Stochastic Variational Analysis

Roger J-B Wets
Mathematics, University of California, Davis

May 2012, ETH-Zürich

Why?

$G : E \to \mathbb{R}^d, \quad G^{-1}(0)$ soln's of $G(x) = 0$, approximations?

Why?

$G : E \to \mathbb{R}^d, \quad G^{-1}(0)$ soln's of $G(x) = 0$, approximations?

$$EG(x) = \mathbb{E}\{G(\boldsymbol{\xi}, x)\} = 0 \ \text{ “approximated” by } G^\nu(x) = 0$$
$$\xi^1, \ldots, \xi^\nu \text{ sample, } \ G^\nu(x) = \tfrac{1}{\nu} \textstyle\sum_{l=1}^{\nu} G(\xi^l, x)$$

Why?

$G : E \to \mathbb{R}^d, \quad G^{-1}(0)$ soln's of $G(x) = 0$, approximations?

$$EG(x) = \mathbb{E}\{G(\boldsymbol{\xi}, x)\} = 0 \text{ "approximated" by } G^\nu(x) = 0$$
$$\xi^1, \ldots, \xi^\nu \text{ sample, } G^\nu(x) = \tfrac{1}{\nu} \textstyle\sum_{l=1}^{\nu} G(\xi^l, x)$$

$$G : \Xi \times D \rightrightarrows E, \text{ set-valued } G(\xi, x) \subset E, \text{ inclusion } \mathbb{E}\{G(\boldsymbol{\xi}, x)\} \ni 0$$
$$\xi^1, \ldots, \xi^\nu \text{ sample, approximation } \tfrac{1}{\nu} \textstyle\sum_{l=1}^{\nu} G(\xi^l, x) \ni 0$$

Why?

$G : E \to \mathbb{R}^d, \quad G^{-1}(0)$ soln's of $G(x) = 0$, approximations?

$EG(x) = \mathbb{E}\{G(\boldsymbol{\xi}, x)\} = 0$ "approximated" by $G^\nu(x) = 0$
$\xi^1, \dots, \xi^\nu$ sample, $G^\nu(x) = \frac{1}{\nu}\sum_{l=1}^{\nu} G(\xi^l, x)$

$G : \Xi \times D \rightrightarrows E$, set-valued $G(\xi, x) \subset E$, inclusion $\mathbb{E}\{G(\boldsymbol{\xi}, x)\} \ni 0$
$\xi^1, \dots, \xi^\nu$ sample, approximation $\frac{1}{\nu}\sum_{l=1}^{\nu} G(\xi^l, x) \ni 0$

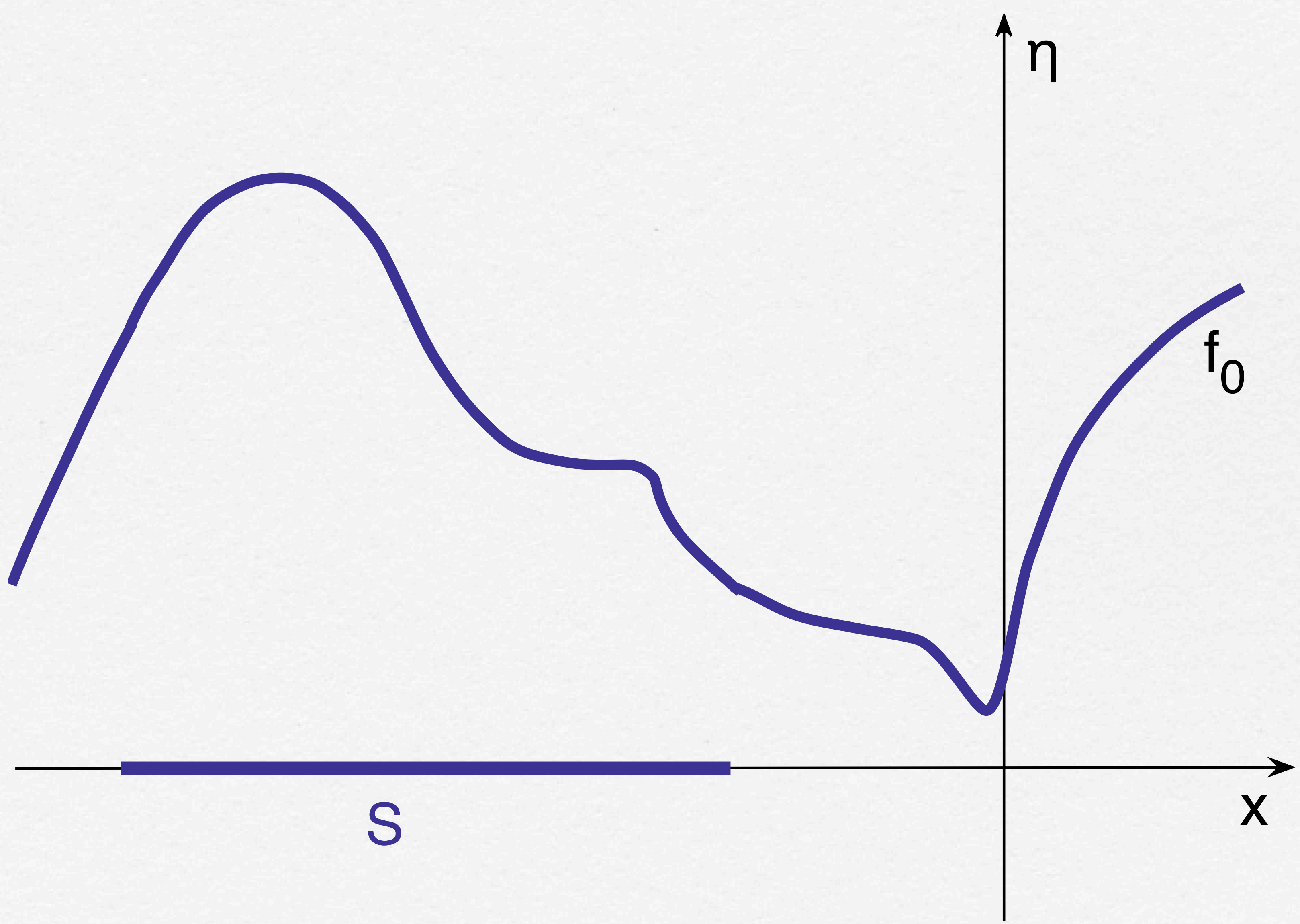
$\min \mathbb{E}\{f(\boldsymbol{\xi}, x)\},\ x \in C, \quad \mathbb{E}\{f(\boldsymbol{\xi}, x)\} = Ef(x) = \int_\Xi f(\xi, x)\, P(d\xi)$
$\xi^1, \dots \xi^\nu$ sample P^ν (random) empirical measure
approx.: $\min \mathbb{E}^\nu\{f(\boldsymbol{\xi}, x)\} = \frac{1}{\nu}\sum_{l=1}^{\nu} f(\xi^l, x),\ x \in C$

Preliminaries (unavoidable)

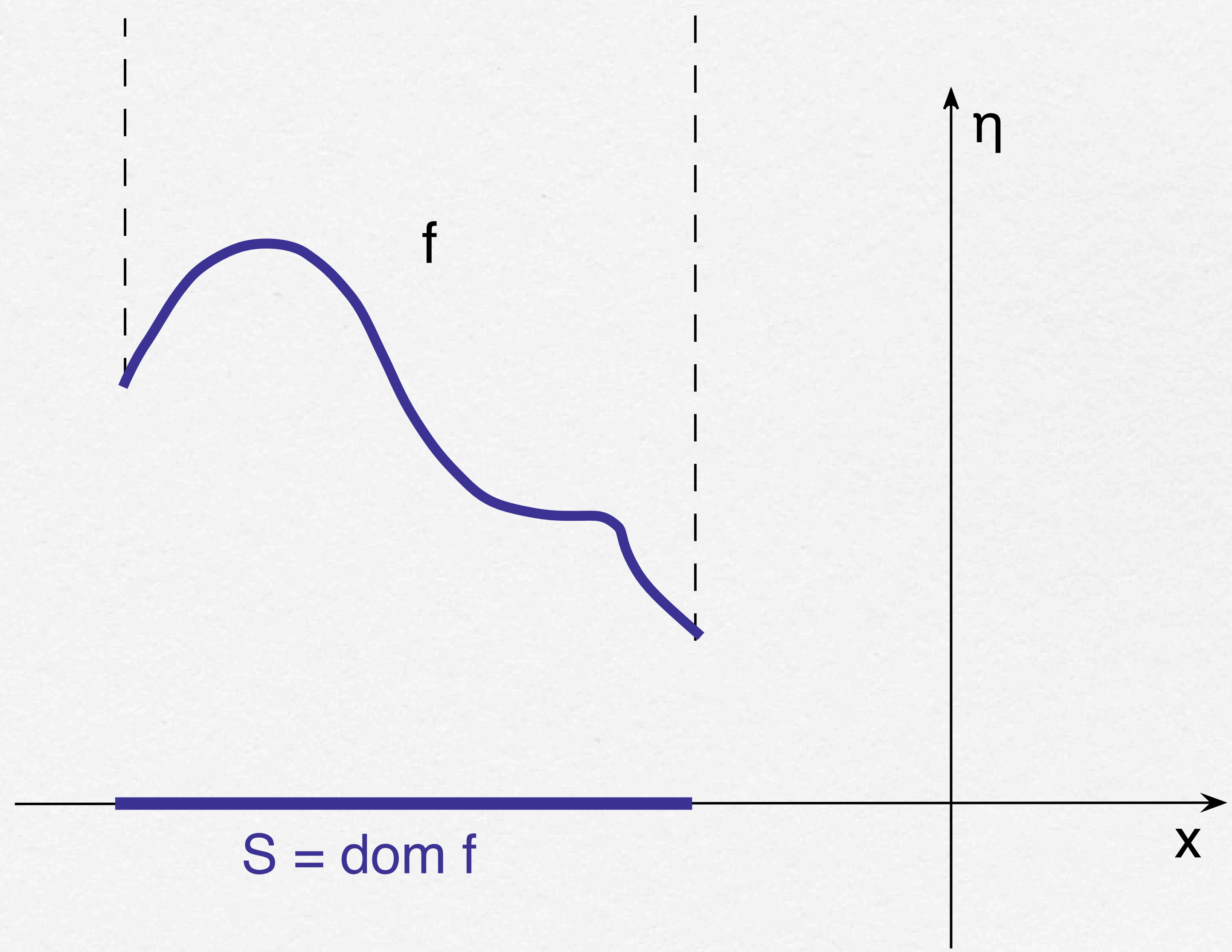
Optimization problem

$$\min f_0(x),\ \ x \in S,$$

$$S = \left\{ x \in \mathbb{R}^n \middle| f_i(x) \le 0,\ i = 1 \to s,\ \ f_i(x) = 0,\ i = s+1 \to m \right\}$$

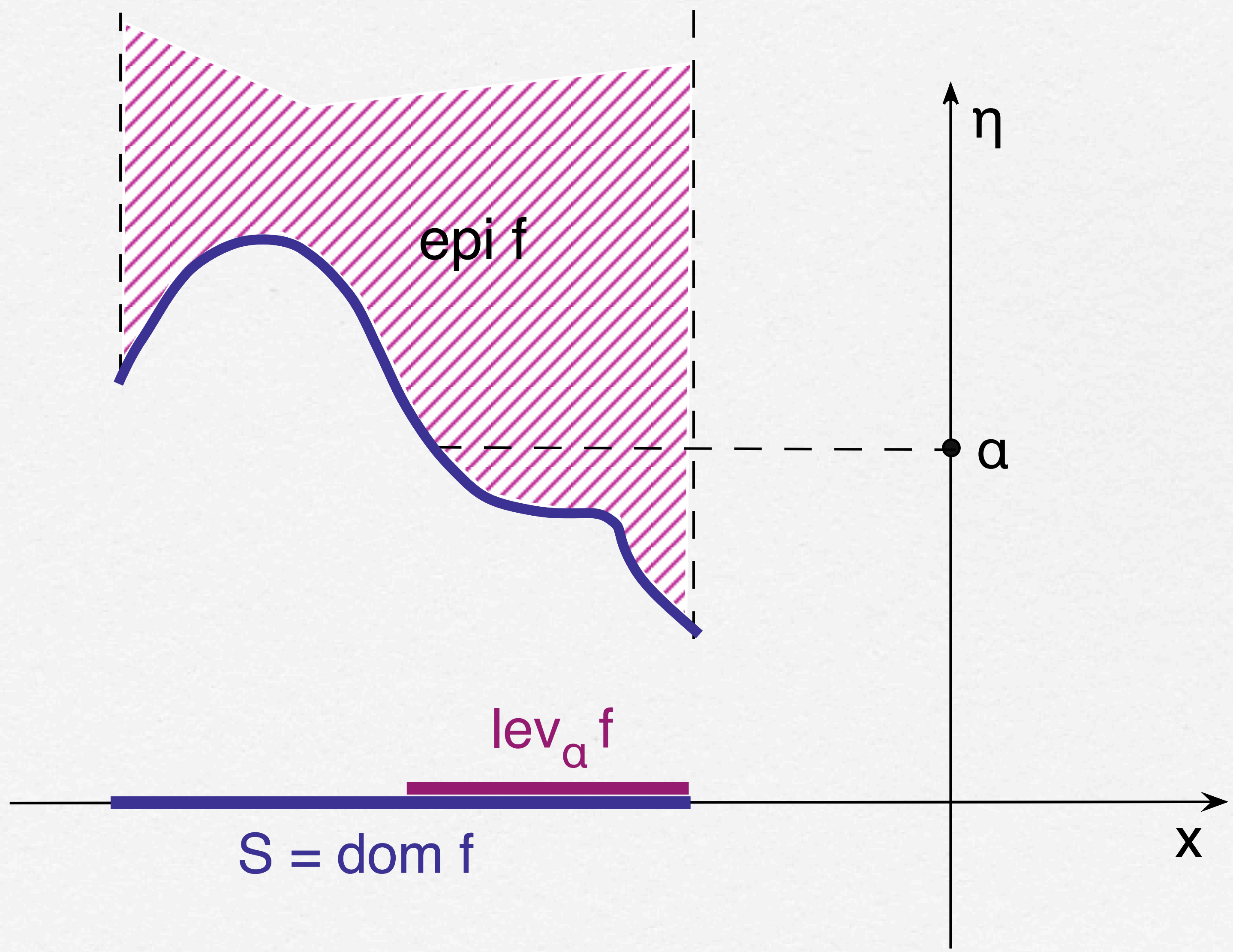


$$\min f \text{ on } E, \quad f = f_0 + \iota_S(x), \quad \iota_S \text{ indicator function of } S$$



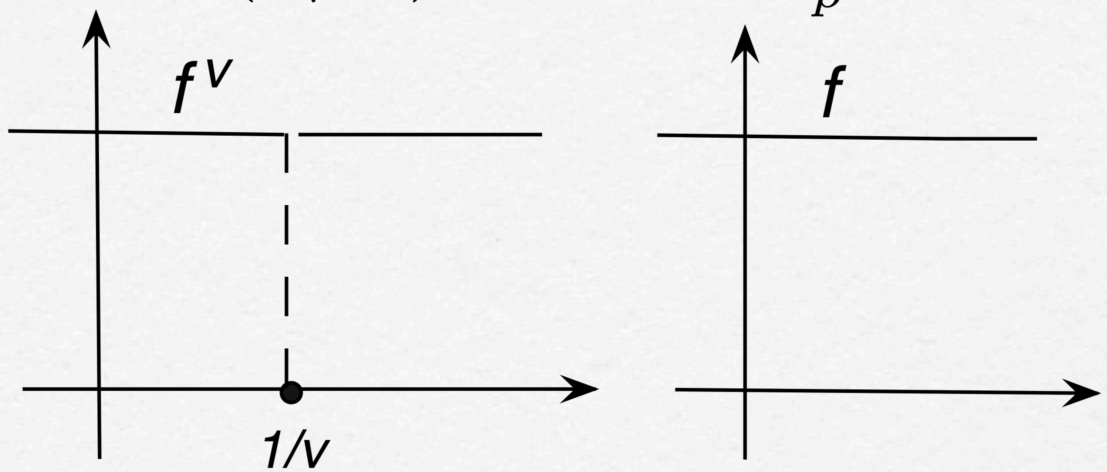
$$\min f \text{ on } E, \quad f = f_0 + \iota_S(x), \quad \iota_S \text{ indicator function of } S$$

$$\operatorname{epi} f = \{(x,\alpha) \in E \times R \,|\, f(x) \le \alpha\}, \quad \operatorname{lev}_\alpha f = \{x \in E \,|\, f(x) \le \alpha\}$$



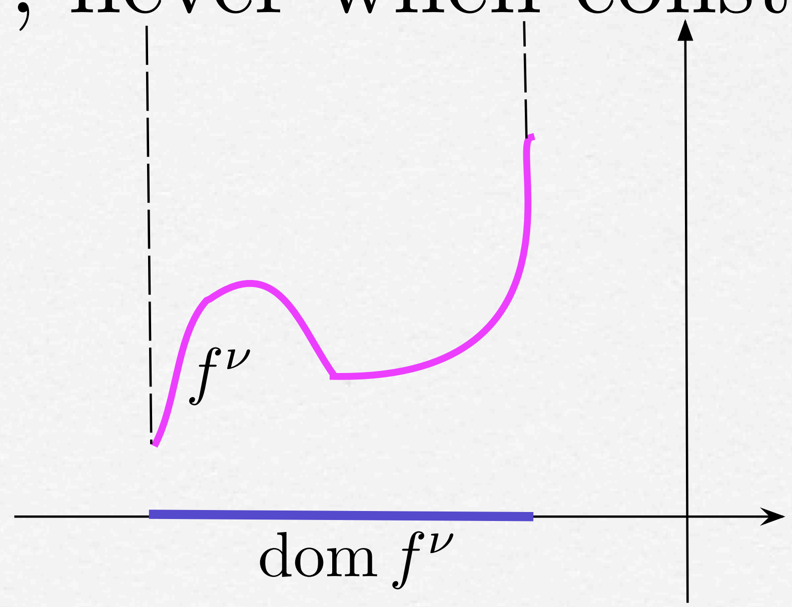
1. pointwise convergence $\not\Rightarrow$ convergence of minimizers

   $f^\nu \equiv 1$ except $f(1/\nu) = 0$, $f^\nu \underset{p}{\to} f \equiv 1$



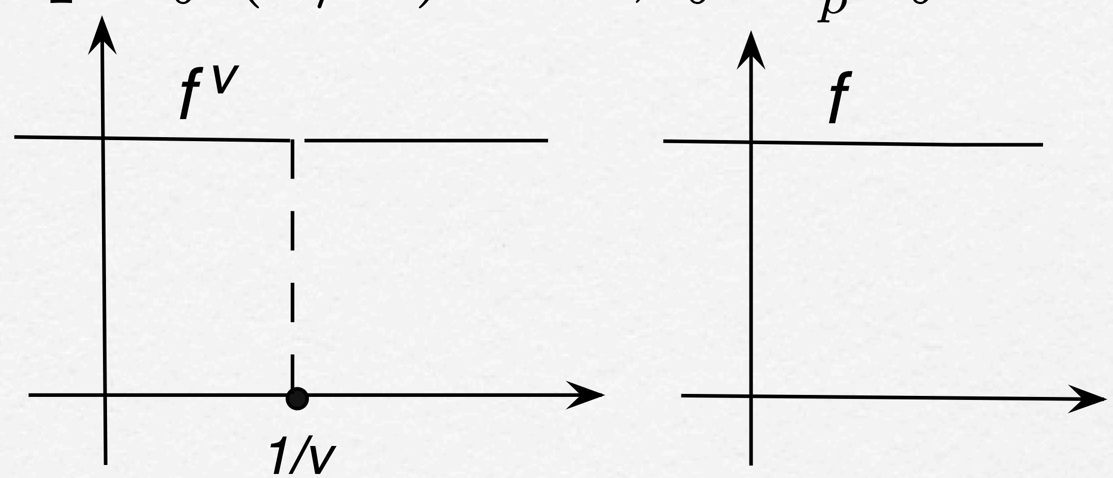
2. uniform convergence implies convergence of minimizers

   but applies rarely, never when constraints depend on ν



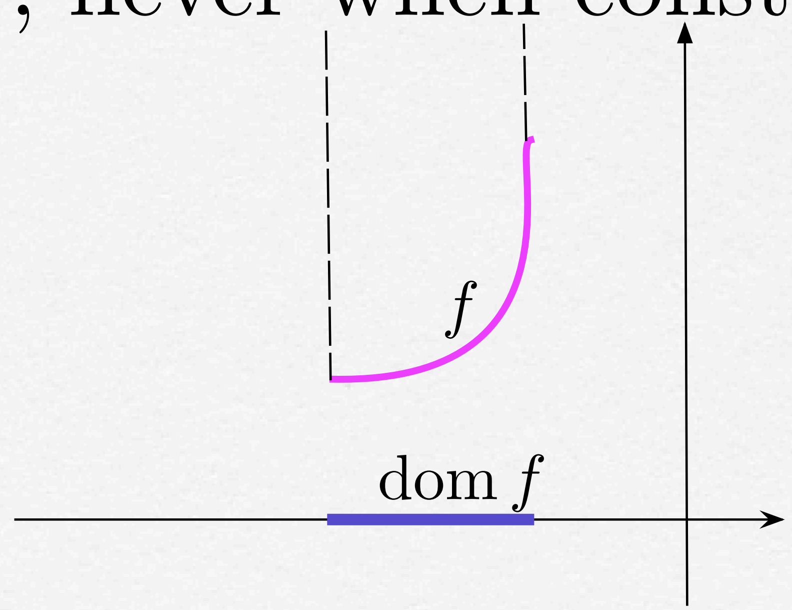
1. pointwise convergence $\not\Rightarrow$ convergence of minimizers

$f^\nu \equiv 1$ except $f(1/\nu) = 0$, $f^\nu \underset{p}{\to} f \equiv 1$



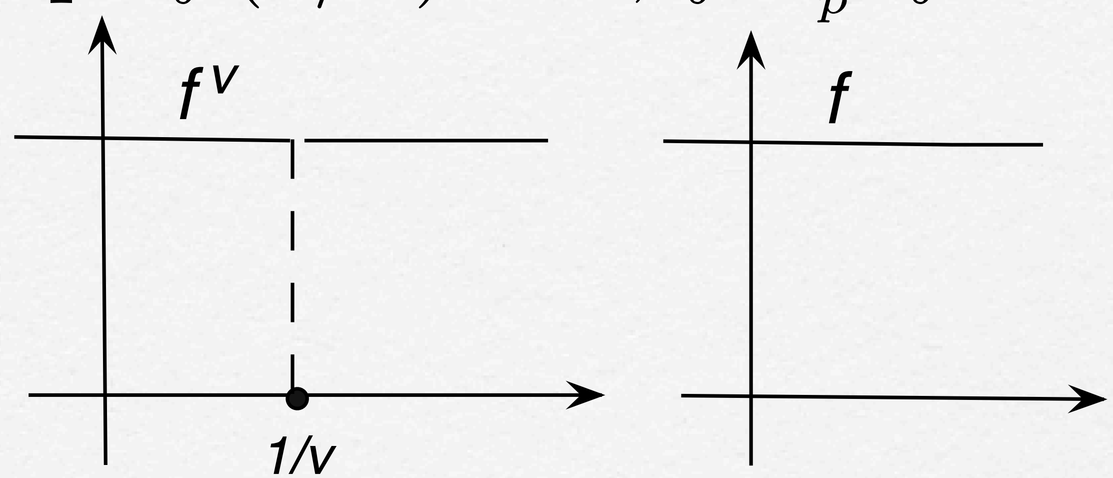
2. uniform convergence implies convergence of minimizers

but applies rarely, never when constraints depend on ν



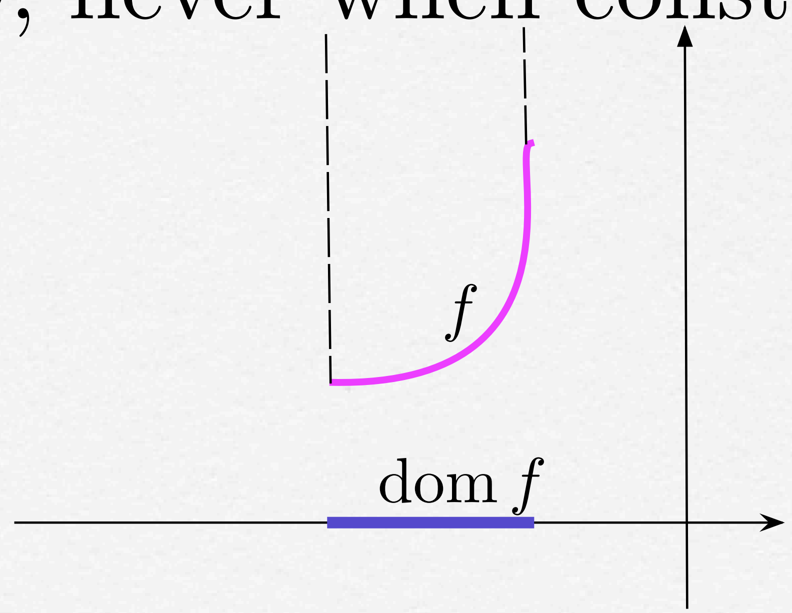
1. pointwise convergence $\not\Rightarrow$ convergence of minimizers

   $f^\nu \equiv 1$ except $f(1/\nu) = 0$, $f^\nu \underset{p}{\rightarrow} f \equiv 1$



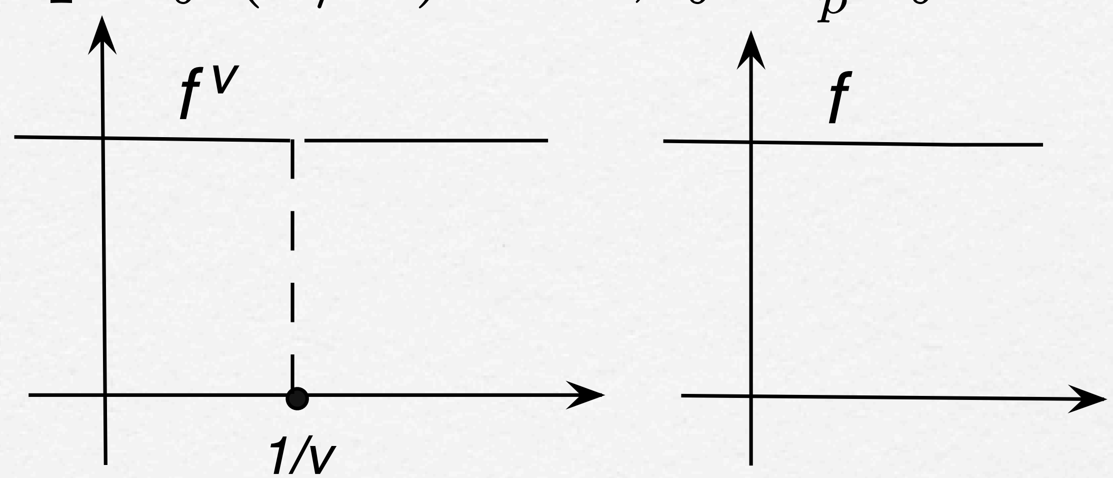
2. uniform convergence implies convergence of minimizers

   but applies rarely, never when constraints depend on ν

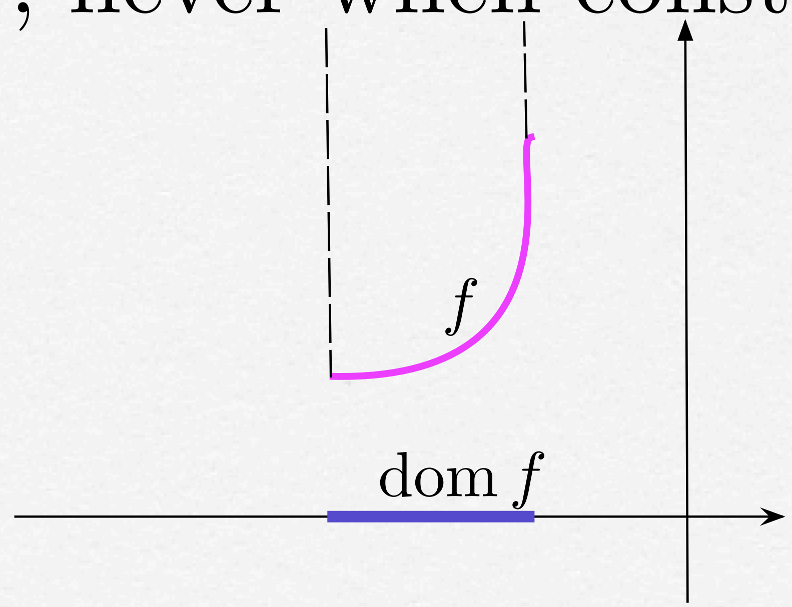


One-sided uniform convergence

1. pointwise convergence $\not\Rightarrow$ convergence of minimizers
$f^\nu \equiv 1$ except $f(1/\nu) = 0$, $f^\nu \underset{p}{\to} f \equiv 1$



2. uniform convergence implies convergence of minimizers
but applies rarely, never when constraints depend on ν



variational epi-convergence

Epi-Convergence

$f^\nu \xrightarrow{e} f$ if for all $x \in E$,

1. $\forall\, x^\nu \to x,\ \liminf_\nu f^\nu(x^\nu) \geq f(x)$

2. $\exists\, x^\nu \to x,\ \limsup_\nu f^\nu(x^\nu) \leq f(x)$

"Geometrically": $\text{epi}\, f^\nu \to \text{epi}\, f$ (later)

Pointwise:

$$\liminf_\nu f^\nu(x) \geq f(x), \qquad \limsup_\nu f^\nu(x) \leq f(x)$$

Continuous: $\forall x^\nu \to x$,

$$\liminf_\nu f^\nu(x^\nu) \geq f(x), \qquad \limsup_\nu f^\nu(x^\nu) \leq f(x)$$

Epi-Convergence $\Rightarrow$

$A^\nu = \arg\min f^\nu$, ε-A^ν : $\varepsilon > 0$ approximate minimizers,

$A = \arg\min f$ of limit problem, ε-A approx. minimizers

A^ν v-converges to A, written $A^\nu \Rightarrow_\nu A$, if

a) $\bar{x} \in$ cluster-points$\{x^\nu \in A^\nu\} \Rightarrow \bar{x} \in A$

b) $\bar{x} \in A \Rightarrow \exists\ \varepsilon_\nu \searrow 0, \mathrm{x}^\nu \in \varepsilon_\nu$-$A^\nu \to \bar{x}$

Epi-Convergence $\rightrightarrows$

$A^\nu = \arg\min f^\nu$, ε-A^ν : $\varepsilon > 0$ approximate minimizers,

$A = \arg\min f$ of limit problem, ε-A approx. minimizers

A^ν v-converges to A, written $A^\nu \rightrightarrows_v A$, if

a) $\bar{x} \in$ cluster-points$\{x^\nu \in A^\nu\} \Rightarrow \bar{x} \in A$

b) $\bar{x} \in A \Rightarrow \exists\ \varepsilon_\nu \searrow 0$, $x^\nu \in \varepsilon_\nu$-$A^\nu \to \bar{x}$

$f^\nu \xrightarrow{e} f$ implies ε-$A^\nu \rightrightarrows_v \varepsilon$-$A$, $\forall \varepsilon \geq 0$
A unique minimizer, ε^ν-$A^\nu \rightrightarrows A$ as $\varepsilon^\nu \searrow 0$.

($\inf f > -\infty$)

Stochastic Optimization

1. **Stochastic Programming** (recourse model)

$$f(\xi, x) = \begin{cases} f_{01}(x) + Q(\xi, x) & \text{if } x \in C_1 \\ \infty & \text{otherwise} \end{cases}$$

$$Q(\xi, x) = \inf_y \left\{ f_{02}(\xi, y) \,\middle|\, y \in C_2(\xi, x) \right\}$$

$$\min Ef(x) = \mathbb{E}\{f(\xi, x)\},$$

SAA-problem: $\min f^\nu(\vec{\xi}^{\,\nu}, x) = \frac{1}{\nu} \sum_{l=1}^{\nu} f(\xi^l, x)$

2. **Statistical Estimation** (fusion of hard & soft information)

$$L(\xi, h) = \begin{cases} -\ln h(\xi) & \text{if } h \geq 0, \int h = 1, h \in A^{\text{soft}} \subset E \\ \infty & \text{otherwise} \end{cases}$$

$$EL(h) = \mathbb{E}\{L(\xi, h)\},\ h^{\text{true}} = \operatorname{argmin}_E \mathbb{E}\{L(\boldsymbol{\xi}, h)\}$$

estimate: $h^\nu \in \operatorname{argmin}_E \mathbb{E}^\nu\{L(\boldsymbol{\xi}, h)\} = \frac{1}{\nu} \sum_{l=1}^{\nu} L(\xi^l, h)$

A^{soft} : constraints on support, moments, shape, smoothness, ...

Pricing financial instruments

3. A contingent claim: environment process: $\left\{ \boldsymbol{\xi}^t \in \mathbb{R}^d \right\}_{t=0}^T$

history: $\overset{\rightarrow t}{\xi}$, $\xi = \xi^T$, price process: $S^t(\overset{\rightarrow t}{\xi}) \in \mathbb{R}^n$; numéraire (risk-free): $S_1^t \equiv 1$

claims: $\left\{ G^t(\overset{\rightarrow t}{\xi}) \right\}_{t=1}^T$; i-strategy: $\left\{ X^t(\overset{\rightarrow t}{\xi}) \right\}_{t=0}^T$; value @ t : $\langle S^t(\overset{\rightarrow t}{\xi}), X^t(\overset{\rightarrow t}{\xi}) \rangle$

Instruments: T-bonds, options, swaps, insurace contracts, mortgages, ...

Pricing financial instruments

3. A contingent claim: environment process: $\{\boldsymbol{\xi}^t \in \mathbb{R}^d\}_{t=0}^T$

history: $\vec{\xi}^t$, $\xi = \xi^T$, price process: $S^t(\vec{\xi}^t) \in \mathbb{R}^n$; numéraire (risk-free): $S_1^t \equiv 1$

claims: $\left\{G^t(\vec{\xi}^t)\right\}_{t=1}^T$; i-strategy: $\left\{X^t(\vec{\xi}^t)\right\}_{t=0}^T$; value @ t : $\langle S^t(\vec{\xi}^t), X^t(\vec{\xi}^t)\rangle$

Instruments: T-bonds, options, swaps, insurace contracts, mortgages, ...

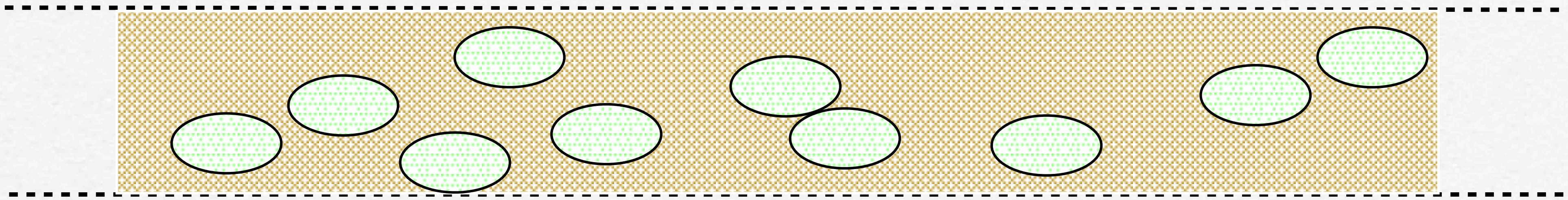
$$\max \mathbb{E}\left\{\langle S^T, X^T\rangle\right\} \text{ such that } \langle S^t, X^t\rangle \le G^t + \langle S^t, X^{t-1}\rangle,\ t = 1 \to T$$

$$\langle S^0, X^0\rangle \le G^0,\ \langle S^T, X^T\rangle \le G^T \text{ a.s.}$$

feasible if $G^0 + \cdots + G^T \ge 0 \ \forall \xi$; arbitrage $\Rightarrow$ unbounded

prob$[\boldsymbol{\xi} = \xi] = p_\xi$ (finite sample?): $\max \sum_{\xi \in \Xi} p_\xi \langle S^T(\xi), X^T(\xi)\rangle \ldots$

4. Stochastic homogenization, ...



$$-\nabla\cdot\big(a(\boldsymbol{\xi},x)\nabla u(\boldsymbol{\xi},x)\big) = h(x) \text{ for } x\in\Omega, \qquad u(\boldsymbol{\xi},x)=0 \text{ on bdry } \Omega$$

Variational formulation: $\forall\,\xi, \quad g(\xi,u) := \frac{1}{2}\int_\Omega a(\xi,x)|\nabla u|^2 dx - \langle h,u\rangle$
find $u(\xi,x) \in \operatorname{argmin}_{u\in H_0^1(\Omega)} g(\xi,u), \quad g(\xi,\cdot): L^2 \to (-\infty,\infty]$. convex

$\mathbb{E}\{u(\boldsymbol{\xi},x)\} \in \operatorname{argmin}_{u\in H_0^1(\Omega)} G(u)$ where $\operatorname{epi} G = \mathbb{E}\{\operatorname{epi} g(\boldsymbol{\xi},\cdot)\}$
$G(u) = \inf_z \big\{\mathbb{E}\{g\big(\boldsymbol{\xi},z(\boldsymbol{\xi})\big)\,\big|\,\mathbb{E}\{z(\boldsymbol{\xi})\} = u\big\}$
$G^* = \mathbb{E}\{g^*(\boldsymbol{\xi},\cdot)\}, \quad g^*(\xi,v) = \sup_u \big\{\langle v,u\rangle - g(\xi,u)\big\}$, conjugate fcn
$\xi^1,\xi^2,\ldots$ stationary, use Ergodic Theorem for random lsc functions

$$G = g^{\text{hom}} = \big(\operatorname{epi}_{\text{w}}\text{-}\lim_\nu \tfrac{1}{\nu}\textstyle\sum_{l=1}^\nu g^*(\boldsymbol{\xi}^l,\cdot)\big)^* \quad\Longrightarrow\quad \text{values of } a^{\text{hom}}(x)$$

Expectation Functionals

$$Ef = \mathbb{E}\{f(\boldsymbol{\xi},\cdot)\}$$

$f:\Xi\times E\to\overline{\mathbb{R}}$, random lsc function, $f(\xi,x) = f_0(\xi,x)$ when $x\in C(\xi)$

$E\subset\mathcal{M}(\Xi,\mathcal{A};\mathbb{R}^n): \mathcal{L}^p(\Xi,\mathcal{A},P;\mathbb{R}^n),\ldots$

others: $C\big((\Xi,\tau);\mathbb{R}^n\big)$,Orlicz, Sobolev, lsc-fcns(E)

$$Ef(x) = \int_\Xi f(\xi,x(\xi))\,P(d\xi) = \mathbb{E}\{f(\boldsymbol{\xi},x(\boldsymbol{\xi}))\}$$

$= \infty$ whenever $\int_\Xi f_+(\xi,x(\xi))\,P(d\xi) = \infty$

$Ef: E\to\overline{\mathbb{R}}$ always defined

Regression: (E is not a linear space)

$$\min\left\{\int_{y\in\mathbb{R}}\int_{x\in[0,1]^n}\phi(y-h(x))\,P(dx,dy)\,\middle|\, h\in \text{lsc-fcns}(\mathbb{R}^n)\cap\mathcal{H}\right\}$$

$\mathcal{H}$ shape restrictions (convex, unimodal, ...)

Random lsc functions

$f:\Xi\times E\to\overline{\mathbb{R}}$ a random lsc function, ξ values in $(\Xi,\mathcal{A},P)$

(a) lsc (lower semicontinuous) in x, $(\forall\xi\in\Xi)$

(b) (ξ,x)-measurable $\quad$ $(\mathcal{A}\times B_E)$-measurable

recall: $f(\xi,x)=f_0(\xi,x)$ when $x\in C(\xi)$ -- stochastic constraints

$$f^\nu(\xi,x)=\begin{cases}\frac{1}{\nu}\sum_{l=1}^{\nu}\left(f(\xi^l,x)\text{ if } x\in C(\xi^l)\right) & \text{(typically)}\\ \infty\text{ otherwise} & (\sim\text{SAA of optimisation problems})\end{cases}$$

Question: Do the $f^\nu(\xi,\cdot)$ epi-converge to $\mathbb{E}\{f(\xi,h)\}$ P-a.s.?

does $x^\nu\in\arg\min f^\nu \Rightarrow_\nu x^*\in\arg\min\mathbb{E}\{f(\xi,x)\}$ P-a.s.?

Random lsc functions

$f : \Xi \times E \to \overline{\mathbb{R}}$ a random lsc function, ξ values in $(\Xi, \mathcal{A}, P)$

(a) lsc (lower semicontinuous) in x, $(\forall \xi \in \Xi)$

(b) (ξ, x)-measurable $\qquad (\mathcal{A} \times B_E)$-measurable

recall: $f(\xi, x) = f_0(\xi, x)$ when $x \in C(\xi)$ -- stochastic constraints

$$f^\nu(\xi, x) = \begin{cases} \frac{1}{\nu} \sum_{l=1}^{\nu} \left(f(\xi^l, x) \text{ if } x \in C(\xi^l) \right) & \text{(typically)} \\ \infty \text{ otherwise} & (\sim \text{SAA of optimisation problems}) \end{cases}$$

Question: Do the $f^\nu(\xi, \cdot)$ epi-converge to $\mathbb{E}\{f(\xi, h)\}$ P-a.s.?

does $x^\nu \in \arg\min f^\nu \Rightarrow_\nu x^* \in \arg\min \mathbb{E}\{f(\xi, x)\}$ P-a.s.?

Law of Large Numbers for random lsc functions
$\sim$ LLN for Stochastic Optimization Problems.

Random lsc functions

$f : \Xi \times E \to \overline{\mathbb{R}}$ a random lsc function, ξ values in $(\Xi, \mathcal{A}, P)$

(a) lsc (lower semicontinuous) in x, $(\forall \xi \in \Xi)$

(b) (ξ, x)-measurable $\quad (\mathcal{A} \times B_E)$-measurable

recall: $f(\xi,x) = f_0(\xi,x)$ when $x \in C(\xi)$ -- stochastic constraints

$$f^\nu(\xi,x) = \begin{cases} \frac{1}{\nu}\sum_{l=1}^{\nu}\left(f(\xi^l,x) \text{ if } x \in C(\xi^l)\right) & \text{(typically)} \\ \infty \text{ otherwise} & (\sim \text{SAA of optimisation problems}) \end{cases}$$

Question: Do the $f^\nu(\xi,\cdot)$ epi-converge to $\mathbb{E}\{f(\xi,h)\}$ P-a.s.?

does $x^\nu \in \arg\min f^\nu \Rightarrow_\nu x^* \in \arg\min \mathbb{E}\{f(\xi,x)\}$ P-a.s.?

$$E^\nu f \xrightarrow{e} Ef \; a.s., \quad E^\nu f(x) = \frac{1}{\nu}\sum_{l=1}^{\nu} f(\boldsymbol{\xi}^l, x)$$

Random lsc functions

(via inf-projections)

D countable dense subset of E

$f : E \to \overline{\mathbb{R}}$, lsc fcn completely identified by

$$\left\{ o_{x\delta} = \inf_{\mathbb{B}^o(x,\delta)} f \,\middle|\, x \in D, \delta \in \mathbb{Q}_+ \right\}, \quad \text{countable}$$

$$\text{or} \quad \left\{ c_{x\delta} = \inf_{\mathbb{B}(x,\delta)} f \,\middle|\, x \in D, \delta \in \mathbb{Q}_+ \right\}$$

$$f(\bar{x}) = \sup_{V \in \mathcal{N}(\bar{x})} \left[\inf_{x \in V} f(x) \right], \quad f \text{ lsc}, \; f(\bar{x}) = \liminf_{x \to \bar{x}} f(x)$$

$$= \sup_{V \in \mathcal{Q}(\bar{x})} \left[\inf_{x \in V} f(x) \right], \quad E \text{ separable} \quad \text{(Polish)}$$

$$\mathcal{Q}(\bar{x}) = \left\{ \mathbb{B}^o(x,\delta) \,\middle|\, x \in D, \delta \in \mathbb{Q}_+, \bar{x} \in \mathbb{B}^o(x,\delta) \right\}$$

$$= \sup_{\delta \in \mathbb{Q}_+} \inf_{\{x \,|\, \mathbb{B}^o(x\delta) \in \mathcal{Q}(\bar{x})\}} o_{x,\delta}$$

$\{c_{x,\delta}\}$ same argument

Epi-convergence

(via inf-projections)

$$f^\nu : \mathbb{R}^n \to \overline{\mathbb{R}},\ f^\nu \xrightarrow{e} f,\ f \text{ lsc},\ \iff \forall \delta \in \mathbb{Q}_+,\ x \in D$$

$$\limsup_\nu c^\nu_{x\delta} \le c_{x\delta}, \quad \liminf_\nu o^\nu_{x\delta} \ge o_{x\delta}$$

for $x \in D, \delta \in \mathbb{Q}_+$: $o^\nu_{x\delta} = \inf_{\mathbb{B}^o(x,\delta)} f^\nu$, $c^\nu_{x\delta} = \inf_{\mathbb{B}(x,\delta)} f^\nu$

(fundamental) **Theorem.** $f^\nu : E \to \overline{\mathbb{R}}$ & f lsc (necessarily)

1. $\text{e-}\liminf_\nu f^\nu \iff \liminf_\nu (\inf_B f^\nu) \ge \inf_B f$ for all compact B
2. $\text{e-}\limsup_\nu f^\nu \iff \limsup_\nu (\inf_O f^\nu) \le \inf_O f$ for all open O

□ Hit-and-miss topology on the space of epigraphs, (later?). □

Scalarization of random lsc fcns

$f : \Xi \times E \to \overline{\mathbb{R}}$, random lsc fcn, completely identified by

$\{o_{x\delta}(\xi) = \inf_{\mathbb{B}^o(x,\delta)} f(\xi,\cdot) \,|\, x \in D, \delta \in \mathbb{Q}_+\}$, countable

or $\{c_{x\delta}(\xi) = \inf_{\mathbb{B}(x,\delta)} f(\xi,\cdot) \,|\, x \in D, \delta \in \mathbb{Q}_+\}$, $\forall \xi \in \Xi$

Scalarization of random lsc fcns

$f : \Xi \times E \to \overline{\mathbb{R}}$, random lsc fcn, completely identified by

$\{o_{x\delta}(\xi) = \inf_{\mathbb{B}^o(x,\delta)} f(\xi,\cdot) \,|\, x \in D, \delta \in \mathbb{Q}_+\}$, countable

or $\{c_{x\delta}(\xi) = \inf_{\mathbb{B}(x,\delta)} f(\xi,\cdot) \,|\, x \in D, \delta \in \mathbb{Q}_+\}$, $\forall \xi \in \Xi$

COUNTABLE

Scalarization of random lsc fcns

$f : \Xi \times E \to \overline{\mathbb{R}}$, random lsc fcn, completely identified by

$\big\{ o_{x\delta}(\xi) = \inf_{\mathbb{B}^o(x,\delta)} f(\xi, \cdot) \,\big|\, x \in D, \delta \in \mathbb{Q}_+ \big\}$, countable

or $\big\{ c_{x\delta}(\xi) = \inf_{\mathbb{B}(x,\delta)} f(\xi, \cdot) \,\big|\, x \in D, \delta \in \mathbb{Q}_+ \big\}$, $\forall \xi \in \Xi$

COUNTABLE

$\forall\, x \in \mathbb{R}^n, \delta > 0$

$\xi \mapsto o_{x\delta} : \Xi \to \overline{\mathbb{R}}$ are measurable,

$o_{x\delta}(\boldsymbol{\xi})$ extended real-valued random variable

$\xi \mapsto c_{x\delta} : \Xi \to \overline{\mathbb{R}}$ are measurable,

$c_{x\delta}(\boldsymbol{\xi})$ extended real-valued random variable.

Scalarization of random lsc fcns

$f : \Xi \times E \to \overline{\mathbb{R}}$, random lsc fcn, completely identified by

$\{o_{x\delta}(\xi) = \inf_{\mathbb{B}^o(x,\delta)} f(\xi,\cdot) \,\big|\, x \in D, \delta \in \mathbb{Q}_+\}$, countable

or $\{c_{x\delta}(\xi) = \inf_{\mathbb{B}(x,\delta)} f(\xi,\cdot) \,\big|\, x \in D, \delta \in \mathbb{Q}_+\}$, $\forall \xi \in \Xi$

COUNTABLE

$\forall\, x \in \mathbb{R}^n, \delta > 0$

$\xi \mapsto o_{x\delta} : \Xi \to \overline{\mathbb{R}}$ are measurable,

$o_{x\delta}(\boldsymbol{\xi})$ extended real-valued random variable

$\xi \mapsto c_{x\delta} : \Xi \to \overline{\mathbb{R}}$ are measurable,

$c_{x\delta}(\boldsymbol{\xi})$ extended real-valued random variable.

- f random lsc fcn $\Rightarrow f + \iota_{\mathbb{B}(x,\delta)}$ random lsc fcn
- f random lsc fcn $\Rightarrow \xi \mapsto \alpha(\xi) = \inf_x f(x,\xi)$ measurable

Probabilistic properties

f random lsc fcn: $\{f(\boldsymbol{\xi}^\nu, \cdot\,)\}_{\nu\in\mathbb{N}}$ iid whenever $\{\boldsymbol{\xi}^\nu\}_{\nu\in\mathbb{N}}$ iid

$$\text{Effös field on lsc-fcns}(E) = \sigma\text{-}\left\{f \in \text{lsc-fcns}(E) \,\middle|\, \inf\nolimits_O < \alpha\right\},\ \ O \text{ open}, \alpha \in \mathbb{R}$$
$$= \mathcal{B}\big(\text{lsc-fcns}(E)\big),\ E \text{ Polish}$$

1. $\{f(\boldsymbol{\xi}^\nu, \cdot\,)\}_{\nu\in\mathbb{N}}$ "i" $\iff$ $\{o_{x\delta}(\boldsymbol{\xi}^\nu),\, \nu \in \mathbb{N}\}$ "i", $\forall x \in \mathbb{Q}^n, \delta \in \mathbb{Q}_+$
2. $f(\boldsymbol{\xi}^1, \cdot\,), f(\boldsymbol{\xi}^2, \cdot\,)$ "id" $\iff$ $o_{x\delta}(\boldsymbol{\xi}^1), o_{x\delta}(\boldsymbol{\xi}^2)$ "id", $\forall x \in \mathbb{Q}^n, \delta \in \mathbb{Q}_+$

the same holds for $\{c_{x\delta}(\,\cdot\,)\}$

Summary

$f : \Xi \times E \to \overline{\mathbb{R}}$ random lsc fcn,
$\quad \boldsymbol{\xi}, \{\boldsymbol{\xi}^\nu\}_{\nu\in\mathbb{N}}$ iid
$\Rightarrow f(\boldsymbol{\xi}, \cdot), \{f(\boldsymbol{\xi}^\nu, \cdot)\}_{\nu\in\mathbb{N}}$ iid

$\Rightarrow \{o_{x\delta}(\boldsymbol{\xi}^\nu) : \Xi \to \overline{\mathbb{R}}\}_{\nu\in\mathbb{N}}$, iid,
countable, identify $f(\boldsymbol{\xi}^\nu \cdot)$

$\Rightarrow \{c_{x\delta}(\boldsymbol{\xi}^\nu) : \Xi \to \overline{\mathbb{R}}\}_{\nu\in\mathbb{N}}$, iid,
countable, identify $f(\boldsymbol{\xi}^\nu \cdot)$

Countable ⇒ a.s.

Lemma. $f, g : E \to \overline{\mathbb{R}}$, lsc. $D = \mathrm{prj}_E$ countable dense subset of $\mathrm{epi}\, f$.

$$f \le g \text{ on } D \implies f \le g \text{ on } E.$$

Proof. $f \le g$ on D only if $\{(x, \alpha) \,|\, \alpha \ge g(x), x \in R\} \subset \mathrm{epi}\, f$.

Taking closure on both sides $\implies$ $\mathrm{epi}\, g \subset \mathrm{epi}\, f$. $\square$

Implication. To check $f(\boldsymbol{\xi}, \cdot) \le g(\boldsymbol{\xi}, \cdot)$ $a.s.$ on E only needs $f(\boldsymbol{\xi}, \cdot) \le g(\boldsymbol{\xi}, \cdot)$ $a.s.$ on D a countable dense subset of E. Restrict $\boldsymbol{\xi}$ to a set of P-measure 1, say Ξ itself (from now on), and $f(\boldsymbol{\xi}, \cdot) \le g(\boldsymbol{\xi}, \cdot)$ on $D \implies f(\boldsymbol{\xi}, \cdot) \le g(\boldsymbol{\xi}, \cdot)$ on E.

LLN: random lsc functions?

$\forall\, x \in D,\ \delta \in \mathbb{Q}_+$

1. $\frac{1}{\nu}\sum_{l=1}^{\nu} o_{x\delta}(\boldsymbol{\xi}^l) \to \mathbb{E}\{o_{x\delta}(\boldsymbol{\xi})\},\ (P^\infty\text{-}a.s.)$

2. $\frac{1}{\nu}\sum_{l=1}^{\nu} c_{x\delta}(\boldsymbol{\xi}^l) \to \mathbb{E}\{c_{x\delta}(\boldsymbol{\xi})\},\ (P^\infty\text{-}a.s.)$

LLN: random lsc functions?

$\forall\, x \in D,\ \delta \in \mathbb{Q}_+$

1. $\frac{1}{\nu}\sum_{l=1}^{\nu} o_{x\delta}(\boldsymbol{\xi}^l) \to \mathbb{E}\{o_{x\delta}(\boldsymbol{\xi})\},\ (P^\infty\text{-}a.s.)$
2. $\frac{1}{\nu}\sum_{l=1}^{\nu} c_{x\delta}(\boldsymbol{\xi}^l) \to \mathbb{E}\{c_{x\delta}(\boldsymbol{\xi})\},\ (P^\infty\text{-}a.s.)$

$\not\Longrightarrow \sum_{l=1}^{\nu} f(\boldsymbol{\xi}^l, \cdot\,) \xrightarrow{e} \mathbb{E}\{f(\boldsymbol{\xi}, \cdot\,)$ because

$$\min\big\{\mathbb{E}\{f(\boldsymbol{\xi}, z)\}\,\big|\, z \in \mathbb{B}(x,\delta)\big\} \neq \mathbb{E}\big\{\min\{f(\boldsymbol{\xi}, z)\}\,\big|\, z \in \mathbb{B}(x,\delta)\big\}$$

in general

Law of Large Numbers: Random lsc functions

LLN: Proof

1. $\exists\, x^\nu \to x : \limsup_\nu E^\nu f \leq Ef$
   for any $x \in E$ and any sample $\xi^\infty = (\xi^1, \xi^2, \ldots)$
   $\lim_\nu \frac{1}{\nu}\sum_{l=1}^{\nu} f(\xi^l, x) \;\sim \lim_\nu \mathbb{E}^\nu\{f(\boldsymbol{\xi}^\infty, x)\} = Ef(x).$

2. $\forall x^\nu \to x, \liminf_\nu E^\nu f \geq Ef$
   for any $x \in E$ and any $\xi^\infty = (\xi^1, \xi^2, \ldots) \in \Xi^\infty$

$$\text{e-}\liminf_{\nu\to\infty} f^\nu(\xi^\infty, x) = \sup_{\delta\searrow 0}\liminf_{\nu\to\infty}\inf_{\mathbb{B}^o(x,\delta)} E^\nu f \geq \sup_{\delta^l\searrow 0}\liminf_{\nu\to\infty}\frac{1}{\nu}\sum_{l=1}^{\nu} o^l_{x^l\delta^l}(\xi^l)$$

where $x^l \in D \to x$, $\delta^l \in \mathbb{Q}_+ \searrow 0$: $\quad x \in \mathbb{B}^o(x^l, \delta^l)$ & $\{\mathbb{B}^o(x^l, \delta^l)\}\searrow$
$\frac{1}{\nu}\sum_{l=1}^{\nu} o^l_{x^l\delta^l}(\xi^l) \to \mathbb{E}\{o^l_{x^l\delta^l}(\boldsymbol{\xi})\}$ & $\mathbb{E}\{o^l_{x^l\delta^l}(\boldsymbol{\xi})\}\nearrow Ef(x)$

$$\implies \quad \text{e-}\liminf_{\nu\to\infty} E^\nu f(x) \geq Ef(x) \quad \square$$

Theorem

$f : \Xi \times E \to \overline{\mathbb{R}}$, locally inf-integrable random lsc function
$\{\boldsymbol{\xi}, \boldsymbol{\xi}^1, \ldots, \}$ are iid Ξ-valued random variables. Then,

$$E^\nu f = \mathbb{E}^\nu\{f(\boldsymbol{\xi}, \cdot\,) = \frac{1}{\nu} \sum_{l=1}^{\nu} f(\boldsymbol{\xi}^l, \cdot\,) \xrightarrow{e} Ef = \mathbb{E}\{f(\boldsymbol{\xi}, \cdot\}$$

which means ε-argmin $E^\nu f \Rightarrow_v \varepsilon$-argmin Ef, $\forall\, \varepsilon \geq 0$

Ef unique minimizer, ε^ν-argmin $E^\nu f \rightrightarrows$ argmin Ef as $\varepsilon^\nu \searrow 0$.

SAA-applies without ‘any’ restrictions

loc.inf-integrable: $\int \inf\{f(\xi, \cdot) \,\big|\, \mathbb{B}(x, \delta)\} > \infty$ for some $\delta > 0$,
irrelevant in applications

Ergodic Theorem

(E, d) Polish, $(\Xi, \mathcal{A}, P)$ & $\mathcal{A}$ P-complete
$f : \Xi \times E \to \overline{\mathbb{R}}$ a random lsc function, locally inf-integrable
$\varphi : \Xi \to \Xi$ ergodic measure preserving transformation. Then,

$$\frac{1}{\nu} \sum_{l=1}^{\nu} f(\varphi^l(\boldsymbol{\xi}, \cdot\,) \xrightarrow{e} Ef \quad a.s.$$

allows for stationary rather than iid.

Application: "samples" coming from dynamic systems,
time series, SDE, etc.