Kernel Machines with Shape Constraints

Zoltán Szabó @ LSE

Joint work with: Pierre-Cyril Aubin-Frankowski @ INRIA

BIRS workshop on New Interfaces of Stochastic Analysis and Rough Paths
Sept. 8, 2022

Shape constraints

Pattern

$$0 \leq Df(\mathbf{x}) \quad \forall \mathbf{x}.$$

Shape constraints

Pattern

$$0 \leq Df(\mathbf{x}) \quad \forall \mathbf{x}.$$

Examples:

1. non-negativity: $0 \leq f(x)$,

Shape constraints

Pattern

$$0 \leq Df(\mathbf{x}) \quad \forall \mathbf{x}.$$

Examples:

1. non-negativity: $0 \leq f(x)$,
2. monotonicity ($\nearrow$): $0 \leq f'(x)$,

Shape constraints

Pattern

$$0 \leq Df(\mathbf{x}) \quad \forall \mathbf{x}.$$

Examples:

1. non-negativity: $0 \leq f(x)$,
2. monotonicity ($\nearrow$): $0 \leq f'(x)$,
3. convexity: $0 \leq f''(x)$,

Shape constraints

Pattern

$$0 \leq Df(\mathbf{x}) \quad \forall \mathbf{x}.$$

Examples:

1. non-negativity: $0 \leq f(x)$,
2. monotonicity ($\nearrow$): $0 \leq f'(x)$,
3. convexity: $0 \leq f''(x)$,
4. *n*-monotonicity: $0 \leq f^{(n)}(x)$,

Shape constraints

Pattern

$$0 \leq Df(\mathbf{x}) \quad \forall \mathbf{x}.$$

Examples:

1. non-negativity: $0 \leq f(x)$,
2. monotonicity ($\nearrow$): $0 \leq f'(x)$,
3. convexity: $0 \leq f''(x)$,
4. n-monotonicity: $0 \leq f^{(n)}(x)$,
5. $(n-1)$-alternating monotonicity: for $n \geq 2$

$$(-1)^j f^{(j)} : \geq 0, \nearrow \text{ and convex } \forall j \in [\![0, n-2]\!].$$

Example: generator of a d-variate Archimedean copula is $(d-2)$-alternating monotone.

Examples continued

6. Monotonicity w.r.t. partial ordering ($\mathbf{u} \preccurlyeq \mathbf{v} \Rightarrow f(\mathbf{u}) \leq f(\mathbf{v})$):

$\mathbf{u} \preccurlyeq \mathbf{v}$ iff

- $u_i \leq v_i$ ($\forall i$; product ordering),
- $\sum_{j \in [i]} u_j \leq \sum_{j \in [i]} v_j$ ($\forall i$; unordered weak majorization).

Examples continued

6. Monotonicity w.r.t. partial ordering ($\mathbf{u} \preccurlyeq \mathbf{v} \Rightarrow f(\mathbf{u}) \le f(\mathbf{v})$):

$$0 \le \partial^{\mathbf{e}_j} f(\mathbf{x}), \quad (\forall j \in [d], \forall \mathbf{x}),$$

$$0 \le \partial^{\mathbf{e}_d} f(\mathbf{x}) \le \ldots \le \partial^{\mathbf{e}_1} f(\mathbf{x}) \quad (\forall \mathbf{x}).$$

$\mathbf{u} \preccurlyeq \mathbf{v}$ iff

- $u_i \le v_i$ ($\forall i$; product ordering),
- $\sum_{j\in[i]} u_j \le \sum_{j\in[i]} v_j$ ($\forall i$; unordered weak majorization).

Examples continued

6. Monotonicity w.r.t. partial ordering ($\mathbf{u} \preccurlyeq \mathbf{v} \Rightarrow f(\mathbf{u}) \leq f(\mathbf{v})$):

$$0 \leq \partial^{\mathbf{e}_j} f(\mathbf{x}), \quad (\forall j \in [d], \forall \mathbf{x}),$$

$$0 \leq \partial^{\mathbf{e}_d} f(\mathbf{x}) \leq \ldots \leq \partial^{\mathbf{e}_1} f(\mathbf{x}) \quad (\forall \mathbf{x}).$$

$\mathbf{u} \preccurlyeq \mathbf{v}$ iff

- $u_i \leq v_i$ ($\forall i$; product ordering),
- $\sum_{j \in [i]} u_j \leq \sum_{j \in [i]} v_j$ ($\forall i$; unordered weak majorization).

7. Supermodularity:

$$0 \leq \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \quad (\forall i \neq j \in [d], \forall \mathbf{x}),$$

i.e. $f(\mathbf{u} \vee \mathbf{v}) + f(\mathbf{u} \wedge \mathbf{v}) \geq f(\mathbf{u}) + f(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$.

Shape constraints are omnipresent [Johnson and Jiang, 2018, Guntuboyina and Sen, 2018, Chetverikov et al., 2018]

- Economics:
  - utility functions are $\nearrow$ and concave [Matzkin, 1991].

Shape constraints are omnipresent [Johnson and Jiang, 2018, Guntuboyina and Sen, 2018, Chetverikov et al., 2018]

- Economics:
  - utility functions are $\nearrow$ and concave [Matzkin, 1991].
  - demand functions of normal goods are downward sloping [Lewbel, 2010, Blundell et al., 2012],

Shape constraints are omnipresent [Johnson and Jiang, 2018, Guntuboyina and Sen, 2018, Chetverikov et al., 2018]

- Economics:
  - utility functions are $\nearrow$ and concave [Matzkin, 1991].
  - demand functions of normal goods are downward sloping [Lewbel, 2010, Blundell et al., 2012],
  - production functions are concave [Varian, 1984].

Shape constraints are omnipresent [Johnson and Jiang, 2018, Guntuboyina and Sen, 2018, Chetverikov et al., 2018]

- Economics:
  - utility functions are $\nearrow$ and concave [Matzkin, 1991].
  - demand functions of normal goods are downward sloping [Lewbel, 2010, Blundell et al., 2012],
  - production functions are concave [Varian, 1984].
- Statistics: quantile function $\nearrow$ w.r.t. the quantile level, pdfs are non-negative and often log-concave.

Shape constraints are omnipresent [Johnson and Jiang, 2018, Guntuboyina and Sen, 2018, Chetverikov et al., 2018]

- Economics:
  - utility functions are $\nearrow$ and concave [Matzkin, 1991].
  - demand functions of normal goods are downward sloping [Lewbel, 2010, Blundell et al., 2012],
  - production functions are concave [Varian, 1984].
- Statistics: quantile function $\nearrow$ w.r.t. the quantile level, pdfs are non-negative and often log-concave.
- Finance:
  - European and American call option prices: convex & monotone in the underlying stock price and $\nearrow$ in volatility [Aït-Sahalia and Duarte, 2003].

Shape constraints are omnipresent [Johnson and Jiang, 2018, Guntuboyina and Sen, 2018, Chetverikov et al., 2018]

- Economics:
  - utility functions are $\nearrow$ and concave [Matzkin, 1991].
  - demand functions of normal goods are downward sloping [Lewbel, 2010, Blundell et al., 2012],
  - production functions are concave [Varian, 1984].
- Statistics: quantile function $\nearrow$ w.r.t. the quantile level, pdfs are non-negative and often log-concave.
- Finance:
  - European and American call option prices: convex & monotone in the underlying stock price and $\nearrow$ in volatility [Aït-Sahalia and Duarte, 2003].
- RL and stochastic optimization: value functions are often convex [Keshavarz et al., 2011, Shapiro et al., 2014].

Shape constraints are omnipresent [Johnson and Jiang, 2018, Guntuboyina and Sen, 2018, Chetverikov et al., 2018]

- Economics:
  - utility functions are $\nearrow$ and concave [Matzkin, 1991].
  - demand functions of normal goods are downward sloping [Lewbel, 2010, Blundell et al., 2012],
  - production functions are concave [Varian, 1984].
- Statistics: quantile function $\nearrow$ w.r.t. the quantile level, pdfs are non-negative and often log-concave.
- Finance:
  - European and American call option prices: convex & monotone in the underlying stock price and $\nearrow$ in volatility [Aït-Sahalia and Duarte, 2003].
- RL and stochastic optimization: value functions are often convex [Keshavarz et al., 2011, Shapiro et al., 2014].
- Supply chain models, game theory: supermodularity [Topkis, 1998, Simchi-Levi et al., 2014].

Focus

- Find $f \in \mathcal{H}$ such that $f(\mathbf{x}_n) \approx y_n$, $0 \leq Df(\mathbf{x})$ $\forall \mathbf{x} \in K$.

Focus

- Find $f \in \mathcal{H}$ such that $f(\mathbf{x}_n) \approx y_n$, $0 \leq Df(\mathbf{x})$ $\forall \mathbf{x} \in K$.
- Various exciting approaches with asymptotic guarantees [Han and Wellner, 2016, Chen and Samworth, 2016, Freyberger and Reeves, 2018, Lim, 2020, Deng and Zhang, 2020, Kur et al., 2020], but
  1. they are often 'soft': restriction at finite many points,

Focus

- Find $f \in \mathcal{H}$ such that $f(\mathbf{x}_n) \approx y_n$, $0 \leq Df(\mathbf{x})$ $\forall \mathbf{x} \in K$.
- Various exciting approaches with asymptotic guarantees [Han and Wellner, 2016, Chen and Samworth, 2016, Freyberger and Reeves, 2018, Lim, 2020, Deng and Zhang, 2020, Kur et al., 2020], but
  1. they are often 'soft': restriction at finite many points,
  2. use simplistic function classes: polynomials, polynomial splines,

Focus

- Find $f \in \mathcal{H}$ such that $f(\mathbf{x}_n) \approx y_n$, $0 \leq Df(\mathbf{x})$ $\forall \mathbf{x} \in K$.
- Various exciting approaches with asymptotic guarantees [Han and Wellner, 2016, Chen and Samworth, 2016, Freyberger and Reeves, 2018, Lim, 2020, Deng and Zhang, 2020, Kur et al., 2020], but
  1. they are often 'soft': restriction at finite many points,
  2. use simplistic function classes: polynomials, polynomial splines,
  3. apply hard-wired parameterizations: exponential, quadratic, or

Focus

- Find $f \in \mathcal{H}$ such that $f(\mathbf{x}_n) \approx y_n$, $0 \leq Df(\mathbf{x})$ $\forall \mathbf{x} \in K$.
- Various exciting approaches with asymptotic guarantees [Han and Wellner, 2016, Chen and Samworth, 2016, Freyberger and Reeves, 2018, Lim, 2020, Deng and Zhang, 2020, Kur et al., 2020], <u>but</u>
  1. they are often 'soft': restriction at finite many points,
  2. use simplistic function classes: polynomials, polynomial splines,
  3. apply hard-wired parameterizations: exponential, quadratic, or
  4. only work for (a few) fixed Ds.

Focus

- Find $f \in \mathcal{H}$ such that $f(\mathbf{x}_n) \approx y_n$, $0 \leq Df(\mathbf{x})$ $\forall \mathbf{x} \in K$.
- Various exciting approaches with asymptotic guarantees [Han and Wellner, 2016, Chen and Samworth, 2016, Freyberger and Reeves, 2018, Lim, 2020, Deng and Zhang, 2020, Kur et al., 2020], but
  1. they are often 'soft': restriction at finite many points,
  2. use simplistic function classes: polynomials, polynomial splines,
  3. apply hard-wired parameterizations: exponential, quadratic, or
  4. only work for (a few) fixed Ds.

Today: optimization framework

rich $\mathcal{H}$, hard ($\forall \mathbf{x} \in K$) shape constraints, modularity in D.

Focus

- Find $f \in \mathcal{H}$ such that $f(\mathbf{x}_n) \approx y_n$, $0 \leq Df(\mathbf{x})$ $\forall \mathbf{x} \in K$.
- Various exciting approaches with asymptotic guarantees [Han and Wellner, 2016, Chen and Samworth, 2016, Freyberger and Reeves, 2018, Lim, 2020, Deng and Zhang, 2020, Kur et al., 2020], but
  1. they are often 'soft': restriction at finite many points,
  2. use simplistic function classes: polynomials, polynomial splines,
  3. apply hard-wired parameterizations: exponential, quadratic, or
  4. only work for (a few) fixed Ds.

Today: optimization framework

rich $\mathcal{H}$, hard ($\forall \mathbf{x} \in K$) shape constraints, modularity in D.

Towards flexible $\mathcal{H}$-s . . .

Kernel

- Def-1 (feature space): $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ kernel if

$$k(x, y) = \langle \varphi(x), \varphi(y) \rangle_{\mathcal{H}} .$$

- Examples ($\gamma > 0$, $c \geq 0$, $p \in \mathbb{Z}^+$):

$$k_p(\mathbf{x}, \mathbf{y}) = (\langle \mathbf{x}, \mathbf{y} \rangle + c)^p, \qquad k_G(\mathbf{x}, \mathbf{y}) = e^{-\gamma \|\mathbf{x} - \mathbf{y}\|_2^2},$$
$$k_L(\mathbf{x}, \mathbf{y}) = e^{-\gamma \|\mathbf{x} - \mathbf{y}\|_1}, \qquad k_e(\mathbf{x}, \mathbf{y}) = e^{\gamma \langle \mathbf{x}, \mathbf{y} \rangle}.$$

Kernel, RKHS

- Def-1 (feature space): $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ kernel if

$$k(x, y) = \langle \varphi(x), \varphi(y) \rangle_{\mathcal{H}}.$$

- Def-2 (reproducing kernel):

$$k(\cdot, x) := [x' \mapsto k(x', x)] \in \mathcal{H}, \qquad f(x) = \langle f, k(\cdot, x) \rangle_{\mathcal{H}}.$$

Constructively, $\mathcal{H}_k = \overline{\{\sum_{i=1}^{n} \alpha_i k(\cdot, x_i) \,:\, \alpha_i \in \mathbb{R}, x_i \in \mathcal{X}, n \in \mathbb{N}^*\}}$.

- Examples ($\gamma > 0$, $c \geq 0$, $p \in \mathbb{Z}^+$):

$$k_p(\mathbf{x}, \mathbf{y}) = (\langle \mathbf{x}, \mathbf{y} \rangle + c)^p, \qquad k_G(\mathbf{x}, \mathbf{y}) = e^{-\gamma \|\mathbf{x}-\mathbf{y}\|_2^2},$$

$$k_L(\mathbf{x}, \mathbf{y}) = e^{-\gamma \|\mathbf{x}-\mathbf{y}\|_1}, \qquad k_e(\mathbf{x}, \mathbf{y}) = e^{\gamma \langle \mathbf{x}, \mathbf{y} \rangle}.$$

Kernel, RKHS

- Def-1 (feature space): $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ kernel if

$$k(x, y) = \langle \varphi(x), \varphi(y) \rangle_{\mathcal{H}}.$$

- Def-2 (reproducing kernel):

$$k(\cdot, x) := [x' \mapsto k(x', x)] \in \mathcal{H}, \qquad f(x) = \langle f, k(\cdot, x) \rangle_{\mathcal{H}}.$$

Constructively, $\mathcal{H}_k = \overline{\{\sum_{i=1}^{n} \alpha_i k(\cdot, x_i) \, : \, \alpha_i \in \mathbb{R}, x_i \in \mathcal{X}, n \in \mathbb{N}^*\}}$.

- Equivalent definitions, $k \overset{1:1}{\leftrightarrow} \mathcal{H}_k$.
- Included: Fourier analysis, polynomials, splines, ...
- Examples ($\gamma > 0$, $c \geq 0$, $p \in \mathbb{Z}^+$):

$$k_p(\mathbf{x}, \mathbf{y}) = (\langle \mathbf{x}, \mathbf{y} \rangle + c)^p, \qquad k_G(\mathbf{x}, \mathbf{y}) = e^{-\gamma \|\mathbf{x} - \mathbf{y}\|_2^2},$$

$$k_L(\mathbf{x}, \mathbf{y}) = e^{-\gamma \|\mathbf{x} - \mathbf{y}\|_1}, \qquad k_e(\mathbf{x}, \mathbf{y}) = e^{\gamma \langle \mathbf{x}, \mathbf{y} \rangle}.$$

Task-1: joint quantile regression (JQR)

- Given: $(\tau_q)_{q\in[Q]} \subset (0,1)$ levels $\nearrow$, $\{(\mathbf{x}_n, y_n)\}_{n\in[N]}$ samples.
- Estimate jointly the τ_q-quantiles of $\mathbb{P}(Y|X=\mathbf{x})$.

Task-1: joint quantile regression (JQR)

- Given: $(\tau_q)_{q\in[Q]} \subset (0,1)$ levels $\nearrow$, $\{(\mathbf{x}_n, y_n)\}_{n\in[N]}$ samples.
- Estimate jointly the τ_q-quantiles of $\mathbb{P}(Y|X=\mathbf{x})$ [Sangnier et al., 2016].
- Objective:

$$\mathcal{L}(\mathbf{f},\mathbf{b}) = \underbrace{\frac{1}{N}\sum_{q\in[Q]}\sum_{n\in[N]} l_{\tau_q}\left(y_n - [f_q(\mathbf{x}_n) + b_q]\right)}_{\text{quantile property}} + \underbrace{\lambda_{\mathbf{b}}\|\mathbf{b}\|_2^2 + \lambda_{\mathbf{f}}\sum_{q\in[Q]}\|f_q\|_k^2}_{\text{regularization}},$$

$$l_\tau(e) = \max(\tau e, (\tau-1)e).$$

Task-1: joint quantile regression (JQR)

- Given: $(\tau_q)_{q\in[Q]} \subset (0,1)$ levels $\nearrow$, $\{(\mathbf{x}_n, y_n)\}_{n\in[N]}$ samples.
- Estimate jointly the τ_q-quantiles of $\mathbb{P}(Y|X=\mathbf{x})$ [Sangnier et al., 2016].
- Objective:

$$\mathcal{L}(\mathbf{f},\mathbf{b}) = \underbrace{\frac{1}{N}\sum_{q\in[Q]}\sum_{n\in[N]} l_{\tau_q}\left(y_n - [f_q(\mathbf{x}_n)+b_q]\right)}_{\text{quantile property}} + \underbrace{\lambda_{\mathbf{b}}\|\mathbf{b}\|_2^2 + \lambda_{\mathbf{f}}\sum_{q\in[Q]}\|f_q\|_k^2}_{\text{regularization}},$$

$$l_\tau(e) = \max(\tau e, (\tau-1)e).$$

- Constraint (non-crossing): $K :=$ smallest rectangle containing $\{\mathbf{x}_n\}_{n\in[N]}$,

$$f_q(\mathbf{x}) + b_q \le f_{q+1}(\mathbf{x}) + b_{q+1},\ \forall q\in[Q-1],\ \forall \mathbf{x}\in K.$$

Task-1: joint quantile regression (JQR)

- Given: $(\tau_q)_{q\in[Q]} \subset (0,1)$ levels $\nearrow$, $\{(\mathbf{x}_n, y_n)\}_{n\in[N]}$ samples.
- Estimate jointly the τ_q-quantiles of $\mathbb{P}(Y|X=\mathbf{x})$ [Sangnier et al., 2016].
- Objective:

$$\mathcal{L}(\mathbf{f},\mathbf{b}) = \underbrace{\frac{1}{N}\sum_{q\in[Q]}\sum_{n\in[N]} l_{\tau_q}\left(y_n - [f_q(\mathbf{x}_n) + b_q]\right)}_{\text{quantile property}} + \underbrace{\lambda_{\mathbf{b}}\|\mathbf{b}\|_2^2 + \lambda_{\mathbf{f}}\sum_{q\in[Q]}\|f_q\|_k^2}_{\text{regularization}},$$

$$l_\tau(e) = \max(\tau e, (\tau-1)e).$$

- Constraint (non-crossing): $K :=$ smallest rectangle containing $\{\mathbf{x}_n\}_{n\in[N]}$,

$$f_q(\mathbf{x}) + b_q \le f_{q+1}(\mathbf{x}) + b_{q+1},\ \forall q \in [Q-1],\ \forall \mathbf{x} \in K.$$

Constraints

function values (f_q) with interaction ($f_{q+1} - f_q$), bias terms (b_q) with interaction ($b_q - b_{q+1}$).

Task-2: convoy localization, one vehicle ($Q = 1$)

- Given: noisy time-location samples $\{(t_n, x_n)\}_{n\in[N]} \subset \underbrace{[0, T]}_{=:\mathcal{T}} \times \mathbb{R}$.
- Goal: learn the (t, x) relation.
- Constraint: lower bound on speed (v_{min}).

Task-2: convoy localization, one vehicle ($Q = 1$)

- Given: noisy time-location samples $\{(t_n, x_n)\}_{n\in[N]} \subset \underbrace{[0, T]}_{=:\mathcal{T}} \times \mathbb{R}$.
- Goal: learn the (t, x) relation.
- Constraint: lower bound on speed (v_{min}).
- Objective:

$$\min_{b \in \mathbb{R}, f \in \mathcal{H}_k} \left[\frac{1}{N} \sum_{n\in[N]} |x_n - [b + f(t_n)]|^2 + \lambda \|f\|^2_{\mathcal{H}_k} \right]$$

$$\text{s.t.}$$

$$v_{\text{min}} \leq f'(t), \quad \forall t \in \mathcal{T}.$$

Task-2b: convoy localization, multiple vehicles ($Q \geq 1$)

- Data: $\left\{(t_{q,n}, x_{q,n})_{n\in[N_q]}\right\}_{q\in[Q]} \subseteq \mathcal{T} \times \mathbb{R}$.
- Constraints: speed ($v_{\min}$), inter-vehicular distance ($d_{\min}$).
- Objective:

$$\min_{\substack{f_1,\ldots,f_Q \in \mathcal{H}_k,\\ b_1,\ldots,b_Q \in \mathbb{R}}} \frac{1}{Q}\sum_{q=1}^{Q}\left[\left(\frac{1}{N_q}\sum_{n=1}^{N_q}\left|x_{q,n} - (b_q + f_q(t_{q,n}))\right|^2\right) + \lambda\|f_q\|_{\mathcal{H}_k}^2\right]$$

$$\text{s.t.}$$

$$d_{\min} + b_{q+1} + f_{q+1}(t) \leq b_q + f_q(t), \forall q \in [Q-1],\ t \in \mathcal{T},$$

$$v_{\min} \leq f_q'(t), \qquad \forall q \in [Q],\ t \in \mathcal{T}.$$

Task-2b: convoy localization, multiple vehicles ($Q \geq 1$)

- Data: $\left\{(t_{q,n}, x_{q,n})_{n\in[N_q]}\right\}_{q\in[Q]} \subseteq \mathcal{T} \times \mathbb{R}$.
- Constraints: speed ($v_{\min}$), inter-vehicular distance ($d_{\min}$).
- Objective:

$$\min_{\substack{f_1,\ldots,f_Q \in \mathcal{H}_k,\\ b_1,\ldots,b_Q \in \mathbb{R}}} \frac{1}{Q}\sum_{q=1}^{Q}\left[\left(\frac{1}{N_q}\sum_{n=1}^{N_q}\left|x_{q,n} - (b_q + f_q(t_{q,n}))\right|^2\right) + \lambda \|f_q\|_{\mathcal{H}_k}^2\right]$$

s.t.

$$d_{\min} + b_{q+1} + f_{q+1}(t) \leq b_q + f_q(t), \forall q \in [Q-1],\ t \in \mathcal{T},$$

$$v_{\min} \leq f_q'(t), \qquad \forall q \in [Q],\ t \in \mathcal{T}.$$

Constraints

function values (f_q) and derivatives (f_q') with interaction ($f_q - f_{q+1}$), bias terms (b_q) with interaction ($b_{q+1} - b_q$).

Our task

$$\left(\bar{\mathbf{f}}, \bar{\mathbf{b}}\right) = \underset{\substack{\mathbf{f}=(f_q)_{q\in[Q]} \in (\mathcal{H}_k)^Q, \\ \mathbf{b}=(b_q)_{q\in[Q]} \in \mathcal{B}, \\ (\mathbf{f},\mathbf{b}) \in C}}{\arg\min} \mathcal{L}(\mathbf{f}, \mathbf{b}),$$

Our task

$$\left(\bar{\mathbf{f}}, \bar{\mathbf{b}}\right) = \operatorname*{arg\,min}_{\substack{\mathbf{f}=(f_q)_{q\in[Q]}\in(\mathcal{H}_k)^Q,\\ \mathbf{b}=(b_q)_{q\in[Q]}\in\mathcal{B},\\ (\mathbf{f},\mathbf{b})\in C}} \mathcal{L}(\mathbf{f}, \mathbf{b}),$$

$$\mathcal{L}(\mathbf{f}, \mathbf{b}) = L\left(\mathbf{b}, \left(\mathbf{x}_n, y_n, (f_q(\mathbf{x}_n))_{q\in[Q]}\right)_{n\in[N]}\right) + \Omega\left(\left(\|f_q\|_{\mathcal{H}_k}\right)_{q\in[Q]}\right),$$

Our task

$$\left(\bar{\mathbf{f}}, \bar{\mathbf{b}}\right) = \underset{\substack{\mathbf{f}=(f_q)_{q\in[Q]} \in (\mathcal{H}_k)^Q,\\ \mathbf{b}=(b_q)_{q\in[Q]} \in \mathcal{B},\\ (\mathbf{f},\mathbf{b}) \in C}}{\arg\min} \mathcal{L}(\mathbf{f}, \mathbf{b}),$$

$$\mathcal{L}(\mathbf{f}, \mathbf{b}) = L\left(\mathbf{b}, \left(\mathbf{x}_n, y_n, (f_q(\mathbf{x}_n))_{q\in[Q]}\right)_{n\in[N]}\right) + \Omega\left((\|f_q\|_{\mathcal{H}_k})_{q\in[Q]}\right),$$

$$C = \{(\mathbf{f}, \mathbf{b}) \,|\, (\mathbf{b}_0 - \mathbf{U}\mathbf{b})_i \leq D_i(\mathbf{W}\mathbf{f} - \mathbf{f}_0)_i(\mathbf{x}), \quad \forall \mathbf{x} \in K_i, \forall i \in [I]\},$$

Our task

$$\left(\bar{\mathbf{f}}, \bar{\mathbf{b}}\right) = \underset{\substack{\mathbf{f}=(f_q)_{q\in[Q]} \in (\mathcal{H}_k)^Q,\\ \mathbf{b}=(b_q)_{q\in[Q]} \in \mathcal{B},\\ (\mathbf{f},\mathbf{b}) \in C}}{\arg\min} \mathcal{L}(\mathbf{f}, \mathbf{b}),$$

$$\mathcal{L}(\mathbf{f}, \mathbf{b}) = L\left(\mathbf{b}, \left(\mathbf{x}_n, y_n, (f_q(\mathbf{x}_n))_{q\in[Q]}\right)_{n\in[N]}\right) + \Omega\left((\|f_q\|_{\mathcal{H}_k})_{q\in[Q]}\right),$$

$$C = \{(\mathbf{f}, \mathbf{b}) \,|\, (\mathbf{b}_0 - \mathbf{U}\mathbf{b})_i \le D_i(\mathbf{W}\mathbf{f} - \mathbf{f}_0)_i(\mathbf{x}), \quad \forall \mathbf{x} \in K_i, \forall i \in [I]\},$$

$$(\mathbf{W}\mathbf{f})_i = \sum_{q\in[Q]} W_{i,q} f_q,$$

Our task

$$\left(\bar{\mathbf{f}}, \bar{\mathbf{b}}\right) = \operatorname*{arg\,min}_{\substack{\mathbf{f}=(f_q)_{q\in[Q]} \in (\mathcal{H}_k)^Q, \\ \mathbf{b}=(b_q)_{q\in[Q]} \in \mathcal{B}, \\ (\mathbf{f},\mathbf{b}) \in C}} \mathcal{L}(\mathbf{f}, \mathbf{b}),$$

$$\mathcal{L}(\mathbf{f}, \mathbf{b}) = L\left(\mathbf{b}, \left(\mathbf{x}_n, y_n, (f_q(\mathbf{x}_n))_{q\in[Q]}\right)_{n\in[N]}\right) + \Omega\left(\left(\|f_q\|_{\mathcal{H}_k}\right)_{q\in[Q]}\right),$$

$$C = \left\{(\mathbf{f}, \mathbf{b}) \,|\, (\mathbf{b}_0 - \mathbf{U}\mathbf{b})_i \le D_i(\mathbf{W}\mathbf{f} - \mathbf{f}_0)_i(\mathbf{x}), \quad \forall \mathbf{x} \in K_i, \forall i \in [I]\right\},$$

$$(\mathbf{W}\mathbf{f})_i = \sum_{q\in[Q]} W_{i,q} f_q,$$

$$D_i = \sum_{j\in[n_{i,j}]} \gamma_{i,j}\partial^{\mathbf{r}_{i,j}}, \; |\mathbf{r}_{i,j}| \le s, \; \gamma_{i,j} \in \mathbb{R}, \; \partial^{\mathbf{r}} f(\mathbf{x}) = \frac{\partial^{|\mathbf{r}|} f(\mathbf{x})}{\partial^{r_1}_{x_1} \cdots \partial^{r_d}_{x_d}}.$$

Blanket assumptions

1. Domain $\mathcal{X} \subseteq \mathbb{R}^d$: open. Kernel $k \in \mathcal{C}^s(\mathcal{X} \times \mathcal{X})$.
2. $K_i \subset \mathcal{X}$: compact, $\forall i$.
3. $\mathbf{f}_{0,i} \in \mathcal{H}_k$ for $\forall i$.
4. Bias domain $\mathcal{B} \subseteq \mathbb{R}^Q$: convex.
5. Loss L restricted to $\mathcal{B}$: strictly convex in $\mathbf{b}$.
6. Regularizer Ω: strictly increasing in each of its argument.

Our strenghtened SOC-constrained formulation

$$(\mathbf{f}_{\boldsymbol{\eta}}, \mathbf{b}_{\boldsymbol{\eta}}) = \underset{\mathbf{f} \in (\mathcal{H}_k)^Q,\, \mathbf{b} \in \mathcal{B}}{\arg\min} \quad \mathcal{L}(\mathbf{f}, \mathbf{b}) \tag{$\mathcal{P}_{\boldsymbol{\eta}}$}$$

$$\text{s.t.}$$

$$(\mathbf{b}_0 - \mathbf{U}\mathbf{b})_i + \eta_i \|(\mathbf{W}\mathbf{f} - \mathbf{f}_0)_i\|_{\mathcal{H}_k} \leq \min_{m \in [M_i]} D_i(\mathbf{W}\mathbf{f} - \mathbf{f}_0)_i \left(\tilde{\mathbf{x}}_{i,m}\right), \; \forall i \in [I], \tag{$\mathcal{C}_{\boldsymbol{\eta}}$}$$

where

- $\{\tilde{\mathbf{x}}_{i,m}\}_{m \in [M_i]}$: a δ_i-net of K_i in $\|\cdot\|_{\mathcal{X}}$,
- $\eta_i = \sup_{m \in [M_i], \mathbf{u} \in \mathbb{B}_{\|\cdot\|_{\mathcal{X}}}(\mathbf{0},1)} \|D_{i,\mathbf{x}} k(\tilde{\mathbf{x}}_{i,m}, \cdot) - D_{i,\mathbf{x}} k(\tilde{\mathbf{x}}_{i,m} + \delta_i \mathbf{u}, \cdot)\|_{\mathcal{H}_k}$,
- $D_{i,\mathbf{x}} k(\mathbf{x}_0, \cdot) := \mathbf{y} \mapsto D_i(\mathbf{x} \mapsto k(\mathbf{x}, \mathbf{y}))(\mathbf{x}_0)$.

Tightening idea

Let $s = 0$, $l = 1$. Recall constraint $(\mathcal{C})$:

$$\left\{(\mathbf{f}, \mathbf{b}) \mid \underbrace{(b_0 - \mathbf{Ub})}_{\beta} \le \underbrace{\underbrace{(\mathbf{Wf} - f_0)}_{\phi}(\mathbf{x})}_{\langle \phi, k(\mathbf{x},\cdot)\rangle_{\mathcal{H}_k}}, \quad \forall \mathbf{x} \in K\right\}$$

Tightening idea

Let $s = 0$, $l = 1$. Recall constraint $(\mathcal{C})$:

$$\left\{(\mathbf{f},\mathbf{b}) \mid \underbrace{(b_0 - \mathbf{U}\mathbf{b})}_{\beta} \leq \underbrace{\underbrace{(\mathbf{W}\mathbf{f} - f_0)}_{\phi}(\mathbf{x})}_{\langle \phi, k(\mathbf{x},\cdot)\rangle_{\mathcal{H}_k}}, \quad \forall \mathbf{x} \in K\right\}, \text{i.e.}$$

$$\Phi(K) := \{k(\mathbf{x},\cdot) : \mathbf{x} \in K\} \subseteq H^{+}_{\phi,\beta} := \left\{g \in \mathcal{H}_k \mid \beta \leq \langle \phi, g\rangle_{\mathcal{H}_k}\right\}$$

Tightening idea

Let $s = 0$, $l = 1$. Recall constraint $(\mathcal{C})$:

$$\Big\{(\mathbf{f}, \mathbf{b}) \,|\, \underbrace{(b_0 - \mathbf{U}\mathbf{b})}_{\beta} \le \underbrace{\underbrace{(\mathbf{W}\mathbf{f} - f_0)}_{\phi}(\mathbf{x})}_{\langle \phi, k(\mathbf{x},\cdot)\rangle_{\mathcal{H}_k}}, \quad \forall \mathbf{x} \in K\Big\}, \text{ i.e.}$$

$$\Phi(K) := \{k(\mathbf{x}, \cdot) : \mathbf{x} \in K\} \subseteq H^{+}_{\phi,\beta} := \Big\{g \in \mathcal{H}_k \,|\, \beta \le \langle \phi, g\rangle_{\mathcal{H}_k}\Big\}$$

- $(\mathcal{C}_{\eta})$ means: covering of $\Phi(K)$ by balls with η-radius centered at the $k(\tilde{\mathbf{x}}_m, \cdot)$ is in the halfspace $H^{+}_{\phi,\beta}$; hence it is tightening.

Tightening idea

Let $s = 0$, $l = 1$. Recall constraint $(\mathcal{C})$:

$$\Big\{(\mathbf{f}, \mathbf{b}) \,|\, \underbrace{(b_0 - \mathbf{U}\mathbf{b})}_{\beta} \leq \underbrace{\underbrace{(\mathbf{W}\mathbf{f} - f_0)}_{\phi}(\mathbf{x})}_{\langle \phi, k(\mathbf{x},\cdot)\rangle_{\mathcal{H}_k}}, \quad \forall \mathbf{x} \in K\Big\}, \text{ i.e.}$$

$$\Phi(K) := \{k(\mathbf{x}, \cdot) : \mathbf{x} \in K\} \subseteq H^+_{\phi,\beta} := \Big\{g \in \mathcal{H}_k \,|\, \beta \leq \langle \phi, g\rangle_{\mathcal{H}_k}\Big\}$$

- $(\mathcal{C}_{\boldsymbol{\eta}})$ means: covering of $\Phi(K)$ by balls with η-radius centered at the $k(\tilde{\mathbf{x}}_m, \cdot)$ is in the halfspace $H^+_{\phi,\beta}$; hence it is tightening.
- η is obtained as the minimal radius.

Theorem

- Minimal values: v_{disc} = value of $(\mathcal{P}_{\boldsymbol{\eta}})$ with '$\boldsymbol{\eta} = \mathbf{0}$', $\bar{v} = \mathcal{L}\left(\bar{\mathbf{f}}, \bar{\mathbf{b}}\right)$, $v_{\boldsymbol{\eta}} = \mathcal{L}\left(\mathbf{f}_{\boldsymbol{\eta}}, \mathbf{b}_{\boldsymbol{\eta}}\right)$.
- Let $\mathbf{f}_{\boldsymbol{\eta}} = \left(f_{\boldsymbol{\eta},q}\right)_{q\in[Q]}$.

Theorem

- Minimal values: v_{disc} = value of $(\mathcal{P}_{\boldsymbol{\eta}})$ with '$\boldsymbol{\eta} = \mathbf{0}$', $\bar{v} = \mathcal{L}\left(\bar{\mathbf{f}}, \bar{\mathbf{b}}\right)$, $v_{\boldsymbol{\eta}} = \mathcal{L}\left(\mathbf{f}_{\boldsymbol{\eta}}, \mathbf{b}_{\boldsymbol{\eta}}\right)$.
- Let $\mathbf{f}_{\boldsymbol{\eta}} = \left(f_{\boldsymbol{\eta},q}\right)_{q\in[Q]}$.

Then,

- (i) Tightening: any $(\mathbf{f}, \mathbf{b})$ satisfying $(\mathcal{C}_{\boldsymbol{\eta}})$ also satisfies $(\mathcal{C})$, hence

$$v_{\text{disc}} \leq \bar{v} \leq v_{\boldsymbol{\eta}}.$$

Theorem

- Minimal values: v_{disc} = value of $(\mathcal{P}_{\boldsymbol{\eta}})$ with '$\boldsymbol{\eta} = \mathbf{0}$', $\bar{v} = \mathcal{L}\left(\bar{\mathbf{f}}, \bar{\mathbf{b}}\right)$, $v_{\boldsymbol{\eta}} = \mathcal{L}\left(\mathbf{f}_{\boldsymbol{\eta}}, \mathbf{b}_{\boldsymbol{\eta}}\right)$.
- Let $\mathbf{f}_{\boldsymbol{\eta}} = (f_{\boldsymbol{\eta},q})_{q\in[Q]}$.

Then,

- (i) Tightening: any $(\mathbf{f}, \mathbf{b})$ satisfying $(\mathcal{C}_{\boldsymbol{\eta}})$ also satisfies $(\mathcal{C})$, hence

$$v_{\text{disc}} \leq \bar{v} \leq v_{\boldsymbol{\eta}}.$$

- (ii) Representer theorem: For $\forall q \in [Q]$, $\exists \tilde{a}_{i,0,q}, \tilde{a}_{i,m,q}, a_{n,q} \in \mathbb{R}$ s.t.

$$f_{\boldsymbol{\eta},q} = \sum_{i\in[I]} \left[\tilde{a}_{i,0,q} f_{0,i} + \sum_{m\in[M_i]} \tilde{a}_{i,m,q} D_{i,\mathbf{x}} k\left(\tilde{\mathbf{x}}_{i,m}, \cdot\right) \right] + \sum_{n\in[N]} a_{n,q} k(\mathbf{x}_n, \cdot).$$

Theorem – continued

- (iii) Performance guarantee: if $\mathcal{L}$ is $(\mu_{f_q}, \mu_{\mathbf{b}})$-strongly convex w.r.t. $(f_q, \mathbf{b})$ for any $q \in [Q]$, then

$$\|f_{\boldsymbol{\eta},q} - \bar{f}_q\|_{\mathcal{H}_k} \leq \sqrt{\frac{2(v_{\eta} - v_{\text{disc}})}{\mu_{f_q}}}, \quad \|\mathbf{b}_{\boldsymbol{\eta}} - \bar{\mathbf{b}}\|_2 \leq \sqrt{\frac{2(v_{\eta} - v_{\text{disc}})}{\mu_{\mathbf{b}}}}.$$

Theorem – continued

- (iii) Performance guarantee: if $\mathcal{L}$ is $(\mu_{f_q}, \mu_{\mathbf{b}})$-strongly convex w.r.t. $(f_q, \mathbf{b})$ for any $q \in [Q]$, then

$$\|f_{\boldsymbol{\eta},q} - \bar{f}_q\|_{\mathcal{H}_k} \leq \sqrt{\frac{2(v_{\eta} - v_{\text{disc}})}{\mu_{f_q}}}, \quad \|\mathbf{b}_{\eta} - \bar{\mathbf{b}}\|_2 \leq \sqrt{\frac{2(v_{\eta} - v_{\text{disc}})}{\mu_{\mathbf{b}}}}.$$

If in addition $\mathbf{U}$ is surjective, $\mathcal{B} = \mathbb{R}^Q$, and $\mathcal{L}(\bar{\mathbf{f}}, \cdot)$ is L_b–Lipschitz continuous on $\mathbb{B}_{\|\cdot\|_2}\left(\bar{\mathbf{b}}, c_f\|\boldsymbol{\eta}\|_\infty\right)$ where $c_f = \sqrt{d}\left\|\left(\mathbf{U}^T\mathbf{U}\right)^{-1}\mathbf{U}^T\right\| \max_{i\in[I]}\left\|(\mathbf{W}\bar{\mathbf{f}} - \mathbf{f}_0)_i\right\|_{\mathcal{H}_k}$, then

$$\|f_{\boldsymbol{\eta},q} - \bar{f}_q\|_{\mathcal{H}_k} \leq \sqrt{\frac{2L_b c_f\|\boldsymbol{\eta}\|_\infty}{\mu_{f_q}}}, \|\mathbf{b}_{\eta} - \bar{\mathbf{b}}\|_2 \leq \sqrt{\frac{2L_b c_f\|\boldsymbol{\eta}\|_\infty}{\mu_{\mathbf{b}}}}.$$

Theorem – continued

- (iii) Performance guarantee: if $\mathcal{L}$ is $(\mu_{f_q}, \mu_{\mathbf{b}})$-strongly convex w.r.t. $(f_q, \mathbf{b})$ for any $q \in [Q]$, then

$$\|f_{\boldsymbol{\eta},q} - \bar{f}_q\|_{\mathcal{H}_k} \leq \sqrt{\frac{2(v_{\eta} - v_{\text{disc}})}{\mu_{f_q}}}, \quad \|\mathbf{b}_{\boldsymbol{\eta}} - \bar{\mathbf{b}}\|_2 \leq \sqrt{\frac{2(v_{\eta} - v_{\text{disc}})}{\mu_{\mathbf{b}}}}.$$

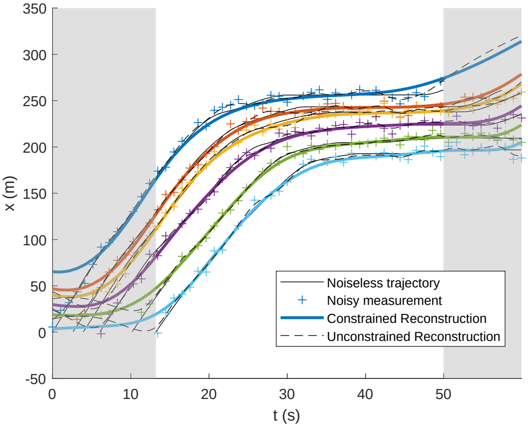
If in addition $\mathbf{U}$ is surjective, $\mathcal{B} = \mathbb{R}^Q$, and $\mathcal{L}(\bar{\mathbf{f}}, \cdot)$ is L_b−Lipschitz continuous on $\mathbb{B}_{\|\cdot\|_2}\left(\bar{\mathbf{b}}, c_f\|\boldsymbol{\eta}\|_\infty\right)$ where $c_f = \sqrt{d}\left\|\left(\mathbf{U}^T\mathbf{U}\right)^{-1}\mathbf{U}^T\right\| \max_{i\in[I]}\left\|(\mathbf{W}\bar{\mathbf{f}} - \mathbf{f}_0)_i\right\|_{\mathcal{H}_k}$, then

$$\|f_{\boldsymbol{\eta},q} - \bar{f}_q\|_{\mathcal{H}_k} \leq \sqrt{\frac{2L_b c_f\|\boldsymbol{\eta}\|_\infty}{\mu_{f_q}}}, \|\mathbf{b}_{\boldsymbol{\eta}} - \bar{\mathbf{b}}\|_2 \leq \sqrt{\frac{2L_b c_f\|\boldsymbol{\eta}\|_\infty}{\mu_{\mathbf{b}}}}.$$

1st bound: computable. 2nd: Larger $M_i \Rightarrow$ smaller $\delta_i \Rightarrow$ smaller η_i $\Rightarrow$ tighter bound.

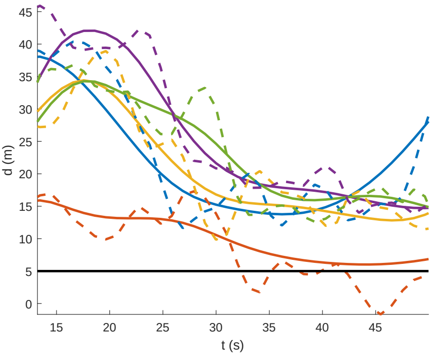
Demo (task-1): convoy localization with traffic jam

Setting: $Q = 6$, $d_{\min} = 5m$, $v_{\min} = 0$.



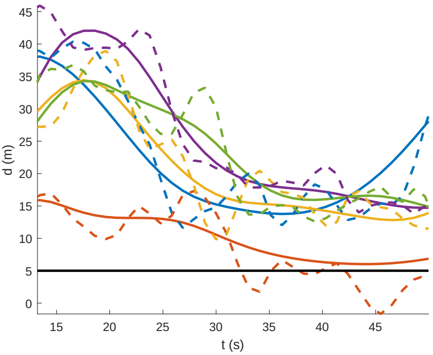
Demo (task-1): continued

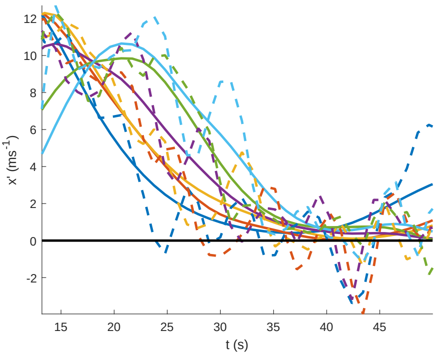
Pairwise distances: $t \mapsto f_q(t) - f_{q+1}(t)$



Demo (task-1): continued

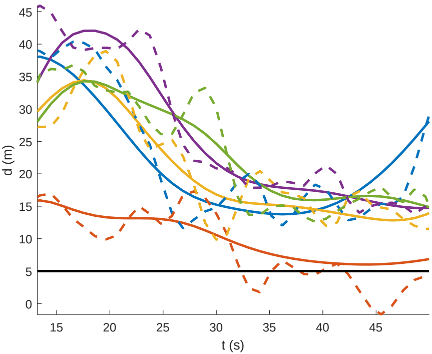
Pairwise distances: $t \mapsto f_q(t) - f_{q+1}(t)$ Speed: $t \mapsto f'_q(t)$

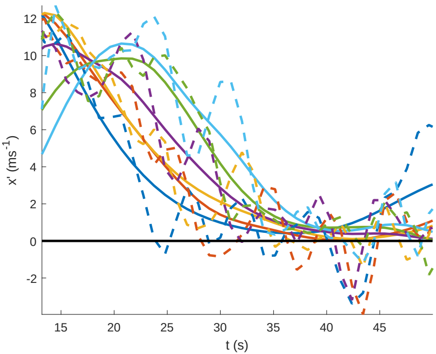




Demo (task-1): continued

Pairwise distances: $t \mapsto f_q(t) - f_{q+1}(t)$ Speed: $t \mapsto f'_q(t)$



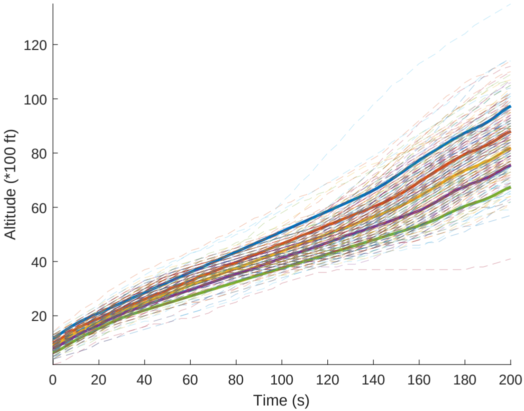


Shape constraints: especially relevant in noisy situations.

Demo (task-2): joint quantile regression

Analysis of aircraft trajectories, ENAC [Nicol, 2013]

- y: radar-measured altitude of aircrafts flying between two cities (Paris & Toulouse); x: time. $d = 1$, $N = 15657$.
- Demo: $\tau_q \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$.
- Constraint: non-crossing, $\nearrow$ (takeoff).



Summary

- Focus: hard affine shape constraints on derivatives & RKHS.
- Proposed framework: SOC-based tightening.
- Applications:
  - convoy localization,
  - joint quantile regression: aircraft trajectories.

References & acknowledgements

- Transportation systems [Aubin-Frankowski et al., 2020].
- Method:
  - $dim(y) = 1$: [Aubin-Frankowski and Szabó, 2020]. Code @ GitHub.
  - $dim(y) \geq 1$ (ex: safety-critical control) and SDP constraints (ex: production functions $\rightarrow$ joint convexity): [Aubin-Frankowski and Szabó, 2022].

Acknowledgements: ZSz benefited from the support of the Europlace Institute of Finance and that of the Chair Stress Test, RISK Management and Financial Steering, led by the French École Polytechnique and its Foundation and sponsored by BNP Paribas.

References & acknowledgements

- Transportation systems [Aubin-Frankowski et al., 2020].
- Method:
  - $dim(y) = 1$: [Aubin-Frankowski and Szabó, 2020]. Code @ GitHub.
  - $dim(y) \geq 1$ (ex: safety-critical control) and SDP constraints (ex: production functions $\rightarrow$ joint convexity): [Aubin-Frankowski and Szabó, 2022].

Acknowledgements: ZSz benefited from the support of the Europlace Institute of Finance and that of the Chair Stress Test, RISK Management and Financial Steering, led by the French École Polytechnique and its Foundation and sponsored by BNP Paribas.

INSTITUT Louis Bachelier

Stress Test
RISK Management and Financial Steering

Task-3: safety-critical control

- Trajectory of an underwater vehicle:

$$t \in \mathcal{T} := [0, 1] \mapsto [x(t); z(t)] \in \mathbb{R}^2.$$

Task-3: safety-critical control

- Trajectory of an underwater vehicle:

$$t \in \mathcal{T} := [0, 1] \mapsto [x(t); z(t)] \in \mathbb{R}^2.$$

- Simplifying assumption: $x(0) = 0, \dot{x}(t) = 1 \, \forall t \in \mathcal{T} \Rightarrow x(t) = t.$

Task-3: safety-critical control

- Trajectory of an underwater vehicle:

$$t \in \mathcal{T} := [0,1] \mapsto [x(t); z(t)] \in \mathbb{R}^2.$$

- Simplifying assumption: $x(0) = 0, \dot{x}(t) = 1 \, \forall t \in \mathcal{T} \Rightarrow x(t) = t$.
- Requirement: stay between the floor and the ceiling of the cavern

$$z(t) \in [z_{\text{low}}(t), z_{\text{up}}(t)] \;\; \forall t \in \mathcal{T}.$$

Task-3: safety-critical control

- Trajectory of an underwater vehicle:

$$t \in \mathcal{T} := [0,1] \mapsto [x(t); z(t)] \in \mathbb{R}^2.$$

- Simplifying assumption: $x(0) = 0, \dot{x}(t) = 1 \, \forall t \in \mathcal{T} \Rightarrow x(t) = t$.
- Requirement: stay between the floor and the ceiling of the cavern

$$z(t) \in [z_{\text{low}}(t), z_{\text{up}}(t)] \;\; \forall t \in \mathcal{T}.$$

- Initial condition: $z(0) = 0$ and $\dot{z}(0) = 0$.

Task-3: safety-critical control

- Trajectory of an underwater vehicle:

$$t \in \mathcal{T} := [0,1] \mapsto [x(t); z(t)] \in \mathbb{R}^2.$$

- Simplifying assumption: $x(0) = 0, \dot{x}(t) = 1 \, \forall t \in \mathcal{T} \Rightarrow x(t) = t$.
- Requirement: stay between the floor and the ceiling of the cavern

$$z(t) \in [z_{\text{low}}(t), z_{\text{up}}(t)] \;\; \forall t \in \mathcal{T}.$$

- Initial condition: $z(0) = 0$ and $\dot{z}(0) = 0$.
- Control task (LQ = linear dynamics & quadratic cost):

$$\begin{aligned} \min_{u \in L^2(\mathcal{T},\mathbb{R})} \quad & \int_{\mathcal{T}} |u(t)|^2 \mathrm{d}t \\ \text{s.t.} \quad & \\ & z(0) = 0, \quad \dot{z}(0) = 0, \\ & \ddot{z}(t) = -\dot{z}(t) + u(t), \; \forall t \in \mathcal{T}, \\ & z_{\text{low}}(t) \leq z(t) \leq z_{\text{up}}(t), \; \forall \, t \in \mathcal{T}. \end{aligned}$$

Task-3: safety-critical control – continued

- With full state $\mathbf{f}(t) := [z(t); \dot{z}(t)] \in \mathbb{R}^2$

$$\dot{\mathbf{f}}(t) = \mathbf{A}\mathbf{f}(t) + \mathbf{B}u(t),\ \mathbf{f}(0) = \mathbf{0},\ \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2},\ \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{R}^2$$

Task-3: safety-critical control – continued

- With full state $\mathbf{f}(t) := [z(t); \dot{z}(t)] \in \mathbb{R}^2$

$$\dot{\mathbf{f}}(t) = \mathbf{A}\mathbf{f}(t) + \mathbf{B}u(t),\ \mathbf{f}(0) = \mathbf{0},\ \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 2},\ \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{R}^2$$

- The controlled trajectories $\mathbf{f}$ belong to a $\mathbb{R}^2$-valued RKHS with kernel

$$k(s,t) := \int_0^{\min(s,t)} e^{(s-\tau)\mathbf{A}}\mathbf{B}\mathbf{B}^\top e^{(t-\tau)\mathbf{A}^\top} \mathrm{d}\tau, \quad s,t \in \mathcal{T},$$

and the task is

$$\begin{aligned} &\min_{\mathbf{f}=[f_1;f_2]\in\mathcal{H}_k} \quad \|\mathbf{f}\|^2_{\mathcal{H}_k} \\ &\text{s.t.} \\ &z_{\text{low}}(t) \le f_1(t) \le z_{\text{up}}(t),\ \forall\, t \in \mathcal{T}. \end{aligned}$$

Task-3: safety-critical control – finished

- Assume for simplicity: z_{low} and z_{up} are piece-wise constant.
- Task:

$$\min_{\mathbf{f}=[f_1;f_2]\in\mathcal{H}_k} \quad \|\mathbf{f}\|^2_{\mathcal{H}_k}$$
$$\text{s.t.}$$
$$z_{\text{low},m} \le f_1(t) \le z_{\text{up},m},\ \forall\, t \in \mathcal{T}_m,\ \forall m \in [M].$$

Task-3: safety-critical control – finished

- Assume for simplicity: z_{low} and z_{up} are piece-wise constant.
- Task:

$$\min_{\mathbf{f}=[f_1;f_2]\in\mathcal{H}_k} \quad \|\mathbf{f}\|_{\mathcal{H}_k}^2$$

$$\text{s.t.}$$

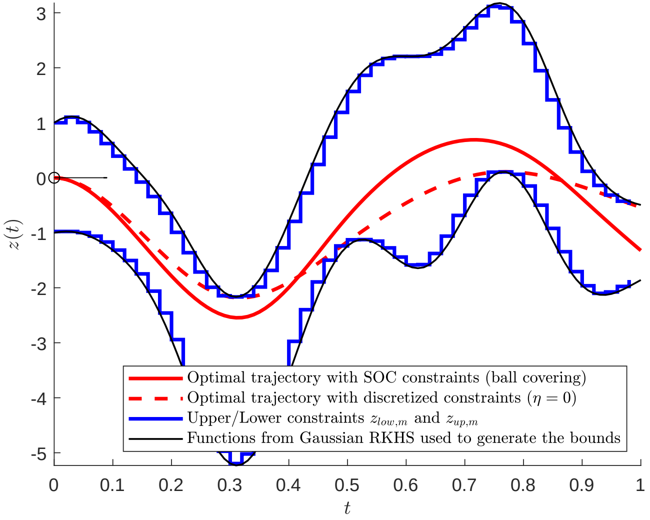
$$z_{\text{low},m} \le f_1(t) \le z_{\text{up},m}, \ \forall\, t \in \mathcal{T}_m, \forall m \in [M].$$

Constraints

linear transformation of functions (f_1), with matrix-valued kernel.

Demo (task-3): control of underwater vehicle

Vs discretization-based approach (which might crash):



Aït-Sahalia, Y. and Duarte, J. (2003).
Nonparametric option pricing under shape restrictions.
Journal of Econometrics, 116(1-2):9–47.

Aubin-Frankowski, P.-C., Petit, N., and Szabó, Z. (2020).
Kernel regression for vehicle trajectory reconstruction under speed and inter-vehicular distance constraints.
In *IFAC World Congress (IFAC WC)*, Berlin, Germany.

Aubin-Frankowski, P.-C. and Szabó, Z. (2020).
Hard shape-constrained kernel machines.
In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 384–395.

Aubin-Frankowski, P.-C. and Szabó, Z. (2022).
Handling hard affine SDP shape constraints in RKHSs.
Technical report.
(`http://arxiv.org/abs/2101.01519`; submitted to JMLR).

Blundell, R., Horowitz, J. L., and Parey, M. (2012).

Measuring the price responsiveness of gasoline demand: economic shape restrictions and nonparametric demand estimation.
Quantitative Economics, 3:29–51.

Chen, Y. and Samworth, R. J. (2016).
Generalized additive and index models with shape constraints.
Journal of the Royal Statistical Society – Statistical Methodology, Series B, 78(4):729–754.

Chetverikov, D., Santos, A., and Shaikh, A. M. (2018).
The econometrics of shape restrictions.
Annual Review of Economics, 10(1):31–63.

Deng, H. and Zhang, C.-H. (2020).
Isotonic regression in multi-dimensional spaces and graphs.
Annals of Statistics, 48(6):3672–3698.

Freyberger, J. and Reeves, B. (2018).
Inference under shape restrictions.
Technical report, University of Wisconsin-Madison.

(https://www.ssc.wisc.edu/~jfreyberger/Shape_Inference_Freyberger_Reeves.pdf).

Guntuboyina, A. and Sen, B. (2018).
Nonparametric shape-restricted regression.
Statistical Science, 33(4):568–594.

Han, Q. and Wellner, J. A. (2016).
Multivariate convex regression: global risk bounds and adaptation.
Technical report.
(https://arxiv.org/abs/1601.06844).

Johnson, A. L. and Jiang, D. R. (2018).
Shape constraints in economics and operations research.
Statistical Science, 33(4):527–546.

Keshavarz, A., Wang, Y., and Boyd, S. (2011).
Imputing a convex objective function.
In *IEEE Multi-Conference on Systems and Control*, pages 613–619.

Kur, G., Dagan, Y., and Rakhlin, A. (2020).
Optimality of maximum likelihood for log-concave density estimation and bounded convex regression.
Technical report.
(https://arxiv.org/abs/1903.05315).

Lewbel, A. (2010).
Shape-invariant demand functions.
The Review of Economics and Statistics, 92(3):549–556.

Lim, E. (2020).
The limiting behavior of isotonic and convex regression estimators when the model is misspecified.
Electronic Journal of Statistics, 14:2053–2097.

Matzkin, R. L. (1991).
Semiparametric estimation of monotone and concave utility functions for polychotomous choice models.
Econometrica, 59(5):1315–1327.

Nicol, F. (2013).
Functional principal component analysis of aircraft trajectories.

In *International Conference on Interdisciplinary Science for Innovative Air Traffic Management (ISIATM)*.

Sangnier, M., Fercoq, O., and d'Alché Buc, F. (2016).
Joint quantile regression in vector-valued RKHSs.
Advances in Neural Information Processing Systems (NIPS), pages 3693–3701.

Shapiro, A., Dentcheva, D., and Ruszczynski, A. (2014).
Lectures on Stochastic Programming: Modeling and Theory.
SIAM - Society for Industrial and Applied Mathematics.

Simchi-Levi, D., Chen, X., and Bramel, J. (2014).
The Logic of Logistics: Theory, Algorithms, and Applications for Logistics Management.
Springer.

Topkis, D. M. (1998).
Supermodularity and complementarity.
Princeton University Press.

Varian, H. R. (1984).
The nonparametric approach to production analysis.
Econometrica, 52(3):579–597.