Control and Inverse Problems Conference

A mollifier approach to nonparametric instrumental regression

Anne VANHEMS

co-written with

Pierre MARÉCHAL and Walter Cedric SIMO TAO LEE

Monastir, May 09-11, 2022

$$Y = h(Z) + \varepsilon$$

$$Y = h(Z) + \varepsilon$$

- $Y, Z, \varepsilon : \Omega \to \mathbb{R}$

$$Y = h(Z) + \varepsilon$$

- $Y, Z, \varepsilon \colon \Omega \to \mathbb{R}$
- $h \colon \mathbb{R} \to \mathbb{R}$ is the unknown *regression function*

$$Y = h(Z) + \varepsilon$$

- $Y, Z, \varepsilon : \Omega \to \mathbb{R}$
- $h : \mathbb{R} \to \mathbb{R}$ is the unknown *regression function*

Problem: identify h under the assumption that Z and ε are linked: $E(\varepsilon|Z) \neq 0$.

$$Y = h(Z) + \varepsilon$$

- $Y, Z, \varepsilon : \Omega \to \mathbb{R}$
- $h : \mathbb{R} \to \mathbb{R}$ is the unknown *regression function*

Problem: identify h under the assumption that Z and ε are linked: $E(\varepsilon|Z) \neq 0$. A standard strategy: introduce an *instrumental variable* $W : \Omega \to \mathbb{R}^p$, that is linked to Z and not to ε.

$$Y = h(Z) + \varepsilon$$

- $Y, Z, \varepsilon : \Omega \to \mathbb{R}$
- $h : \mathbb{R} \to \mathbb{R}$ is the unknown *regression function*

Problem: identify h under the assumption that Z and ε are linked: $E(\varepsilon|Z) \neq 0$. A standard strategy: introduce an *instrumental variable* $W : \Omega \to \mathbb{R}^p$, that is linked to Z and not to ε.

$$Y = h(Z) + \varepsilon, \quad E(\varepsilon|W) = 0$$

$$Y = h(Z) + \varepsilon$$

- $Y, Z, \varepsilon : \Omega \to \mathbb{R}$
- $h : \mathbb{R} \to \mathbb{R}$ is the unknown *regression function*

Problem: identify h under the assumption that Z and ε are linked: $E(\varepsilon|Z) \neq 0$. A standard strategy: introduce an *instrumental variable* $W : \Omega \to \mathbb{R}^p$, that is linked to Z and not to ε.

$$Y = h(Z) + \varepsilon, \quad E(\varepsilon|W) = 0$$

$$E(Y|W) = E(h(Z)|W)$$

Assumption 1. The laws P_Z, P_W, P_Y are absolutely continuous with respect to the Lebesgue measure λ.

Assumption 1. The laws P_Z, P_W, P_Y are absolutely continuous with respect to the Lebesgue measure λ.

Then, the equation $E(Y|W) = E(h(Z)|W)$ reduces to the functional integral equation

$$\int f_{YW}(y,w)y\,\mathrm{d}y = \int f_{ZW}(z,w)h(z)\,\mathrm{d}z, \quad w \in \{x | f_W(x) \neq 0\}.$$

Assumption 1. The laws P_Z, P_W, P_Y are absolutely continuous with respect to the Lebesgue measure λ.

Then, the equation $E(Y|W) = E(h(Z)|W)$ reduces to the functional integral equation

$$\int f_{YW}(y,w)y\,\mathrm{d}y = \int f_{ZW}(z,w)h(z)\,\mathrm{d}z, \quad w \in \{x | f_W(x) \neq 0\}.$$

Assumption 2. The kernel f_{ZW} is $\lambda \otimes \lambda$-square integrable.

Assumption 1. The laws P_Z, P_W, P_Y are absolutely continuous with respect to the Lebesgue measure λ.

Then, the equation $E(Y|W) = E(h(Z)|W)$ reduces to the functional integral equation

$$\int f_{YW}(y,w)y\,\mathrm{d}y = \int f_{ZW}(z,w)h(z)\,\mathrm{d}z, \quad w \in \{x | f_W(x) \neq 0\}.$$

Assumption 2. The kernel f_{ZW} is $\lambda \otimes \lambda$-square integrable.

Then, the above equation can be written as $g = Th$, in which T is a Hilbert-Schmidt operator.

Assumption 1. The laws P_Z, P_W, P_Y are absolutely continuous with respect to the Lebesgue measure λ.

Then, the equation $E(Y|W) = E(h(Z)|W)$ reduces to the functional integral equation

$$\int f_{YW}(y,w)y\,\mathrm{d}y = \int f_{ZW}(z,w)h(z)\,\mathrm{d}z, \quad w \in \{x | f_W(x) \neq 0\}.$$

Assumption 2. The kernel f_{ZW} is $\lambda \otimes \lambda$-square integrable.

Then, the above equation can be written as $g = Th$, in which T is a Hilbert-Schmidt operator. Notice that for this equation to have a solution, it is necessary that $g \in L^2(\mathbb{R})$.

Assumption 1. The laws P_Z, P_W, P_Y are absolutely continuous with respect to the Lebesgue measure λ.

Then, the equation $E(Y|W) = E(h(Z)|W)$ reduces to the functional integral equation

$$\int f_{YW}(y,w)y\,\mathrm{d}y = \int f_{ZW}(z,w)h(z)\,\mathrm{d}z, \quad w \in \{x | f_W(x) \neq 0\}.$$

Assumption 2. The kernel f_{ZW} is $\lambda \otimes \lambda$-square integrable.

Then, the above equation can be written as $g = Th$, in which T is a Hilbert-Schmidt operator. Notice that for this equation to have a solution, it is necessary that $g \in L^2(\mathbb{R})$.

Assumption 3. The function $g(w) = \int f_{YW}(y,w)y\,\mathrm{d}y$ belongs to $L^2(\mathbb{R})$.

Remark

The last assumption is satisfied in particular if $E[Y^2] < \infty$ and f_W is bounded.

Remark

The last assumption is satisfied in particular if $E[Y^2] < \infty$ and f_W is bounded.

Remark

In practice, g is estimated from observed sample, and the constraint that $g \in L^2(\mathbb{R})$ may be incorporated in the estimation process.

Remark

The last assumption is satisfied in particular if $E[Y^2] < \infty$ and f_W is bounded.

Remark

In practice, g is estimated from observed sample, and the constraint that $g \in L^2(\mathbb{R})$ may be incorporated in the estimation process.

Remark

Note also that the operator T is unknown and it also needs to be estimated from observed sample:

$$T : h \to \int f_{ZW}(z,w)h(z)\,\mathrm{d}z.$$

Mollifiers in approximation theory

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 $h \in L^p(\mathbb{R}^n)$,

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

Mollifiers in approximation theory

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 $h \in L^p(\mathbb{R}^n)$,

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

Remark

The family of operators (C_β) given by $C_\beta h = \varphi_\beta * h$ is referred to as an *approximation of unity*.

Mollifiers for inverse problems

Approximate inverses and *Variational mollification*

- 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.
- N. ALIBAUD, P. MARÉCHAL and Y. SAESOR, *A variational approach to the inversion of truncated Fourier operators*, Inverse Problems, 2009.
- X. BONNEFOND and P. MARÉCHAL, *A variational approach to the inversion of some compact operators*, Pacific Journal of Optimization, 2009.

Overview of approximate inverses

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

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

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

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

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

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

Overview of approximate inverses

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

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

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

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

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

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

Assume the existence of a family of functions $\left(v_\beta(\cdot, y)\right)$ such that

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

Overview of approximate inverses

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

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

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

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

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

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

Assume the existence of a family of functions $\big(v_\beta(\cdot,y)\big)$ such that

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

Then h_β is given by

$$h_\beta(y) = \langle h, T^* v_\beta(\cdot,y)\rangle = \langle Th, v_\beta(\cdot,y)\rangle = \langle g, v_\beta(\cdot,y)\rangle$$

Overview of approximate inverses

We may use the minimum norm least square solution to

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

Overview of approximate inverses

We may use the minimum norm least square solution to

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

In this context, the family of mappings

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

is called an *approximate inverse* of T

Overview of approximate inverses

We may use the minimum norm least square solution to

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

In this context, the family of mappings

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

is called an *approximate inverse* of T

If $\psi_\beta(x,y) = \varphi_\beta(y-x)$, the function h_β is then a convolution of h:

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

Mollification in variational form

$$Th = g$$

- $T\colon L^2(V) \to L^2(\mathbb{R}^p)$ is bounded linear and injective. where

$$L^2(V) = \{\, h \in L^2(\mathbb{R}) \mid \operatorname{supp} h \subset V \,\}, \quad V \quad \text{bounded}.$$

Mollification in variational form

$$Th = g$$

- $T : L^2(V) \to L^2(\mathbb{R}^p)$ is bounded linear and injective. where

$$L^2(V) = \left\{ h \in L^2(\mathbb{R}) \mid \operatorname{supp} h \subset V \right\}, \quad V \quad \text{bounded}.$$

- $\inf \left\{ \|Th\| \mid h \in (\ker T)^{\perp},\ \|h\| = 1 \right\} = 0$

Mollification in variational form

$$T h = g$$

- $T : L^2(V) \to L^2(\mathbb{R}^p)$ is bounded linear and injective. where

$$L^2(V) = \left\{ h \in L^2(\mathbb{R}) \;\middle|\; \operatorname{supp} h \subset V \right\}, \quad V \quad \text{bounded}.$$

- $\inf \left\{ \|T h\| \;\middle|\; h \in (\ker T)^{\perp},\ \|h\| = 1 \right\} = 0$

Principle: *Due to the ill-posedness of the problem, give up recovering the true object $h^{\dagger}$ but instead try to recover a smooth version of $h^{\dagger}$, namely*

$$C_{\beta} h^{\dagger} = \varphi_{\beta} * h^{\dagger}$$

.

Heuristics

- $h_\circ = C_\beta h^\dagger + (I - C_\beta) h^\dagger$

Heuristics

- $h_\circ = C_\beta h^\dagger + (I - C_\beta)h^\dagger$
- Undesired component: $(I - C_\beta)h^\dagger$

Heuristics

- $h_\circ = C_\beta h^\dagger + (I - C_\beta)h^\dagger$
- Undesired component: $(I - C_\beta)h^\dagger$
- Penalty term: $\|(I - C_\beta)h\|^2$

Heuristics

- $h_\circ = C_\beta h^\dagger + (I - C_\beta) h^\dagger$
- Undesired component: $(I - C_\beta) h^\dagger$
- Penalty term: $\|(I - C_\beta) h\|^2$
- A natural choice for the fit term is $\|g - Th\|^2$

Regularization scheme

- Define the *target object* to be $C_\beta h^\dagger$

Regularization scheme

- Define the *target object* to be $C_\beta h^\dagger$
- Define the *reconstructed object* h_β as the solution of

$$\underset{h\in L^2(V)}{\mathrm{Min}} \quad \frac{1}{2}\left\|g - Th\right\|_{L^2(\mathbb{R}^p)} + \frac{1}{2}\left\|(I - C_\beta)h\right\|_{L^2(\mathbb{R})}$$

Regularization scheme

- Define the *target object* to be $C_\beta h^\dagger$
- Define the *reconstructed object* h_β as the solution of

$$\underset{h \in L^2(V)}{\mathrm{Min}} \quad \frac{1}{2}\|g - Th\|_{L^2(\mathbb{R}^p)} + \frac{1}{2}\|(I - C_\beta)h\|_{L^2(\mathbb{R})}$$

$$h_\beta := \left(T^*T + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T^* g$$

Regularization scheme

- Define the *target object* to be $C_\beta h^\dagger$
- Define the *reconstructed object* h_β as the solution of

$$\underset{h \in L^2(V)}{\text{Min}} \quad \frac{1}{2}\left\|g - Th\right\|_{L^2(\mathbb{R}^p)} + \frac{1}{2}\left\|(I - C_\beta)h\right\|_{L^2(\mathbb{R})}$$

$$h_\beta := \left(T^*T + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T^* g$$

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

Main issues

- Wellposedness for fixed $\beta > 0$

Main issues

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

Main issues

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

Theorem (Consistency)

Assume that $h^\dagger \in L^2(V) \cap H^s(\mathbb{R})$, that $g = Th^\dagger$ and let

$$h_\beta := \left(T^*T + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T^* g.$$

Then $\left(h_\beta\right)_{\beta \in (0,1]}$ is bounded and weakly compact in $L^2(V)$.

Theorem (Consistency)

Assume that $h^\dagger \in L^2(V) \cap H^s(\mathbb{R})$, that $g = Th^\dagger$ and let

$$h_\beta := \left(T^*T + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T^* g.$$

Then $\left(h_\beta\right)_{\beta \in (0,1]}$ is bounded and weakly compact in $L^2(V)$.

Moreover, for every sequence $(\beta_n)_n$ converging to 0,

- $h_{\beta_n} \rightharpoonup h^\dagger$;
- $\lim_{R \to \infty} \sup_{n \in \mathbb{N}} \int_{\|x\| > R} |h_{\beta_n}(x)|^2 \, dx = 0$;
- $\sup_{n \in \mathbb{N}} \| \mathscr{T}_\delta h_{\beta_n} - h_{\beta_n} \|_{L^2(\mathbb{R})} \to 0$ as $\delta \to 0$.

Theorem (Consistency)

Assume that $h^\dagger \in L^2(V) \cap H^s(\mathbb{R})$, that $g = Th^\dagger$ and let

$$h_\beta := \left(T^*T + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T^* g.$$

Then $\left(h_\beta\right)_{\beta\in(0,1]}$ is bounded and weakly compact in $L^2(V)$.

Moreover, for every sequence $(\beta_n)_n$ converging to 0,

- $h_{\beta_n} \rightharpoonup h^\dagger$;
- $\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \int_{\|x\|>R} |h_{\beta_n}(x)|^2 \,\mathrm{d}x = 0$;
- $\sup_{n\in\mathbb{N}} \|\mathscr{T}_\delta h_{\beta_n} - h_{\beta_n}\|_{L^2(\mathbb{R})} \to 0$ as $\delta \to 0$.

Consequently, by the Fréchet-Kolmogorov theorem,

$$h_\beta \to h^\dagger \text{ as } \beta \to 0.$$

Theorem (Consistency under approximate setting)

With the notation and assumptions of the previous theorem, let T_n and g_n be approximations of T and g:

$$T_n \to T \text{ and } g_n \to g \text{ as } n \to \infty.$$

Let

$$\begin{aligned} h_{\beta,n} &:= \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* g_n \\ h_{\beta,n}^\dagger &:= \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* T_n h^\dagger. \end{aligned}$$

Theorem (Consistency under approximate setting)

With the notation and assumptions of the previous theorem, let T_n and g_n be approximations of T and g:

$$T_n \to T \text{ and } g_n \to g \text{ as } n \to \infty.$$

Let

$$\begin{aligned} h_{\beta,n} &:= \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* g_n \\ h_{\beta,n}^\dagger &:= \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* T_n h^\dagger. \end{aligned}$$

Then

(i) $\|h^\dagger - h_{\beta,n}^\dagger\|_{L^2} \to 0$ as $\beta \downarrow 0$;

(ii) $\|h_{\beta,n}^\dagger - h_{\beta,n}\|_{L^2} \leq \bar{C}\beta^{-2s}\left(\|(T_n - T)h^\dagger\|_{L^2} + \|g - g_n\|_{L^2}\right)$.

Corollary

In the above setting, there exists a parameter choice rule $\beta(n) \to 0$ as $n \to \infty$ such that

$$\|h^{\dagger} - h_{\beta,n}\|_{L^2} \to 0 \text{ as } n \to \infty.$$

Corollary

In the above setting, there exists a parameter choice rule $\beta(n) \to 0$ as $n \to \infty$ such that

$$\|h^{\dagger} - h_{\beta,n}\|_{L^2} \to 0 \text{ as } n \to \infty.$$

For example, if $\|(T - T_n)h^{\dagger}\| = \|g - g_n\| = \mathscr{O}(1/n)$, then $\beta(n) = n^{-\tau/(2s)}$ with $\tau < 1$ is a converging a priori selection rule.

In the simulations, we compare 5 methods:

- Mollification: $h_{\beta,n} = \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* g_n$;

In the simulations, we compare 5 methods:

- Mollification: $h_{\beta,n} = \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* g_n$;
- Tikhonov: $h_{\alpha,n} = \left(T_n^* T_n + \alpha I\right)^{-1} T_n^* g_n$;

In the simulations, we compare 5 methods:

- Mollification: $h_{\beta,n} = \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* g_n$;
- Tikhonov: $h_{\alpha,n} = \left(T_n^* T_n + \alpha I\right)^{-1} T_n^* g_n$;
- Generalized Tikhonov: $h_{\alpha,n} = \left(T_n^* T_n + \alpha D_2^* D_2\right)^{-1} T_n^* g_n$;

In the simulations, we compare 5 methods:

- Mollification: $h_{\beta,n} = \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* g_n$;
- Tikhonov: $h_{\alpha,n} = \left(T_n^* T_n + \alpha I\right)^{-1} T_n^* g_n$;
- Generalized Tikhonov: $h_{\alpha,n} = \left(T_n^* T_n + \alpha D_2^* D_2\right)^{-1} T_n^* g_n$;
- Landweber: $h_{k+1,n} = T_n^*(g_n - \mu T_n h_{k,n}) + h_{k,n}$, $\mu < 1$, $k = 1, 2, \ldots$;

In the simulations, we compare 5 methods:

- Mollification: $h_{\beta,n} = \left(T_n^* T_n + (I - C_\beta)^*(I - C_\beta)\right)^{-1} T_n^* g_n$;
- Tikhonov: $h_{\alpha,n} = \left(T_n^* T_n + \alpha I\right)^{-1} T_n^* g_n$;
- Generalized Tikhonov: $h_{\alpha,n} = \left(T_n^* T_n + \alpha D_2^* D_2\right)^{-1} T_n^* g_n$;
- Landweber: $h_{k+1,n} = T_n^*(g_n - \mu T_n h_{k,n}) + h_{k,n}$, $\mu < 1$, $k = 1, 2, \ldots$;
- Spectral cut-off: $h_{k,n} = \sum_{j=1}^{k} \frac{1}{\sigma_j} \langle v_j, g_n \rangle u_j, \quad k = 1, 2, \ldots, N.$

For each regularization method, we computed the reconstruction error

$$\left\| h^{\dagger} - h_{\text{reg.par},n} \right\|$$

For each regularization method, we computed the reconstruction error

$$\left\| h^{\dagger} - h_{\text{reg.par},n} \right\|$$

- Mollification: $\|h^{\dagger} - h_{\beta,n}\|$
- Tikhonov and Generalized Tikhonov: $\|h^{\dagger} - h_{\alpha,n}\|$
- Landweber: $\|h^{\dagger} - h_{k,n}\|$
- Spectral Cutoff: $\|h^{\dagger} - h_{k,n}\|$

We consider a sample size $n = 3000$ and we generate the data as follows:

We consider a sample size $n = 3000$ and we generate the data as follows:

- we generate independently (W, ε) from a uniform distribution $\mathscr{U}(-0.5, 0.5)$

We consider a sample size $n = 3000$ and we generate the data as follows:

- we generate independently (W, ε) from a uniform distribution $\mathscr{U}(-0.5, 0.5)$
- We generate Z as:
  $Z = m(W) + 0.5\,\varepsilon, \quad m(W) = 0.25 + 0.5\,W$

We consider a sample size $n = 3000$ and we generate the data as follows:

- we generate independently (W, ε) from a uniform distribution $\mathscr{U}(-0.5, 0.5)$
- We generate Z as:
  $Z = m(W) + 0.5\,\varepsilon, \quad m(W) = 0.25 + 0.5\,W$
- Y is defined as $Y = h_l(Z) + \varepsilon, \quad l = 1, 2$

We consider a sample size $n = 3000$ and we generate the data as follows:

- we generate independently (W, ε) from a uniform distribution $\mathscr{U}(-0.5, 0.5)$
- We generate Z as:
  $Z = m(W) + 0.5\,\varepsilon, \quad m(W) = 0.25 + 0.5\,W$
- Y is defined as $Y = h_l(Z) + \varepsilon, \quad l = 1, 2$
  - **Smooth case**: $h_1 = f(0.4, 0.1) + f(0.65, 0.075)$ truncated to the interval $[0, 1]$, where the function $f(\mu, \sigma)$ is the p.d.f of the gaussian of mean μ and standard deviation σ.

We consider a sample size $n = 3000$ and we generate the data as follows:

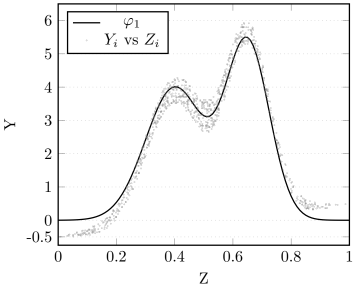
- we generate independently (W, ε) from a uniform distribution $\mathscr{U}(-0.5, 0.5)$
- We generate Z as:
  $Z = m(W) + 0.5\,\varepsilon, \quad m(W) = 0.25 + 0.5\,W$
- Y is defined as $Y = h_l(Z) + \varepsilon, \quad l = 1, 2$
  - **Smooth case**: $h_1 = f(0.4, 0.1) + f(0.65, 0.075)$ truncated to the interval $[0, 1]$, where the function $f(\mu, \sigma)$ is the p.d.f of the gaussian of mean μ and standard deviation σ.
  - **Nonsmooth case**: $h_2(t) = \exp(-|x - 0.5|)$

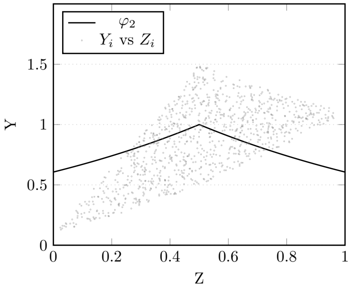
- The mollifier φ_β is the centered gaussian kernel with standard deviation β;

- The mollifier φ_β is the centered gaussian kernel with standard deviation β;
- The discretization of the problem is done by projection onto 100 finite dimensional basis of gate functions on $[0,1]$. The finite dimensional equation is given by

$$\begin{aligned} M_{ji} &= \mathbb{E}[\phi_i(Z)\psi_j(W)], \text{ and} \\ \bar{r}_j &= \mathbb{E}[Y\psi_j(W)],\ i,j = 1,\dots,100. \end{aligned}$$

The data


φ_1
Y_i vs Z_i
Y
Z
6
5
4
3
2
1
0
-0.5
0
0.2
0.4
0.6
0.8
1


φ_2
Y_i vs Z_i
Y
Z
1.5
1
0.5
0
0.2
0.4
0.6
0.8
1

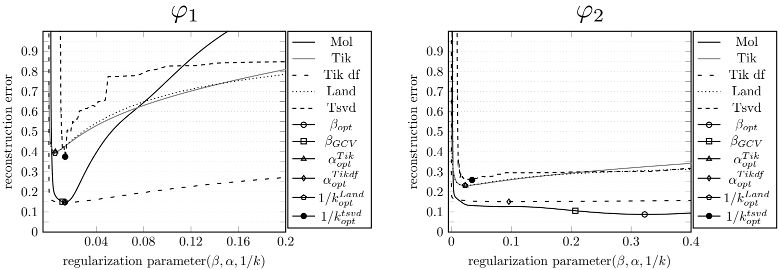


Figure: Error versus regularization parameter

Best approximation (smooth case)

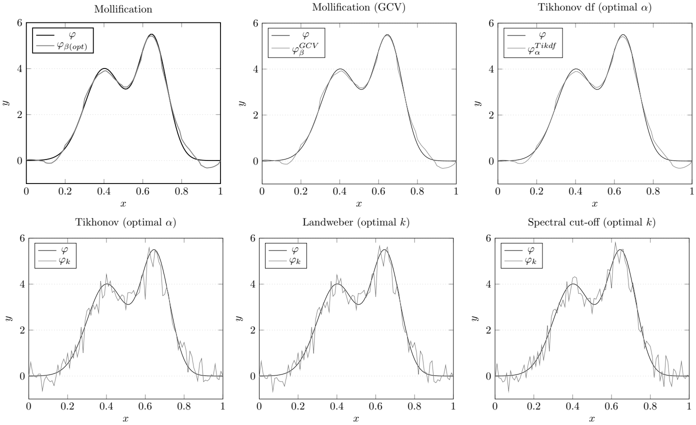


Figure: Comparison best approximation of each regularization method for the function φ_1 in case of projection onto gate functions.

Best approximation (nonsmooth case)

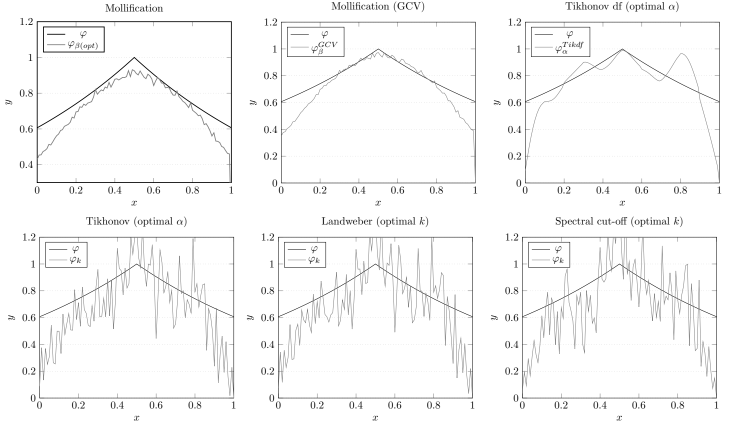


Figure: Comparison best approximation of each regularization method for the function φ_2 in case of projection onto gate functions.

A Monte-Carlo experiment

Monte-Carlo performances of the 5 methods with $M = 1000$

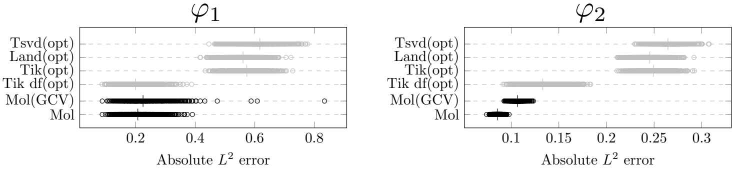


Figure: Results of Monte Carlo simulation for the functions φ_1 and φ_2.

Thank you for your attention!

N. ALIBAUD, P. MARÉCHAL and Y. SAESOR, *A variational approach to the inversion of truncated Fourier operators*, Inverse Problems, 2009

X. BONNEFOND and P. MARÉCHAL, *A variational approach to the inversion of some compact operators*, Pacific Journal of Optimization, 2009

M. CARRASCO, J.P. FLORENS and E. RENAULT, *Linear inverse problems in structural econometrics estimation based on spectral decomposition and regularization*, Handbook of econometrics **6**, pp. 5633-5751, 2007.

P. HALL and J.L. HOROWITZ, *Nonparametric methods for inference in the presence of instrumental variables*, The Annals of Statistics, 2005

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.