-D_{m''}(u+v))|^2 dudv} \sqrt{\int \int |\tilde{V}_{m''}(v)|^4 |\tilde{p}(-D_{m''}(u+v))|^2 dudv} \\ &\leq \frac{D_{m''}^2}{4\pi^2} \sqrt{\frac{1}{D_{m''}} \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} |\tilde{V}_{m''}(u)|^4 |\tilde{p}(x)|^2 dudx} \sqrt{\frac{1}{D_{m''}} \int_{-\pi}^{\pi} \int_{-\infty}^{\infty} |\tilde{V}_{m''}(v)|^4 |\tilde{p}(y)|^2 dydv} \\ &\leq \frac{D_{m''}}{4\pi^2} \int_{-\pi}^{\pi} |\tilde{V}_{m''}(u)|^4 du \int |\tilde{p}(x)|^2 dx \end{aligned}$$

Or $|\tilde{f}(x)| \leq \frac{1}{2\pi} \int |e^{ixu} f(u)| du \leq \frac{1}{2\pi}$. Donc

$$\int |\tilde{p}(x)|^2 dx \leq \frac{1}{4\pi^2} \int |\tilde{q}(x)|^2 dx \leq \frac{\|q\|^2}{2\pi}$$

$$\text{et } \sum_{j,k \in \mathbb{Z}} |\rho_{j,k}|^2 \leq \frac{\|q\|^2}{8\pi^3} D_{m''} \Delta_4(m'')$$

où $\Delta_4(m) = \int_{-\pi}^{\pi} |\tilde{q}(D_m u)|^{-4} du = \frac{1}{D_m} \int_{-\pi D_m}^{\pi D_m} |\tilde{q}(v)|^{-4} dv$.

On remarque aussi que $\forall u \quad |\tilde{p}(u)| \leq \frac{1}{2\pi}$. L'égalité (6) donne donc :

$$\begin{aligned} \sum_{j,k \in \mathbb{Z}} |\rho_{j,k}|^2 &\leq \frac{1}{(4\pi^2)} \int_{-\pi}^{\pi} D_{m''}^2 |\tilde{V}_{m''}(u) \tilde{V}_{m''}(v)|^2 dudv \\ &\leq \frac{D_{m''}^2}{(4\pi^2)} \left(\int_{-\pi}^{\pi} |\tilde{V}_{m''}(u)|^2 du \right)^2 \\ &\leq \frac{1}{4} D_{m''}^2 \Delta(m'')^2 \end{aligned}$$

Ainsi

$$v = \min \left\{ \frac{\|q\|}{\sqrt{8\pi^3}} \sqrt{D_{m''} \Delta_4(m'')}, \frac{1}{2} D_{m''} \Delta(m'') \right\}$$

4.3.1 premier cas : $q \in OS(\gamma)$

On a vu que $\Delta(m) = \|K\|_2^2$ avec $h = 1/\pi D_m$, donc pour q ordinary smooth, il existe $Q > 0$ ne dépendant que de Q_1 et γ tel que

$$\Delta(m) \leq Q D_m^{2\gamma}.$$

Lemme 5 Si $q \in OS(\gamma)$ alors il existe $C_0 > 0$ ne dépendant que de Q_1, Q_2 et γ tel que

$$\Delta(m) \geq C_0 \sqrt{\Delta_4(m)}$$

Démonstration :

$$\Delta_4(m) \leq \frac{1}{D_m Q_1^2} \int_{-\pi D_m}^{\pi D_m} (1+v^2)^{2\gamma} dv \leq M(\gamma, Q_1) D_m^{4\gamma}$$

D'autre part,

$$\begin{aligned} \Delta(m) &\geq \frac{1}{\pi^2 D_m Q_2} \int_0^{\pi D_m} (1+v^2)^\gamma dv \\ &\geq \frac{1}{\pi^2 D_m Q_2} \int_0^{\pi D_m} \frac{v}{\pi D_m} (1+v^2)^\gamma dv \\ &\geq \frac{D_m^{-2}}{2\pi^3 Q_2} \left[\frac{(1+v^2)^{\gamma+1}}{\gamma+1} \right]_0^{\pi D_m} \geq M'(\gamma, Q_2) D_m^{2\gamma} \end{aligned}$$

Finalement $\Delta_4(m) \leq M D_m^{4\gamma} \leq \frac{M}{M'^2} \Delta(m)^2$ donc il existe $C_0 = \frac{M'}{\sqrt{M}}$ telle que $C_0 \sqrt{\Delta_4(m)} \leq \Delta(m)$. $\square$

Donc

$$v = \frac{\|q\|}{\sqrt{8\pi^3}} \sqrt{D_{m''} \Delta_4(m'')} \leq \frac{\|q\|}{\sqrt{8\pi^3} C_0} \sqrt{D_{m''} \Delta(m'')}$$

On prend $\xi = 1$.

On peut maintenant appliquer l'inégalité de Talagrand :

$$\mathbb{E} \left(\sup_{t \in \mathcal{I}(m, m')} \nu_n^2(t) - 2(4 + \xi^2) H^2 \right)_+ \leq K' [U(m'') + V(m'')]$$

où $U(m'') = \frac{\sqrt{D_{m''}}\Delta(m'')}{e^{-K_3\sqrt{D_{m''}}}}$
 $V(m'') = \frac{D_{m''}\Delta(m'')}{n^2}e^{-K_2\sqrt{n}}$
 K_2 est une constante numérique
 K_3 et K' sont des constantes dépendant de Q_1, Q_2, γ et $\|q\|$.

$$V(m'') \leq \frac{1}{\pi n}e^{-K_2\sqrt{n}}$$

$$\sum_{m'' \in \mathcal{M}_n} V(m'') \leq \frac{|\mathcal{M}_n|}{\pi n}e^{-K_2\sqrt{n}}$$

Or

$$m \in \mathcal{M}_m \Rightarrow D_m \Delta(m) \leq \frac{n}{\pi}$$

$$\Rightarrow D_m^{2\gamma+1} \leq \frac{n}{\pi M'}$$

Donc $\exists \kappa |\mathcal{M}_n| \leq \kappa \ln n$ et il existe $C_1 > 0$ telle que $\sum_{m'' \in \mathcal{M}_n} V(m'') \leq \frac{C_1}{n}$

$$\begin{aligned} U(m'') &\leq \frac{Q}{n} D_{m''}^{2\gamma+1/2} e^{-K_3\sqrt{D_{m''}}} \\ &\leq \frac{Q}{n} (D_{m'}^{2\gamma+1/2} + D_m^{2\gamma+1/2}) e^{-\frac{K_3}{2}(\sqrt{D_{m'}} + \sqrt{D_m})} \\ \sum_{m'' \in \mathcal{M}_n} U(m'') &\leq \frac{Q}{n} e^{-\frac{K_3}{2}\sqrt{D_m}} \sum_{m'} D_{m'}^{2\gamma+1/2} e^{-\frac{K_3}{2}\sqrt{D_{m'}}} \\ &\quad + \frac{Q}{n} D_m^{2\gamma+1/2} e^{-\frac{K_3}{2}\sqrt{D_m}} \sum_{m'} e^{-\frac{K_3}{2}\sqrt{D_{m'}}} \\ &\leq \frac{C_2}{n} \end{aligned}$$

Donc

$$\mathbb{E}(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - 10H^2)_+ \leq \frac{C_3}{n}$$

où C_3 est une constante qui dépend de Q_1, Q_2, γ et $\|q\|$. On pose

$$p(m, m') = 10H^2 = \frac{10\pi}{n} D_{m''} \Delta(m'')$$

et

$$pen(m) = 60\pi Q \frac{D_m^{2\gamma+1}}{n}$$

de telle sorte que (5) soit vérifiée. On remarque que pen ne dépend que de q et donc $\hat{m} = \arg \min_{m \in \mathcal{M}_n} \{\gamma_n(\hat{f}_m) + pen(m)\}$ est bien calculable. Avec cette pénalisation, on a

$$\mathbb{E}(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - p(m, m'))_+ \leq \frac{C_3}{n}$$

Donc

$$\mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 \leq 2\mathbb{E}\|f_m - f\|^2 + \frac{9C_3}{n} + 180\pi Q \frac{D_m^{2\gamma+1}}{n}$$

$$\mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 \leq C_4 \left\{ \mathbb{E}\|f_m - f\|^2 + \frac{D_m^{2\gamma+1}}{n} \right\}$$

Théorème 1 On considère le modèle (1) avec les hypothèses H1, H2, H3 et H4. Soit

$$\gamma_n(t) = \|t\|^2 - \frac{2}{n} \sum_{i=1}^n \psi_t(Y_i) \quad \text{où} \quad \psi_t = \bar{\mathcal{F}}\left(\frac{\tilde{t}}{\tilde{q}}\right).$$

Soit $\hat{f}_m = \arg \min_{t \in S_m} \gamma_n(t)$ et $\hat{m} = \arg \min_{m \in \mathcal{M}_n} \{\gamma_n(\hat{f}_m) + \text{pen}(m)\}$ où $\text{pen}(m) = 60\pi Q \frac{D_m^{2\gamma+1}}{n}$ et Q est une constante ne dépendant que de Q_1 et γ . L'estimateur $\hat{f}_{\hat{m}}$ vérifie

$$\mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 \leq C_4 \inf_{m \in \mathcal{M}_n} \left\{ \mathbb{E}\|f_m - f\|^2 + \frac{D_m^{2\gamma+1}}{n} \right\}$$

où C_4 dépend de Q_1, Q_2, γ et $\|q\|$.

Cet estimateur converge donc vers f avec les vitesses obtenues dans la section 3 :

$$\begin{aligned} \mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 &= O(n^{-\frac{2\delta}{2\delta+2\gamma+1}}) \quad \text{si } f \in OS(\delta) \\ &= O\left(\frac{(\ln n)^{\frac{2\gamma+1}{r}}}{n}\right) \quad \text{si } f \in SS(A, r) \end{aligned}$$

4.3.2 deuxième cas : $q \in SS(B, s)$

On a à nouveau besoin d'un lemme :

Lemme 6 Si $q \in SS(B, s)$ alors il existe C_0 ne dépendant que de Q_1, Q_2, B et s telle que

$$\Delta(m) \geq C_0 D_m^{-s/2} \sqrt{\Delta_4(m)}$$

Démonstration :

$$\Delta_4(m) \leq \frac{2}{D_m Q_1^2} \int_0^{\pi D_m} e^{4Bv^s} dv$$

D'après le lemme 2, $\Delta_4(m) \leq M(Q_1, B, s) D_m^{-s} e^{4B\pi^s D_m^s}$.

Par ailleurs, d'après la remarque 3, $\Delta(m) \geq M'(Q_2, B, s) D_m^{-s} e^{2B\pi^s D_m^s}$. Finalement $D_m^{-s/2} \sqrt{\Delta_4(m)} \leq \frac{\sqrt{M}}{M'} \Delta(m)$ $\square$

On sait également qu'il existe une constante Q dépendant de Q_1, B et s telle que

$$\Delta(m) \leq Q D_m^{-s} e^{2B\pi^s D_m^s}$$

Ainsi

$$v \leq \min \left\{ \frac{\|q\|Q}{\sqrt{8\pi^3}C_0} D_{m''}^{\frac{1-s}{2}} e^{2B\pi^s D_{m''}^s}, \frac{Q}{2} D_{m''}^{1-s} e^{2B\pi^s D_{m''}^s} \right\}$$

Deux cas se présentent donc.

- si $s \geq 1$

Dans ce cas $v \leq \frac{Q}{2} D_{m''}^{1-s} e^{2B\pi^s D_{m''}^s}$

On applique l'inégalité de Talagrand :

$$\mathbb{E} \left(\sup_{t \in \mathcal{T}(m, m'')} \nu_n^2(t) - 2(4 + \xi^2)H^2 \right)_+ \leq K' [U(m'') + V(m'')]$$

$$\begin{aligned} \text{où} \quad U(m'') &= \frac{D_{m''}^{1-s}}{n} e^{2B\pi^s D_{m''}^s} e^{-K_3 \xi^2} \\ V(m'') &= \frac{D_{m''} \Delta(m'')}{n^2} e^{-K_2 \sqrt{n} \xi} \end{aligned}$$

K_2 est une constante numérique

K_3 et K' sont des constantes dépendant de Q_1, Q_2, B et s .

Soit $\xi^2 = \frac{4B\pi^s}{K_3} D_{m''}^s$ de telle sorte que

$$\begin{aligned}
2B\pi^s D_{m''}^s - K_3 \xi^2 &= -2B\pi^s D_{m''}^s \\
U(m'') &\leq \frac{D_{m''}^{1-s}}{n} e^{-2B\pi^s D_{m''}^s} \\
&\leq \frac{D_{m'}^{1-s} + D_m^{1-s}}{n} e^{-B\pi^s (D_m^s + D_{m'}^s)} \\
\sum_{m'' \in \mathcal{M}_n} U(m'') &\leq \frac{e^{-B\pi^s D_m^s}}{n} \sum_{m'} D_{m'}^{1-s} e^{-B\pi^s D_m^s} \\
&\quad + \frac{D_m^{1-s} e^{-B\pi^s D_m^s}}{n} \sum_{m'} e^{-B\pi^s D_m^s} \\
&\leq \frac{C_1}{n} \\
V(m'') &\leq \frac{1}{\pi n} e^{-K_2 \sqrt{n} \frac{4B\pi^s}{K_3} D_{m''}^s} \leq \frac{1}{\pi n} e^{-K_4 \sqrt{n} (D_m^s + D_{m'}^s)} \\
\sum_{m'' \in \mathcal{M}_n} V(m'') &\leq \frac{|\mathcal{M}_n|}{\pi n} e^{-K_4 \sqrt{n} D_m^s} + \frac{1}{\pi n} \sum_{m'} e^{-K_4 \sqrt{n} D_{m'}^s}
\end{aligned}$$

Or

$$\begin{aligned}
m \in \mathcal{M}_m &\Rightarrow D_m \Delta(m) \leq \frac{n}{\pi} \\
&\Rightarrow D_m^{1-s} e^{2B\pi^s D_m^s} \leq \frac{n}{\pi M'}
\end{aligned}$$

Donc $\exists \kappa |\mathcal{M}_n| \leq \kappa \ln \ln n$ et il existe $C_2 > 0$ telle que $\sum_{m'' \in \mathcal{M}_n} V(m'') \leq \frac{C_2}{n}$ Donc

$$\mathbb{E} \left(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - 2(4 + \xi^2) H^2 \right)_+ \leq \frac{C_3}{n}$$

où C_3 est une constante qui dépend de Q_1, Q_2, B et s . On pose

$$p(m, m') = 2(4 + \frac{4B\pi^s}{K_3} D_{m''}^s) \frac{\pi}{n} D_{m''} \Delta(m'')$$

et

$$pen(m) = a \frac{D_m}{n} e^{2B\pi^s D_m^s}$$

où a est une constante qui dépend de B, s, Q_1 et Q_2 de telle sorte que (5) soit vérifiée.

- si $s < 1$

Dans ce cas $v \leq \frac{\|q\|Q}{\sqrt{8\pi^s} C_0} D_{m''}^{\frac{1-s}{2}} e^{2B\pi^s D_{m''}^s}$

On applique l'inégalité de Talagrand :

$$\mathbb{E} \left(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - 2(4 + \xi^2) H^2 \right)_+ \leq K' [U(m'') + V(m'')]$$

$$\begin{aligned}
\text{où } U(m'') &= \frac{D_{m''}^{\frac{1-s}{2}}}{n} e^{2B\pi^s D_{m''}^s} e^{-K_3 D_{m''}^{\frac{1-s}{2}} \xi^2} \\
V(m'') &= \frac{D_{m''} \Delta(m'')}{n^2} e^{-K_2 \sqrt{n} \xi}
\end{aligned}$$

K_2 est une constante numérique

K_3 et K' sont des constantes dépendant de Q_1, Q_2, B, s et $\|q\|$.

Il faut à nouveau distinguer deux cas.

– si $s < \frac{1}{3}$

On a alors $\frac{1-s}{2} > s$ et on peut poser $\xi^2 = 1$

$$U(m'') \leq \frac{D_{m''}^{\frac{1-s}{2}}}{n} e^{2B\pi^s D_{m''}^s - K_3 D_{m''}^{\frac{1-s}{2}}}$$

$$\sum_{m' \in \mathcal{M}_n} U(m'') \leq \frac{C_1}{n}$$

$$V(m'') \leq \frac{1}{\pi n} e^{-K_2 \sqrt{n}}$$

$$\sum_{m' \in \mathcal{M}_n} V(m'') \leq \frac{C_2}{n}$$

Donc

$$\mathbb{E} \left(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - 10H^2 \right)_+ \leq \frac{C_3}{n}$$

où C_3 est une constante qui dépend de Q_1, Q_2, B, s et $\|q\|$. On pose

$$p(m, m') = \frac{10\pi}{n} D_{m''} \Delta(m'')$$

et

$$pen(m) = a \frac{D_m^{1-s}}{n} e^{2B\pi^s D_m^s}$$

où a est une constante qui dépend de B, s et Q_1 de telle sorte que (5) soit vérifiée.

– si $s \geq \frac{1}{3}$

On a alors $\frac{1-s}{2} \leq s$ et on pose $\xi^2 = \frac{4B\pi^s}{K_3} D_{m''}^{\frac{3s-1}{2}}$

$$U(m'') \leq \frac{D_{m''}^{\frac{1-s}{2}}}{n} e^{-2B\pi^s D_{m''}^s}$$

$$\sum_{m' \in \mathcal{M}_n} U(m'') \leq \frac{C_1}{n}$$

$$V(m'') \leq \frac{1}{\pi n} e^{-K_4 \sqrt{n} D_{m''}^{\frac{3s-1}{4}}}$$

$$\sum_{m' \in \mathcal{M}_n} V(m'') \leq \frac{C_2}{n}$$

Donc

$$\mathbb{E} \left(\sup_{t \in \mathcal{T}(m, m')} \nu_n^2(t) - 2(4 + \xi^2)H^2 \right)_+ \leq \frac{C_3}{n}$$

où C_3 est une constante qui dépend de Q_1, Q_2, B, s et $\|q\|$. On pose

$$p(m, m') = 2 \left(4 + \frac{4B\pi^s}{K_3} D_{m''}^{\frac{3s-1}{2}} \right) \frac{\pi D_{m''}}{n} \Delta(m'')$$

et

$$pen(m) = a \frac{D_m^{\frac{1+s}{2}}}{n} e^{2B\pi^s D_m^s}$$

où a est une constante qui dépend de B, s et Q_1 et $\|q\|$ de telle sorte que (5) soit vérifiée.

Ainsi, dans tous les cas

$$\mathbb{E}(\sup_{t \in T(m, m')} \nu_n^2(t) - p(m, m'))_+ \leq \frac{C_3}{n}$$

mais

$$\begin{aligned} \text{pen}(m) &= \frac{a}{n} D_m^{1-s} e^{2B\pi^s D_m^s} & \text{si } s < \frac{1}{3} \\ & \frac{a}{n} D_m^{\frac{1+s}{2}} e^{2B\pi^s D_m^s} & \text{si } \frac{1}{3} \leq s < 1 \\ & \frac{a}{n} D_m e^{2B\pi^s D_m^s} & \text{si } s \geq 1 \end{aligned}$$

On peut écrire $\text{pen}(m)$ de la façon suivante :

$$\text{pen}(m) = \frac{a}{n} D_m^{1-[s-(\frac{3s-1}{2})_+]_+} e^{2B\pi^s D_m^s}$$

Théorème 2 *On considère le modèle (1) avec les hypothèses H1, H2, H3 et H4. Soit*

$$\gamma_n(t) = \|t\|^2 - \frac{2}{n} \sum_{i=1}^n \psi_t(Y_i) \quad \text{où } \psi_t = \bar{\mathcal{F}}\left(\frac{\tilde{t}}{q}\right).$$

Soit $\hat{f}_m = \arg \min_{t \in S_m} \gamma_n(t)$ et $\hat{m} = \arg \min_{m \in \mathcal{M}_n} \{\gamma_n(\hat{f}_m) + \text{pen}(m)\}$

où $\text{pen}(m) = \frac{a}{n} D_m^{1-[s-(\frac{3s-1}{2})_+]_+} e^{2B\pi^s D_m^s}$ et a est une constante ne dépendant que de Q_1, Q_2, B, s et $\|q\|$. L'estimateur $\hat{f}_{\hat{m}}$ vérifie

$$\mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 \leq C_4 \inf_{m \in \mathcal{M}_n} \{\mathbb{E}\|f_m - f\|^2 + \frac{D_m^{1-[s-(\frac{3s-1}{2})_+]_+} e^{2B\pi^s D_m^s}}{n}\}$$

où C_4 dépend de Q_1, Q_2, B, s et C .

C'est seulement dans le cas où $s < \frac{1}{3}$ qu'on retrouve le même ordre de la variance que précédemment.

Si $f \in OS(\delta)$, on peut remarquer que la puissance de D_m ne modifie pas le résultat final (elle ne change que le m qui réalise l'inf).

$$\mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 = O((\ln n)^{-\frac{2\delta}{s}})$$

Si $f \in SS(A, r)$, le résultat n'est modifié que si $r \geq s$. Avec les notations de la section 2.4.4, on a

- si $r = s$

$$\mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 = O(n^{-\frac{A}{A+B}} (\ln n)^{\frac{A}{A+B} \frac{1-[s-(\frac{3s-1}{2})_+]_+}{s}})$$

- si $r < s$

$$\mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 = O(\exp[\sum_{i=0}^k b_i (\ln n)^{(i+1)\lambda-i}])$$

- si $r > s$

$$\mathbb{E}\|\hat{f}_{\hat{m}} - f\|^2 = O\left(\frac{(\ln n)^{\frac{1-[s-(\frac{3s-1}{2})_+]_+}{r}}}{n} \exp[-\sum_{i=0}^k a_i (\ln n)^{(i+1)\mu-i}]\right)$$

Remarque 7 Dans le cas $r = s = 1$, on obtient la vitesse $(\frac{\ln n}{n})^{\frac{A}{A+B}}$: on a donc une perte logarithmique lorsqu'on passe à l'estimation adaptative comme l'a remarqué Tsybakov (2000).

References

- [1] BARRON,A.R., BIRGÉ,L. et MASSART,P. (1999) Risk bounds for model selection via penalization. *Probab. Theory Relat. Fields* **113**, 301-413.
- [2] BIRGÉ,L. et MASSART,P. (1997) From model selection to adaptative estimation. In : D.Pollard, E.Torgersen et G.Yang, eds. *Festschrift for Lucien Le Cam: Research Papers in Probability and Statistics* (Springer-Verlag,New-York) pp. 55-87.
- [3] BIRGÉ,L. et MASSART,P. (1998) Minimum contrast estimators on sieves : exponential bounds and rates of convergence. *Bernoulli* **4**, 329-375.
- [4] BUTUCEA,C. (2003) Deconvolution of analytic densities with smooth noise. *Working Paper*.
- [5] CARROL,R.J. et HALL,P. (1988) Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.* **83**, 1184-1186.
- [6] DEVROYE,L. (1989) Consistent deconvolution in density estimation. *Canad. Journ. Statist.* **17**, 235-239.
- [7] FAN,J. (1991) On the optimal rates of convergence for nonparametric deconvolution problem. *Ann. Statist.* **19**, 1257-1272.
- [8] FAN,J. (1993) Adaptatively local one-dimensional subproblems with application to a deconvolution problem. *Ann. Statist.* **21**, 600-610.
- [9] LIU,M.C. et TAYLOR,R.L.(1989) A consistent non-parametric density estimator for the deconvolution problem. *Canad. Journ. Statist.* **17**, 427-438.
- [10] GOLDENSCHLUGER,A. (1999) On pointwise adaptative nonparametric deconvolution. *Bernoulli* **5**, 907-926.
- [11] MASRY,E. (1993a) Strong consistency and rates for deconvolution of multivariate densities of stationary processes. *Stoch. Processes Appl.* **47**, 53-74.
- [12] MASRY,E. (1993b) Asymptotic normality for deconvolution estimators of multivariate densities of stationary processes. *J. Multivariate Anal.* **44**, 47-68.
- [13] PENSKY,M. et VIDAKOVIC,B. (1999) Adaptative wavelet estimator for non-parametric density deconvolution. *Ann. Statist.* **27**, 2033-2053.
- [14] PENSKY,M. (2002) Density deconvolution based on wavelets with bounded supports. *Stat. Probab. Lett.* **56**, 261-269.
- [15] PITERBARG,V. et PENSKAYA,M.(1993) On asymptotic distribution of integrated squared error of an estimate of a component of a convolution. *Math. Methods Statistics* **2**, 30-41.
- [16] STEFANSKI,L. (1990) Rates of convergence of some estimators in a class of deconvolution problems. *Statist. Prob. Lett.* **9**, 229-235.

- [17] STEFANSKI,L. et CARROL,R.J. (1990) Deconvoluting kernel density estimators. *Statistics* **21**, 169-184.
- [18] TALAGRAND,M. (1996) New concentration inequalities in product spaces. *Invent. Math.* **126**, 505-563.
- [19] TAYLOR,R.L. et ZHANG,H.M. (1990) On a strongly consistent non-parametric density estimator for deconvolution problem. *Commun. Statist Theory Meth.* **19**, 3325-3342.
- [20] TSYBAKOV,A.B. (2000) On the best rate of adaptative estimation in some inverse problems. *C. R. Acad. Sci. Paris* **330**,Série I, 835-840.
- [21] ZHANG,C.H. (1990) Fourier methods for estimating mixing densities and distributions. *Ann. Statist.* **18**, 806-831.