Some ideas

Oleg V. Lepski

Institut de Mathématiques de Marseille
Aix-Marseille Université

CIRM, February 8, 2016

Problem formulation.

- $(\mathcal{X}^{(n)}, \mathfrak{T}^{(n)}, \mathbb{P}_f^{(n)}, f \in \mathbb{F})$ is the statistical experiment generated by the observation $X^{(n)}$.
- $A : \mathbb{F} \to \mathfrak{S}$ and $B : \mathbb{F} \to \mathfrak{S}$ are two mappings to be estimated and $\mathfrak{S}$ is a set endowed with semi-metrics ℓ and ρ.
- For any $X^{(n)}$-measurable $\mathfrak{S}$-valued map $\widetilde{Q}$, $f \in \mathbb{F}$, $q \geq 1$

$$\mathcal{R}_A^q\big[\widetilde{Q}, f\big] = \mathbb{E}_f^{(n)}\Big[\ell\big(\widetilde{Q}, A(f)\big)\Big]^q$$

$$\mathcal{R}_B^q\big[\widetilde{Q}, f\big] = \mathbb{E}_f^{(n)}\Big[\rho\big(\widetilde{Q}, B(f)\big)\Big]^q$$

Problem formulation.

- $(\mathcal{X}^{(n)}, \mathfrak{T}^{(n)}, \mathbb{P}_f^{(n)}, f \in \mathbb{F})$ is the statistical experiment generated by the observation $X^{(n)}$.

- $A : \mathbb{F} \to \mathfrak{S}$ and $B : \mathbb{F} \to \mathfrak{S}$ are two mappings to be estimated and $\mathfrak{S}$ is a set endowed with semi-metrics ℓ and ρ

- For any $X^{(n)}$-measurable $\mathfrak{S}$-valued map $\widetilde{Q}$, $f \in \mathbb{F}$, $q \geq 1$

$$\mathcal{R}_A^q\big[\widetilde{Q}, f\big] = \mathbb{E}_f^{(n)}\Big[\ell\big(\widetilde{Q}, A(f)\big)\Big]^q, \quad \mathcal{R}_B^q\big[\widetilde{Q}, f\big] = \mathbb{E}_f^{(n)}\Big[\rho\big(\widetilde{Q}, B(f)\big)\Big]^q$$

Important! We will assume that $X^{(n)} = \big(X_1^{(n)}, X_2^{(n)}\big)$, where $X_1^{(n)}, X_2^{(n)}$ are independent random elements.

- $\mathbb{P}_{1,f}^{(n)}$ and $\mathbb{P}_{2,f}^{(n)}$ denote marginal laws of $X_1^{(n)}$ and $X_2^{(n)}$;

- $\mathbb{E}_{i,f}^{(n)}, i = 1, 2,$ – mathematical expectation w.r.t. $\mathbb{P}_{i,f}^{(n)}$.

- $\mathfrak{H}$ is a set, $\mathfrak{H}_n \subseteq \mathfrak{H}$, $n \in \mathbb{N}^*$, are countable of subsets.
- $\{\widehat{A}_h, \mathfrak{h} \in \mathfrak{H}\}$, $\{\widehat{A}_{\mathfrak{h},\eta}, \mathfrak{h}, \eta \in \mathfrak{H}\}$ – $X_1^{(n)}$-measurable $\mathfrak{S}$-valued;
- $\varepsilon_n \to 0$, $n \to \infty$ be a given sequence.

$\mathbf{A}^{\text{permute}}$. $\widehat{A}_{\mathfrak{h},\eta} \equiv \widehat{A}_{\eta,\mathfrak{h}}$, **for any** $\eta, \mathfrak{h} \in \mathfrak{H}$.

$\mathbf{A}^{\text{upper}}$. **For any** $n \geq 1$

$$\sup_{f\in\mathbb{F}} \mathbb{E}_{1,f}^{(n)} \left(\sup_{\mathfrak{h}\in\mathfrak{H}_n} \left[\ell\left(\widehat{A}_{\mathfrak{h}}, \Lambda_{\mathfrak{h}}(f)\right) - \Delta_n(\mathfrak{h}) \right]_+^q \right) \leq \varepsilon_n^q;$$

$$\sup_{f\in\mathbb{F}} \mathbb{E}_{1,f}^{(n)} \left(\sup_{\mathfrak{h},\eta\in\mathfrak{H}_n} \left[\ell\left(\widehat{A}_{\mathfrak{h},\eta}, \Lambda_{\mathfrak{h},\eta}(f)\right) - \left\{ \Delta_n(\mathfrak{h}) \wedge \Delta_n(\eta) \right\} \right]_+^q \right) \leq \varepsilon_n^q.$$

- $\{\Lambda_{\mathfrak{h}}(f),\ \mathfrak{h} \in \mathfrak{H}\}$, $\{\Lambda_{\mathfrak{h},\eta}(f),\ \mathfrak{h}, \eta \in \mathfrak{H}\}$ – $\mathfrak{S}$-valued;
- $\Delta_n = \{\Delta_n(\mathfrak{h}),\ \mathfrak{h} \in \mathfrak{H}\}$ – $X_1^{(n)}$-measurable **positive** variables.

(Ψ_n, ℓ)-selection rule for estimating A.

- Fix $\boldsymbol{\Psi}_n \in \mathfrak{M}_n$, where

$$\mathfrak{M}_n = \left\{ \boldsymbol{\Delta}'_n : \ \boldsymbol{\Delta}_n(\cdot) \le \boldsymbol{\Delta}'_n(\cdot), \ \boldsymbol{\Delta}'_n(\cdot) \text{ is } X_1^{(n)} - \text{measurable} \right\}$$

- **Define for any $\mathfrak{h} \in \mathfrak{H}_n$**

$$\widehat{R}_n(\mathfrak{h}) = \sup_{\eta \in \mathfrak{H}_n} \left[\ell\big(\widehat{A}_{\mathfrak{h},\eta}, \widehat{A}_\eta\big) - 2\boldsymbol{\Psi}_n(\eta) \right]_+$$

- **Let $\hat{\mathfrak{h}}^{(n)} \in \mathfrak{H}_n$ be an arbitrary $X_1^{(n)}$ – measurable element:**

$$\widehat{R}\big(\hat{\mathfrak{h}}^{(n)}\big) + 2\boldsymbol{\Psi}_n\big(\hat{\mathfrak{h}}^{(n)}\big) \le \inf_{\mathfrak{h} \in \mathfrak{H}_n} \left\{ \widehat{R}(\mathfrak{h}) + 2\boldsymbol{\Psi}_n(\mathfrak{h}) \right\} + \varepsilon_n$$

- Set for any $f \in \mathbb{F}$, $\mathfrak{h} \in \mathfrak{H}_n$ and $n \ge 1$

$$\mathcal{B}_A^{(n)}(f, \mathfrak{h}) = \ell\big(\Lambda_{\mathfrak{h}}(f), A(f)\big) + 2 \sup_{\eta \in \mathfrak{H}_n} \ell\big(\Lambda_{\mathfrak{h},\eta}(f), \Lambda_\eta(f)\big)$$

$$\psi_n(f, \mathfrak{h}) = \left[\mathbb{E}_{1,f}^{(n)} \left\{ \boldsymbol{\Psi}_n^q(\mathfrak{h}) \right\} \right]^{\frac{1}{q}}$$

(Ψ_n, ℓ)-selection rule for A. Oracle inequality

$$\widehat{R}_n(\mathfrak{h}) = \sup_{\eta \in \mathfrak{H}_n} \left[\ell\left(\widehat{A}_{\mathfrak{h},\eta}, \widehat{A}_\eta\right) - 2\Psi_n(\eta)\right]_+$$

$$\widehat{R}\left(\hat{\mathfrak{h}}^{(n)}\right) + 2\Psi_n\left(\hat{\mathfrak{h}}^{(n)}\right) \le \inf_{\mathfrak{h} \in \mathfrak{H}_n} \left\{\widehat{R}(\mathfrak{h}) + 2\Psi_n(\mathfrak{h})\right\} + \varepsilon_n$$

Theorem 1. Let $\mathbf{A}^{\text{permute}}$ and $\mathbf{A}^{\text{upper}}$ be fulfilled.

Then, for any $f \in \mathbb{F}, n \ge 1$ and $\Psi_n \in \mathfrak{M}_n$

$$\mathcal{R}_A\left[\widehat{A}_{\hat{\mathfrak{h}}^{(n)}}, f\right] \le \inf_{\mathfrak{h} \in \mathfrak{H}_n} \left\{\mathcal{B}_A^{(n)}(f, \mathfrak{h}) + 5\psi_n(f, \mathfrak{h})\right\} + 6\varepsilon_n.$$

- $\mathcal{B}_A^{(n)}(f, \mathfrak{h}) = \ell\left(\Lambda_{\mathfrak{h}}(f), A(f)\right) + 2\sup_{\eta \in \mathfrak{H}_n} \ell\left(\Lambda_{\mathfrak{h},\eta}(f), \Lambda_\eta(f)\right)$
- $\psi_n(f, \mathfrak{h}) = \left[\mathbb{E}_{1,f}^{(n)}\left\{\Psi_n^q(\mathfrak{h})\right\}\right]^{\frac{1}{q}}$

Estimation of B. Hypotheses.

- $\{\widehat{B}_{\mathfrak{h}}, \mathfrak{h} \in \mathfrak{H}\}$ – family of $X_2^{(n)}$-measurable $\mathfrak{S}$-valued mappings;

Objective: To bound from above the risk of the "*plug-in*" estimator $\widehat{B}_{\hat{\mathfrak{h}}(n)}$.

- $\{\Upsilon_{\mathfrak{h}},\ \mathfrak{h} \in \mathfrak{H}\}$ – a collection of $\mathfrak{S}$-valued functionals.

$\mathbf{B}^{\text{main}}$. $\exists C_\ell$ such that for any $f \in \mathbb{F}$, $n \geq 1$ and $\mathfrak{h} \in \mathfrak{H}_n$

$$\rho\big(\Upsilon_{\mathfrak{h}}(f), B(f)\big) \leq C_\ell\, \ell\big(\Lambda_{\mathfrak{h}}(f), A(f)\big)$$

$\mathbf{B}^{\text{upper}}$. $\Phi_n \in \mathfrak{M}_n$ and for any $n \geq 1$

$$\sup_{f\in\mathbb{F}} \mathbb{E}_{1,f}^{(n)}\bigg(\sup_{\mathfrak{h}\in\mathfrak{H}_n}\Big[\mathcal{E}_n(\mathfrak{h}, f) - \Phi_n(\mathfrak{h})\Big]_+^q\bigg) \leq \varepsilon_n^q;$$

- $\mathcal{E}_n(\mathfrak{h}, f) = \Big(\mathbb{E}_{2,f}^{(n)}\big\{\rho^q\big(\widehat{B}_{\mathfrak{h}}, \Upsilon_{\mathfrak{h}}(f)\big)\big\}\Big)^{\frac{1}{q}}$;

- $\Phi_n = \{\Phi_n(\mathfrak{h}),\ \mathfrak{h} \in \mathfrak{H}\}$ – $X_1^{(n)}$-measurable **positive** variables.

B^main. $\exists C_\ell$ such that for any $f \in \mathbb{F}$, $n \geq 1$ and $\mathfrak{h} \in \mathfrak{H}_n$

$$\rho\big(\Upsilon_{\mathfrak{h}}(f), B(f)\big) \leq C_\ell\, \ell\big(\Lambda_{\mathfrak{h}}(f), A(f)\big)$$

B^upper. $\Phi_n \in \mathfrak{M}_n$ and for any $n \geq 1$

$$\sup_{f \in \mathbb{F}} \mathbb{E}^{(n)}_{1,f}\left(\sup_{\mathfrak{h} \in \mathfrak{H}_n} \Big[\mathcal{E}_n(\mathfrak{h}, f) - \Phi_n(\mathfrak{h})\Big]_+^q \right) \leq \varepsilon_n^q;$$

- $\mathcal{E}_n(\mathfrak{h}, f) = \left(\mathbb{E}^{(n)}_{2,f}\left\{\rho^q\big(\widehat{B}_{\mathfrak{h}}, \Upsilon_{\mathfrak{h}}(f)\big)\right\}\right)^{\frac{1}{q}};$
- $\Phi_n = \{\Phi_n(\mathfrak{h}),\ \mathfrak{h} \in \mathfrak{H}\}$ – $X_1^{(n)}$-measurable **positive** variables.

Remark. Set $\mathcal{E}_n(\mathfrak{h}) = \sup_{f \in \mathbb{F}} \mathcal{E}_n(\mathfrak{h}, f)$ and note that if $\mathcal{E}_n \in \mathfrak{M}_n$ the hypothesis **B**^upper is obviously fulfilled with $\Phi_n = \mathcal{E}_n$. This choice of Φ_n is reasonable for the statistical models in which $\mathcal{E}_n(\mathfrak{h}, f)$ is independent or depends "weakly" on f.

First procedure and oracle inequality.

B^main. $\exists C_\ell$ such that for any $f \in \mathbb{F}$, $n \geq 1$ and $\mathfrak{h} \in \mathfrak{H}_n$

$$\rho\big(\Upsilon_{\mathfrak{h}}(f), B(f)\big) \leq C_\ell\, \ell\big(\Lambda_{\mathfrak{h}}(f), A(f)\big)$$

B^upper. $\Phi_n \in \mathfrak{M}_n$ and for any $n \geq 1$

$$\sup_{f\in\mathbb{F}} \mathbb{E}^{(n)}_{1,f}\left(\sup_{\mathfrak{h}\in\mathfrak{H}_n}\Big[\mathcal{E}_n(\mathfrak{h}, f) - \Phi_n(\mathfrak{h})\Big]_+^q\right) \leq \varepsilon_n^q;$$

Theorem 2. Let A^{permute}, A^{upper}, B^{main} and B^{upper} be fulfilled.

Let $\hat{\mathfrak{h}}^{(n)}$ is obtained by (Φ_n, ℓ)-selection rule. Then, for any $f \in \mathbb{F}$ and $n \geq 1$

$$\mathcal{R}_B\left[\widehat{B}_{\hat{\mathfrak{h}}^{(n)}}, f\right] \leq C_1 \inf_{\mathfrak{h}\in\mathfrak{H}_n}\left\{\mathcal{B}^{(n)}_A(f,\mathfrak{h}) + \phi_n(f,\mathfrak{h})\right\} + C_2\varepsilon_n.$$

- $\phi_n(f,\mathfrak{h}) = \left[\mathbb{E}^{(n)}_{1,f}\left\{\Phi_n^q(\mathfrak{h})\right\}\right]^{\frac{1}{q}}$, $C_1 = 7C_\ell + 2$, $C_2 = 10C_\ell + 4$.

Second procedure and oracle inequality.

Problem. We note that the use (Φ_n,ℓ)-selection rule if $\Phi_n \neq \Psi_n$ does not allow to solve the problems of estimating $A(\cdot)$ and $B(\cdot)$ simultaneously, i.e. by use of the same (Ψ_n,ℓ)-selection rule in both problems.

Objective. To prove an analog of Theorem 2 when $\hat{\mathfrak{h}}^{(n)}$ is obtained by (Ψ_n,ℓ)-selection rule with $\Psi_n \in \mathfrak{M}_n(\delta_n)$.

- $\delta_n \to 0, n \to \infty$ - given sequence and

$$\mathfrak{M}_n(\delta_n) := \left\{\Psi_n \in \mathfrak{M}_n : \ \inf_{\mathfrak{h}\in\mathfrak{H}_n} \Psi_n(\mathfrak{h}) \geq \delta_n\right\}$$

Remark. $\mathfrak{M}_n(\delta_n) \supset \left\{\Psi_n^* : \ \Psi_n^* \equiv \Psi_n + \delta_n, \ \Psi_n \in \mathfrak{M}_n\right\}$.

Notations. $\tau_n(f) = \left[\mathbb{E}^{(n)}_{1,f}\left\{\sup_{\mathfrak{h}\in\mathfrak{H}_n(f)} \Phi_n^q(\mathfrak{h})\right\}\right]^{\frac{1}{q}} + \mathfrak{r}_n(f)$

$$\mathfrak{H}_n(f) = \left\{\mathfrak{h} \in \mathfrak{H}_n : \Psi_n(\mathfrak{h}) < 2\inf_{\mathfrak{h}\in\mathfrak{H}_n}\left[\mathcal{B}_A^{(n)}(f,\mathfrak{h}) + 2\Psi_n(\mathfrak{h})\right]\right\}$$

- $\mathfrak{r}_n(f) = \left[\mathbb{E}^{(n)}_{1,f}\left\{\sup_{\mathfrak{h}\in\mathfrak{H}_n} \Phi_n^{2q}(\mathfrak{h})\right\}\right]^{\frac{1}{2q}} (8\varepsilon_n/\delta_n)^{\frac{1}{2}}$

Second procedure and oracle inequality.

Notations. $\tau_n(f) = \left[\mathbb{E}_{1,f}^{(n)}\left\{\sup_{\mathfrak{h}\in\mathfrak{H}_n(f)} \Phi_n^q(\mathfrak{h})\right\}\right]^{\frac{1}{q}} + \mathfrak{r}_n(f)$

$$\mathfrak{H}_n(f) = \left\{\mathfrak{h} \in \mathfrak{H}_n : \Psi_n(\mathfrak{h}) < 2\inf_{\mathfrak{h}\in\mathfrak{H}_n}\left[\mathcal{B}_A^{(n)}(f,\mathfrak{h}) + 2\Psi_n(\mathfrak{h})\right]\right\}$$

- $\mathfrak{r}_n(f) = \left[\mathbb{E}_{1,f}^{(n)}\left\{\sup_{\mathfrak{h}\in\mathfrak{H}_n} \Phi_n^{2q}(\mathfrak{h})\right\}\right]^{\frac{1}{2q}} (8\varepsilon_n/\delta_n)^{\frac{1}{2}}$

Theorem 3. Let $\mathrm{A}^{\mathrm{permute}}$, $\mathrm{A}^{\mathrm{upper}}$, $\mathrm{B}^{\mathrm{main}}$ and $\mathrm{B}^{\mathrm{upper}}$ be fulfilled.

Let $\hat{\mathfrak{h}}^{(n)}$ is obtained by (Ψ_n, ℓ)-selection rule. Then, for any $f \in \mathbb{F}$, $n \geq 1$ and $\Psi_n \in \mathfrak{M}_n(\delta_n)$

$$\mathcal{R}_B\left[\widehat{B}_{\hat{\mathfrak{h}}^{(n)}}, f\right] \leq C_3 \inf_{\mathfrak{h}\in\mathfrak{H}_n}\left\{\mathcal{B}_A^{(n)}(f,\mathfrak{h}) + \psi_n(f,\mathfrak{h})\right\} + \tau_n(f) + C_4\varepsilon_n$$

- $C_3 = 7C_\ell + 1$, $C_4 = 10C_\ell + 1$.

Adaptive estimation of B.

Problem. We would like to emphasize that the hypothesis $\mathbf{B}^{\mathrm{main}}$ is quite restrictive since it can be checked for any $f \in \mathbb{F}$.

Objective. To weaken $\mathbf{B}^{\mathrm{main}}$ in the case of adaptive estimation.

- Let $\{\mathbb{F}_\alpha, \alpha \in \mathfrak{A}\}$ be a given collection of subsets of $\mathbb{F}$.

$\underline{\mathbf{B}^{\mathrm{adap}}.}$ $\forall \alpha \in \mathfrak{A}\ \exists \pi_\alpha : \mathbb{R}_+ \to \mathbb{R}_+$, **nondecreasing, concave and such that for any** $f \in \mathbb{F}_\alpha$, $n \geq 1$ **and** $\mathfrak{h} \in \mathfrak{H}_n$

$$\rho^q\big(\Upsilon_{\mathfrak{h}}(f), B(f)\big) \leq \pi_\alpha\Big(\ell^q\big(\Lambda_{\mathfrak{h}}(f), A(f)\big)\Big)$$

Adaptive estimation of B.

- Let $\{\mathbb{F}_\alpha, \alpha \in \mathfrak{A}\}$ be a given collection of subsets of $\mathbb{F}$.

B$^{\text{adap}}$. $\forall \alpha \in \mathfrak{A}\ \exists \pi_\alpha : \mathbb{R}_+ \to \mathbb{R}_+$, **nondecreasing, concave and such that for any** $f \in \mathbb{F}_\alpha$, $n \geq 1$ **and** $\mathfrak{h} \in \mathfrak{H}_n$

$$\rho^q\big(\Upsilon_{\mathfrak{h}}(f), B(f)\big) \leq \pi_\alpha\Big(\ell^q\big(\Lambda_{\mathfrak{h}}(f), A(f)\big)\Big)$$

Theorem 4. Let A$^{\text{permute}}$, A$^{\text{upper}}$, B$^{\text{adap}}$ and B$^{\text{upper}}$ be fulfilled.

Let $\hat{\mathfrak{h}}^{(n)}$ is obtained by (Ψ_n, ℓ)-selection rule. Then, for any $\alpha \in \mathfrak{A}$, $n \geq 1$ and $\Psi_n \in \mathfrak{M}_n(\delta_n)$

$$\sup_{f \in \mathbb{F}_\alpha} \mathcal{R}_B\Big[\widehat{B}_{\hat{\mathfrak{h}}^{(n)}}, f\Big] \leq \pi_\alpha^{1/q}\Big(\big[\varphi_n(\mathbb{F}_\alpha) + 10\varepsilon_n\big]^q\Big) + \varphi_n^*(\mathbb{F}_\alpha) + \varepsilon_n$$

- $\varphi_n(\mathbb{F}_\alpha) = \sup\limits_{f \in \mathbb{F}_\alpha}\Big[\inf\limits_{\mathfrak{h} \in \mathfrak{H}_n}\big\{\mathcal{B}_A^{(n)}(f, \mathfrak{h}) + \psi_n(f, \mathfrak{h})\big\}\Big].$

- $\varphi_n^*(\mathbb{F}_\alpha) = \sup_{f \in \mathbb{F}_\alpha} \tau_n(f).$

Examples $\rho = \ell$. Generalized deconvolution model.

- Observation $Z^{(n)} = (Z_1, \dots, Z_n)$

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

- $X_i \in \mathbb{R}^d, i = 1, \dots, n$ are **i.i.d.** random vectors with common density f **to be estimated**;
- The noise variables $Y_i \in \mathbb{R}^d, i = 1, \dots, n$, are **i.i.d.** random vectors with **known** common density g;
- $\varepsilon_i \in \{0, 1\}, i = 1, \dots, n$, are **i.i.d.** Bernoulli random variables with $\mathbb{P}(\varepsilon_1 = 1) = \alpha$, $\alpha \in [0, 1]$ **is supposed to be known**;
- The sequences $\{X_i, i = 1, \dots, n\}, \{Y_i, i = 1, \dots, n\}$ and $\{\varepsilon_i, i = 1, \dots, n\}$ are supposed to be mutually independent.

Goal: estimation of $B(f) = f$ under $\mathbb{L}_p$-loss, i.e

$$\ell(\cdot) = \|\cdot\|_p, 1 \le p \le \infty.$$

- For any $\vec{h} \in \mathcal{H}^d$ let $M\big(\cdot, \vec{h}\big)$ satisfy the operator equation

$$\boxed{K_{\vec{h}}(\cdot) = (1-\alpha)M\big(\cdot, \vec{h}\big) + \alpha \int_{\mathbb{R}^d} g(t-\cdot)M\big(t, \vec{h}\big)\mathrm{d}t}$$

  - $\mathcal{H}^d$ is the diadic grid in $(0,\infty)^d$;
  - $K_{\vec{h}}(y) = \big[\prod_{j=1}^d h_j^{-1}\big] K\big(y_1/h_1, \dots, y_d/h_d\big),\ y \in \mathbb{R}^d$

- Estimator for $B(f) = f$

$$\boxed{\widehat{B}_{\vec{h}}(x) = n^{-1}\sum_{i=1}^n M\big(Z_i - x, \vec{h}\big)}$$

Objective: to propose a data-driven selection rule from the family

$$\mathcal{F}\big(\mathcal{H}^d\big) = \big\{\widehat{B}_{\vec{h}}(\cdot),\ \vec{h} \in \mathcal{H}^d\big\}$$

Examples. Generalized deconvolution model.

Idea: to estimate first $A(f) = g \star f$ which is the density of Z_1 using the selection rule from the family of usual kernel estimators

$$\boxed{\widehat{A}_{\vec{h}}(x) = n^{-1} \sum_{i=1}^{n} K_{\vec{h}}\big(Z_i - x\big), \ \vec{h} \in \mathcal{H}^d}$$

and then to use the estimator $\widehat{B}_{\vec{\mathfrak{h}}}$, where $\vec{\mathfrak{h}}$ is the selected multi-bandwidth.

- If $p = 2$ the hypothesis $\mathbf{B}^{\text{main}}$ is verified for any $\alpha \in (0, 1)$ with $C_\ell = \nu^{-1}$ under the following assumption

$$\boxed{\text{There exists } \nu > 0 : \big|1 - \alpha + \alpha \check{g}(t)\big| \geq \nu, \ \forall t \in \mathbb{R}^d}$$

- If $p \neq 2$ the hypothesis $\mathbf{B}^{\text{main}}$ is verified for all $\alpha \in (1/2, 1)$ with $C_\ell = (2\alpha - 1)^{-1}$ without any assumption imposed on g.

Examples. Adaptive estimation of derivatives.

- $X_i \in \mathbb{R}, i = 1, \ldots, n$ are **i.i.d.** random variables with common density f **to be estimated**;

Goal: estimation of $B(f) = f^{(m)}, m \in \mathbb{N}^*$ under $\mathbb{L}_p$-loss, i.e

$$\ell(\cdot) = \|\cdot\|_p, 1 \leq p \leq \infty.$$

- $\mathbb{F}_\alpha = W_s^k(L), \alpha = (k, L), k > m, L > 0$, where

$$W_s^k(L) = \left\{ w : \mathbb{R} \to \mathbb{R} : \|w\|_s + \|w^{(k)}\|_s \leq L \right\}$$

Idea: to estimate first $A(f) = f$ using the selection rule from the family of usual kernel estimators

$$\boxed{\widehat{A}_h(x) = n^{-1} \sum_{i=1}^{n} K_h\big(X_i - x\big), \;\; h \in \mathcal{H}}$$

and then to use the estimator $\widehat{B}_{\mathfrak{h}} = \widehat{A}_{\mathfrak{h}}^{(m)}$, where $\mathfrak{h}$ is the selected bandwidth.

Examples. Adaptive estimation of derivatives.

Goal: estimation of $B(f) = f^{(m)}, m \in \mathbb{N}^*$ under $\mathbb{L}_p$-loss.

- $\mathbb{F}_\alpha = W_s^k(L), \alpha = (k, L), k > m, L > 0$, where

$$W_s^k(L) = \left\{ w : \mathbb{R} \to \mathbb{R} : \|w\|_s + \|w^{(k)}\|_s \le L \right\}$$

Idea: to estimate first $A(f) = f$ using the selection rule from the family of usual kernel estimators

$$\boxed{\widehat{A}_h(x) = n^{-1} \sum_{i=1}^n K_h\big(X_i - x\big), \;\; h \in \mathcal{H}}$$

and then to use the estimator $\widehat{B}_{\mathfrak{h}} = \widehat{A}_{\mathfrak{h}}^{(m)}$, where $\mathfrak{h}$ is the selected bandwidth.

- Hypothesis $\mathbf{B}^{\text{adap}}$ is verified if $s \le p$ with

$$\pi_\alpha(z) = \kappa L^{\frac{m}{k-1/s+1/p}} z^{\frac{k-m-1/s+1/p}{k-1/s+1/p}}, \quad \alpha = (k, L)$$

  - κ is the universal constant appeared in Kolmogorov inequality.