La ρ-estimation appliquée à l'estimation de la densité de transition d'une chaîne de Markov

Mathieu Sart

Université Jean Monnet Saint-Etienne
Université de Lyon
Institut Camille Jordan

2-3 février 2017

Cadre statistique

- On considère une chaine de Markov $(X_i)_{i\in\mathbb{N}}$ à valeurs dans $\mathbb{R}^d$.
- On suppose que pour tout $x \in \mathbb{R}^d$, la loi conditionnelle $\mathcal{L}(X_{i+1} \mid X_i = x)$ admet une densité $s(x,\cdot)$ par rapport à la mesure de Lebesgue μ.
- Le but est d'estimer la densité de transition $(x,y) \mapsto s(x,y)$ sur $[0,1]^{2d}$ à partir des observations $X_0, \ldots, X_n$.

Définition du risque

- Soit M la mesure aléatoire $M = n^{-1} \sum_{i=0}^{n-1} \delta_{X_i} \otimes \mu$.
- Soit $\mathbb{L}^1_+([0,1]^{2d}, M)$ le cone des fonctions positives et intégrables sur $\mathbb{R}^{2d}$ de support inclus dans $[0,1]^{2d}$ par rapport à M.
- Soit H la distance définie pour tout $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ par

$$
\begin{aligned}
H^2(f, f') &= \frac{1}{2n} \sum_{i=0}^{n-1} \int_{[0,1]^d} \left(\sqrt{f(X_i, y)} - \sqrt{f'(X_i, y)} \right)^2 \mathbb{1}_{[0,1]^d}(X_i) \, \mathrm{d}y \\
&= \frac{1}{2} \int_{[0,1]^{2d}} \left(\sqrt{f(x, y)} - \sqrt{f'(x, y)} \right)^2 \mathrm{d}M(x, y)
\end{aligned}
$$

- La qualité d'un estimateur $\hat{s}$ est mesurée par $H^2(s, \hat{s})$.

Idée de la ρ-estimation

- Pour estimer s, il faut se donner un modèle $S \subset \mathbb{L}^1_+([0,1]^{2d}, M)$.
- On construit alors un estimateur $\hat{s} \in S$.

Idée de la ρ-estimation

- Pour estimer s, il faut se donner un modèle $S \subset \mathbb{L}^1_+([0,1]^{2d}, M)$.
- On construit alors un estimateur $\hat{s} \in S$.
- Si $s \in S$, on espère un résultat du type $H^2(s, \hat{s}) \leq D_S/n$ où D_S représente la "dimension" du modèle.

Idée de la ρ-estimation

- Pour estimer s, il faut se donner un modèle $S \subset \mathbb{L}^1_+([0,1]^{2d}, M)$.
- On construit alors un estimateur $\hat{s} \in S$.
- Si $s \in S$, on espère un résultat du type $H^2(s, \hat{s}) \leq D_S/n$ où D_S représente la "dimension" du modèle.
- Si $s \notin S$, alors quelque soit $\hat{s} \in S$, $H^2(s, \hat{s}) \geq H^2(s, S)$.

Idée de la ρ-estimation

- Pour estimer s, il faut se donner un modèle $S \subset \mathbb{L}^1_+([0,1]^{2d}, M)$.
- On construit alors un estimateur $\hat{s} \in S$.
- Si $s \in S$, on espère un résultat du type $H^2(s, \hat{s}) \leq D_S/n$ où D_S représente la "dimension" du modèle.
- Si $s \notin S$, alors quelque soit $\hat{s} \in S$, $H^2(s, \hat{s}) \geq H^2(s, S)$.
- On souhaite donc construire un estimateur $\hat{s}$ tel que

$$H^2(s, \hat{s}) \leq C \left[H^2(s, S) + \frac{D_S}{n} \right]$$

Idée de la ρ-estimation

- Soit $f \in \mathbb{L}^1_+([0,1]^{2d}, M)$.
- Problème: comment savoir si $H^2(s, f)$ est petit ?

Idée de la ρ-estimation

- Soit $f \in \mathbb{L}^1_+([0,1]^{2d}, M)$.
- Problème: comment savoir si $H^2(s,f)$ est petit ?
- Idée: pour f et f' construire une fonction mesurable des observations $T(f,f') \approx H^2(s,f) - H^2(s,f')$

Idée de la ρ-estimation

- Soit $f \in \mathbb{L}^1_+([0,1]^{2d}, M)$.
- Problème: comment savoir si $H^2(s, f)$ est petit ?
- Idée: pour f et f' construire une fonction mesurable des observations $T(f, f') \approx H^2(s, f) - H^2(s, f')$ de sorte que

$$\sup_{f' \in S} T(f, f') \approx H^2(s, f) - H^2(s, S).$$

Idée de la ρ-estimation

- Soit $f \in \mathbb{L}^1_+([0,1]^{2d}, M)$.
- Problème: comment savoir si $H^2(s, f)$ est petit ?
- Idée: pour f et f' construire une fonction mesurable des observations $T(f, f') \approx H^2(s, f) - H^2(s, f')$ de sorte que

$$\sup_{f' \in S} T(f, f') \approx H^2(s, f) - H^2(s, S).$$

On minimise alors sur S, $f \mapsto \sup_{f' \in S} T(f, f')$.

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \,\mathrm{d}M + \frac{1}{2} \int \left(\sqrt{\frac{f + f'}{2}} - \left(\sqrt{f} + \sqrt{f'} \right) \right) \left(\sqrt{f'} - \sqrt{f} \right) \mathrm{d}M$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \,\mathrm{d}M + R(f, f')$$

Alors,

$$T_E(f, f') = H^2(s, f) - H^2(s, f') + \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} \left(\sqrt{s} - \sqrt{\frac{f + f'}{2}} \right)^2 \mathrm{d}M$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Alors,

$$T_E(f, f') = H^2(s, f) - H^2(s, f') + \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} \left(\sqrt{s} - \sqrt{\frac{f + f'}{2}} \right)^2 \mathrm{d}M$$

On peut voir que

$$\frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} \left(\sqrt{s} - \sqrt{\frac{f + f'}{2}} \right)^2 \mathrm{d}M \leq \frac{1}{\sqrt{2}} (H^2(s, f) + H^2(s, f'))$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Alors,

$$T_E(f, f') \quad \leq \quad (1 + 1/\sqrt{2}) H^2(s, f) - (1 - 1/\sqrt{2}) H^2(s, f')$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Alors,

$$\begin{aligned} T_E(f, f') &\leq (1 + 1/\sqrt{2}) H^2(s, f) - (1 - 1/\sqrt{2}) H^2(s, f') \\ T_E(f', f) &\leq (1 + 1/\sqrt{2}) H^2(s, f') - (1 - 1/\sqrt{2}) H^2(s, f) \end{aligned}$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Alors,

$$\begin{aligned} T_E(f, f') &\leq (1 + 1/\sqrt{2}) H^2(s, f) - (1 - 1/\sqrt{2}) H^2(s, f') \\ T_E(f, f') &\geq (1 - 1/\sqrt{2}) H^2(s, f) - (1 + 1/\sqrt{2}) H^2(s, f') \end{aligned}$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Alors,

$$\begin{aligned} \sup_{f' \in S} T_E(f, f') &\leq (1 + 1/\sqrt{2}) H^2(s, f) - (1 - 1/\sqrt{2}) H^2(s, S) \\ \sup_{f' \in S} T_E(f, f') &\geq (1 - 1/\sqrt{2}) H^2(s, f) - (1 + 1/\sqrt{2}) H^2(s, S) \end{aligned}$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Alors,

$$\begin{aligned} \inf_{f \in S} \sup_{f' \in S} T_E(f, f') &\leq \sqrt{2} H^2(s, S) \\ \sup_{f' \in S} T_E(f, f') &\geq (1 - 1/\sqrt{2}) H^2(s, f) - (1 + 1/\sqrt{2}) H^2(s, S) \end{aligned}$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \,\mathrm{d}M + R(f, f')$$

Alors,

$$\begin{aligned} \inf_{f \in S} \sup_{f' \in S} T_E(f, f') &\leq \sqrt{2} H^2(s, S) \\ \sup_{f' \in S} T_E(f, f') &\geq (1 - 1/\sqrt{2}) H^2(s, f) - (1 + 1/\sqrt{2}) H^2(s, S) \end{aligned}$$

Donc, tout minimiseur $\bar{s}$ de $f \mapsto \sup_{f' \in S} T_E(f, f')$ vérifie

$$\sup_{f' \in S} T_E(\bar{s}, f') = \inf_{f \in S} \sup_{f' \in S} T_E(f, f') \leq \sqrt{2} H^2(s, S)$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Alors,

$$\begin{aligned} \inf_{f \in S} \sup_{f' \in S} T_E(f, f') &\leq \sqrt{2} H^2(s, S) \\ \sup_{f' \in S} T_E(f, f') &\geq (1 - 1/\sqrt{2}) H^2(s, f) - (1 + 1/\sqrt{2}) H^2(s, S) \end{aligned}$$

Donc, tout minimiseur $\bar{s}$ de $f \mapsto \sup_{f' \in S} T_E(f, f')$ vérifie

$$\begin{aligned} \sup_{f' \in S} T_E(\bar{s}, f') = \inf_{f \in S} \sup_{f' \in S} T_E(f, f') &\leq \sqrt{2} H^2(s, S) \\ \sup_{f' \in S} T_E(\bar{s}, f') &\geq (1 - 1/\sqrt{2}) H^2(s, \bar{s}) - (1 + 1/\sqrt{2}) H^2(s, S) \end{aligned}$$

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}} \int \frac{\sqrt{f'} - \sqrt{f}}{\sqrt{f + f'}} s \, \mathrm{d}M + R(f, f')$$

Tout minimiseur $\bar{s}$ de $f \mapsto \sup_{f' \in S} T_E(f, f')$ vérifie

$$H^2(s, \bar{s}) \leq (5 + 4\sqrt{2}) H^2(s, S)$$

En particulier, si $s \in S$, $\bar{s} = s$.

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}n} \sum_{i=0}^{n-1} \int \frac{\sqrt{f'(X_i, y)} - \sqrt{f(X_i, y)}}{\sqrt{f(X_i, y) + f'(X_i, y)}} s(X_i, y) \, \mathrm{d}y + R(f, f')$$

Tout minimiseur $\bar{s}$ de $f \mapsto \sup_{f' \in S} T_E(f, f')$ vérifie

$$H^2(s, \bar{s}) \leq (5 + 4\sqrt{2}) H^2(s, S)$$

En particulier, si $s \in S$, $\bar{s} = s$.

Construction de $T(f, f')$

Soient $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ et soit

$$T_E(f, f') = \frac{1}{\sqrt{2}n} \sum_{i=0}^{n-1} \int_{[0,1]^d} \frac{\sqrt{f'(X_i, y)} - \sqrt{f(X_i, y)}}{\sqrt{f(X_i, y) + f'(X_i, y)}} s(X_i, y) \, \mathrm{d}y + R(f, f')$$

Tout minimiseur $\bar{s}$ de $f \mapsto \sup_{f' \in S} T_E(f, f')$ vérifie

$$H^2(s, \bar{s}) \leq (5 + 4\sqrt{2}) H^2(s, S)$$

En particulier, si $s \in S$, $\bar{s} = s$.

Construction de $T(f, f')$

Pour $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ on définit

$$T(f, f') = \frac{1}{\sqrt{2}n} \sum_{i=0}^{n-1} \frac{\sqrt{f'(X_i, X_{i+1})} - \sqrt{f(X_i, X_{i+1})}}{\sqrt{f(X_i, X_{i+1}) + f'(X_i, X_{i+1})}} + R(f, f')$$

et on définit alors le ρ-estimateur $\hat{s}$ comme étant n'importe quel minimiseur sur S de

$$f \mapsto \sup_{f' \in S} T(f, f').$$

Construction de $T(f, f')$

Pour $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ on définit

$$T(f, f') = \frac{1}{\sqrt{2}n} \sum_{i=0}^{n-1} \frac{\sqrt{f'(X_i, X_{i+1})} - \sqrt{f(X_i, X_{i+1})}}{\sqrt{f(X_i, X_{i+1}) + f'(X_i, X_{i+1})}} + R(f, f')$$

et on définit alors le ρ-estimateur $\hat{s}$ comme étant n'importe quel minimiseur sur S de

$$f \mapsto \sup_{f' \in S} T(f, f').$$

Il faut néanmoins contrôler l'erreur d'estimation donnée par le processus

$$(f, f') \mapsto T(f, f') - T_E(f, f')$$

Erreur d'approximation

Pour $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ on a

$$T_E(f,f') = \frac{1}{\sqrt{2}n}\sum_{i=0}^{n-1}\int \frac{\sqrt{f'(X_i,y)}-\sqrt{f(X_i,y)}}{\sqrt{f(X_i,y)+f'(X_i,y)}} s(X_i,y)\,\mathrm{d}y + R(f,f')$$

$$T(f,f') = \frac{1}{\sqrt{2}n}\sum_{i=0}^{n-1}\frac{\sqrt{f'(X_i,X_{i+1})}-\sqrt{f(X_i,X_{i+1})}}{\sqrt{f(X_i,X_{i+1})+f'(X_i,X_{i+1})}} + R(f,f')$$

Erreur d'approximation

Pour $f, f' \in \mathbb{L}^1_+([0,1]^{2d}, M)$ on a

$$T_E(f,f') = \frac{1}{\sqrt{2}n}\sum_{i=0}^{n-1}\int \frac{\sqrt{f'(X_i,y)}-\sqrt{f(X_i,y)}}{\sqrt{f(X_i,y)+f'(X_i,y)}}s(X_i,y)\,\mathrm{d}y + R(f,f')$$

$$T(f,f') = \frac{1}{\sqrt{2}n}\sum_{i=0}^{n-1}\frac{\sqrt{f'(X_i,X_{i+1})}-\sqrt{f(X_i,X_{i+1})}}{\sqrt{f(X_i,X_{i+1})+f'(X_i,X_{i+1})}} + R(f,f')$$

Soit $Z(f,f') = T(f,f') - T_E(f,f')$ l'erreur d'approximation. Alors,

$$Z(f,f') = \frac{1}{n}\sum_{i=0}^{n-1}\left\{\psi\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) - \mathbb{E}\left[\psi\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) \mid X_i\right]\right\}$$

avec $\psi(x) = \frac{1}{\sqrt{2}}\frac{\sqrt{x}-1}{\sqrt{x+1}}$.

Erreur d'approximation

- Soient $f, f' \in S$ et

$$Z(f,f') = \frac{1}{n}\sum_{i=0}^{n-1}\left\{\psi\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) - \mathbb{E}\left[\psi\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) \mid X_i\right]\right\}$$

avec $\psi(x) = \frac{1}{\sqrt{2}}\frac{\sqrt{x}-1}{\sqrt{x+1}}$.

Erreur d'approximation

- Soient $f, f' \in S$ et

$$Z(f,f') = \frac{1}{n}\sum_{i=0}^{n-1}\left\{\psi\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) - \mathbb{E}\left[\psi\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) \mid X_i\right]\right\}$$

avec $\psi(x) = \frac{1}{\sqrt{2}}\frac{\sqrt{x}-1}{\sqrt{x+1}}$.

- Soit

$$\vartheta(f,f') = \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[\psi^2\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) \mid X_i\right]$$

Erreur d'approximation

- Soient $f, f' \in S$ et

$$Z(f,f') = \frac{1}{n}\sum_{i=0}^{n-1}\left\{\psi\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) - \mathbb{E}\left[\psi\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) \mid X_i\right]\right\}$$

avec $\psi(x) = \frac{1}{\sqrt{2}}\frac{\sqrt{x}-1}{\sqrt{x+1}}$.

- Soit

$$\vartheta(f,f') = \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[\psi^2\left(\frac{f'(X_i,X_{i+1})}{f(X_i,X_{i+1})}\right) \mid X_i\right]$$

- Supposons qu'il existe c et D tel que pour tout $f, f' \in S$, sur un événement

$$Z(f,f') \leq c\left\{\vartheta(f,f') + \frac{D}{n}\right\}$$

alors sur cet évenement,

$$H^2(s,\hat{s}) \leq C\left\{H^2(s,S) + \frac{D}{n}\right\}$$

Erreur d'approximation

- Problématique: soit $\mathcal{G} = \{\psi(f'/f), f, f' \in S\}$ une collection de fonctions bornées par $1/\sqrt{2}$,

$$
\begin{aligned}
Z(g) &= \frac{1}{n}\sum_{i=0}^{n-1}\{g(X_i, X_{i+1}) - \mathbb{E}[g(X_i, X_{i+1}) \mid X_i]\} \\
\vartheta(g) &= \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[g^2(X_i, X_{i+1}) \mid X_i\right]
\end{aligned}
$$

Erreur d'approximation

- Problématique: soit $\mathcal{G} = \{\psi(f'/f), f, f' \in S\}$ une collection de fonction bornées par $1/\sqrt{2}$,

$$
\begin{aligned}
Z(g) &= \frac{1}{n}\sum_{i=0}^{n-1}\{g(X_i, X_{i+1}) - \mathbb{E}[g(X_i, X_{i+1}) \mid X_i]\} \\
\vartheta(g) &= \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[g^2(X_i, X_{i+1}) \mid X_i\right]
\end{aligned}
$$

- Sous quelles conditions sur $\mathcal{G}$ a t-on pour tout $g \in \mathcal{G}$ et grande probabilité:

$$
Z(g) \le c\left\{\vartheta(g) + \frac{D}{n}\right\}
$$

Erreur d'approximation

- Problématique: soit $\mathcal{G} = \{\psi(f'/f), f, f' \in S\}$ une collection de fonction bornées par $1/\sqrt{2}$,

$$\begin{aligned} Z(g) &= \frac{1}{n}\sum_{i=0}^{n-1}\{g(X_i, X_{i+1}) - \mathbb{E}\left[g(X_i, X_{i+1}) \mid X_i\right]\} \\ \vartheta(g) &= \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[g^2(X_i, X_{i+1}) \mid X_i\right] \end{aligned}$$

- Supposons $\mathcal{G}$ **finie**: alors, pour tout $\xi > 0$,

$$\mathbb{P}\left[Z(g) \le c\left\{\vartheta(g) + \frac{\log|\mathcal{G}|}{n} + \xi\right\}\right] \ge 1 - e^{-n\xi}$$

Un résultat

Theorem

Soit $S \subset \mathbb{L}^1_+([0,1]^{2d}, M)$ *un ensemble fini. Alors, il existe un* ρ*-estimateur tel que pour tout* $\xi > 0$,

$$\mathbb{P}\left[H^2(s,\hat{s}) \leq C\left\{H^2(s,S) + \frac{\log|S|}{n} + \xi\right\}\right] \geq 1 - e^{-n\xi},$$

où C *est une constante universelle. En particulier,*

$$\mathbb{E}\left[H^2(s,\hat{s})\right] \leq C'\left\{\mathbb{E}\left[H^2(s,S)\right] + \frac{\log|S|}{n}\right\}.$$

Un résultat

- Soit V un sous espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$.

Un résultat

- Soit V un sous espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$.
- Soit $S_V(\eta)$ un réseau de la boule $\mathcal{B}(0,2)$ de V de pas η.
- On pose alors $S = \left\{ f_+^2,\ f \in S_V(\eta) \right\}$

Un résultat

- Soit V un sous espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$.
- Soit $S_V(\eta)$ un réseau de la boule $\mathcal{B}(0,2)$ de V de pas η.
- On pose alors $S = \left\{f_+^2,\ f \in S_V(\eta)\right\}$
- Cela donne un estimateur $\hat{s}$ tel que

$$\mathbb{E}\left[H^2(s,\hat{s})\right] \leq C\left\{\inf_{f\in S_V(\eta)} \mathbb{E}\left[H^2(s,f_+^2)\right] + \frac{\log|S_V(\eta)|}{n}\right\}$$

Un résultat

- Soit V un sous espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$.
- Soit $S_V(\eta)$ un réseau de la boule $\mathcal{B}(0,2)$ de V de pas η.
- On pose alors $S = \left\{f_+^2,\ f \in S_V(\eta)\right\}$
- Cela donne un estimateur $\hat{s}$ tel que

$$\mathbb{E}\left[H^2(s,\hat{s})\right] \leq C\left\{\inf_{f\in S_V(\eta)} \mathbb{E}\left[H^2(s,f_+^2)\right] + \frac{\log|S_V(\eta)|}{n}\right\}$$

- Si les variables X_i admettent une densité φ_i majorée par κ,

$$\begin{aligned}\mathbb{E}\left[H^2(s,f_+^2)\right] &\leq \frac{1}{2n}\sum_{i=0}^{n-1}\int_{[0,1]^{2d}}\left(\sqrt{s(x,y)} - f(x,y)\right)^2 \varphi_i(x)\,\mathrm{d}x\,\mathrm{d}y \\ &\leq \frac{\kappa}{2} d_2^2(\sqrt{s}|_{[0,1]^{2d}}, f)\end{aligned}$$

Un résultat

- Soit V un sous espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$.
- Soit $S_V(\eta)$ un réseau de la boule $\mathcal{B}(0,2)$ de V de pas η.
- On pose alors $S = \left\{f_+^2,\ f \in S_V(\eta)\right\}$
- Cela donne un estimateur $\hat{s}$ tel que

$$\mathbb{E}\left[H^2(s,\hat{s})\right] \leq C\left\{\inf_{f \in S_V(\eta)} \mathbb{E}\left[H^2(s,f_+^2)\right] + \frac{\log|S_V(\eta)|}{n}\right\}$$

- Si les variables X_i admettent une densité φ_i majorée par κ,

$$\inf_{f \in S_V(\eta)} \mathbb{E}\left[H^2(s,f_+^2)\right] \leq \frac{\kappa}{2} d_2^2(\sqrt{s}|_{[0,1]^{2d}}, S_V(\eta))$$

Un résultat

- Soit V un sous espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$.
- Soit $S_V(\eta)$ un réseau de la boule $\mathcal{B}(0,2)$ de V de pas η.
- On pose alors $S = \left\{f_+^2,\ f \in S_V(\eta)\right\}$
- Cela donne un estimateur $\hat{s}$ tel que

$$\mathbb{E}\left[H^2(s,\hat{s})\right] \leq C\left\{\inf_{f\in S_V(\eta)} \mathbb{E}\left[H^2(s,f_+^2)\right] + \frac{\log|S_V(\eta)|}{n}\right\}$$

- Si les variables X_i admettent une densité φ_i majorée par κ,

$$\inf_{f\in S_V(\eta)} \mathbb{E}\left[H^2(s,f_+^2)\right] \leq \frac{\kappa}{2}\left(d_2\left(\sqrt{s}|_{[0,1]^{2d}}, V\cap\mathcal{B}(0,2)\right) + \eta\right)^2$$

Un résultat

- Soit V un sous espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$.
- Soit $S_V(\eta)$ un réseau de la boule $\mathcal{B}(0,2)$ de V de pas η.
- On pose alors $S = \left\{f_+^2,\ f \in S_V(\eta)\right\}$
- Cela donne un estimateur $\hat{s}$ tel que

$$\mathbb{E}\left[H^2(s,\hat{s})\right] \leq C\left\{\inf_{f \in S_V(\eta)} \mathbb{E}\left[H^2(s,f_+^2)\right] + \frac{\log|S_V(\eta)|}{n}\right\}$$

- Si les variables X_i admettent une densité φ_i majorée par κ,

$$\inf_{f \in S_V(\eta)} \mathbb{E}\left[H^2(s,f_+^2)\right] \leq \frac{\kappa}{2}\left(d_2\left(\sqrt{s}|_{[0,1]^{2d}}, V\right) + \eta\right)^2$$

Un résultat

- Soit V un sous espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$.
- Soit $S_V(\eta)$ un réseau de la boule $\mathcal{B}(0,2)$ de V de pas η.
- On pose alors $S = \left\{f_+^2,\ f \in S_V(\eta)\right\}$
- Cela donne un estimateur $\hat{s}$ tel que

$$\mathbb{E}\left[H^2(s,\hat{s})\right] \leq C \left\{ \inf_{f \in S_V(\eta)} \mathbb{E}\left[H^2(s, f_+^2)\right] + \frac{\log |S_V(\eta)|}{n} \right\}$$

- Si les variables X_i admettent une densité φ_i majorée par κ,

$$\inf_{f \in S_V(\eta)} \mathbb{E}\left[H^2(s, f_+^2)\right] \leq \frac{\kappa}{2} \left(d_2 \left(\sqrt{s}|_{[0,1]^{2d}}, V \right) + \eta \right)^2$$

- En choisissant bien η:

$$\mathbb{E}\left[H^2(s,\hat{s})\right] \leq C' \left[d_2^2(\sqrt{s}|_{[0,1]^{2d}}, V) + \frac{\dim V \log (n/\dim V)}{n} \right].$$

où C' dépend de κ.

Un théorème de sélection de modèle

Assumption

Pour tout $i \in \{0, \dots, n-1\}$, X_i admet une densité par rapport à la mesure de Lebesgue majorée par κ.

Theorem

Soit $\mathbb{V}$ une collection au plus dénombrable d'espace vectoriel de dimension finie de $\mathbb{L}^2([0,1]^{2d})$ et Δ une application positive sur $\mathbb{V}$ telle que $\sum_{V \in \mathbb{V}} e^{-\Delta(V)} \leq 1$. Sous l'hypothèse ci dessus, il existe un ρ-estimateur pénalisé tel que

$$\mathbb{E}\left[H^2(s, \hat{s})\right] \leq C \inf_{V \in \mathbb{V}} \left\{ d_2^2\left(\sqrt{s}|_{[0,1]^{2d}}, V\right) + \frac{\Delta(V) + \dim(V) \log\left(n/\dim V\right)}{n} \right\}$$

où $C > 0$ dépend de κ

Modèle de fonctions constantes par morceaux

Soit

$$V_m = \left\{ \sum_{K \in m} a_K \mathbb{1}_K, \ \forall K \in m, \ a_K \in [0, +\infty) \right\}.$$

Le ρ-estimateur définie sur $S = V_m$ est l'histogramme

$$\hat{s}_m = \sum_{K \in m} \frac{\sum_{i=0}^{n-1} \mathbb{1}_K(X_i, X_{i+1})}{\sum_{i=0}^{n-1} \int_{[0,1]^d} \mathbb{1}_K(X_i, y) \, \mathrm{d}y} \mathbb{1}_K.$$

Modèle de fonctions constantes par morceaux

Soit

$$V_m = \left\{ \sum_{K \in m} a_K \mathbb{1}_K,\ \forall K \in m,\ a_K \in [0, +\infty) \right\}.$$

Le ρ-estimateur définie sur $S = V_m$ est l'histogramme

$$\hat{s}_m = \sum_{K \in m} \frac{\sum_{i=0}^{n-1} \mathbb{1}_K(X_i, X_{i+1})}{\sum_{i=0}^{n-1} \int_{[0,1]^d} \mathbb{1}_K(X_i, y)\, \mathrm{d}y} \mathbb{1}_K.$$

Proposition

Pour toute partition m finie de $[0,1]^{2d}$,

$$\mathbb{E}\left[H^2(s, \hat{s}_m)\right] \leq C \left\{ \mathbb{E}\left[H^2(s, V_m)\right] + \frac{1 + \log n}{n} |m| \right\}$$

où $C = 4 + \log 2$.

Objectif

Soit $\mathcal{M}$ une collection (finie) de partitions. Le but est de sélectionner un estimateur parmi la collection

$$\{\hat{s}_m,\ m \in \mathcal{M}\}.$$

Partitions

- On considère des partitions issues de l'algorithme récursif de Devore et Yu (1990)
- Pour tout $\ell \in \mathbb{N}^\star$, $\mathcal{M}_\ell$ désigne l'ensemble des partitions construits par l'algorithme en au plus ℓ étapes.
- $\mathcal{M}_\infty = \cup_{\ell \geq 1} \mathcal{M}_\ell$

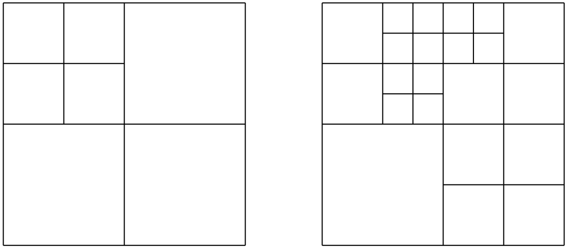


Figure: Gauche: exemple de partition de $\mathcal{M}_2$. Droite: exemple de partition de $\mathcal{M}_3$. ($d = 1$)

Critère

Soit pour $m \in \mathcal{M}_\ell$,

$$\gamma(m) = \left\{ \sum_{K \in m} \sup_{m' \in \mathcal{M}_\ell} \left[T\left(\hat{s}_m \mathbb{1}_K, \hat{s}_{m'} \mathbb{1}_K\right) - \mathrm{pen}(m' \vee K) \right] \right\} + 2\mathrm{pen}(m).$$

où $m' \vee K$ est la partition de K définie par

$$m' \vee K = \{K' \cap K,\ K' \in m',\ K \cap K' \neq \emptyset\}$$

et où

$$\mathrm{pen}\left(m' \vee K\right) = L \frac{|m' \vee K| \log n}{n}.$$

Critère

Soit pour $m \in \mathcal{M}_\ell$,

$$\gamma(m) = \left\{ \sum_{K \in m} \sup_{m' \in \mathcal{M}_\ell} \left[T\left(\hat{s}_m \mathbb{1}_K, \hat{s}_{m'} \mathbb{1}_K\right) - \mathrm{pen}(m' \vee K) \right] \right\} + 2\mathrm{pen}(m).$$

où $m' \vee K$ est la partition de K définie par

$$m' \vee K = \{K' \cap K,\ K' \in m',\ K \cap K' \neq \emptyset\}$$

et où

$$\mathrm{pen}\left(m' \vee K\right) = L \frac{|m' \vee K| \log n}{n}.$$

La partition $\hat{m}$ sélectionnée est alors celle qui minimise le critère γ.
Le ρ-estimateur pénalisé est alors $\hat{s}_{\hat{m}}$

Une inégalité oracle

Theorem

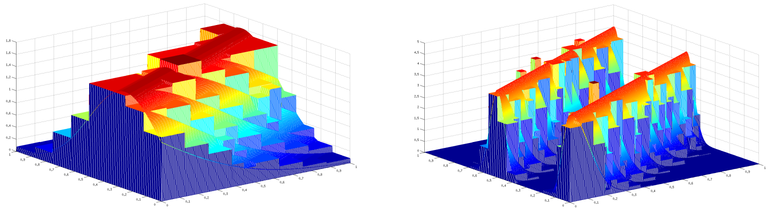
Il existe une constante universelle $L_0 > 0$ *telle que pour tout* $L \geq L_0$, $\ell \in \mathbb{N}^\star \cup \{\infty\}$, *l'estimateur* $\hat{s}$ *vérifie*

$$\mathbb{E}\left[H^2(s, \hat{s})\right] \leq C \inf_{m \in \mathcal{M}_\ell} \left\{\mathbb{E}\left[H^2(s, V_m)\right] + \mathrm{pen}(m)\right\}$$

où $C > 0$ *est universelle et où*

$$\mathrm{pen}(m) = L \frac{|m| \log n}{n}.$$

Simulations



$X_{k+1} = 0.5X_k + 0.25(1 + \mathcal{N}_k)$ $\qquad$ $X_{k+1} = 0.5X_k + 0.5U_k.$

Figure: Densité de transition et ρ-estimateur

$\mathcal{N}_k \sim \mathcal{N}(0, 1)$
U_k mélange de deux gaussiennes $0.5\mathcal{N}(0, 0.1) + 0.5\mathcal{N}(1, 0.1)$

Simulations

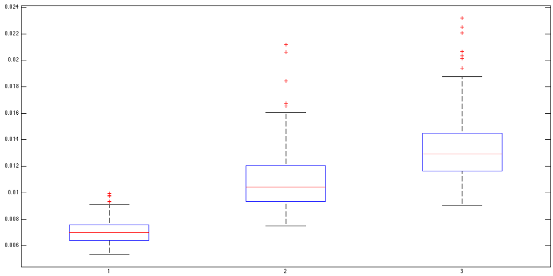


Figure: Risque Hellinger, cas 1

1: L'estimateur oracle, meilleur estimateur de $\{\hat{s}_m,\ m \in \mathcal{M}_\ell\}$.
2: Le ρ-estimateur pénalisé
3: L'estimateur basé sur un critère $\mathbb{L}^2$, proposé par Akakpo and Lacour (2011).

Simulations

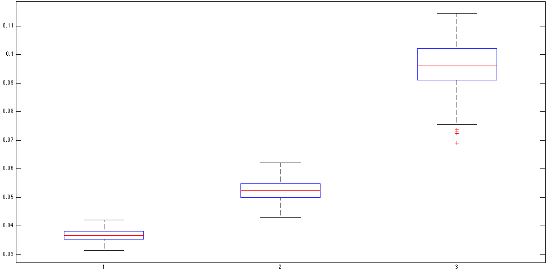


Figure: Risque Hellinger, cas 2

1: L'estimateur oracle, meilleur estimateur de $\{\hat{s}_m,\ m \in \mathcal{M}_\ell\}$.
2: Le ρ-estimateur pénalisé
3: L'estimateur basé sur un critère $\mathbb{L}^2$, proposé par Akakpo and Lacour (2011).

Références

- Estimation of the transition density of a Markov chain, Sart, 2014, *Annales de l'Institut Henri Poincaré. Probabilités et Statistique*
- A new method for estimation and model selection: ρ-estimation, Baraud, Birgé, Sart, 2016, *Inventiones Mathematicae*

Merci de votre attention