An inverse problem point of view for adaptive estimation in a shifted curves model

Jérémie Bigot & Sébastien Gadat

Insitut de Mathématiques de Toulouse
Université de Toulouse

5èmes Rencontre Statistiques Mathématiques
Bordeaux-Santander-Toulouse-Valladolid

Parc Ornithologique du Teich, June 2009

Grenander's pattern theory (1993)

Data : a set of n similar curves or images obtained through the deformation of the same template

A general model : observation of $Y_m : \Omega \to \mathbb{R}, \; m = 1, \ldots, n$ where $\Omega \subset \mathbb{R}^d$ with $d = 1, 2, 3$ such that

$$Y_m(x) = f(\phi_m(x)) + W_m(x), \; \text{for } x \in \Omega,$$

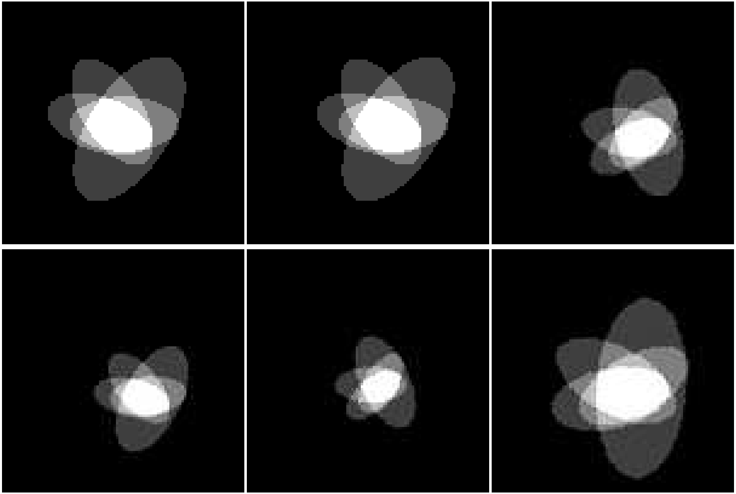
where

- $f : \mathbb{R}^d \to \mathbb{R}$ is a common unknown template
- $\phi_m : \mathbb{R}^d \to \mathbb{R}^d$ are unknown deformations, possibly random
- W_m some additive noise

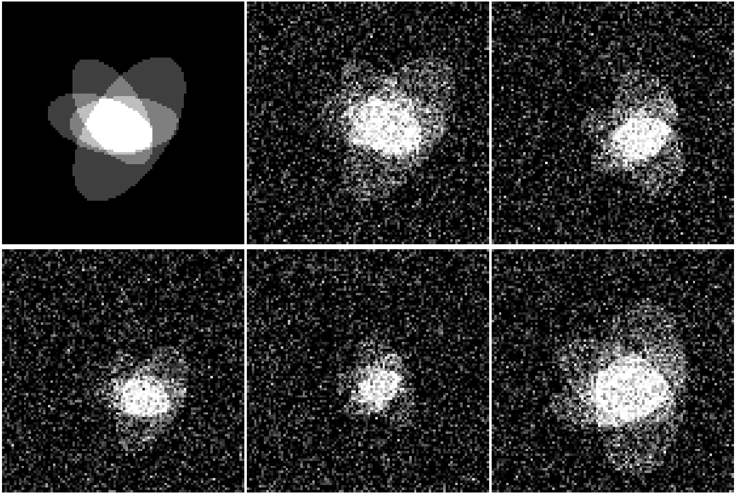
Problem : to recover f as $n \to +\infty$

- In statistics work by Gasser, Kneip, Silverman, Ramsay, mainly on curve alignment...
- In image processing work by Amit, Grenander, Joshi, Miller, Trouvé, Younes...
- Recently work by Gamboa, Loubes, Maza, Vimond, Bigot, Gadat

A first example of a template estimation problem



A first example of a template estimation problem



A second example of a template estimation problem

Estimating f a deconvolution problem ?

Direct mean of the observed images - blurring effect

Different models for the deformations

Rigid deformations

- Translation : $\phi(x) = x - b$ where $b \in \mathbb{R}^d$
- Rotation + scaling (in $\mathbb{R}^2$) : $\phi(x) = \frac{1}{a} A_\theta x$ where $a \in \mathbb{R}^+$ and

$$A_\theta = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

- Affine (Translation + rotation + scaling) : $\phi(x) = \frac{1}{a} A_\theta (x - b)$, either 2D or 3D

Different models for the deformations

Non-rigid deformations

- Small deformations : $\phi(x) = x + h(x)$ where $h : \mathbb{R}^d \to \mathbb{R}^d$ is an unconstrained function . Problem ϕ is not necessarly invertible if h is large. (Work by Faugeras, Amit,...)
- Large deformations (i.e. **diffeomorphisms**) : $\phi(x)$ is an invertible and smooth deformation from $\mathbb{R}^d$ to $\mathbb{R}^d$ (Work by Grenander, Trouvé, Younes, Miller,...)

Simplest model : shifted 1D curves

Observations : independent realizations of n noisy and shifted curves $Y_1, \ldots Y_n$ coming from the model

$$dY_m(x) = f(x - \tau_m)dx + \epsilon dW_m(x),\ x \in [0,1],\ m = 1, \ldots, n$$

where

- $f : [0,1] \to \mathbb{R}$ is the unknown common shape of the curves (with period 1)
- W_m are independent standard Brownian motions on $[0,1]$
- ϵ level of noise in each curve

Remark : $\epsilon \to 0$ corresponds to $N \to +\infty$ in the model (with $\epsilon = \frac{\sigma}{\sqrt{N}}$)

$$Y_{m,i} = f(x_i - \tau_m) + \sigma z_{m,i},\ x_i = \frac{i}{N},\ i = 1, \ldots, N,\ \text{and } z_{m,i} \sim_{i.i.d.} N(0,1)$$

Simplest model : shifted 1D curves

Observations : independent realizations of n noisy and shifted curves $Y_1, \ldots Y_n$ coming from the model

$$dY_m(x) = f(x - \tau_m)dx + \epsilon dW_m(x),\ x \in [0,1],\ m = 1, \ldots, n$$

Different models for the shifts τ_m :

- Deterministic shifts : the τ_m are **fixed parameters** to estimate : semi-parameteric estimation in the setting n **fixed and** $\epsilon \to 0$ (Gamboa, Loubes & Maza (2007), Vimond (2008), extension to 2D images by Bigot, Gamboa & Vimond (2009))
- τ_m's are unknown **random shifts** independent of the W_m's such that

$$\tau_m \sim_{i.i.d} g \quad m = 1, \ldots, n,$$

where g is a unknown density on $\mathbb{R}$

Simplest model : shifted 1D curves

Observations : independent realizations of n noisy and shifted curves $Y_1, \dots Y_n$ coming from the model

$$dY_m(x) = f(x - \tau_m)dx + \epsilon dW_m(x), \; x \in [0,1], \; m = 1, \dots, n$$

This talk : case of random shifts $\tau_m \sim_{i.i.d} g$, $m = 1, \dots, n$, with **known** density g.

Problem : estimation of f in the asymptotic setting :

- $n \to +\infty$ and ϵ **is fixed** (This talk)
- $n \to +\infty$ and $\epsilon \to 0$ (Work in progress...)

A simple model for randomly shifted curves

Observations : independent realizations of n noisy and randomly shifted curves $Y_1, \ldots Y_n$ coming from the model

$$dY_m(x) = f(x - \tau_m)dx + \epsilon dW_m(x),\ x \in [0,1],\ m = 1, \ldots, n$$

Main objectives : estimating the function f and to derive asymptotic (as $n \to +\infty$) upper and lower bounds for the minimax risk

$$\mathcal{R}_n(\mathcal{F}) = \inf_{\hat{f}_n} \sup_{f \in \mathcal{F}} \mathcal{R}(\hat{f}_n, f),\ \text{where}$$

- $\mathcal{R}(\hat{f}_n, f) = \mathbb{E}\|\hat{f}_n - f\|^2 = \mathbb{E}\int_0^1 |\hat{f}_n(x) - f(x)|^2 dx$
- $\mathcal{F} \subset L^2([0,1])$ e.g a Sobolev or a Besov ball
- $\hat{f}_n$ a measurable function of the processes $\{Y_m,\ m = 1, \ldots, n\}$

Simplest case : no shifts

Observations : independent realizations of n noisy and curves $Y_1, \ldots Y_n$

$$dY_m(x) = f(x)dx + \epsilon dW_m(x),\ x \in [0,1],\ m = 1, \ldots, n$$

Classical result : if $\mathcal{F} = H^s(A)$ (Sobolev ball of radius A) or $\mathcal{F} = B^s_{p,q}(A)$ (Besov ball of radius A) with smoothness index s ("number of derivatives") then

$$\mathcal{R}_n(\mathcal{F}) = \inf_{\hat{f}_n} \sup_{f \in \mathcal{F}} \mathcal{R}(\hat{f}_n, f) \sim C n^{-\frac{2s}{2s+1}}$$

A connexion with a deconvolution problem

Model : $dY_m(x) = f(x - \tau_m)dx + \epsilon dW_m(x), \; x \in [0,1], \; m = 1, \ldots, n$

A deconvolution problem ? The expectation of each oberved curve is given by $\mathbb{E}\left[f(x - \tau_m)\right] = \int_{\mathbb{R}} f(x-\tau)g(\tau)d\tau = f \star g(x)$

Define

$$\xi_m(x) = f(x - \tau_m) - \int_{\mathbb{R}} f(x-\tau)g(\tau)d\tau,$$

$\xi(x) = \frac{1}{n}\sum_{m=1}^{n} \xi_m(x)$, and taking the mean of the n curves yields

$$dY(x) = \int_0^1 f(x-\tau)g(\tau)d\tau dx + \underbrace{\xi(x)dx}_{\text{Non-Gaussian Error}} + \underbrace{\frac{\epsilon}{\sqrt{n}}dW(x)}_{\text{Standard Gaussian Error}}, \; x \in [0,1],$$

A connexion with a deconvolution problem

Case of standard deconvolution with a Gaussian error :

$$dY(x) = \int_0^1 f(x-\tau)g(\tau)d\tau dx + \frac{\epsilon}{\sqrt{n}}dW(x) \; x \in [0,1],$$

Minimax rate of convergence : let $\gamma_\ell = \int_{-\infty}^{+\infty} e^{-i2\pi\ell x} g(x)dx$. Assume that for some real $\nu > 0$

$$C_{min}|\ell|^{-\nu} \leq |\gamma_\ell| \leq C_{max}|\ell|^{-\nu}.$$

for all $\ell \in \mathbb{Z}$.

Then for $\mathcal{F} = H^s(A)$ (Sobolev ball) or $\mathcal{F} = B^s_{p,q}(A)$ (Besov ball) with smoothness index s ("number of derivatives") then

$\mathcal{R}_n(\mathcal{F}) = \inf_{\hat{f}_n} \sup_{f \in \mathcal{F}} \mathcal{R}(\hat{f}_n, f) \sim Cn^{-\frac{2s}{2s+2\nu+1}}$ (instead of $n^{-\frac{2s}{2s+1}}$ in the direct case)

Model in the Fourier domain

For $\ell \in \mathbb{Z}$, let $\theta_\ell = \int_0^1 e^{-2i\ell\pi x} f(x)dx$ and $c_{m,\ell} = \int_0^1 e^{-2i\ell\pi x} dY_m(x)$. Then

$$\begin{aligned} c_{m,\ell} &= \theta_\ell e^{-i2\pi\ell\tau_m} + \epsilon_m z_{\ell,m} \text{ with } z_{\ell,m} \sim_{i.i.d.} N_{\mathbb{C}}(0,1) \\ &= \theta_\ell \gamma_\ell + \xi_{\ell,m} + \epsilon_m z_{\ell,m} \text{ with } \xi_{\ell,m} = \theta_\ell e^{-i2\pi\ell\tau_m} - \theta_\ell\gamma_\ell, \end{aligned}$$

where with $\gamma_\ell = \mathbb{E}\left(e^{-i2\pi\ell\tau}\right) = \int_{-\infty}^{+\infty} e^{-i2\pi\ell x} g(x)dx$.

Then, average the Fourier coefficients over the n curves

$$\tilde{c}_\ell = \frac{1}{n}\sum_{m=1}^{n} c_{\ell,m} = \theta_\ell\gamma_\ell + \underbrace{\xi_\ell}_{\text{Non-Gaussian Error}} + \underbrace{\frac{\epsilon}{\sqrt{n}}\eta_\ell}_{\text{Standard Gaussian Error}}, \text{ with } \eta_\ell \sim_{i.i.d.} N_{\mathbb{C}}(0,1)$$

with $\xi_\ell = \frac{1}{n}\sum_{m=1}^{n} \xi_{\ell,m}$.

Note that

$$\mathbb{E}|\xi_\ell|^2 = \frac{1}{n}|\theta_\ell|^2(1 - |\gamma_\ell|^2)$$

Problem : the variance of ξ_ℓ depends on the unknown $|\theta_\ell|^2$

Deconvolution in the Fourier domain

Assuming that the density g of the shifts is known, an estimation of θ_ℓ is given by

$$\hat{\theta}_\ell = \frac{\tilde{c}_\ell}{\gamma_\ell} = \theta_\ell + \frac{\xi_\ell}{\gamma_\ell} + \frac{\epsilon}{\sqrt{n}} \frac{\eta_\ell}{\gamma_\ell}$$

with $\gamma_\ell = \mathbb{E}\left(e^{-i2\pi\ell\tau}\right) = \int_{-\infty}^{+\infty} e^{-i2\pi\ell x} g(x) dx$.

Main assumption on g : polynomial decay of the γ_ℓ's i.e for some real $\nu > 0$,

$$C_{min}|\ell|^{-\nu} \leq |\gamma_\ell| \leq C_{max}|\ell|^{-\nu}.$$

for all $\ell \in \mathbb{Z}$.

Filtering in the Fourier domain

Linear estimator for f by spectra cut-off : take

$$\hat{\theta}_\ell^M = \frac{\tilde{c}_\ell}{\gamma_\ell}, \text{ for all } |\ell| \leq M$$

and

$$\hat{\theta}_\ell^M = 0, \text{ for all } |\ell| > M$$

where M is some integer to be chosen. For $\hat{f}_{n,M}(x) = \sum_{\ell \in \mathbb{Z}} \hat{\theta}_\ell^M e^{-i2\pi\ell x}$, one has

$$\mathcal{R}(\hat{f}_{n,M}, f) = \mathbb{E} \sum_{\ell \in \mathbb{Z}} |\hat{\theta}_\ell - \theta_\ell|^2.$$

Bias-variance decomposition of the risk

$$\mathcal{R}(\hat{f}_{n,M}, f) = \underbrace{\sum_{|\ell|>M} |\theta_\ell|^2}_{Bias} + \underbrace{\frac{1}{n} \sum_{|\ell|\leq M} \left[|\theta_\ell|^2 \left(\frac{1}{|\gamma_\ell|^2} - 1 \right) + \frac{\epsilon^2}{|\gamma_\ell|^2} \right]}_{Variance}.$$

Filtering in the Fourier domain

Define the following Sobolev ball of radius A :

$$H_s(A) = \left\{ f \in L^2([0,1]) \; ; \sum_{\ell \in \mathbb{Z}} (1 + |\ell|^{2s}) |\theta_\ell|^2 \leq A, \right\} \text{ with } A > 0, s > 0$$

Proposition

If $M = M_{n,s} \sim n^{\frac{1}{2s+2\nu+1}}$, *then* $\sup_{f \in H_s(A)} \mathcal{R}(\hat{f}_{n,M_{n,s}}, f) = \mathcal{O}(n^{-\frac{2s}{2s+2\nu+1}})$

Problem :

- $\hat{f}_{n,M_{n,s}}$ depends on the unknown regularity s (non-adaptive estimator)
- if f is piecewise C^s with s large, then $f \notin H_\alpha(A)$ for $\alpha > 1/2$. So,

$$\sup_{f \in \text{Piece-wise } C^s} \mathcal{R}(\hat{f}_{n,M_{n,s}}, f) = \mathcal{O}(n^{-\frac{1}{1+2\nu+1}})$$

(non-optimal estimator in standard deconvolution)

Meyer wavelets

Let $(\phi_{j_0,k}, \psi_{j,k})_{j \geq j_0, 0 \leq k \leq 2^j - 1}$ be the periodized Meyer wavelet basis of $L^2([0,1])$.

Advantages : Meyer wavelets are band-limited functions since for

$$\psi_\ell^{j,k} = \int_0^1 e^{-i2\pi \ell x} \psi_{j,k}(x) dx, \; \ell \in \mathbb{Z},$$

the set $C_j = \{\ell \in \mathbb{Z}; \psi_\ell^{j,k} \neq 0\}$ is finite with $\#\{C_j = c2^j\}$.

Then, **wavelet coefficients of f can be computed from its Fourier coefficients as**

$$\beta_{j,k} = \int_0^1 f(x) \psi_{j,k}(x) dx = \sum_{\ell \in C_j} \psi_\ell^{j,k} \theta_\ell, \text{ where } \theta_\ell = \int_0^1 e^{-2i\ell\pi x} f(x) dx.$$

Meyer wavelets = usefull tool for deconvolution (work by Johnstone *et al.* (2004), Pensky & Sapatinas (2008), and fast WaveD algorithm by Raimondo (2006))

Estimation by hard thresholding

Recall that

$$\hat{\theta}_\ell = \frac{\tilde{c}_\ell}{\gamma_\ell} = \theta_\ell + \frac{\xi_\ell}{\gamma_\ell} + \frac{\epsilon}{\sqrt{n}} \frac{\eta_\ell}{\gamma_\ell}$$

and estimation of the wavelet coefficients of f is then given by

$$\hat{\beta}_{j,k} = \sum_{\ell \in C_j} \psi_\ell^{j,k} \hat{\theta}_\ell \text{ and } \hat{c}_{j_0,k} = \sum_{\ell \in C_{j_0}} \phi_\ell^{j_0,k} \hat{\theta}_\ell.$$

Non-linear estimation by hard-thresholding

$$\hat{f}_n^h = \sum_{k=0}^{2^{j_0}-1} \hat{c}_{j_0,k} \phi_{j_0,k} + \sum_{j=j_0}^{j_1} \sum_{k=0}^{2^j-1} \hat{\beta}_{j,k} \mathbb{1}_{\{|\hat{\beta}_{j,k}| \geqslant \lambda_{j,k}\}} \psi_{j,k}$$

where $\lambda_{j,k}$ is a threshold to be calibrated.

Probability of deviation

Proposition

For $0 \leq k \leq 2^j - 1$, *define*

$$\sigma_j^2 = 2^{-j}\epsilon^2 \sum_{\ell \in \Omega_j} |\gamma_\ell|^{-2}, V_j^2 = \|g\|_\infty 2^{-j} \sum_{\ell \in \Omega_j} \frac{|\theta_\ell|^2}{|\gamma_\ell|^2} \text{ and } \delta_j = 2^{-j/2} \sum_{\ell \in \Omega_j} \frac{|\theta_\ell|}{|\gamma_\ell|}.$$

For any $t > 0$, *define*

$$\alpha_{j,k}(t) = 2 \max \left(\sigma_j \sqrt{\frac{2t}{n}}, \sqrt{\frac{2V_j^2 t}{n}} + \delta_j \frac{t}{3n} \right)$$

Then,

$$\mathbb{P}\left(|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \alpha_{j,k}(t) \right) \leq 2\exp(-t).$$

Definition of the threshold

Estimation of $|\theta_l|^2$ is given by

$$|\hat{\theta}|_\ell^2 = \frac{1}{n}\sum_{m=1}^{n}(|c_{m,\ell}|^2 - \epsilon^2) \text{ where } |c_{m,\ell}|^2 = |\theta_\ell|^2 + \epsilon^2 |z_{\ell,m}|^2 - 2\mathcal{R}\left(e^{i2\pi\ell\tau_m} z_{\ell,m}\right)$$

For some contant $\eta \geq 2$, use random thresholds of the form

$$\lambda_{j,k} = \lambda_j = 2\max\left(\sigma_j\sqrt{\frac{2\eta\log(n)}{n}}, \sqrt{\frac{2\hat{V}_j^2\eta\log(n)}{n}} + \hat{\delta}_j\frac{\eta\log(n)}{3n}\right)$$

with $\hat{V}_j^2 = \|g\|_\infty 2^{-j}\sum_{\ell\in\Omega_j}\frac{|\tilde{\theta}_\ell|^2}{|\gamma_\ell|^2}$ and $\hat{\delta}_j = 2^{-j/2}\sum_{\ell\in\Omega_j}\frac{|\tilde{\theta}_\ell|}{|\gamma_\ell|}$, where

$$|\tilde{\theta}_\ell| = \sqrt{|\hat{\theta}|_\ell^2 + \epsilon^2} + \epsilon\sqrt{\frac{2\log(n^2 2^j)}{n}}.$$

Then $\mathbb{P}\left(|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \lambda_j\right) \leq 2n^{-\eta}$

Adaptive estimation over Besov spaces

Theorem

Assume that $2^{j_1} \sim \left(\frac{n}{\log(n)}\right)^{\frac{1}{2\nu+1}}$ *and* $2^{j_0} \sim \log(n)$. *Recall that*

$$\hat{f}_n^h = \sum_{k=0}^{2^{j_0}-1} \hat{c}_{j_0,k}\phi_{j_0,k} + \sum_{j=j_0}^{j_1}\sum_{k=0}^{2^j-1} \hat{\beta}_{j,k}\mathbb{1}_{\{|\hat{\beta}_j| \geqslant \lambda_{j,k}\}}\psi_{j,k}$$

Then, for $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $A > 0$

$$\sup_{f \in B^s_{p,q}(A)} \|\hat{f}_n^h - f\|^2 = \mathcal{O}\left(\left(\frac{n}{\log(n)}\right)^{-\frac{2s}{2s+2\nu+1}}\right),$$

with $s > 1/p'$, $(s + 1/2 - 1/p')p > \nu(2-p)$ *with* $p' = \min(2,p)$

Asymptotic lower bound

Theorem

Let $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $s \geq 1/p$ *and* $A > 0$. *Then, if*

$$s > \nu + 1/2 \text{ and } \nu > 1/2,$$

there exists a constant $C > 0$ *depending only on* A, s, p, q *such that*

$$\lim_{n \to +\infty} n^{\frac{2s}{2s+2\nu+1}} \mathcal{R}_n(B^s_{p,q}(A)) \geq C$$

Likelihood ratio of the model

Hypothesis $H_0 : dY_m(x) = \epsilon dW_m(x),\ m = 1,\dots,n$

Hypothesis $H_f : dY_m(x) = f(x - \tau_m)dx + \epsilon dW_m(x),\ m = 1,\dots,n$

By conditionning with respect to the τ_m 's and using the Girsanov formula, one has that

$$
\begin{aligned}
\Lambda_n(f,h) &= \frac{d\mathbb{P}_{H_f}}{d\mathbb{P}_{H_h}} \\
&= \prod_{i=1}^{n} \exp\left(\int_0^1 (f(x-\tau_i) - h(x-\tau_i))dY_i(x) + \frac{1}{2}\|h\|^2 - \frac{1}{2}\|f\|^2 \right).
\end{aligned}
$$

Assouad's cube technique (e.g HKPT (1998))

Let $j \geq 0$ and consider for any $\eta = (\eta_k)_{k \in \{0 \dots 2^j - 1\}} \in \{\pm 1\}^{2^j}$ the function $f_{j,\eta}$ defined as

$$f_{j,\eta} = \gamma_j \sum_{k=0}^{2^j - 1} \eta_k \psi_{j,k} \text{ with } \gamma_j = c2^{-j(s+1/2)}, \text{ and } c \text{ such that } f_{j,\eta} \in B^s_{p,q}(A)$$

Let also the vector $\eta^i \in \{\pm 1\}^{2^j}$ with components equal to those of η except the i^{th} one.

Classically a lower bound for the minimax rate is obtained by controlling (e.g HKPT (1998))

$$\mathbb{P}_{f_{j,\eta}} \left(\Lambda_n(f_{j,\eta^i}, f_{j,\eta}) \geq e^{-\lambda} \right) \geq \pi_0$$

Problem : in this way, the inverse problem does not appear

Adaptation of a classical Lemma (e.g HKPT (1998))

Let

$$\mathbb{E}_\tau\left(\Lambda_n(f_{j,\eta^i}, f_0)\right) = \int_{\mathbb{R}^n} \prod_{i=1}^n e^{\int_0^1 (f(x-\tau_i) - h(x-\tau_i))dY_i(x) + \frac{1}{2}\|h\|^2 - \frac{1}{2}\|f\|^2} g(\tau_1)\dots g(\tau_n) d\tau_1 \dots d\tau_n$$

Lemma

Suppose there exists $\lambda > 0$ and $\pi_0 > 0$ such that for all sufficiently large n

$$\mathbb{P}_{f_{j,\eta}}\left(\frac{\mathbb{E}_\tau\left(\Lambda_n(f_{j,\eta^i}, f_0)\right)}{\mathbb{E}_\tau\left(\Lambda_n(f_{j,\eta}, f_0)\right)} \geq e^{-\lambda}\right) \geq \pi_0,$$

for all $f_{j,\eta}$ and all $i \in \{0 \dots 2^j - 1\}$. Then, there exists a positive constant C, such that for all sufficiently large n and any estimator $\hat{f}_n$

$$\max_{\eta \in \{\pm 1\}^{2^j}} \mathbb{E}_{f_{j,\eta}} \|\hat{f}_n - f_{j,\eta}\|^2 \geq C\pi_0 e^{-\lambda} 2^j \gamma_j^2$$

Control of the expectation over the random shifts of the likelihood ratio

Lemma

If $j = j(n)$ such that $2^{j(n)} \sim n^{\frac{1}{2s+2\nu+1}}$, then there exists $\lambda > 0$ and $\pi_0 > 0$ such that for all sufficiently large n

$$\mathbb{P}_{f_{j,\eta}}\left(\frac{\mathbb{E}_{\tau}\left(\Lambda_n(f_{j,\eta^i}, f_0)\right)}{\mathbb{E}_{\tau}\left(\Lambda_n(f_{j,\eta}, f_0)\right)} \geq e^{-\lambda}\right) \geq \pi_0.$$

provided $s > \nu + 1/2$ and $\nu > 1/2$.

Since $\gamma_j = c2^{-j(s+1/2)}$, the choice $2^{j(n)} \sim n^{\frac{1}{2s+2\nu+1}}$ implies that

$$\max_{\eta \in \{\pm 1\}^{2^j}} \mathbb{E}_{f_{j,\eta}} \|\hat{f}_n - f_{j,\eta}\|^2 \geq C\pi_0 e^{-\lambda} 2^j \gamma_j^2 \sim n^{-\frac{2s}{2s+2\nu+1}}$$

Comparison with Procrustean mean

Iterative procedure (Kneip & Gasser (1988), Wang & Gasser (1997))

- Initialisation : $\hat{f}_0 = \frac{1}{n} \sum_{m=1}^{n} Y_m$
- For $1 \leq i \leq i_{\max}$ do
  - For $1 \leq m \leq n$ compute

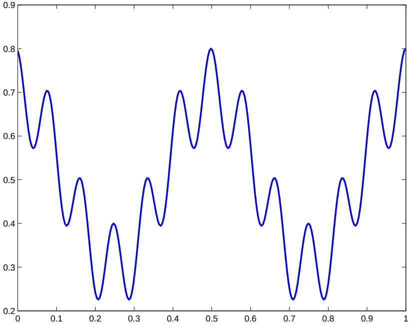
$$\hat{\tau}_{m,i} = \arg\min_{\tau \in \mathbb{R}} \left\| Y_m(\cdot + \tau) - \hat{f}_{i-1} \right\|^2$$

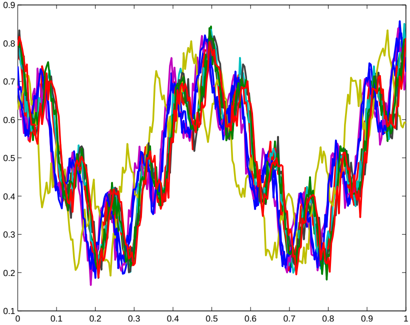
  - Then take $\hat{f}_i(x) = \frac{1}{n} \sum_{m=1}^{n} Y_m(x + \hat{\tau}_{m,i})$

Fast convergence ($i_{\max} = 3$ is enough) but it highly depends on the initialisation $\hat{f}_0$

Wave example

Laplace distribution $g(x) = \frac{1}{\sqrt{2}\sigma} \exp\left(-\sqrt{2}\frac{|x|}{\sigma}\right)$ for $x \in \mathbb{R}$, and
$\gamma_\ell = \frac{1}{1+2\sigma^2\pi^2\ell^2}$ i.e $\nu = 2$

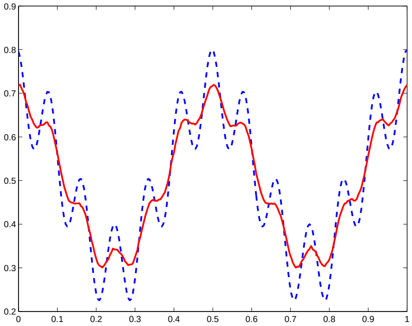




True f and a sample of 10 noisy curves out of $n = 200$
Curves are sampled at $N = 256$ equally spaced points on $[0, 1]$

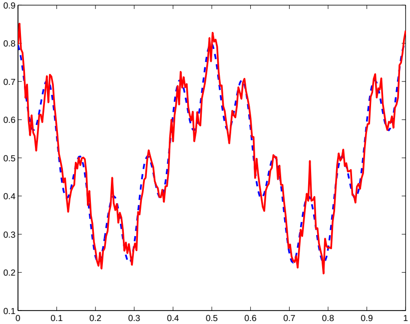
Wave example - Direct mean

Direct mean of the $n = 200$ observed curves



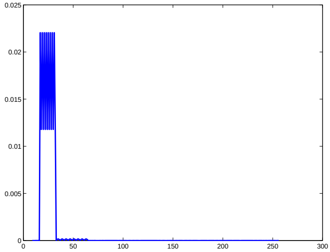
Wave example - Deconvolution step

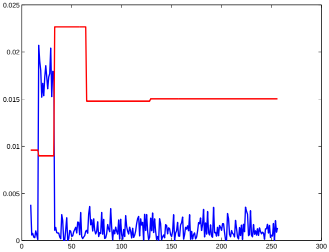
Deconvolution of the direct mean without any smoothing/thresholding



Wave example - Wavelet thresholding

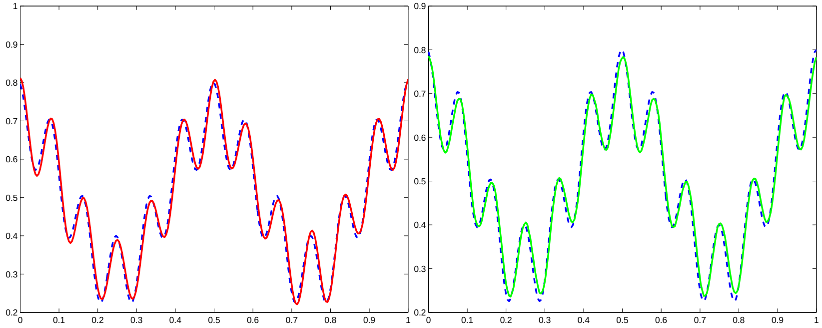
Choice for the resolution levels : $j_0 = 3$ and $j_1 = 7 = \log 2(N) - 1$





True wavelet coefficients - Noisy wavelet coefficients and threshold λ_j

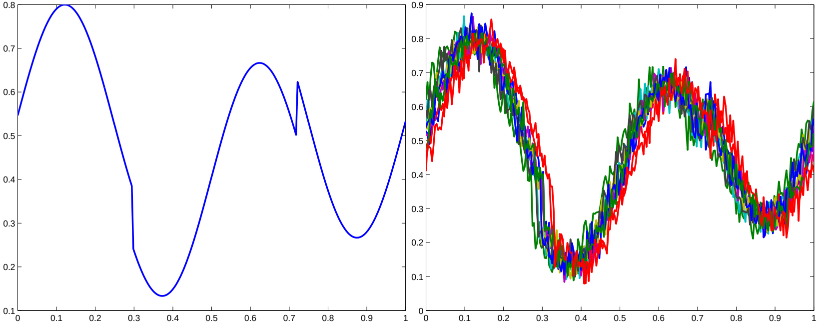
Wave example - Comparison with Procrustean mean



Wavelet thresholding - Procrustean mean

HeaviSine example

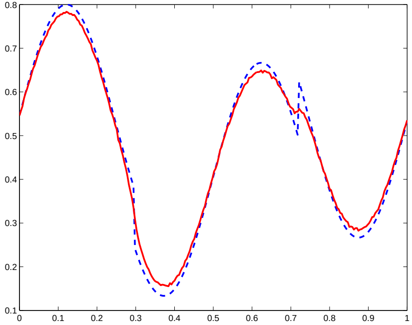
Laplace distribution $g(x) = \frac{1}{\sqrt{2}\sigma} \exp\left(-\sqrt{2}\frac{|x|}{\sigma}\right)$ for $x \in \mathbb{R}$, and $\gamma_\ell = \frac{1}{1+2\sigma^2\pi^2\ell^2}$ i.e $\nu = 2$



True f and a sample of 10 noisy curves out of $n = 200$

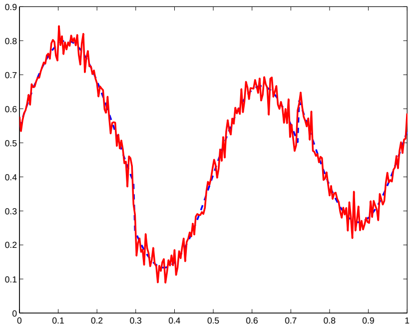
HeaviSine example - Direct mean

Direct mean of the $n = 200$ observed curves



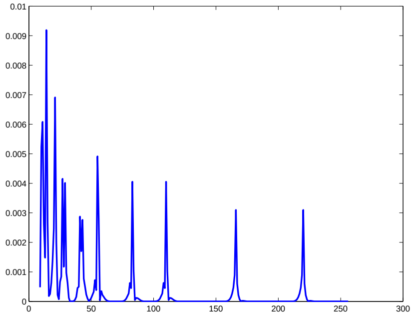
HeaviSine example - Deconvolution step

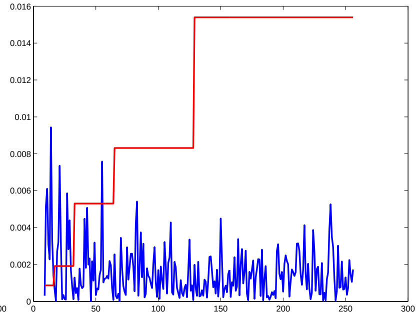
Deconvolution of the direct mean without any smoothing/thresholding



HeaviSine example - Wavelet thresholding

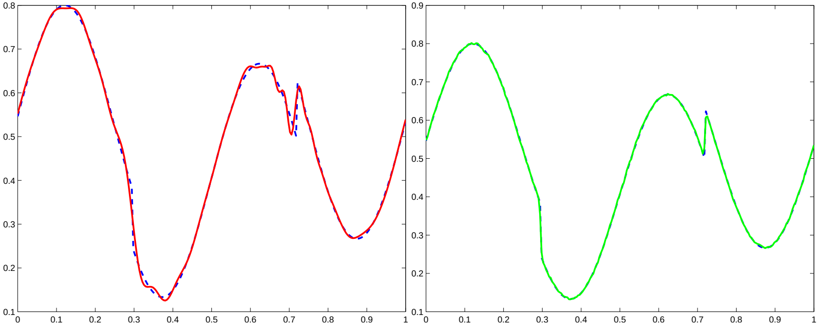
Choice for the resolution levels : $j_0 = 3$ and $j_1 = 7 = \log 2(N) - 1$





True wavelet coefficients - Noisy wavelet coefficients and threshold λ_j

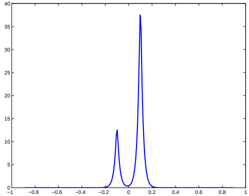
HeaviSine example - Comparison with Procrustean mean

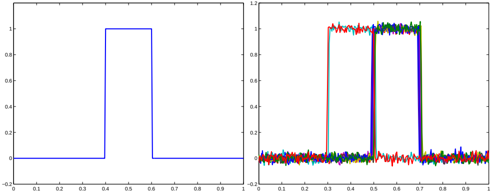


Wavelet thresholding - Procrustean mean

Step example

Density of the shifts g : mixture of two Laplace densities

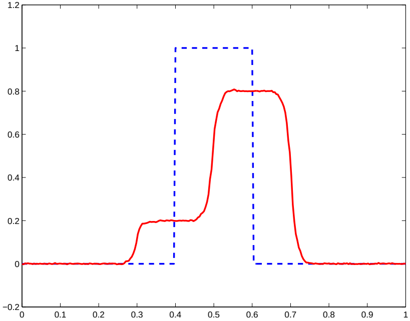




True f and a sample of 10 noisy curves out of $n = 200$

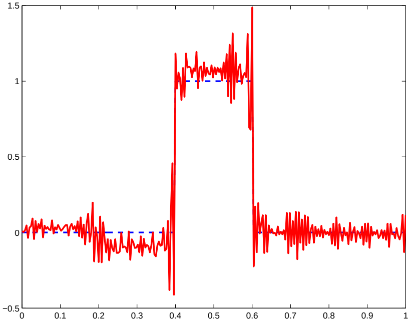
Step example - Direct mean

Direct mean of the $n = 200$ observed curves

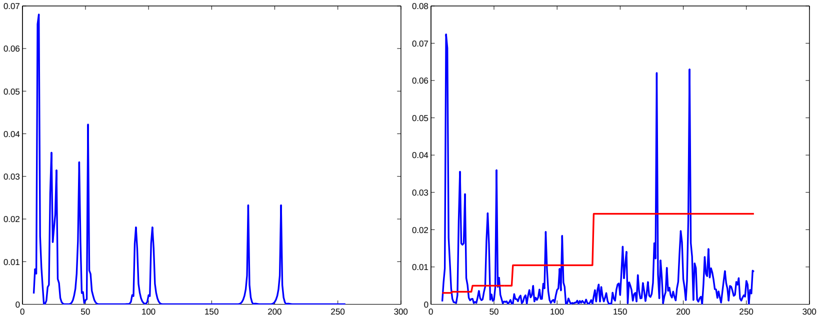


Step example - Deconvolution step

Deconvolution of the direct mean without any smoothing/thresholding

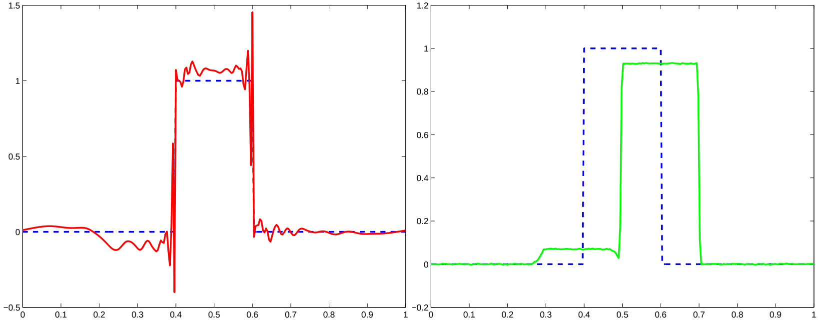


Step example - Wavelet thresholding



True wavelet coefficients - Noisy wavelet coefficients and threshold λ_j

Step example - Comparison with Procrustean mean



Wavelet thresholding - Procrustean mean

Some perspectives

- Consider an asymptotic setting with $n \to +\infty$ **and** $\epsilon \to 0$ (work in progress) : in this case
  - estimation of the shifts τ_m is possible
  - there is no inverse problem and the minimax risk becomes (conjecture...)

$$\mathcal{R}_{n,\epsilon}(\mathcal{F}) = \inf_{\hat{f}_n} \sup_{f \in \mathcal{F}} \mathcal{R}(\hat{f}_{n,\epsilon}, f) \sim C \left(\frac{\epsilon^2}{n} \right)^{\frac{2s}{2s+1}}$$

- Estimation of the density g
- Extension to images and more complex deformations (first steps in this direction by Bigot, Gamboa & Vimond (2009), Bigot, Loubes & Vimond (2008), Bigot, Gadat & Loubes (2009))