Supplementary Material for “Adaptive kernel estimation of the baseline function in the Cox model, with high-dimensional covariates”

Agathe Guilloux
Laboratoire de Statistique Théorique et Appliquée,
Université Pierre et Marie Curie - Paris 6
e-mail : agathe.guilloux@upmc.fr

Sarah Lemler
Laboratoire de Mathématiques et Modélisation d'Évry UMR CNRS 8071- USC INRA,
Université d'Évry Val d'Essonne, France
e-mail : sarah.lemler@genopole.cnrs.fr

Marie-Luce Taupin
Laboratoire de Mathématiques et Modélisation d'Évry UMR CNRS 8071- USC INRA,
Université d'Évry Val d'Essonne, France
e-mail : marie-luce.taupin@genopole.cnrs.fr
Unité MaIAGE, INRA Jouy-En-Josas, France

1 Technical lemmas

1.1 Proof of Lemma 6.2

From Definition (20), we have $\alpha_h(t) = K_h * \alpha_0(t)$ for all $t \in [0, \tau]$ so that

$$\mathbb{E}[||\bar{\alpha}_h - \alpha_0||_2^2] = \mathbb{E}[||\bar{\alpha}_h - \alpha_h||_2^2] + ||\alpha_h - \alpha_0||_2^2.$$

The first term of the right part of this equality can be rewritten as

$$\mathbb{E}[||\bar{\alpha}_h - \alpha_h||_2^2] = \int_0^\tau \mathrm{Var}[\bar{\alpha}_h(t)]\mathrm{d}t.$$

It remains to bound $\mathrm{Var}[\bar{\alpha}_h(t)]$:

$$\begin{aligned}\mathrm{Var}[\bar{\alpha}_h(t)] &= \frac{1}{n}\,\mathrm{Var}\left[\int_0^\tau K_h(t-u)\frac{1}{S(u,\boldsymbol{\beta_0})}\mathrm{d}N_1(u)\right] \\ &\leq \frac{1}{n}\mathbb{E}\left[\left(\int_0^\tau K_h(t-u)\frac{1}{S(u,\boldsymbol{\beta_0})}\mathrm{d}N_1(u)\right)^2\right].\end{aligned}$$

We apply the Doob-Meyer decomposition $N_1 = M_1 + \Lambda_1$:

$$\begin{aligned}\mathrm{Var}[\bar{\alpha}_h(t)] \leq& \frac{2}{n}\mathbb{E}\left[\left(\int_0^\tau \frac{K_h(t-u)}{S(u,\boldsymbol{\beta_0})}\mathrm{d}M_1(u)\right)^2\right] \\ &+ \frac{2}{n}\mathbb{E}\left[\left(\int_0^\tau \frac{K_h(t-u)}{S(u,\boldsymbol{\beta_0})}\alpha_0(u)\mathrm{e}^{\boldsymbol{\beta_0^T Z_1}}Y_1(u)\mathrm{d}u\right)^2\right]\end{aligned}$$

Then, we have

$$\mathbb{E}\left[\left(\int_0^\tau \frac{K_h(t-u)}{S(u,\boldsymbol{\beta_0})}\mathrm{d}M_1\right)^2\right] \leq \mathbb{E}\left[\int_0^\tau \frac{K_h^2(t-u)}{S^2(u,\boldsymbol{\beta_0})}\alpha_0(u)\mathrm{e}^{\boldsymbol{\beta_0^T Z_1}}Y_1(u)\mathrm{d}u\right],$$

and we finally get that

$$\mathbb{E}\left[\left(\int_0^\tau \frac{K_h(t-u)}{S(u,\boldsymbol{\beta_0})}\mathrm{d}M_1\right)^2\right] \leq \frac{\mathbb{E}[\mathrm{e}^{\boldsymbol{\beta_0^T Z_1}}]||\alpha_0||_{\infty,\tau}\tau}{c_S^2}\frac{||K||^2_{\mathbb{L}^2(\mathbb{R})}}{h}.$$

Thus, the integrated variance term of the pseudo-estimator is bounded by

$$\int_0^\tau \mathrm{Var}[\bar{\alpha}_h(t)]\mathrm{d}t \leq \frac{2||\alpha_0||_{\infty,\tau}\tau}{c_S^2}(\mathbb{E}[\mathrm{e}^{\boldsymbol{\beta_0^T Z_1}}] + ||\alpha_0||_{\infty,\tau}\mathbb{E}[\mathrm{e}^{2\boldsymbol{\beta_0^T Z_1}}]\tau)\frac{||K||^2_{\mathbb{L}^2(\mathbb{R})}}{nh},$$

which gives a bound of order $1/nh$.

Gathering the bias term $||\alpha_h - \alpha_0||_2^2$ and the bound on variance term gives Inequality (22) in Lemma 6.2.

□

1.2 Proof of Lemma 6.3

The proof of Lemma 6.3 relies on an additional lemmas. First, write

$$\hat{\alpha}_h^{\hat{\beta}}(t) - \bar{\alpha}_h(t) = \frac{1}{nh}\sum_{i=1}^n \int_0^\tau \frac{S(u,\boldsymbol{\beta_0}) - S_n(u,\hat{\boldsymbol{\beta}})}{S(u,\boldsymbol{\beta_0})S_n(u,\hat{\boldsymbol{\beta}})}K\left(\frac{t-u}{h}\right)\mathbb{1}_{\{\bar{Y}(u)>0\}}\mathrm{d}N_i(u).$$

We study the difference process $\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h$ on Ω_k, defined by (32) and on its complement. From Lemma 6.6, the process $\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h$ is controled on Ω_k^c. The following lemma allows to bound the difference process on Ω_k.

Lemma 1.1. *Under Assumptions 3.4.(ii)-(iii) , 3.5.(i)-(iii), 3.1 and 3.2, for any $k \in \mathbb{N}$, we have*

$$\mathbb{E}[||\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h||_2^2 \mathbb{1}(\Omega_k)] \leq c_4(s)\frac{\log(n^k p)}{n},$$

where c_4 is a constant depending on $B, |\boldsymbol{\beta_0}|_1, R, ||\alpha_0||_{\infty,\tau}, c_S, ||K||_{\mathbb{L}^2(\mathbb{R})}, \tau$ and s the sparsity index of $\boldsymbol{\beta_0}$.

Gathering Lemmas 6.6 and 1.1, we finally get that, for a fixed k

$$\mathbb{E}[||\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h||_2^2] \leq c(s)\frac{\log(n^k p)}{n},$$

with $c(s)$ a constant depending on $B, |\boldsymbol{\beta_0}|_1, R, ||\alpha_0||_{\infty,\tau}, c_S, ||K||_{\mathbb{L}^2(\mathbb{R})}, \tau$ and s the sparsity index of $\boldsymbol{\beta_0}$. and Lemma 6.3 is then proved. Let now prove the Lemmas 6.6 and 1.1.

1.3 Proof of Lemma 6.6 :

We have to bound $\mathbb{E}[||\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h||_2^2 \mathbb{1}(\Omega_k^c)]$, which is equal to

$$\mathbb{E}[||\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h||_2^2 \mathbb{1}(\Omega_k^c)] = \int_0^\tau \mathbb{E}[(\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h)^2(t)\mathbb{1}(\Omega_k^c)]\mathrm{d}t.$$

First, let us focus on $\mathbb{E}[(\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h)^2(t)\mathbb{1}(\Omega_k^c)]$ defined by

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n \int_0^\tau K_h(t-u)\mathbb{1}_{\{\bar{Y}(u)>0\}}\frac{S(u,\boldsymbol{\beta_0}) - S_n(u,\hat{\boldsymbol{\beta}})}{S(u,\boldsymbol{\beta_0})S_n(u,\hat{\boldsymbol{\beta}})}\mathrm{d}N_i(t)\right)^2 \mathbb{1}(\Omega_k^c)\right].$$

From Assumptions 3.1, $\hat{\boldsymbol{\beta}}$ belongs to a ball $\boldsymbol{\mathcal{B}}(0,R)$ and $|\boldsymbol{\beta_0}|_1 < \infty$, so we have the following bound

$$S_n(u,\hat{\boldsymbol{\beta}}) - S(u,\boldsymbol{\beta_0}) \leq \frac{1}{n}\sum_{i=1}^n \mathrm{e}^{\boldsymbol{\beta_0^T Z_i}}\mathrm{e}^{\boldsymbol{(\hat{\beta}-\beta_0)^T Z_i}} \leq \mathrm{e}^{2B|\boldsymbol{\beta_0}|_1}\mathrm{e}^{BR}. \tag{1}$$

For sake of simplicity, let us denote $C(B,R,|\boldsymbol{\beta_0}|_1)$ the bound in (1). From $\mathbb{1}_{\{\bar{Y}(u)>0\}}$ in the definition of $\hat{\alpha}_h^{\hat{\beta}}$, there exists $i_0 \in \{1,...,n\}$ such that $Y_{i_0} \neq 0$, so that from Assumption 3.4.(i)

$$S_n(u,\hat{\boldsymbol{\beta}}) \geq \frac{1}{n}\mathrm{e}^{-B|\boldsymbol{\beta_0}|_1}\mathrm{e}^{-B|\hat{\boldsymbol{\beta}}-\boldsymbol{\beta_0}|_1} \geq \frac{1}{n}e^{-2B|\boldsymbol{\beta_0}|_1}\mathrm{e}^{-BR}. \tag{2}$$

Combining (1) and (2), for $\tilde{C}(B,R,|\boldsymbol{\beta_0}|_1) = \mathrm{e}^{8B|\boldsymbol{\beta_0}|_1}\mathrm{e}^{4BR}$, we obtain the following bound

$$\mathbb{E}[(\hat{\alpha}_h^{\hat{\beta}} - \bar{\alpha}_h)^2(t)\mathbb{1}(\Omega_k^c)] \leq \tilde{C}(B,R,|\boldsymbol{\beta_0}|_1)\frac{n^2}{c_S^2}\mathbb{E}\Big[\Big(\int_0^\tau K_h(t-u)\mathrm{d}N_1(u)\Big)^2\mathbb{1}(\Omega_k^c)\Big]. \tag{3}$$

From Assumption 3.5.(i) and the Cauchy-Schwarz Inequality, we get

$$\mathbb{E}\Big[\Big(\int_0^\tau K_h(t-u)\mathrm{d}N_1(u)\Big)^2\mathbb{1}(\Omega_k^c)\Big] \leq \frac{||K||_\infty^2}{h^2}\mathbb{E}[N_1(\tau)^2\mathbb{1}(\Omega_k^c)]$$

$$\leq \frac{||K||_\infty^2}{h^2}\mathbb{E}[N_1(\tau)^4]^{1/2}\mathbb{P}(\Omega_k^c)^{1/2}. \qquad (4)$$

From the Doob-Meyer decomposition and the Bürkholder Inequality 2.4, we deduce that $\mathbb{E}[N_1(\tau)] < \infty$.

Now we focus on $\int_0^\tau \mathbb{E}[(\hat{\alpha}_h^{\hat{\boldsymbol{\beta}}} - \bar{\alpha}_h)^2(t)\mathbb{1}(\Omega_k^c)]\mathrm{d}t$. From the two bounds (3) and (4) obtained above, we have

$$\int_0^\tau \mathbb{E}[(\hat{\alpha}_h^{\hat{\boldsymbol{\beta}}} - \bar{\alpha}_h)^2(t)\mathbb{1}(\Omega_k^c)]\mathrm{d}t \leq \tilde{C}(B,R,|\boldsymbol{\beta_0}|_1)\frac{n^2}{c_S^2}\Big(\frac{||K||_\infty^2}{h^2}\mathbb{E}^{1/2}[N_1(\tau)^4]\sqrt{\mathbb{P}(\Omega_k^c)}\Big).$$

Let introduce the following lemma that gives a bound for $\mathbb{P}(\Omega_k^c)$.

Lemma 1.2. *Under Assumptions 3.1 and 3.2, for all $k \in \mathbb{N}$, there exists $n_0 \in \mathbb{N}$, such that for $n > n_0$ we have*

$$\mathbb{P}[\Omega_k^c] \leq c_2 n^{-k}, \qquad (5)$$

where c_2 is a constant depending on B, $|\boldsymbol{\beta_0}|_1$ and s.

From Lemma 1.2 and the fact that $h^{-1} \leq n$ from Assumption 3.5.(ii), we get that

$$\int_0^\tau \mathbb{E}[(\hat{\alpha}_h^{\hat{\boldsymbol{\beta}}} - \bar{\alpha}_h)^2(t)\mathbb{1}(\Omega_k^c)]\mathrm{d}t \leq C(B,|\boldsymbol{\beta_0}|_1,R,||\alpha_0||_{\infty,\tau},c_S,\tau,||K||_\infty)n^{4-k/2},$$

where $C(B,|\boldsymbol{\beta_0}|_1,R,||\alpha_0||_{\infty,\tau},c_S,\tau,||K||_\infty)$ is a constant depending on elements in brackets. Finally, we obtain

$$\mathbb{E}[||\hat{\alpha}_h^{\hat{\boldsymbol{\beta}}} - \bar{\alpha}_h||_2^2\mathbb{1}(\Omega_k^c)] \leq C(B,|\boldsymbol{\beta_0}|_1,R,||\alpha_0||_{\infty,\tau},c_S,\tau,||K||_\infty)n^{4-k/2},$$

which ends the proof of Lemma 6.6. □

1.4 Proof of Lemma 1.1 :

On Ω_k, we have

$$\mathbb{E}[(\hat{\alpha}_h^{\hat{\boldsymbol{\beta}}} - \bar{\alpha}_h)^2(t)\mathbb{1}(\Omega_k)] \leq \frac{16B^2\mathrm{e}^{4B|\boldsymbol{\beta_0}|_1}\mathrm{e}^{2BR}}{c_S^2}\frac{C^2(s)\log(pn^k)}{n}\mathbb{E}\Bigg[\Bigg(\frac{1}{n}\sum_{i=1}^n\int_0^\tau\frac{|K_h(t-u)|}{S(u,\boldsymbol{\beta_0})}\mathrm{d}N_i(u)\Bigg)^2\Bigg].$$

Then, from the Doob-Meyer decomposition, we deduce that

$$\mathbb{E}\Bigg[\Bigg(\frac{1}{n}\sum_{i=1}^n\int_0^\tau\frac{|K_h(t-u)|}{S(u,\boldsymbol{\beta_0})}\mathrm{d}N_i(u)\Bigg)^2\Bigg] \leq \frac{2}{n}\mathbb{E}\Bigg[\Bigg(\int_0^\tau\frac{K_h(t-u)}{S(u,\boldsymbol{\beta_0})}\mathrm{d}M_1(u)\Bigg)^2\Bigg]$$

$$+\frac{2}{n}\mathbb{E}\Bigg[\Bigg(\int_0^\tau\frac{K_h(t-u)}{S(u,\boldsymbol{\beta_0})}\alpha_0(u)\mathrm{e}^{\boldsymbol{\beta_0^T Z_1}}Y_1(u)\mathrm{d}u\Bigg)^2\Bigg]$$

$$\leq \frac{2||\alpha_0||_{\infty,\tau}}{c_S^2}(\mathbb{E}[\mathrm{e}^{\beta_0^T Z_1}] + ||\alpha_0||_{\infty,\tau}\mathbb{E}[\mathrm{e}^{2\boldsymbol{\beta_0^T Z_1}}])\frac{||K||_{\mathbb{L}^2(\mathbb{R})}^2}{c_S^2}.$$

We finally obtain

$$\mathbb{E}[||\hat{\alpha}_h^{\hat{\boldsymbol{\beta}}} - \bar{\alpha}_h||_{2,h}^2 \mathbb{1}(\Omega_k)] \leq C(c_S, ||\alpha_0||_{\infty,\tau}, |\boldsymbol{\beta_0}|_1, B, R, ||K||_{\mathbb{L}^2(\mathbb{R})}, s)\frac{\log(pn^k)}{n}. \tag{6}$$

$\square$

1.5 Proof of Lemma 1.2

In order to calculate $\mathbb{P}(\Omega_k^c)$, let us begin by study the set $\Omega_{H,k}$ defined by (30). Let us introduce the two following sets

$$\Omega_1 := \left\{ \omega : \forall u \in [0,\tau], |S_n(u, \hat{\boldsymbol{\beta}}) - S_n(u, \boldsymbol{\beta_0})| \leq B\mathrm{e}^{BR}\mathrm{e}^{2B|\boldsymbol{\beta_0}|_1} C(s)\sqrt{\frac{\log(pn^k)}{n}} \right\},$$
$$\Omega_2 := \left\{ \omega : \forall u \in [0,\tau], |S_n(u, \boldsymbol{\beta_0}) - S(u, \boldsymbol{\beta_0})| \leq B\mathrm{e}^{BR}\mathrm{e}^{2B|\boldsymbol{\beta_0}|_1} C(s)\sqrt{\frac{\log(pn^k)}{n}} \right\}.$$

We have $\Omega_{H,k} \supset \Omega_1 \cap \Omega_2$. We begin to calculate $\mathbb{P}(\Omega_1^c)$. By definition, we have

$$\begin{aligned}|S_n(u, \hat{\boldsymbol{\beta}}) - S_n(u, \boldsymbol{\beta_0})| &= \left| \frac{1}{n} \sum_{i=1}^n (\mathrm{e}^{\hat{\boldsymbol{\beta}}^{\boldsymbol{T}} \boldsymbol{Z_i}} - \mathrm{e}^{\boldsymbol{\beta_0^T Z_i}}) Y_i(u) \right| \\ &\leq \mathrm{e}^{B|\boldsymbol{\beta_0}|_1} |\mathrm{e}^{B|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta_0}|_1} - 1|\end{aligned}$$

Under Assumptions 3.1 and 3.2, from Proposition 3.3, there exists a constant $c > 0$ such that, with probability larger than $1 - cn^{-k}$,

$$|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta_0}|_1 \leq C(s)\sqrt{\frac{\log(pn^k)}{n}}.$$

So, with probability larger than $1 - cn^{-k}$, using that $|\mathrm{e}^x - \mathrm{e}^y| \leq |x - y|\mathrm{e}^{x \vee y}$ for all x, y, we have

$$\begin{aligned}|S_n(u, \hat{\boldsymbol{\beta}}) - S_n(u, \boldsymbol{\beta_0})| &\leq \mathrm{e}^{B|\boldsymbol{\beta_0}|_1} B|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta_0}|_1 \mathrm{e}^{B|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta_0}|_1} \\ &\leq B\mathrm{e}^{BR}\mathrm{e}^{2B|\boldsymbol{\beta_0}|_1} C(s)\sqrt{\frac{\log(pn^k)}{n}}.\end{aligned}$$

We deduce that

$$\mathbb{P}(\Omega_1^c) \leq cn^{-k}. \tag{7}$$

To calculate $\mathbb{P}(\Omega_2^c)$, we remark that

$$n(S_n(u, \boldsymbol{\beta_0}) - S(u, \boldsymbol{\beta_0})) = \sum_{i=1}^n \left(\mathrm{e}^{\boldsymbol{\beta_0^T Z_i}} Y_i(u) - \mathbb{E}[\mathrm{e}^{\boldsymbol{\beta_0^T Z_i}} Y_i(u)] \right).$$

As $0 \leq \mathrm{e}^{\boldsymbol{\beta_0^T Z_1}} Y_1(u) \leq \mathrm{e}^{B|\boldsymbol{\beta_0}|_1}$, we apply a Hoeffding inequality:

$$\mathbb{P}\left(|S_n(u,\boldsymbol{\beta_0}) - S(u,\boldsymbol{\beta_0})| \geq \frac{y}{n}\right) \leq 2\exp\left(-\frac{2y^2}{n\mathrm{e}^{2B|\boldsymbol{\beta_0}|_1}}\right),$$

and with $y = B\mathrm{e}^{BR}\mathrm{e}^{2B|\boldsymbol{\beta_0}|_1} C(s)\sqrt{n\log(pn^k)}/2$, we finally get

$$\begin{aligned}\mathbb{P}\left(|S_n(u,\boldsymbol{\beta_0}) - S(u,\boldsymbol{\beta_0})| \geq C(s)B\mathrm{e}^{BR}\mathrm{e}^{2B|\boldsymbol{\beta_0}|_1}\sqrt{\frac{\log(pn^k)}{n}}\right)&\\ \leq 2\exp\left(-\frac{2B^2\mathrm{e}^{2BR}\mathrm{e}^{4B|\boldsymbol{\beta_0}|_1}C^2(s)\log(pn^k)}{\mathrm{e}^{2B|\boldsymbol{\beta_0}|_1}}\right)&\\ \leq \frac{2}{pn^k}.&\end{aligned}$$

We conclude that there exists a constant $c_7 > 0$ such that

$$\mathbb{P}(\Omega_2^c) \leq c_7 n^{-k}. \tag{8}$$

Gathering (7) and (8), we obtain

$$\mathbb{P}(\Omega_{H,k}^c) \leq \mathbb{P}(\Omega_1^c) + \mathbb{P}(\Omega_2^c) \leq \tilde{c}n^{-k}, \tag{9}$$

where $\tilde{c} > 0$ is a constant. It remains to calculate $\mathbb{P}(\Omega_{S_n}^c)$, with $\Omega_{S_n}^c$ defined by (31), to obtain $\mathbb{P}(\Omega_k^c)$. We decompose

$$S_n(u,\hat{\boldsymbol{\beta}}) - S(u,\boldsymbol{\beta_0}) = S_n(u,\hat{\boldsymbol{\beta}}) - S_n(u,\boldsymbol{\beta_0}) + S_n(u,\boldsymbol{\beta_0}) - S(u,\boldsymbol{\beta_0}).$$

On $\Omega_1 \cap \Omega_2$,

$$S_n(u,\hat{\boldsymbol{\beta}}) - S_n(u,\boldsymbol{\beta_0}) \geq -2B\mathrm{e}^{BR}\mathrm{e}^{2B|\boldsymbol{\beta_0}|_1}C(s)\sqrt{\frac{\log(pn^k)}{n}} \in (-\infty, 0)$$

So for n large enough, we have that $S_n(u,\hat{\boldsymbol{\beta}}) - S_n(u,\boldsymbol{\beta_0}) \geq -c_S/2$. For n large enough, $\Omega_1 \cap \Omega_2 \subset \Omega_{S_n}$, and

$$\mathbb{P}(\Omega_{S_n}^c) \leq \mathbb{P}(\Omega_1^c) + \mathbb{P}(\Omega_2^c) \leq \tilde{c}n^{-k}. \tag{10}$$

Gathering (9) and (10), we finally obtain for n large enough that $\mathbb{P}(\Omega_k^c) \leq c_2 n^{-k}$, where c_2 is a constant depending on B, $|\boldsymbol{\beta_0}|_1$ and s. □

2 Classical results

In this appendix, some classical technical lemmas and a theorem needed for the proofs of the two main theorems of the chapter, are listed. We do not give the proofs of these well-known results but we give the references where to find their proofs.

2.1 A Cauchy-Schwarz Inequality

The following lemma gives a useful inequality concerning integrals with respect to the counting process N.

Lemma 2.1 (Cauchy-Schwarz). *For all function g bounded on $[0,\tau]$,*

$$N(\tau)\int_{\tau_1}^{\tau_2} g^2(s)\mathrm{d}N(s) \geq \left(\int_{\tau_1}^{\tau_2} g(s)\mathrm{d}N(s)\right)^2, \quad 0 \leq \tau_1 \leq \tau_2 \leq \tau$$

We refer to Bouaziz et al. (2013) for the proof of this lemma.

2.2 Young Inequality

The following lemma provides an inequality that bounds a norm of the convolution product of two functions by a product of norms of each function.

Lemma 2.2 (Young Inequality). *Let $p, q \in [1, +\infty)$ such that $1/p+1/q \geq 1$. If $s \in \mathbb{L}^p(\mathbb{R})$ and $t \in \mathbb{L}^q(\mathbb{R})$, then s and t are convolable. Moreover, if $1/r = 1/p + 1/q - 1$, then $f * g \in \mathbb{L}^r(\mathbb{R})$ and*

$$||s * t||_r \leq ||s||_p ||t||_q$$

This convolution inequality is proved in Hirsch and Lacombe (1999) (Theorem 3.4 p.149).

2.3 Talagrand Inequality

The following Talagrand Inequality is a concentration inequality that allows to control the supremum of an empirical process.

Theorem 2.3 (Talagrand Inequality). *Let $\xi_1, ..., \xi_n$ be independent random values, and let*

$$\nu_{n,\xi}(f) = \frac{1}{n}\sum_{i=1}^{n}\{f(\xi_i) - \mathbb{E}[f(\xi_i)]\}.$$

Then, for a countable class of functions $\mathcal{F}$ uniformly bounded and $\varepsilon > 0$, we have

$$\mathbb{E}\left[\left\{\sup_{f\in\mathcal{F}}\nu_{n,\xi}^2(f) - 2(1+2\varepsilon^2)H^2\right\}_+\right] \leq \frac{4}{d}\left(\frac{W}{n}\mathrm{e}^{-d\varepsilon^2\frac{nH^2}{W}} + \frac{98M^2}{dn^2\varphi^2(\varepsilon)}\mathrm{e}^{-\frac{2d\varphi(\varepsilon)\varepsilon}{\tau\sqrt{2}}\frac{nH}{M}}\right),$$

with $\varphi(\varepsilon) = \sqrt{1+\varepsilon^2} - 1$, $d = 1/6$ and

$$\sup_{f\in\mathcal{F}}||f||_\infty \leq M, \quad \mathbb{E}\Big[\sup_{f\in\mathcal{F}}|\nu_{n,\xi}(f)|\Big] \leq H, \quad \sup_{f\in\mathcal{F}}\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}[f(\xi)] \leq W.$$

This theorem is a useful corollary from the classical Talagrand established by Talagrand (1996). The proof of Theorem 2.3 can be found in Comte et al. (2008) (Lemma 6.1). The proof of the theorem follows from a concentration Inequality in Klein and Rio (2005) and arguments that can be found in Birgé and Massart (1998).

2.4 A classical inequality: the Bürkholder Inequality

The last technical result is a Bürkholder Inequality that gives a norm relation between a martingale and its optional process. We refer to Liptser and Shiryayev (1989) p.75, for the proof of this result.

Theorem 2.4 (Bürkholder Inequality). *If $M = (M_t, \mathcal{F}_t)_{t\geq 0}$ is a martingale, then there are universal constants γ_b and κ_b (independent of M) such that for every $t \geq 0$*

$$\gamma_b ||\sqrt{[M]_t}||_2 \leq ||M_t||_2 \leq \kappa_b ||\sqrt{[M]_t}||_2,$$

where $[M]_t$ is the quadratic variation of M_t.

This theorem is used to prove Lemma 6.5 and in the oracle inequalities of Theorem 4.1, the constants depend on κ_b.

References

L. Birgé and P. Massart. Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli*, 4(3):329–375, 1998.

O. Bouaziz, F. Comte, and A. Guilloux. Nonparametric estimation of the intensity function of a recurrent event process. *Statistica Sinica*, 23(2):635–665, 2013.

F. Comte, J. Dedecker, and M.L. Taupin. Adaptive density deconvolution with dependent inputs. *Mathematical Methods of Statistics*, 17(2):87–112, 2008.

F. Hirsch and G. Lacombe. *Elements of functional analysis*, volume 192. Springer, 1999.

T. Klein and E. Rio. Concentration around the mean for maxima of empirical processes. *The Annals of Probability*, 33(3):1060–1077, 2005.

R. Sh. Liptser and A. N. Shiryayev. *Theory of martingales*, volume 49 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1989. ISBN 0-7923-0395-4. URL `http://dx.doi.org/10.1007/978-94-009-2438-3`. Translated from the Russian by K. Dzjaparidze [Kacha Dzhaparidze].

M. Talagrand. New concentration inequalities in product spaces. *Inventiones mathematicae*, 126(3):505–563, 1996.