A Variational approach to a Probability Function Estimation Problem Under Stochastic Ambiguity

Julio Deride

joint work with Johannes Royset and Fernanda Urrea

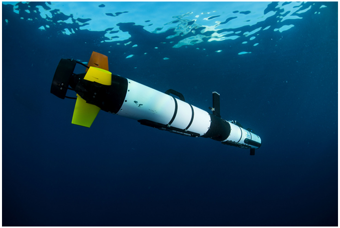
Universidad Técnica Federico Santa María
Departament of Mathematics

International School of Mathematics Guido Stampacchia
73rd Workshop: Robustness and resilience in stochastic optimization and statistical learning: Mathematical foundations

May 23, 2022

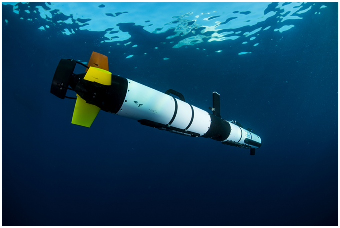
UNIVERSIDAD TECNICA FEDERICO SANTA MARIA
DEPARTAMENTO DE MATEMÁTICA

Motivational example: **U**nmanned **U**nderwater **V**ehicle



Consider an UUV returning to a docking station after a mission.

Motivational example: **U**nmanned **U**nderwater **V**ehicle

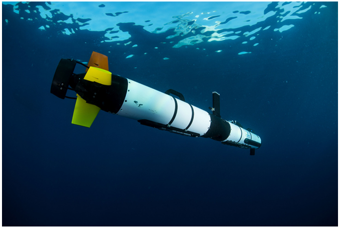


Consider an UUV returning to a docking station after a mission.

- UUV knows the location of the station on the map, but it has only a notion about its own location.

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

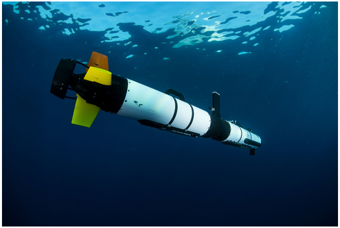
Motivational example: **U**nmanned **U**nderwater **V**ehicle



Consider an UUV returning to a docking station after a mission.

- UUV knows the location of the station on the map, but it has only a notion about its own location.
- The docking station sends out pings that can be picked up by the UUV when close enough.

Motivational example: **U**nmanned **U**nderwater **V**ehicle



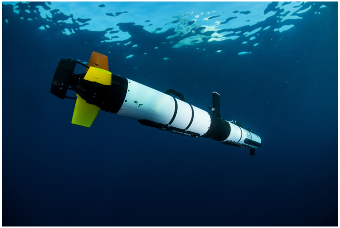
Consider an UUV returning to a docking station after a mission.

- UUV knows the location of the station on the map, but it has only a notion about its own location.
- The docking station sends out pings that can be picked up by the UUV when close enough.

The UUV uses these pings to improve the estimate of its own location.

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Motivational example: **U**nmanned **U**nderwater **V**ehicle



Consider an UUV returning to a docking station after a mission.

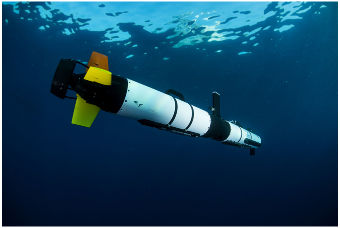
- UUV knows the location of the station on the map, but it has only a notion about its own location.
- The docking station sends out pings that can be picked up by the UUV when close enough.

The UUV uses these pings to improve the estimate of its own location.

- The UUV has an accurate model (in the short term) for its location, given its initial condition.

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Motivational example: **U**nmanned **U**nderwater **V**ehicle



Consider an UUV returning to a docking station after a mission.

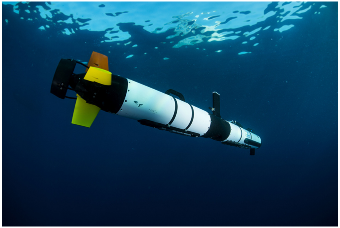
- UUV knows the location of the station on the map, but it has only a notion about its own location.
- The docking station sends out pings that can be picked up by the UUV when close enough.

The UUV uses these pings to improve the estimate of its own location.

- The UUV has an accurate model (in the short term) for its location, given its initial condition.

we are interested in modeling the position of the UUV

Motivational example: **U**nmanned **U**nderwater **V**ehicle



Consider an UUV returning to a docking station after a mission.

- UUV knows the location of the station on the map, but it has only a notion about its own location.
- The docking station sends out pings that can be picked up by the UUV when close enough.

The UUV uses these pings to improve the estimate of its own location.

- The UUV has an accurate model (in the short term) for its location, given its initial condition.

we are interested in modeling the position of the UUV

(note that GPS *does* **not** *work under deep water*)

UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

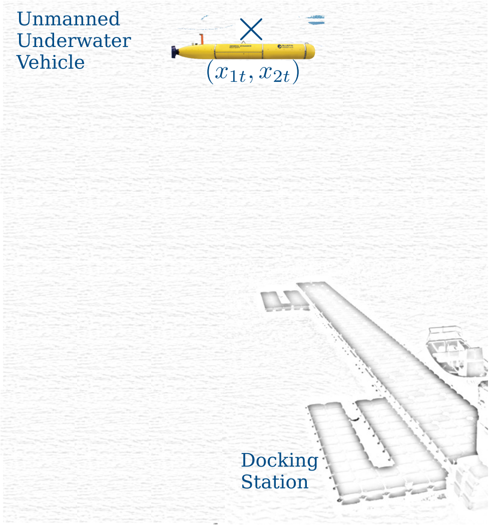


Figure: UUV scheme

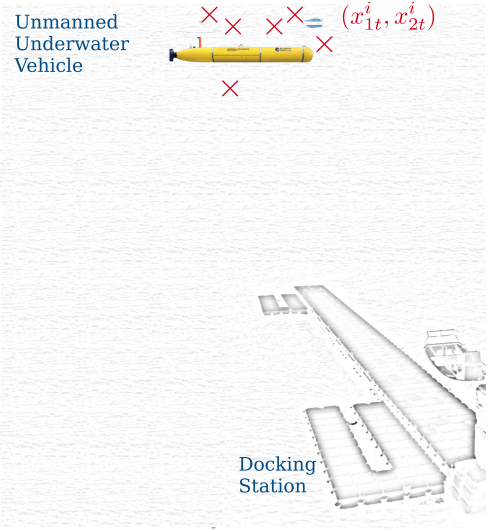


Figure: UUV scheme

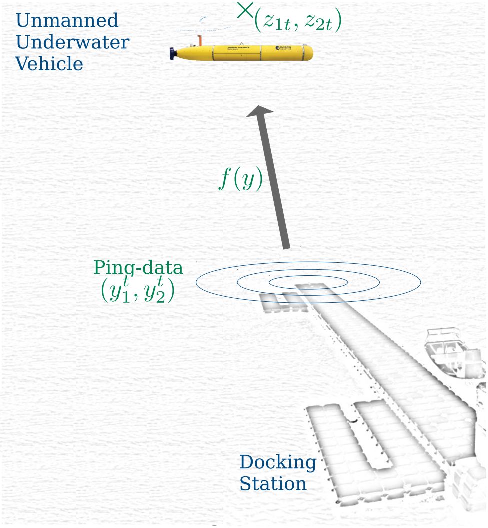


Figure: UUV scheme

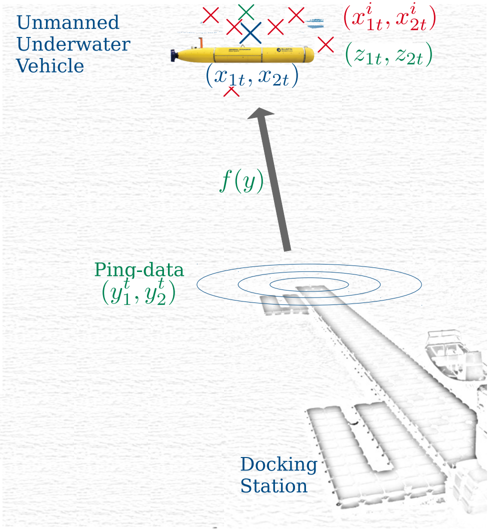


Figure: UUV scheme

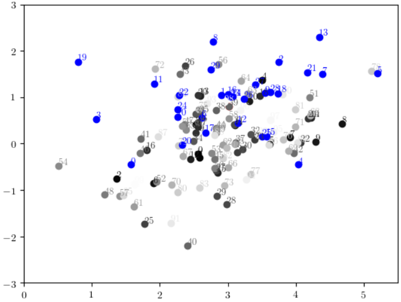
A Statistical Estimation Problem

Statistical estimation

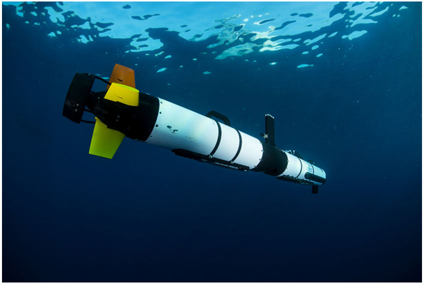
Estimation problems

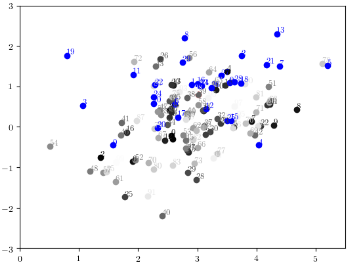
We consider an estimation problem as a decision making process over the probability distribution of a random variable, based on observed information, and, eventually, prior knowledge.

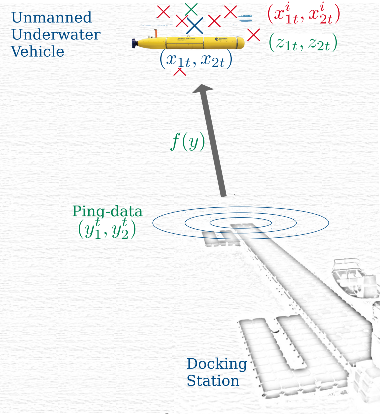
This can be seen as a STATISTICAL ESTIMATION PROBLEM



UNIVERSIDAD TECNICA FEDERICO SANTA MARIA







WHERE IS OUR UUV?

Mathematical program for the UUV position problem

Consider the following 2-dimensional model:

- Let $\{x_T^i\}$ be the sample for the position of the UUV
  Let F_T be the empirical CDF of the uuv location
  Let $\{y^1, \dots, y^T\}$ be the ping data (noisy).
- Let f be the funtion that models the location changes (Dubin's model) given the initial conditions,
  Propagate the position model from y as $z_T^T = y^T$, $z_T^{T-1} = f(y^{T-1})$,...,$z_T^1 = f^{(t)}(y^1)$.
  Let G_T be the empirical CDF of $\{z_T^t : t = 1, \dots, T\}$.

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Mathematical program for the UUV position problem

Consider the following 2-dimensional model:

- Let $\{x_T^i\}$ be the sample for the position of the UUV
  Let F_T be the empirical CDF of the uuv location
  Let $\{y^1, \ldots, y^T\}$ be the ping data (noisy).
- Let f be the funtion that models the location changes (Dubin's model) given the initial conditions,
  Propagate the position model from y as $z_T^T = y^T$, $z_T^{T-1} = f(y^{T-1})$,...,$z_T^1 = f^{(t)}(y^1)$.
  Let G_T be the empirical CDF of $\{z_T^t : t = 1, \ldots, T\}$.

Estimation problem

Find $\hat{F} \in \mathscr{F}$ such that

$$\hat{F} \in \operatorname*{argmin}_{F \in C \subset \mathscr{F}} \left\{ d\!l(F, F_T) \,\middle|\, d\!l(F, G_T) \leq \eta_T \right\}$$

,

Estimation with contextual estimation (1/2)

Following in the tradition of M-estimators,

find best estimate according to criterion (sq.error, likelihood, etc)

$$\hat{F} \in \operatorname*{argmin}_{F \in \mathscr{F}} \left\{ \rho(X, F) \,\middle|\, F \in C \right\},$$

where $X = (X_1, \ldots, X_n)$ are random vector

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Estimation with contextual estimation (1/2)

Following in the tradition of M-estimators,

find best estimate according to criterion (sq.error, likelihood, etc)

$$\hat{F} \in \operatorname*{argmin}_{F \in \mathscr{F}} \left\{ \rho(X,\, F) \,\middle|\, F \in C \right\},$$

where $X = (X_1, \ldots, X_n)$ are random vector

We are addressing two main challenges:

general constraints (C) and abstract spaces ($\mathscr{F}$)

In the past, non parametric statistics has dealt with specific constraints (log-concave or monotone distributions)

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Estimation with contextual estimation (1/2)

Following in the tradition of M-estimators,

find best estimate according to criterion (sq.error, likelihood, etc)

$$\hat{F} \in \operatorname*{argmin}_{F \in \mathscr{F}} \left\{ \rho(X,\ F) \,\middle|\, F \in C \right\},$$

where $X = (X_1, \ldots, X_n)$ are random vector

We are addressing two main challenges:

general constraints (C) and abstract spaces ($\mathscr{F}$)

In the past, non parametric statistics has dealt with specific constraints (log-concave or monotone distributions)

Applications:

Least squares	$\rho(X, y, F) = \|y - F(X)\|^2$
Maximum likelihood	$\rho(X, F) = -\log(F(X))$
Support vector machine	$\rho(X, y, F) = \max\{0, 1 - yF'(X)\}$

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Estimation with contextual estimation (2/2)

We are considering the original constrained M-estimator problem

$$\hat{F} \in \operatorname*{argmin}_{F \in C \subset \mathscr{F}} \rho(X, F),$$

with F in a general metric space (abstract representations, models)

UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Estimation with contextual estimation (2/2)

We are considering the original constrained and robust M-estimator problem

$$\hat{F} \in \operatorname*{argmin}_{F \in C \subset \mathscr{F}} \sup_{G \in \mathscr{G}(F)} \rho(X, F, G),$$

with F in a general metric space (abstract representations, models)

Estimation with contextual estimation (2/2)

We are considering the original constrained and robust M-estimator problem

$$\hat{F} \in \operatorname*{argmin}_{F \in C \subset \mathscr{F}} \sup_{G \in \mathscr{G}(F)} \rho(X, F, G),$$

with F in a general metric space (abstract representations, models)

Examining

existence, approximation, consistency, formulations...

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Estimation with contextual estimation (2/2)

We are considering the original constrained and robust M-estimator problem

$$\hat{F} \in \operatorname*{argmin}_{F \in C \subset \mathscr{F}} \sup_{G \in \mathscr{G}(F)} \rho(X, F, G),$$

with F in a general metric space (abstract representations, models)

Examining

existence, approximation, consistency, formulations...

Approximation?

$$\hat{F}^\nu \in \varepsilon^\nu - \operatorname*{argmin}_{F \in C^\nu \subset \mathscr{F}^\nu} \sup_{G \in \mathscr{G}^\nu(F)} \rho^\nu(X, F, G)$$

when $\nu \to \infty$, $(\varepsilon^\nu \to 0)$, $\hat{F}^\nu \to \hat{F}$?

Lopsided convergence [Royset, Wets 2017]

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Space selection and Approximation

Estimation problem

Find $\hat{F} \in \mathscr{F}$ such that

$$\hat{F} \in \operatorname*{argmin}_{F \in C \subset \mathscr{F}} \left\{ d\!\!l(F, F_T) \,\middle|\, d\!\!l(F, G_T) \leq \eta_T \right\}$$

First step: Space selection $\mathscr{F}$

We need a class functions that is

- flexible, but avoid overfitting and high errors
- simple, but able to identify key characteristics
- incorporate soft (auxiliary) information and assumptions
- computationally tractable
- facilitate analysis

First step: Space selection $\mathscr{F}$

We need a class functions that is

- flexible, but avoid overfitting and high errors
- simple, but able to identify key characteristics
- incorporate soft (auxiliary) information and assumptions
- computationally tractable
- facilitate analysis

Possibilities:

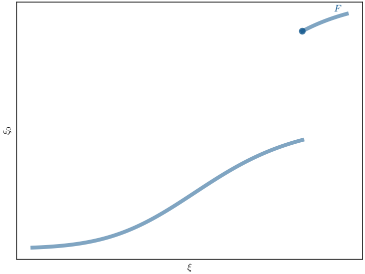
- $F(x) = \langle \alpha, x \rangle + \beta$, (affine)
- $F(x) = \sum c_i \varphi_i(x)$ (kernel)
- L_p, Sobolev, ...
- parametric (finite-dim); nonparametric (∞-dim)

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Upper semi-continuous functions (usc)

Let $\mathscr{F}$ be the space of upper-semi continuous functions, nondecreasing, $[l_1, u_1] \times [l_2, u_2] \subset I\!R^2 \to [0, 1]$.

$$F \text{ usc} \iff \forall x \limsup_{x^\nu \to x} F(x^\nu) \leq F(x)$$



UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Second step: Topology

Why is it important?

Second step: Topology

Why is it important?

We need a precise notion of **proximity** for approximation.

Second step: Topology

Why is it important?

We need a precise notion of **proximity** for approximation.

We base our approach on set convergence.

Second step: Topology

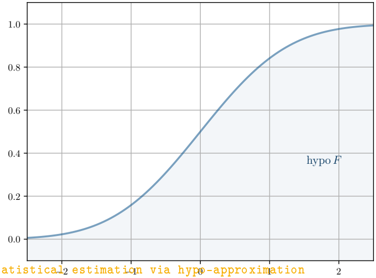
Why is it important?

We need a precise notion of **proximity** for approximation.

We base our approach on set convergence.

Define the hypograph of a function $F : \mathbb{R}^2 \to \mathbb{R}$ as

$$\mathrm{hypo}(F) = \{(x, \alpha) : \alpha \le F(x), x \in \mathbb{R}^d, \alpha \in \mathbb{R}\}.$$



Hypo-Convergence (Attouch-Wets on hypographs)

Hypo-Convergence

$\boxed{\{F^\nu\} \text{ hypo-converges to } F}$ iif $\boxed{\begin{array}{c}\text{hypo}(F^\nu) \text{ converges to hypo}(F)\\ \text{as sets (Painlevé-Kuratowski)}\end{array}}$

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Hypo-Convergence (Attouch-Wets on hypographs)

Hypo-Convergence

$\{F^\nu\}$ hypo-converges to F iif hypo(F^ν) converges to hypo(F) as sets (Painlevé-Kuratowski)

Prop: Weak convergence compatibility

$\{F^\nu\}$ hypo-converges to F then* $\{F^\nu\}$ converges weakly to F

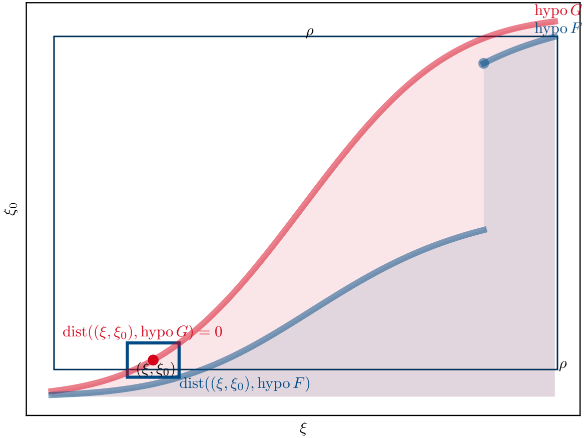
Prop: A metrizable topology

The hypo-convergence can be metrized by the Attouch-Wets distance $d\!l$

For closed and bounded $C \subset \mathscr{F}$, $(C, d\!l)$ is a compact metric space

Attouch-Wets distance in the space of usc-cdf

$$dl(F,G) := \int_0^\infty dl_\rho(F,G)e^{-\rho}\,d\rho, \qquad dl_\rho(F,G) := \max_{\|\xi\|_\infty \le \rho} \{|\text{dist}(\xi, \text{hypo}F) - \text{dist}(\xi, \text{hypo}G)|\}$$



UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Properties

- $(\text{usc}-\text{fcns}(\mathbb{R}^m), d\!l^h)$ is a complete separable metric space.
- $\{g^\nu\} \subseteq \text{usc}-\text{fcns}(\mathbb{R}^m)$ hypo-converge to $g \in \text{usc}-\text{fcns}(\mathbb{R}^m)$ if $d\!l^h(g^\nu, g) \to 0$. Equivalently

$$g^\nu \xrightarrow{h} g \iff \begin{cases} \forall \xi^\nu \to \xi,\ \limsup g^\nu(\xi^\nu) \leq g(\xi) \\ \forall \xi, \exists \xi^\nu \to \xi,\ \text{ such that } \liminf g^\nu(\xi^\nu) \geq g(\xi) \end{cases}$$

- $\text{usc}-\text{fcns}(\mathbb{R}^m)$ is not a linear space, but it is a pointed cone.

A cone $K \subset \mathbb{R}^n$ is *pointed* if it coitains no lines, i.e., if $x \in K$ and $-x \in K$ then $x = 0$.

Hypo-convergence of Empirical distribution

If $\{\xi^\nu\}$ iid rv with distribution F, then

$$F^\nu(\cdot, \omega) = \frac{1}{\nu}\sum_{j=1}^{\nu} I_{\{\xi^j(\omega)\leq \cdot\}} \xrightarrow{h} F \text{ a.s.}$$

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Weak-limit is a distribution?... Tightness

Tightness

A subset $\mathcal{S} \subset \mathscr{D}$ is **tight** of for all $\varepsilon > 0$, there exists a rectangle A such that $\Delta_A F \geq 1 - \varepsilon$ for all $F \in \mathcal{S}$

Convergence to distribution functions

If $\{F^\nu\} \subset \mathscr{D}$ is tight then

- There exists a $\{F^{\nu_k}\}$, F such that $dl^h(F^{\nu_k}, F) \to 0$, as $k \to \infty$.
- If $F : \mathbb{R}^m \to \mathbb{R}$ is the hypo-limit of $\{F^\nu\}$, then $F \in \mathscr{D}$.

Compactness and Tightness

For $\mathcal{S} \subset \mathscr{D}$

- if $\mathcal{S}$ compact then $\mathcal{S}$ tight.
- If $\mathcal{S}$ is tight, then $cl\mathcal{S}$ is compact and contained in $\mathscr{D}$.

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Constraint examples: Moments

Convexity under moment information

The ambiguity set with constrained moments

$$\mathscr{F}(x) = \left\{ F \in \mathscr{D} : \int \xi^k \, dF(\xi) = a_k(x),\ k = 1, \ldots, K \right\}$$

is convex due to linearity of the integral.

Constraint examples: Ambiguity sets

Ambiguity sets under moment information

Given $\mu_1 < \mu_2$ and $s > 0$, the *ambiguity set of distributions*

$$\mathscr{F}_{\mu_1,\mu_2,s} = \left\{ F \in \mathscr{D} : F \text{ has mean in } [\mu_1, \mu_2], \text{ and stdv in } [0, s] \right\}$$

is tight and for every $r > 0$, $\exists \nu_r$ and $F^1, \ldots, F^{\nu_r}$ such that

$$\{F^1, \ldots, F^{\nu_r}\} \subset \mathscr{F}_{\mu_1,\mu_2,s} \subset \bigcup_{\nu=1}^{\nu_r} \mathbb{B}(F^\nu, r).$$

Proof. Thightnes is given by Chebyshev's inequality. This implies compactness.

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

A Robust CDF Estimation Problem

(getting back to the UUV position estimation)

Back to our CDF estimation problem

Estimation problem

Find $\hat{F} \in \mathscr{F}$ such that

$$\hat{F} \in \underset{F \in C \subset \mathscr{F}}{\operatorname{argmin}} \left\{ d\!l(F, F_T) \,\middle|\, d\!l(F, G_T) \leq \eta_T \right\}$$

Back to our CDF estimation problem

Estimation problem

Find $\hat{F} \in \mathscr{F}$ such that

$$\hat{F} \in \operatorname*{argmin}_{F \in C \subset \mathscr{F}} \left\{ dl(F, F_T) \,\middle|\, dl(F, G_T) \leq \eta_T \right\}$$

Some of the difficulties for solving this problem

- Selection of $\mathscr{F}$, continuous? usc
- Set of constraints C to be considered (maintaining **closedness**)
- Explicit computation of dl is hard

$$dl(F, G) := \int_0^\infty dl_\rho(F, G) e^{-\rho} \, d\rho, \qquad dl_\rho(F, G) := \max_{\|\xi\|_S \leq \rho} \{ |d(\xi, h_F) - d(\xi, h_G)| \}$$

Approximation?

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Back to our CDF estimation problem

ESTIMATION PROBLEM

Find $\hat{F} \in \mathscr{F}$ such that

$$\hat{F} \in \underset{F \in C \subset \mathscr{F}}{\operatorname{argmin}} \left\{ d\!l(F, F_T) \,\middle|\, d\!l(F, G_T) \le \eta_T \right\}$$

Some of the difficulties for solving this problem

- Selection of $\mathscr{F}$, continuous? usc ← epi-splines (1)
- Set of constraints C to be considered (maintaining **closedness**) ← so far, $C = \mathscr{F}$ (2)
- Explicit computation of $d\!l$ is hard ← use approximation $\hat{d\!l}$ (3)

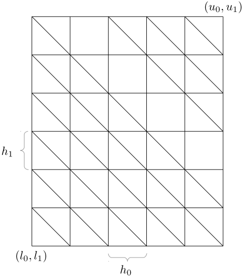
$$d\!l(F, G) := \int_0^\infty d\!l_\rho(F, G) e^{-\rho} \, d\rho, \qquad d\!l_\rho(F, G) := \max_{\|\xi\|_S \le \rho} \{ |\mathrm{d}(\xi, \mathrm{h}_F) - \mathrm{d}(\xi, \mathrm{h}_G)| \}$$

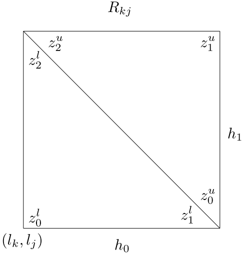
Approximation?

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Approximation ①: Epi-splines

- $R = [l, u] = [l_0, u_0] \times [l_1, u_1]$: bounded rectangular domain
- $\mathscr{F}^\nu$: 1-degree epi-splines (piecewise affine functions) over a triangular partition of the domain.
- $R^\nu = \{R_{kj}\}_{11}^{N_0^\nu N_1^\nu}$, each R_{kj} divided into two triangles,
- Each function $F \in \mathscr{F}^\nu$ is represented by the values that it takes on vertices of each triangular region (z_0^l, z_1^l, z_2^l) and (z_0^u, z_1^u, z_2^u)





Approximation ②: Constraints

We can consider constraints such as

- Distribution
- Concavity, Monotonicity
- Continuity, Lipschitz
- Pointwise bounds
- Stochastic dominance (under study)
- (central) Moments bounds

$\boxed{\text{this family of constraints is closed under } (\mathscr{F}, d\!l)}$

Approximation ②: Constraints

We can consider constraints such as

- Distribution
- Concavity, Monotonicity
- Continuity, Lipschitz
- Pointwise bounds
- Stochastic dominance (under study)
- (central) Moments bounds

$\boxed{\text{this family of constraints is closed under } (\mathscr{F}, d\!l)}$

Thus,

$$\boxed{\begin{array}{c} C \subset \mathscr{F} \text{ is compact, and therefore} \\ \operatorname*{argmin}_{F \in C} d\!l(F, F_0) \neq \emptyset \end{array}}$$

← Existence of solution guaranteed

UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Approximation ③: Distance bounds

Let F be a CDF over R, partitioned into sub-rectangular elements R_k

Bounds for the hypo-distance [Royset,2019]

For $\rho \geq 1$, and multivariate CDF F and G defined over R

$$d\!l(F,G) \leq \underbrace{e^{-\rho} + \inf \left\{ \eta \geq 0 \;\middle|\; \begin{array}{l} G\left(l^k + \eta \mathbf{1}\right) + \eta \geq F(u^k) \\ F\left(l^k + \eta \mathbf{1}\right) + \eta \geq G(u^k),\ \forall k = 1, \dots, N \end{array} \right\}}_{d^{\nu}}$$

where (l^k, u^k), $k = 1, \dots, N$ are points in the box $4\rho I\!B^{\infty}$.

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Approximation ③: Distance bounds

Let F be a CDF over R, partitioned into sub-rectangular elements R_k

Bounds for the hypo-distance [Royset,2019]

For $\rho \geq 1$, and multivariate CDF F and G defined over R

$$dl(F,G) \leq \underbrace{e^{-\rho} + \inf \left\{ \eta \geq 0 \;\middle|\; \begin{array}{l} G\left(l^k + \eta\mathbf{1}\right) + \eta \geq F(u^k) \\ F\left(l^k + \eta\mathbf{1}\right) + \eta \geq G(u^k),\ \forall k = 1,\dots,N \end{array} \right\}}_{dl^\nu}$$

where (l^k, u^k), $k = 1, \dots, N$ are points in the box $4\rho\mathbb{B}^\infty$.

Discrete space approximation

We embed our problem onto the usc functions, endowed with the hypo-topology, i.e.,

$$(\mathscr{F} = \text{usc} - \text{fcn}(\mathbb{R}^d),\ dl) \leftarrow (\mathscr{F}^\nu, dl^\nu)$$

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Approximation - Convergence

True problem

$$\min_{F \in C \subset \mathscr{F}} \{ d\!l(F, F_T) \,|\, d\!l(F, G_T) \le \eta_T \}$$

$\Uparrow$ hypo

Approximate problems

$$\min_{F \in C^\nu \subset \mathscr{F}^\nu} \left\{ \eta \ge 0 \;\middle|\; \begin{aligned} F\left(l^k + \eta \mathbf{1}\right) + \eta &\ge F_T(u^k) \\ F_T\left(l^k + \eta \mathbf{1}\right) + \eta &\ge F(u^k) \\ F\left(l^k + \eta_T \mathbf{1}\right) + \eta_T &\ge G_T(u^k) \\ G_T\left(l^k + \eta_T \mathbf{1}\right) + \eta_T &\ge F(u^k),\ \forall k = 1, \ldots, N \end{aligned} \right\}$$

Also lower bound; computation by bisection search

And modeling of contextual information (continuity, Lipschitz)

UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Computational advantages/challenges

Summarizing, the proposed approximation has the following features:

- We solve a sequence of linear problems
- but the **size of the problem** depends on the **size of the mesh**
- still, we can **solve them efficiently** by using state-of-the-art solvers
- and, more important, we can **add constraints** to the problem formulation.
- Notice that we are still constrained on the dimensionality of the distribution (work in progress)

Numerical examples

Algorithm: Bisection method

UUV Estimation problem

$$\hat{F} \in \operatorname*{argmin}_{F \in C \subset \mathscr{F}} \left\{ d\!l(F, F_T) \,\middle|\, d\!l(F, G_T) \leq \eta_T \right\}$$

- Set $\eta_l = 0, \eta_u = \eta_0$ and $0 < \eta_T < 1$.
- While: $s < \varepsilon_1$ or $|\eta_l - \eta_u| > \varepsilon_2$
  - Set $\eta = \frac{\eta_l + \eta_u}{2}$ and solve

$$\max_{s \geq 0, F} \quad s \tag{1}$$

$$\text{s.t} \quad F\left(x_{lj}^{u,2}\right) \leq F_0\left(x_{lj}^{l,1} + \eta \mathbf{1}\right) + \eta + s \tag{2}$$

$$F_0\left(x_{lj}^{l,2}\right) \leq F\left(x_{lj} + \eta \mathbf{1}\right) + \eta + s \tag{3}$$

$$F\left(x_{lj}^{u,2}\right) \leq G_0\left(x_{lj}^{l,1} + \eta_T \mathbf{1}\right) + \eta_T + s \tag{4}$$

$$G_0\left(x_{lj}^{l,2}\right) \leq F\left(x_{lj} + \eta_T \mathbf{1}\right) + \eta_T + s \tag{5}$$

$$F \in C \tag{6}$$

  - If the problem is infeasible, set $\eta_l = \eta$, else set $\eta_u = \eta$

UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Two uniform distributions

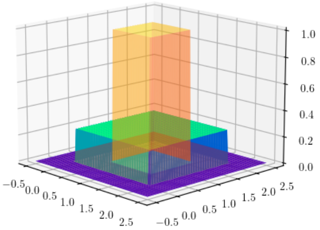


Figure: $F_T \sim U[0,2] \times [0,2]$ (blue)
$G_T \sim U[0.5,1.5] \times [0.5,1.5]$ (orange)

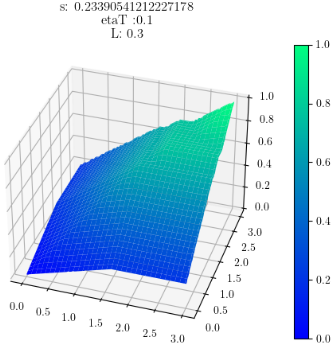


Figure: Solution with Lip = 0.3

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Two uniform distributions

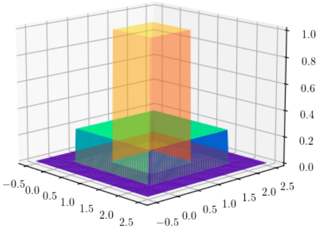


Figure: $F_T \sim U[0, 2] \times [0, 2]$ (blue)
$G_T \sim U[0.5, 1.5] \times [0.5, 1.5]$ (orange)

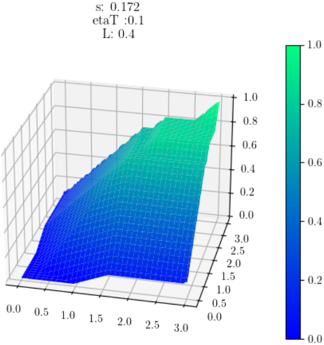


Figure: Solution with Lip $= 0.4$

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Two uniform distributions

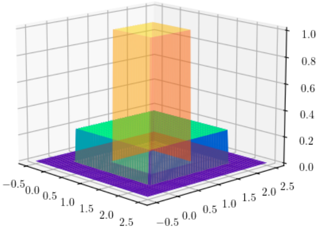


Figure: $F_T \sim U[0,2] \times [0,2]$ (blue) $G_T \sim U[0.5, 1.5] \times [0.5, 1.5]$ (orange)

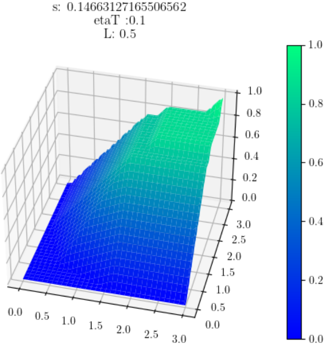


Figure: Solution with Lip = 0.5

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Where is our UUV?

UUV: Sample and estimated position $(\{x_i^T\}_i, \{z_j^T\}_j)$

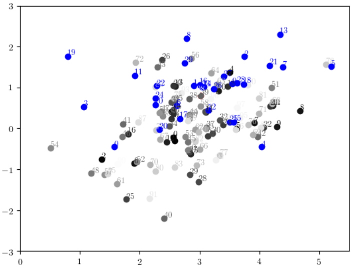


Figure: Sample data ($\{x_i^T\}$ (F_T) and estimated data $\{z_j^T\}$) (G_T)

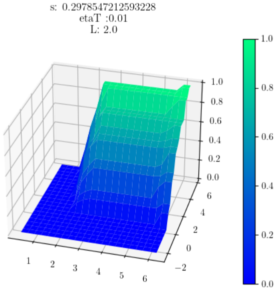


Figure: Solution $\hat{\mathbf{F}}$, $\eta_T = 0.01$

UUV: Sample and estimated position ($\{x_i^T\}_i$, $\{z_j^T\}_j$)

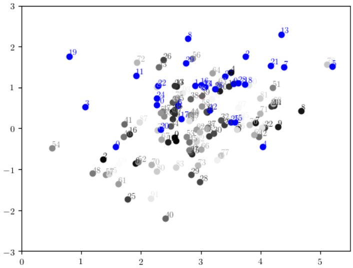


Figure: Sample data ($\{x_i^T\}$ (F_T) and estimated data $\{z_j^T\}$) (G_T)

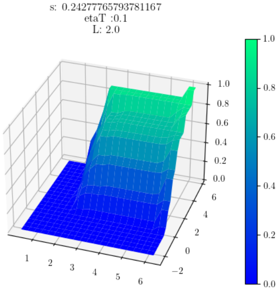


Figure: Solution $\hat{\mathbf{F}}$, $\eta_T = 0.1$

UUV: Sample and estimated position $(\{x_i^T\}_i, \{z_j^T\}_j)$

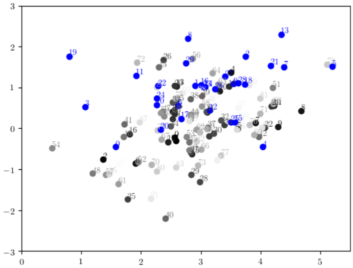


Figure: Sample data $(\{x_i^T\}$ (F_T) and estimated data $\{z_j^T\})$ (G_T)

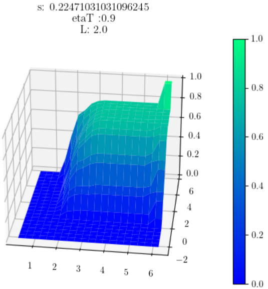


Figure: Solution $\hat{\mathbf{F}}$, $\eta_T = 0.9$

The end

Thank you for your attention

Questions?

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

References

- Salinetti, G., Wets, R.J-B *On the convergence in distribution of measurable multifunctions (random sets) normal integrands, stochastic processes and stochastic infima*, Mathematics of Operations Research, 11(3):385–419, 1986.
- Royset, J., *Approximations of Semicontinuous Functions with Applications to Stochastic Optimization and Statistical Estimation*, 2019, Mathematical Programming.
- Royset, J., Wets, R.J-B, *Variational analysis of constrained M-estimators*, 2020, *Annals of Statistics*, Vol. 48, No. 5, pp. 2759-2790.
- Royset, J., Wets, R. J-B, *Variational theory for optimization under stochastic ambiguity*, SIAM Journal on Optimization, 27(2), 1118-1149, 2017.

More info

https://sites.google.com/view/deride-home/home

UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

A Variational approach to a Probability Function Estimation Problem Under Stochastic Ambiguity

Julio Deride

joint work with Johannes Royset and Fernanda Urrea

Universidad Técnica Federico Santa María
Departament of Mathematics

International School of Mathematics Guido Stampacchia
73rd Workshop: Robustness and resilience in stochastic optimization and statistical learning: Mathematical foundations

May 23, 2022

EX UMBRA IN SOLEM
UNIVERSIDAD TECNICA FEDERICO SANTA MARIA
DEPARTAMENTO DE MATEMÁTICA