Control & Inverse Problems

Deconvolution of probability densities by mollification

Pierre Maréchal

Institut de Mathématiques de Toulouse
Université Paul Sabatier

pr.marechal@gmail.com

Monastir

May 9-11, 2022

Collaborators

THORSTEN HOHAGE
Göttingen University

LÉOPOLD SIMAR
ISBA, Université Catholique de Louvain-la-Neuve

ANNE VANHEMS
TBS and TSE, Université de Toulouse

Consider the equation $Y = X + \varepsilon$ in which:

1. Y is the observed random vector;
2. X is the latent random vector;
3. ε is a random noise vector.

Consider the equation $Y = X + \varepsilon$ in which:

1. Y is the observed random vector;
2. X is the latent random vector;
3. ε is a random noise vector.

Standing assumptions:

(A1) X and ε are independent;

(A2) Y, X and ε have densities with respect to the Lebesgue measure, denoted respectively by g, f and γ;

(A3) both f and g belong to $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

Consider the equation $Y = X + \varepsilon$ in which:

1. Y is the observed random vector;
2. X is the latent random vector;
3. ε is a random noise vector.

Standing assumptions:

(A1) X and ε are independent;

(A2) Y, X and ε have densities with respect to the Lebesgue measure, denoted respectively by g, f and γ;

(A3) both f and g belong to $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

The density f then satisfies the equation $T_\gamma f = g$, in which T_γ is the convolution operator

$$\begin{array}{rccc} T_\gamma : & L^2(\mathbb{R}) & \longrightarrow & L^2(\mathbb{R}) \\ & f & \longmapsto & T_\gamma f := f * \gamma. \end{array}$$

The density g is in fact *unknown*, but estimated from the statistical sample $Y_1, \ldots, Y_n$.

The density g is in fact *unknown*, but estimated from the statistical sample $Y_1, \dots, Y_n$. The unknown density g is then replaced by a *nonparametric* estimator g_n.

The density g is in fact *unknown*, but estimated from the statistical sample $Y_1, \ldots, Y_n$. The unknown density g is then replaced by a *nonparametric* estimator g_n.

Notation

- Fourier transform of an integrable function h on $\mathbb{R}^d$:

$$\hat{h}(\xi) = \int e^{-2i\pi\langle x,\xi\rangle} h(x)\,\mathrm{d}x.$$

The density g is in fact *unknown*, but estimated from the statistical sample $Y_1, \ldots, Y_n$. The unknown density g is then replaced by a *nonparametric* estimator g_n.

Notation

- Fourier transform of an integrable function h on $\mathbb{R}^d$:

$$\hat{h}(\xi) = \int e^{-2i\pi\langle x,\xi\rangle} h(x)\,\mathrm{d}x.$$

- Corresponding Fourier-Plancherel operator on $L^2(\mathbb{R}^d)$: U.

Tikhonov regularization

$$f_\alpha^{\mathrm{TK}} = \left(T_\gamma^* T_\gamma + \alpha I\right)^{-1} T_\gamma^* g \quad \text{or} \quad \hat{f}_\alpha^{\mathrm{TK}} = \frac{\bar{\hat{\gamma}}}{|\hat{\gamma}|^2 + \alpha}\hat{g}$$

Tikhonov regularization

$$f_\alpha^{\mathrm{TK}} = \left(T_\gamma^* T_\gamma + \alpha I\right)^{-1} T_\gamma^* g \quad \text{or} \quad \hat{f}_\alpha^{\mathrm{TK}} = \frac{\bar{\hat{\gamma}}}{|\hat{\gamma}|^2 + \alpha}\hat{g}$$

Remarks

- Variational regularization method:

$$\underset{f}{\mathrm{Min}} \quad \left\|g - T_\gamma f\right\|^2 + \alpha \left\|f\right\|^2$$

Tikhonov regularization

$$f_\alpha^{\mathrm{TK}} = \left(T_\gamma^* T_\gamma + \alpha I\right)^{-1} T_\gamma^* g \quad \text{or} \quad \hat{f}_\alpha^{\mathrm{TK}} = \frac{\bar{\hat{\gamma}}}{|\hat{\gamma}|^2 + \alpha}\hat{g}$$

Remarks

- Variational regularization method:

$$\operatorname*{Min}_f \; \left\|g - T_\gamma f\right\|^2 + \alpha \left\|f\right\|^2$$

- Penalizes *uniformly* $|\hat{f}(\xi)|^2$:

$$\operatorname*{Min}_f \; \left\|\hat{g} - \hat{\gamma}\cdot\hat{f}\right\|^2 + \alpha \left\|\hat{f}\right\|^2$$

Tikhonov regularization

$$f_\alpha^{\mathrm{TK}} = \left(T_\gamma^* T_\gamma + \alpha I\right)^{-1} T_\gamma^* g \quad \text{or} \quad \hat{f}_\alpha^{\mathrm{TK}} = \frac{\bar{\hat{\gamma}}}{|\hat{\gamma}|^2 + \alpha} \hat{g}$$

Remarks

- Variational regularization method:

$$\operatorname*{Min}_f \ \left\|g - T_\gamma f\right\|^2 + \alpha \left\|f\right\|^2$$

- Penalizes *uniformly* $|\hat{f}(\xi)|^2$:

$$\operatorname*{Min}_f \ \left\|\hat{g} - \hat{\gamma} \cdot \hat{f}\right\|^2 + \alpha \left\|\hat{f}\right\|^2$$

- Does not allow to recover a density: $\hat{f}_\alpha^{\mathrm{TK}}(0) \neq 1$

Deconvolution kernels

$$\hat{f}_h^{\mathrm{DK}} = \frac{\hat{\varphi}_h}{\hat{\gamma}}\hat{g} \quad \text{or} \quad f_h^{\mathrm{DK}} = \left(T_\gamma^* T_\gamma\right)^{-1} T_\gamma^* C_h g$$

Deconvolution kernels

$$\hat{f}_h^{\mathrm{DK}} = \frac{\hat{\varphi}_h}{\hat{\gamma}}\hat{g} \quad \text{or} \quad f_h^{\mathrm{DK}} = \left(T_\gamma^* T_\gamma\right)^{-1} T_\gamma^* C_h g$$

Remarks

- Restrictive assumptions on φ_h and γ:

$$\sup_\xi \left|\frac{\hat{\varphi}_h(\xi)}{\hat{\gamma}(\xi)}\right| < \infty \quad \text{and} \quad \int \left|\frac{\hat{\varphi}_h(\xi)}{\hat{\gamma}(\xi)}\right| \mathrm{d}\xi < \infty$$

Deconvolution kernels

$$\hat{f}_h^{\mathrm{DK}} = \frac{\hat{\varphi}_h}{\hat{\gamma}}\hat{g} \quad \text{or} \quad f_h^{\mathrm{DK}} = (T_\gamma^* T_\gamma)^{-1} T_\gamma^* C_h g$$

Remarks

- Restrictive assumptions on φ_h and γ:

$$\sup_\xi \left|\frac{\hat{\varphi}_h(\xi)}{\hat{\gamma}(\xi)}\right| < \infty \quad \text{and} \quad \int \left|\frac{\hat{\varphi}_h(\xi)}{\hat{\gamma}(\xi)}\right| \mathrm{d}\xi < \infty$$

- Not *a priori* a variational method, but solution of:

$$\operatorname*{Min}_f \ \left\|C_h g - T_\gamma f\right\|^2$$

Spectral cut-off

$$\hat{f}_a^{\mathrm{SC}} = \frac{1_{|\hat{\gamma}|^2 \geq a}}{\hat{\gamma}} \hat{g}$$

Spectral cut-off

$$\hat{f}_a^{\mathrm{SC}} = \frac{1_{|\hat{\gamma}|^2 \geq a}}{\hat{\gamma}}\,\hat{g}$$

Remarks

- Particular case of deconvolution kernels method with

$$\hat{\varphi}_a = 1_{|\hat{\gamma}|^2 \geq a}$$

Spectral cut-off

$$\hat{f}_a^{\mathrm{SC}} = \frac{1_{|\hat{\gamma}|^2 \geq a}}{\hat{\gamma}} \hat{g}$$

Remarks

- Particular case of deconvolution kernels method with

$$\hat{\varphi}_a = 1_{|\hat{\gamma}|^2 \geq a}$$

- Gibbs phenomena can be expected: sinc impulse response.

Consider the general ill-posed linear operator equation

$$Tf = g$$

Consider the general ill-posed linear operator equation

$$Tf = g$$

- $T\colon F \to G$ bounded linear operator

Consider the general ill-posed linear operator equation

$$Tf = g$$

- $T\colon F \to G$ bounded linear operator
- F, G separable, infinite dimensional, Hilbert

Consider the general ill-posed linear operator equation

$$Tf = g$$

- $T : F \to G$ bounded linear operator
- F, G separable, infinite dimensional, Hilbert
- $\inf\left\{ \|Tf\| \;\middle|\; f \in (\ker T)^{\perp},\ \|f\| = 1 \right\} = 0$

Consider the general ill-posed linear operator equation

$$Tf = g$$

- $T\colon F \to G$ bounded linear operator
- F, G separable, infinite dimensional, Hilbert
- $\inf\left\{ \|Tf\| \;\middle|\; f \in (\ker T)^\perp,\ \|f\| = 1 \right\} = 0$

$T^\dagger$ is densely defined and unbounded

Consider the general ill-posed linear operator equation

$$Tf = g$$

- $T\colon F \to G$ bounded linear operator
- F, G separable, infinite dimensional, Hilbert
- $\inf\left\{ \|Tf\| \;\middle|\; f \in (\ker T)^{\perp},\ \|f\| = 1 \right\} = 0$

$T^{\dagger}$ is densely defined and unbounded

Remark

If T is injective, the last condition reduces to

$$\inf\left\{ \|Tf\| \;\middle|\; \|f\| = 1 \right\} = 0$$

Highlights

Mollifiers were introduced in the field of Partial Differential Equations by K.O. Friedrichs:

- K.O. FRIEDRICHS, *The identity of weak and strong extensions of differential operators*, Transactions of the AMS, 1944

Highlights

Mollifiers were introduced in the field of Partial Differential Equations by K.O. Friedrichs:

- K.O. FRIEDRICHS, *The identity of weak and strong extensions of differential operators*, Transactions of the AMS, 1944

Let's mention a well-known approximation theorem:

Theorem

Let $\varphi \in L^1(\mathbb{R}^n)$ be such that $\int \varphi(x)\,\mathrm{d}x = 1$. For every $\beta > 0$, let

$$\varphi_\beta(x) := \frac{1}{\beta^n}\varphi\left(\frac{x}{\beta}\right)$$

Let $p \in [1,\infty)$. Then, for every $f \in L^p(\mathbb{R}^n)$,

$$\|\varphi_\beta * f - f\|_p \longrightarrow 0 \quad \text{as} \quad \beta \downarrow 0$$

Highlights

Mollifiers were used in Inverse Problems in various forms:

- Data smoothing
- Hilbert space duality
- Variational formulation

- D.A. Murio, *The mollification method and the numerical solution of ill-posed problems*, John Wiley & Sons, 2011

Highlights

Mollifiers were used in Inverse Problems in various forms:

- Data smoothing
- Hilbert space duality
- Variational formulation

- A. K. Louis & P. Maass, *A mollifier method for linear operator equations of the first kind*, Inverse Problems, 1990
- T. Schuster, *The method of approximate inverse: theory and applications*, Vol. 1906, Berlin: Springer, 2007

Highlights

Mollifiers were used in Inverse Problems in various forms:

- Data smoothing
- Hilbert space duality
- Variational formulation

- A. LANNES *et al.*, *Stabilized reconstruction in signal and image processing*, Journal of Modern Optics, 1987
- N. ALIBAUD, P. M. and Y. SAESOR, *A variational approach to the inversion of truncated Fourier operators*, Inverse Problems, 2009
- X. BONNEFOND and P. M., *A variational approach to the inversion of some compact operators*, Pacific Journal of Optimization, 2009

Variational setting: general case

$$Tf = g$$

- $T\colon F \to G$ bounded linear operator, injective
- F, G separable, infinite dimensional, Hilbert
- $\inf\left\{ \|Tf\| \;\middle|\; f \in (\ker T)^{\perp},\ \|f\| = 1 \right\} = 0$

Variational setting: general case

$$Tf = g$$

- $T\colon F \to G$ bounded linear operator, injective
- F, G separable, infinite dimensional, Hilbert
- $\inf\left\{ \|Tf\| \;\middle|\; f \in (\ker T)^{\perp},\ \|f\| = 1 \right\} = 0$

One must give up recovering the true object $f_\circ$!

Variational setting: general case

$$Tf = g$$

- $T\colon F \to G$ bounded linear operator, injective
- F, G separable, infinite dimensional, Hilbert
- $\inf\left\{ \|Tf\| \;\middle|\; f \in (\ker T)^{\perp},\ \|f\| = 1 \right\} = 0$

One must give up recovering the true object $f_\circ$!

New target: recovering $C_\beta f_\circ := \varphi_\beta * f_\circ$?

Variational setting: general case

$$Tf = g$$

- $T\colon F \to G$ bounded linear operator, injective
- F, G separable, infinite dimensional, Hilbert
- $\inf\left\{ \|Tf\| \;\middle|\; f \in (\ker T)^\perp,\ \|f\| = 1 \right\} = 0$

One must give up recovering the true object $f_\circ$!

New target: recovering $C_\beta f_\circ := \varphi_\beta * f_\circ$?

Our purpose: do this in the framework of variational methods

Heuristics

- $f_\circ = C_\beta f_\circ + (I - C_\beta) f_\circ$

Heuristics

- $f_\circ = C_\beta f_\circ + (I - C_\beta) f_\circ$
- Undesired component: $(I - C_\beta) f_\circ$

Heuristics

- $f_\circ = C_\beta f_\circ + (I - C_\beta)f_\circ$
- Undesired component: $(I - C_\beta)f_\circ$
- Penalty term: $\|(I - C_\beta)f\|^2$

Heuristics

- $f_\circ = C_\beta f_\circ + (I - C_\beta) f_\circ$
- Undesired component: $(I - C_\beta) f_\circ$
- Penalty term: $\| (I - C_\beta) f \|^2$
- Assume we can generate the data g_β corresponding to $C_\beta f_\circ$

Heuristics

- $f_\circ = C_\beta f_\circ + (I - C_\beta) f_\circ$
- Undesired component: $(I - C_\beta) f_\circ$
- Penalty term: $\|(I - C_\beta) f\|^2$
- Assume we can generate the data g_β corresponding to $C_\beta f_\circ$
- Then a natural choice for the fit term is $\|g_\beta - Tf\|^2$

Regularization scheme

- Define the *target object* to be $C_\beta f_\circ$

Regularization scheme

- Define the *target object* to be $C_\beta f_\circ$
- Generate, from $g \simeq Tf_\circ$ an approximation g_β of $TC_\beta f_\circ$

Regularization scheme

- Define the *target object* to be $C_\beta f_\circ$
- Generate, from $g \simeq Tf_\circ$ an approximation g_β of $TC_\beta f_\circ$
- Define the *reconstructed object* f_β as the solution of

$$\operatorname*{Min}_{f \in F} \quad \frac{1}{2}\left\|g_\beta - Tf\right\|_G^2 + \frac{1}{2}\left\|(I - C_\beta)f\right\|_F^2$$

Regularization scheme

- Define the *target object* to be $C_\beta f_\circ$
- Generate, from $g \simeq Tf_\circ$ an approximation g_β of $TC_\beta f_\circ$
- Define the *reconstructed object* f_β as the solution of

$$\underset{f\in F}{\text{Min}} \quad \frac{1}{2}\|g_\beta - Tf\|_G^2 + \frac{1}{2}\|(I - C_\beta)f\|_F^2$$

$$f_\beta := \big(T^*T + (I - C_\beta)^*(I - C_\beta)\big)^{-1} T^* g_\beta$$

Regularization scheme

- Define the *target object* to be $C_\beta f_\circ$
- Generate, from $g \simeq Tf_\circ$ an approximation g_β of $TC_\beta f_\circ$
- Define the *reconstructed object* f_β as the solution of

$$\operatorname*{Min}_{f \in F} \quad \frac{1}{2}\|g_\beta - Tf\|_G^2 + \frac{1}{2}\|(I - C_\beta)f\|_F^2$$

$$f_\beta := \big(T^*T + (I - C_\beta)^*(I - C_\beta)\big)^{-1} T^* g_\beta$$

- Regard $(C_\beta)_{\beta \in (0,1]}$ as an *approximation of unity*, and consider the asymptotic behavior as $\beta \downarrow 0$

Main issues

- Finding the regularized data g_β

Main issues

- Finding the regularized data g_β

This is achieved if we find Φ_β *such that* $\Phi_\beta T = TC_\beta$, *since then*

$$\Phi_\beta g = \Phi_\beta Tf = TC_\beta f$$

Main issues

- Finding the regularized data g_β

This is achieved if we find Φ_β *such that* $\Phi_\beta T = TC_\beta$*, since then*

$$\Phi_\beta g = \Phi_\beta Tf = TC_\beta f$$

intertwining relationship

Main issues

- Finding the regularized data g_β

This is achieved if we find Φ_β *such that* $\Phi_\beta T = TC_\beta$, *since then*

$$\Phi_\beta g = \Phi_\beta Tf = TC_\beta f$$

intertwining relationship

- Wellposedness for fixed $\beta > 0$

Main issues

- Finding the regularized data g_β

This is achieved if we find Φ_β such that $\Phi_\beta T = TC_\beta$, since then

$$\Phi_\beta g = \Phi_\beta T f = TC_\beta f$$

intertwining relationship

- Wellposedness for fixed $\beta > 0$
- Asymptotic behavior as $\beta \downarrow 0$

Main issues

- Finding the regularized data g_β

This is achieved if we find Φ_β *such that* $\Phi_\beta T = TC_\beta$, *since then*

$$\Phi_\beta g = \Phi_\beta Tf = TC_\beta f$$

intertwining relationship

- Wellposedness for fixed $\beta > 0$
- Asymptotic behavior as $\beta \downarrow 0$
- Computational aspects

Intertwining relationship for deconvolution

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

Intertwining relationship for deconvolution

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$$T = U^{-1}\hat{k}U$$

convolution by k

Intertwining relationship for deconvolution

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$$T = U^{-1}\hat{k}U$$

convolution by k

Convolution operators commute

$$TC_\beta = C_\beta T$$

Intertwining relationship for deconvolution

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$$T = U^{-1}\hat{k}U$$

convolution by k

Convolution operators commute

$$TC_\beta = C_\beta T$$

$$\Phi_\beta = C_\beta$$

Another example: Fourier truncation

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

Another example: Fourier truncation

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$T = \mathbb{1}_W U$ with W bounded domain [Fourier truncation]

Another example: Fourier truncation

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$T = \mathbb{1}_W U$ with W bounded domain [Fourier truncation]

$$TC_\beta = \mathbb{1}_W U U^{-1}\hat{\varphi}_\beta U = \hat{\varphi}_\beta \mathbb{1}_W U = \hat{\varphi}_\beta T$$

Another example: Fourier truncation

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$T = \mathbb{1}_W U$ with W bounded domain [Fourier truncation]

$$TC_\beta = \mathbb{1}_W U U^{-1}\hat{\varphi}_\beta U = \hat{\varphi}_\beta \mathbb{1}_W U = \hat{\varphi}_\beta T$$

$$\Phi_\beta = \hat{\varphi}_\beta$$

One more example: the Radon transform

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

One more example: the Radon transform

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$$(Tf)(\theta, s) = \int f(\mathbf{x})\delta(s - \langle \theta, \mathbf{x} \rangle)\, d\mathbf{x} \quad \text{[Radon transformation]}$$

One more example: the Radon transform

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$$(Tf)(\theta, s) = \int f(\mathbf{x})\delta(s - \langle\theta, \mathbf{x}\rangle)\,\mathrm{d}\mathbf{x} \quad \text{[Radon transformation]}$$

$$T(f_1 * f_2) = Tf_1 \circledast Tf_2 \text{ with } \circledast \text{ convolution w.r.t. } s$$

One more example: the Radon transform

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$$(Tf)(\theta, s) = \int f(\mathbf{x})\delta(s - \langle\theta, \mathbf{x}\rangle)\,\mathrm{d}\mathbf{x} \quad \text{[Radon transformation]}$$

$$T(f_1 * f_2) = Tf_1 \circledast Tf_2 \text{ with } \circledast \text{ convolution w.r.t. } s$$

$$TC_\beta f = T(\varphi_\beta * f) = T\phi_\beta \circledast Tf$$

One more example: the Radon transform

$$TC_\beta = \Phi_\beta T \quad \text{with} \quad C_\beta := U^{-1}\hat{\varphi}_\beta U$$

$$(Tf)(\theta, s) = \int f(\mathbf{x})\delta(s - \langle \theta, \mathbf{x} \rangle)\,\mathrm{d}\mathbf{x} \quad \text{[Radon transformation]}$$

$$T(f_1 * f_2) = Tf_1 \circledast Tf_2 \text{ with } \circledast \text{ convolution w.r.t. } s$$

$$TC_\beta f = T(\varphi_\beta * f) = T\phi_\beta \circledast Tf$$

$$\Phi_\beta = (g \mapsto T\varphi_\beta \circledast g)$$

Applications under study:

- Nonparametric instrumental regression
  (with A. VANHEMS and W. SIMO)
- Cauchy problem for the inhomogeneous Helmoltz equation
  (with F. TRIKI and W. SIMO)
- Inversion of the real Laplace Transform
  (with F. TRIKI and W. SIMO)

Consider the inverse problem $Tf = g$ in which $T\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ has unbounded pseudo-inverse as usual.

Consider the inverse problem $Tf = g$ in which $T\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ has unbounded pseudo-inverse as usual.

Definition

A function $\psi_\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is called a *mollifier* if

(1) for every $\beta > 0$ and $\mathbf{y} \in \mathbb{R}^n$, $\psi_\beta(\cdot, \mathbf{y}) \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, and

$$\int_{\mathbb{R}} \psi_\beta(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x} = 1$$

(2) for every $f \in L^2(\mathbb{R}^n)$, the function f_β defined by

$$f_\beta(y) = \langle f, \psi_\beta(\cdot, \mathbf{y}) \rangle = \int_{\mathbb{R}^n} f(\mathbf{x})\, \psi_\beta(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{x}$$

converges to f in $L^2(\mathbb{R}^n)$ as $\beta \downarrow 0$

Now, assuming the existence of a family of functions $\big(v_\beta(\cdot, \mathbf{y})\big)$ such that

$$\forall \beta > 0, \quad \forall \mathbf{y} \in \mathbb{R}^n, \quad T^* v_\beta(\cdot, \mathbf{y}) = \psi_\beta(\cdot, \mathbf{y}) \tag{1}$$

we see that f_β is then given by

$$f_\beta(\mathbf{y}) = \big\langle f\,, T^* v_\beta(\cdot, \mathbf{y}) \big\rangle = \big\langle Tf\,, v_\beta(\cdot, \mathbf{y}) \big\rangle = \big\langle g\,, v_\beta(\cdot, \mathbf{y}) \big\rangle$$

Now, assuming the existence of a family of functions $\big(v_\beta(\cdot,\mathbf{y})\big)$ such that

$$\forall\beta>0,\quad \forall\mathbf{y}\in\mathbb{R}^n,\quad T^*v_\beta(\cdot,\mathbf{y})=\psi_\beta(\cdot,\mathbf{y}) \tag{1}$$

we see that f_β is then given by

$$f_\beta(\mathbf{y})=\big\langle f\,,T^*v_\beta(\cdot,\mathbf{y})\big\rangle=\big\langle Tf\,,v_\beta(\cdot,\mathbf{y})\big\rangle=\big\langle g\,,v_\beta(\cdot,\mathbf{y})\big\rangle$$

More generally, if $\psi_\beta(\cdot,\mathbf{y})$ belongs to $\mathscr{D}\big((T^*)^\dagger\big)=\operatorname{ran}T^*+(\operatorname{ran}T^*)^\perp$, then the minimum norm least square solution to (1) is used instead, and denoted by $v_\beta(\cdot,\mathbf{y})$ again.

Now, assuming the existence of a family of functions $\big(v_\beta(\cdot,\mathbf{y})\big)$ such that

$$\forall \beta > 0, \quad \forall \mathbf{y} \in \mathbb{R}^n, \quad T^* v_\beta(\cdot,\mathbf{y}) = \psi_\beta(\cdot,\mathbf{y}) \tag{1}$$

we see that f_β is then given by

$$f_\beta(\mathbf{y}) = \big\langle f\,,\, T^* v_\beta(\cdot,\mathbf{y})\big\rangle = \big\langle Tf\,,\, v_\beta(\cdot,\mathbf{y})\big\rangle = \big\langle g\,,\, v_\beta(\cdot,\mathbf{y})\big\rangle$$

More generally, if $\psi_\beta(\cdot,\mathbf{y})$ belongs to $\mathscr{D}\big((T^*)^\dagger\big) = \operatorname{ran} T^* + (\operatorname{ran} T^*)^\perp$, then the minimum norm least square solution to (1) is used instead, and denoted by $v_\beta(\cdot,\mathbf{y})$ again. The family of mappings

$$\begin{array}{rccc} \tilde{R}_\beta : & L^2(\mathbb{R}^n) & \longrightarrow & L^2(\mathbb{R}^n) \\ & g & \longmapsto & \langle g, v_\beta(\cdot,\mathbf{y})\rangle \end{array}$$

is then called an *approximate inverse* of T

If $\psi_\beta(\mathbf{x},\mathbf{y}) = \varphi_\beta(\mathbf{y}-\mathbf{x})$, the function f_β is then a convolution of f:

$$f_\beta(\mathbf{y}) = \int_{\mathbb{R}^n} f(\mathbf{x})\psi_\beta(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\varphi_\beta(\mathbf{y}-\mathbf{x})\,\mathrm{d}\mathbf{x} = (\varphi_\beta * f)(\mathbf{y})$$

If $\psi_\beta(\mathbf{x},\mathbf{y}) = \varphi_\beta(\mathbf{y}-\mathbf{x})$, the function f_β is then a convolution of f:

$$f_\beta(\mathbf{y}) = \int_{\mathbb{R}^n} f(\mathbf{x})\psi_\beta(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\varphi_\beta(\mathbf{y}-\mathbf{x})\,\mathrm{d}\mathbf{x} = (\varphi_\beta * f)(\mathbf{y})$$

The family of functions (φ_β), indexed by β in some interval of the form $(0,\beta_\circ]$, emulates the Dirac distribution as $\beta \downarrow 0$.

If $\psi_\beta(\mathbf{x},\mathbf{y}) = \varphi_\beta(\mathbf{y}-\mathbf{x})$, the function f_β is then a convolution of f:

$$f_\beta(\mathbf{y}) = \int_{\mathbb{R}^n} f(\mathbf{x})\psi_\beta(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\varphi_\beta(\mathbf{y}-\mathbf{x})\,\mathrm{d}\mathbf{x} = (\varphi_\beta * f)(\mathbf{y})$$

The family of functions (φ_β), indexed by β in some interval of the form $(0,\beta_\circ]$, emulates the Dirac distribution as $\beta \downarrow 0$. It is referred to as an *approximate unity*.

If $\psi_\beta(\mathbf{x},\mathbf{y}) = \varphi_\beta(\mathbf{y}-\mathbf{x})$, the function f_β is then a convolution of f:

$$f_\beta(\mathbf{y}) = \int_{\mathbb{R}^n} f(\mathbf{x})\psi_\beta(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^n} f(\mathbf{x})\varphi_\beta(\mathbf{y}-\mathbf{x})\,\mathrm{d}\mathbf{x} = (\varphi_\beta * f)(\mathbf{y})$$

The family of functions (φ_β), indexed by β in some interval of the form $(0,\beta_\circ]$, emulates the Dirac distribution as $\beta \downarrow 0$. It is referred to as an *approximate unity*.

A standard way to produce such an approximation of unity is to choose an integrable function φ and to define φ_β by

$$\varphi_\beta(\mathbf{x}) := \frac{1}{\beta^n}\varphi\left(\frac{\mathbf{x}}{\beta}\right), \quad \mathbf{x} \in \mathbb{R}^n$$

Recall the DK solution:

$$\hat{f}_h^{\mathrm{DK}} = \frac{\hat{\varphi}_h}{\hat{\gamma}}\,\hat{g}.$$

Recall the DK solution:

$$\hat{f}_h^{\mathrm{DK}} = \frac{\hat{\varphi}_h}{\hat{\gamma}}\,\hat{g}.$$

We readily see that it correspond to Murio's *data smooting* approach.

Recall the DK solution:

$$\hat{f}_h^{\mathrm{DK}} = \frac{\hat{\varphi}_h}{\hat{\gamma}}\,\hat{g}.$$

We readily see that it correspond to Murio's *data smooting* approach. We will see below that it is also Louis & Maass' *approximate inverse* solution, but that the variational approach differs, and is in fact preferable.

Approximate inverses

- $T = T_\gamma = U^* \hat{\gamma} U$

Approximate inverses

- $T = T_\gamma = U^* \hat{\gamma} U$
- $T^* = U^* \overline{\hat{\gamma}} U$

Approximate inverses

- $T = T_\gamma = U^* \hat{\gamma} U$
- $T^* = U^* \overline{\hat{\gamma}} U$
- $(T^*)^{-1} = U^* \frac{1}{\overline{\hat{\gamma}}} U$

Approximate inverses

- $T = T_\gamma = U^* \hat{\gamma} U$
- $T^* = U^* \overline{\hat{\gamma}} U$
- $(T^*)^{-1} = U^* \frac{1}{\overline{\hat{\gamma}}} U$

Recall: for every $\beta > 0$ and every $\mathbf{y} \in \mathbb{R}^n$,

$$T^* v_\beta(\cdot, \mathbf{y}) = \psi_\beta(\cdot, \mathbf{y})$$

Approximate inverses

- $T = T_\gamma = U^* \hat{\gamma} U$
- $T^* = U^* \overline{\hat{\gamma}} U$
- $(T^*)^{-1} = U^* \frac{1}{\hat{\gamma}} U$

Recall: for every $\beta > 0$ and every $\mathbf{y} \in \mathbb{R}^n$,

$$T^* v_\beta(\cdot, \mathbf{y}) = \psi_\beta(\cdot, \mathbf{y})$$

$$U^* \overline{\hat{\gamma}} U v_\beta(\cdot, \mathbf{y}) = \psi_\beta(\cdot, \mathbf{y})$$

Approximate inverses

- $T = T_\gamma = U^* \hat{\gamma} U$
- $T^* = U^* \overline{\hat{\gamma}} U$
- $(T^*)^{-1} = U^* \frac{1}{\overline{\hat{\gamma}}} U$

Recall: for every $\beta > 0$ and every $\mathbf{y} \in \mathbb{R}^n$,

$$T^* v_\beta(\cdot, \mathbf{y}) = \psi_\beta(\cdot, \mathbf{y})$$

$$U^* \overline{\hat{\gamma}} U v_\beta(\cdot, \mathbf{y}) = \psi_\beta(\cdot, \mathbf{y})$$

$$
\begin{aligned}
Uv_\beta(\cdot,\mathbf{y})(\xi) &= \left[\frac{1}{\overline{\hat{\gamma}}} U\psi_\beta(\cdot,\mathbf{y})\right](\xi) \\
&= \frac{1}{\overline{\hat{\gamma}(\xi)}} \int e^{-2i\pi\langle \mathbf{x},\xi\rangle} \varphi_\beta(\mathbf{y}-\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&= \frac{1}{\overline{\hat{\gamma}(\xi)}} \int e^{-2i\pi\langle (\mathbf{y}-\mathbf{x}'),\xi\rangle} \varphi_\beta(\mathbf{x}')\,\mathrm{d}\mathbf{x}' \\
&= \frac{1}{\overline{\hat{\gamma}(\xi)}} e^{-2i\pi\langle \mathbf{y},\xi\rangle} \hat{\varphi}_\beta(-\xi) \\
&= e^{-2i\pi\langle \mathbf{y},\xi\rangle} \frac{\overline{\hat{\varphi}_\beta(\xi)}}{\overline{\hat{\gamma}(\xi)}}
\end{aligned}
$$

The approximate inverse solution is then given by:

$$f_\beta(\mathbf{y}) \quad = \quad \langle\, g\,,\, v_\beta(\cdot,\mathbf{y})\,\rangle$$

The approximate inverse solution is then given by:

$$\begin{aligned} f_\beta(\mathbf{y}) &= \langle\, g\,,\, v_\beta(\cdot,\mathbf{y})\,\rangle \\ &= \langle\, Ug\,,\, Uv_\beta(\cdot,\mathbf{y})\,\rangle \end{aligned}$$

The approximate inverse solution is then given by:

$$
\begin{aligned}
f_\beta(\mathbf{y}) &= \left\langle g\,, v_\beta(\cdot,\mathbf{y}) \right\rangle \\
&= \left\langle Ug\,, Uv_\beta(\cdot,\mathbf{y}) \right\rangle \\
&= \int \hat{g}(\xi) e^{2i\pi\langle \mathbf{y},\xi\rangle} \frac{\hat{\varphi}_\beta(\xi)}{\hat{\gamma}(\xi)}\, \mathrm{d}\xi
\end{aligned}
$$

The approximate inverse solution is then given by:

$$
\begin{aligned}
f_\beta(\mathbf{y}) &= \left\langle g\,, v_\beta(\cdot,\mathbf{y})\right\rangle \\
&= \left\langle Ug\,, Uv_\beta(\cdot,\mathbf{y})\right\rangle \\
&= \int \hat{g}(\xi)e^{2i\pi\langle \mathbf{y},\xi\rangle}\frac{\hat{\varphi}_\beta(\xi)}{\hat{\gamma}(\xi)}\,\mathrm{d}\xi \\
&= \left(U^*\frac{\hat{\varphi}_\beta}{\hat{\gamma}}Ug\right)(\mathbf{y})
\end{aligned}
$$

The approximate inverse solution is then given by:

$$\begin{aligned} f_\beta(\mathbf{y}) &= \langle g\,, v_\beta(\cdot,\mathbf{y})\rangle \\ &= \langle Ug\,, Uv_\beta(\cdot,\mathbf{y})\rangle \\ &= \int \hat{g}(\xi)e^{2i\pi\langle \mathbf{y},\xi\rangle}\frac{\hat{\varphi}_\beta(\xi)}{\hat{\gamma}(\xi)}\,\mathrm{d}\xi \\ &= \left(U^*\frac{\hat{\varphi}_\beta}{\hat{\gamma}}Ug\right)(\mathbf{y}) \end{aligned}$$

The approximate inverse operator is to be compared with the inverse of $T\colon L^2(\mathbb{R}^n)\to \operatorname{ran} T$, which is given by

$$T^{-1} = U^*\frac{1}{\hat{\gamma}}U$$

$$\operatorname*{Min}_{f \in L^2(\mathbb{R})} \quad \frac{1}{2}\left\|C_\beta g - T_\gamma f\right\|^2 + \frac{1}{2}\left\|(I - C_\beta)f\right\|^2$$

$$\operatorname*{Min}_{f\in L^2(\mathbb{R})} \quad \frac{1}{2}\left\|C_\beta g - T_\gamma f\right\|^2 + \frac{1}{2}\left\|(I-C_\beta)f\right\|^2$$

$$f_\beta := \left(T_\gamma^* T_\gamma + (I-C_\beta)^*(I-C_\beta)\right)^{-1} T_\gamma^* C_\beta g$$

$$\operatorname*{Min}_{f\in L^2(\mathbb{R})} \quad \frac{1}{2}\left\|C_\beta g - T_\gamma f\right\|^2 + \frac{1}{2}\left\|(I-C_\beta)f\right\|^2$$

$$f_\beta := \left(T_\gamma^* T_\gamma + (I-C_\beta)^*(I-C_\beta)\right)^{-1} T_\gamma^* C_\beta g$$

$$\hat{f}_\beta = \frac{\bar{\hat{\gamma}}\hat{\varphi}_\beta}{|\hat{\gamma}|^2 + |1-\hat{\varphi}_\beta|^2}\cdot\hat{g}$$

Stability in $L^2(\mathbb{R})$

$$f_\beta = \left(T_\gamma^* T_\gamma + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_\gamma^* C_\beta g$$

Stability in $L^2(\mathbb{R})$

$$f_\beta = \left(T_\gamma^* T_\gamma + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_\gamma^* C_\beta g$$

Thus f_β depends continuously on g if and only if the operator

$$T_\gamma^* T_\gamma + (I - C_\beta)^*(I - C_\beta)$$

has a bounded inverse

Stability in $L^2(\mathbb{R})$

$$f_\beta = \left(T_\gamma^* T_\gamma + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_\gamma^* C_\beta g$$

Thus f_β depends continuously on g if and only if the operator

$$T_\gamma^* T_\gamma + (I - C_\beta)^*(I - C_\beta)$$

has a bounded inverse

$$T_\gamma^* T_\gamma + \alpha (I - C_\beta)^*(I - C_\beta) = U^*\left(|\hat{\gamma}|^2 + |1 - \hat{\varphi}_\beta|^2\right) U$$

Stability in $L^2(\mathbb{R})$

$$f_\beta = \left(T_\gamma^* T_\gamma + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_\gamma^* C_\beta g$$

Thus f_β depends continuously on g if and only if the operator

$$T_\gamma^* T_\gamma + (I - C_\beta)^*(I - C_\beta)$$

has a bounded inverse

$$T_\gamma^* T_\gamma + \alpha(I - C_\beta)^*(I - C_\beta) = U^*\left(|\hat{\gamma}|^2 + |1 - \hat{\varphi}_\beta|^2\right)U$$

Make sure that $|\hat{\gamma}|^2 + |1 - \hat{\varphi}_\beta|^2$ remains bounded away from zero!

Example: Lévy kernels

$$\hat{\varphi}(\xi) = \exp\left(-|\xi|^s\right) \text{ with } s \in]0,2]$$

Example: Lévy kernels

$$\hat{\varphi}(\xi) = \exp\left(-|\xi|^s\right) \text{ with } s \in]0,2]$$

$$\varphi = U^{-1}\hat{\varphi}$$

Example: Lévy kernels

$$\hat{\varphi}(\xi) = \exp\left(-|\xi|^s\right) \text{ with } s \in]0,2]$$

$$\varphi = U^{-1}\hat{\varphi}$$

Proposition

The above defined function φ is positive, even, decreasing on $\mathbb{R}_+$ and of class $\mathscr{C}^\infty$

Overview

$$\hat{f}_{\mathrm{REG}} = \Phi\,\hat{g}$$

Overview

$$\hat{f}_{\text{REG}} = \Phi\,\hat{g}$$

	Functional to be minimized	Filter Φ
TK	$\frac{1}{2}\|g-\gamma*f\|^2+\frac{\alpha}{2}\|f\|^2$	$\frac{\bar{\hat{\gamma}}}{\|\hat{\gamma}\|^2+\alpha}$
DK	$\frac{1}{2}\|\varphi_h*g-\gamma*f\|^2$	$\frac{\hat{\varphi}_h}{\hat{\gamma}}$
SC	$\frac{1}{2}\|\check{\mathbb{1}}_{\{\|\hat{\gamma}\|^2\geq a\}}*g-\gamma*f\|^2$	$\frac{\mathbb{1}_{\{\|\hat{\gamma}\|^2\geq a\}}}{\hat{\gamma}}$
MO	$\frac{1}{2}\|\varphi_\beta*g-\gamma*f\|^2+\frac{1}{2}\|f-\varphi_\beta*f\|^2$	$\frac{\bar{\hat{\gamma}}\hat{\varphi}_\beta}{\|\hat{\gamma}\|^2+\|1-\hat{\varphi}_\beta\|^2}$

- The density g is unknown and is estimated using a sample of observations $Y_1, \ldots, Y_n$.

- The density g is unknown and is estimated using a sample of observations $Y_1, \ldots, Y_n$.
- We denote by g_n (resp. $\hat{g}_n$) the estimator of g (resp. $\hat{g}$).

- The density g is unknown and is estimated using a sample of observations $Y_1, \ldots, Y_n$.
- We denote by g_n (resp. $\hat{g}_n$) the estimator of g (resp. $\hat{g}$).
- Our objective is now to study the convergence of $f_{\beta,n}$ to f

- The density g is unknown and is estimated using a sample of observations $Y_1, \ldots, Y_n$.
- We denote by g_n (resp. $\hat{g}_n$) the estimator of g (resp. $\hat{g}$).
- Our objective is now to study the convergence of $f_{\beta,n}$ to f

$$f_{\beta,n} = \left(T_\gamma^* T_\gamma + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_\gamma^* C_\beta g_n$$

Theorem (Consistency)

Assume g_n is a consistent nonparametric estimator of g, that is, that $E\|g_n - g\|$ goes to zero as n goes to infinity. Let $f_{\beta,n}$ denote the mollified solution corresponding to data g_n. There then exist a sequence $\beta_n \downarrow 0$ such that

$$E\|f_{\beta_n,n} - f_\circ\| \longrightarrow 0 \quad \text{as} \quad n \to \infty.$$

Assumption (filter shape assumption)

There exists a strictly decreasing differentiable function $\Phi\colon [0,\infty) \to \mathbb{R}$ such that

$$\forall \xi \in \mathbb{R}^d, \quad \hat{\varphi}(\xi) = \Phi(|\xi|)$$

and constants $s, C_\Phi \in \mathbb{R}_+^*$ with the following properties:

$$\forall t \in [0,1], \quad \frac{1}{2} \leq \Phi(t) \leq 1$$

$$\forall t \in [0,1], \quad C_\Phi^{-1} t^s \leq 1 - \Phi(t) \leq C_\Phi t^s$$

$$\forall t \in [0,1], \quad |\Phi'(t)| \leq C_\Phi t^{s-1}$$

$$\int_0^\infty \Phi(t)^2 t^{d-1}\,\mathrm{d}t < \infty.$$

Definition

For a function $f \in L^2(\mathbb{R}^d)$ let

$$e_f(t) := \int_{|\xi|>t} \left|\hat{f}(\xi)\right|^2 \,\mathrm{d}\xi, \qquad t > 0.$$

The *Besov-Nikolskiĭ space $B^u_{2,\infty}(\mathbb{R}^d)$ of smoothness index $u > 0$* is the set of all $f \in L^2(\mathbb{R}^d)$ for which

$$\|f\|_{B^u_{2,\infty}} := \left(\sup_{t>0}(1+t)^{2u} e_f(t)\right)^{1/2}$$

is finite.

Definition

For a function $f \in L^2(\mathbb{R}^d)$ let

$$e_f(t) := \int_{|\xi|>t} \left|\hat{f}(\xi)\right|^2 \mathrm{d}\xi, \quad t > 0.$$

The *Besov-Nikolskiĭ space $B^u_{2,\infty}(\mathbb{R}^d)$ of smoothness index $u > 0$* is the set of all $f \in L^2(\mathbb{R}^d)$ for which

$$\|f\|_{B^u_{2,\infty}} := \left(\sup_{t>0}(1+t)^{2u} e_f(t)\right)^{1/2}$$

is finite.

With the above norm, $B^u_{2,\infty}(\mathbb{R}^d)$ is a Banach space.

Definition

For a function $f \in L^2(\mathbb{R}^d)$ let

$$e_f(t) := \int_{|\xi|>t} \left|\hat{f}(\xi)\right|^2 \mathrm{d}\xi, \quad t > 0.$$

The *Besov-Nikolskiĭ space $B^u_{2,\infty}(\mathbb{R}^d)$ of smoothness index $u > 0$* is the set of all $f \in L^2(\mathbb{R}^d)$ for which

$$\|f\|_{B^u_{2,\infty}} := \left(\sup_{t>0}(1+t)^{2u} e_f(t)\right)^{1/2}$$

is finite.

With the above norm, $B^u_{2,\infty}(\mathbb{R}^d)$ is a Banach space. The Sobolev space $H^u(\mathbb{R}^d)$ is a subspace of $B^u_{2,\infty}(\mathbb{R}^d)$ since

$$(1+t)^{2u} e_f(t) \leq \int_{|\xi|>t} (1+|\xi|)^{2u} |(\hat{f})(\xi)|^2 \mathrm{d}\xi \leq \|f\|^2_{H^u}.$$

Throughout, we use the standard decompostion

$$f_{\beta,n} - f = f_{\beta,n} - f_\beta + f_\beta - f.$$

Throughout, we use the standard decompostion

$$f_{\beta,n} - f = f_{\beta,n} - f_\beta + f_\beta - f.$$

For convenience, we refer to

- the deterministic error $\|f_\beta - f\|$ as the *bias* (or *regularization bias*);

Throughout, we use the standard decompostion

$$f_{\beta,n} - f = f_{\beta,n} - f_\beta + f_\beta - f.$$

For convenience, we refer to

- the deterministic error $\|f_\beta - f\|$ as the *bias* (or *regularization bias*);
- the statistical quadratic error $\mathrm{E}\left(\|f_{\beta,n} - f_\beta\|^2\right)$ as the *variance*.

Ordinary smoothness

We assume here that the density γ of ε satisfies the following *ordinary smoothness condition*:

$$C^{-1}(1+|\xi|)^{-a} \leq |\hat{\gamma}(\xi)| \leq C(1+|\xi|)^{-a}, \quad \xi \in \mathbb{R}^d,$$

for some $a > 0$ and $C \geq 1$.

Ordinary smoothness

We assume here that the density γ of ε satisfies the following *ordinary smoothness condition*:

$$C^{-1}(1+|\xi|)^{-a} \leq |\hat{\gamma}(\xi)| \leq C(1+|\xi|)^{-a}, \quad \xi \in \mathbb{R}^d,$$

for some $a > 0$ and $C \geq 1$. In this case, the problem is mildly ill-posed.

Ordinary smoothness

We assume here that the density γ of ε satisfies the following *ordinary smoothness condition*:

$$C^{-1}(1+|\xi|)^{-a} \leq |\hat{\gamma}(\xi)| \leq C(1+|\xi|)^{-a}, \quad \xi \in \mathbb{R}^d,$$

for some $a > 0$ and $C \geq 1$. In this case, the problem is mildly ill-posed.

Theorem (bound on bias term)

Suppose the above filter shape assumption and the ordinary smoothness condition are satisfied. Then for $0 < u < s+a$ the following statements are equivalent:

- $f \in B^u_{2,\infty}(\mathbb{R}^d)$;
- $\sup_{0<\beta\leq 1} \beta^{-\frac{su}{s+a}} \|f - f_\beta\| < \infty$.

Moreover, $\|f - f_\beta\| = O(\beta^s)$ as $\beta \downarrow 0$ if $f \in H^{s+a}(\mathbb{R}^d)$.

Ordinary smoothness

Proposition (bound on the variance term)

We have

$$\mathrm{E}\left(\|f_{\beta,n}-f_\beta\|^2\right) \leq \frac{2}{n}\|\Phi_\beta\|^2_{L^2(\mathbb{R}^d)}.$$

In particular, if the filter shape assumption and the ordinary smoothness condition are satisfied and $4s \geq d-2a$, then

$$\mathrm{E}\left(\|f_{\beta,n}-f_\beta\|^2\right) = O\left(\frac{1}{n}\beta^{-\frac{s(d+2a)}{s+a}}\right).$$

Ordinary smoothness

Now we can state an order-optimal bound on the mean integrated square error in terms of the sample size.

Ordinary smoothness

Now we can state an order-optimal bound on the mean integrated square error in terms of the sample size. We write $\psi_1(x) \sim \psi_2(x)$ as $x \to x_0$ for two positive functions ψ_1 and ψ_2 if $\liminf_{x\to x_0} \frac{\psi_1(x)}{\psi_2(x)} > 0$ and $\limsup_{x\to x_0} \frac{\psi_1(x)}{\psi_2(x)} < \infty$.

Ordinary smoothness

Now we can state an order-optimal bound on the mean integrated square error in terms of the sample size. We write $\psi_1(x) \sim \psi_2(x)$ as $x \to x_0$ for two positive functions ψ_1 and ψ_2 if $\liminf_{x\to x_0} \frac{\psi_1(x)}{\psi_2(x)} > 0$ and $\limsup_{x\to x_0} \frac{\psi_1(x)}{\psi_2(x)} < \infty$.

Theorem (convergence rate)

Suppose the above filter shape assumption and the ordinary smoothness condition are satisfied, that $4s \geq d - 2a$ and that $f \in B^u_{2,\infty}(\mathbb{R}^d)$ for some $0 < u < s + a$ or $f \in H^{s+a}(\mathbb{R}^d)$ for $u = s + a$. Then, for

$$\beta \sim n^{-\frac{s+a}{2su+s(d+2a)}},$$

we obtain the optimal rate

$$\mathrm{E}\left(\|f_{\beta,n} - f\|^2\right) = O\left(n^{-\frac{u}{u+a+d/2}}\right) \quad \text{as} \quad n \to \infty.$$

Super-smoothness

We assume here that the density γ of ε satisfies the following *super-smoothness condition*:

$$C^{-1}\exp\big(-\kappa|\xi|^a\big) \leq |\widehat{\gamma}(\xi)|^2 \leq C\exp\big(-\kappa|\xi|^a\big), \quad \xi \in \mathbb{R}^d.$$

for some constants $a, \kappa > 0$ and $C \geq 1$.

Super-smoothness

We assume here that the density γ of ε satisfies the following *super-smoothness condition*:

$$C^{-1}\exp\big(-\kappa|\xi|^a\big) \leq |\widehat{\gamma}(\xi)|^2 \leq C\exp\big(-\kappa|\xi|^a\big), \quad \xi \in \mathbb{R}^d.$$

for some constants $a, \kappa > 0$ and $C \geq 1$. In this case, the problem is severely ill-posed.

Super-smoothness

We assume here that the density γ of ε satisfies the following *super-smoothness condition*:

$$C^{-1}\exp\big(-\kappa|\xi|^a\big) \leq |\widehat{\gamma}(\xi)|^2 \leq C\exp\big(-\kappa|\xi|^a\big), \quad \xi \in \mathbb{R}^d.$$

for some constants $a, \kappa > 0$ and $C \geq 1$. In this case, the problem is severely ill-posed. Note that $a = 2$ corresponds to Gaussian errors ε and $a = 1$ to Cauchy errors.

Super-smoothness

We assume here that the density γ of ε satisfies the following *super-smoothness condition*:

$$C^{-1}\exp\big(-\kappa|\xi|^a\big) \leq |\widehat{\gamma}(\xi)|^2 \leq C\exp\big(-\kappa|\xi|^a\big), \quad \xi \in \mathbb{R}^d.$$

for some constants $a, \kappa > 0$ and $C \geq 1$. In this case, the problem is severely ill-posed. Note that $a = 2$ corresponds to Gaussian errors ε and $a = 1$ to Cauchy errors.

Theorem (bound on bias term)

Suppose that the filter shape assumption with $s > \frac{1}{2}$ and the super-smoothness condition are satisfied. Then the following statements are equivalent for $u > 0$:

$$f \in B^u_{2,\infty}(\mathbb{R}^d),$$

$$\sup_{0<\beta<1} (-\ln\beta)^{u/a}\|f - f_\beta\| < \infty.$$

Super-smoothness

Proposition (bound on variance term)

If the super-smoothness condition holds true, then for any $b > 4s$ the statistical error satisfies

$$\mathrm{E}\left(\|f_{\beta,n} - \mathrm{E}\left(f_{\beta,n}\right)\|^2\right) = O\left(\frac{1}{n}\beta^{-2b}\right).$$

Super-smoothness

Proposition (bound on variance term)

If the super-smoothness condition holds true, then for any $b > 4s$ the statistical error satisfies

$$\mathrm{E}\left(\|f_{\beta,n} - \mathrm{E}\left(f_{\beta,n}\right)\|^2\right) = O\left(\frac{1}{n}\beta^{-2b}\right).$$

Combining the previous results yields the following logarithmic convergence rates with respect to the sample size:

Theorem (convergence rate)

Suppose that the filter shape assumption with $s > \frac{1}{2}$ and the super-smoothness condition are satisfied. Let $f \in B^u_{2,\infty}(\mathbb{R}^d)$ for some $u > 0$ and let $\beta = \frac{1}{n}$. Then

$$\mathrm{E}\left(\|f_{\beta,n} - f\|^2\right) = O\left((\ln n)^{-2u/a}\right) \quad \text{as } n \to \infty.$$

Thankyou !