VARIATIONAL ANALYSIS: APPROXIMATION METHODOLOGY

Roger J-B Wets

Mathematics, University of California, Davis

Hanoi, Vietnam -- Spring 2013
Institute of Mathematics & VIASM

A STOCHASTIC LINEAR PROGRAM?

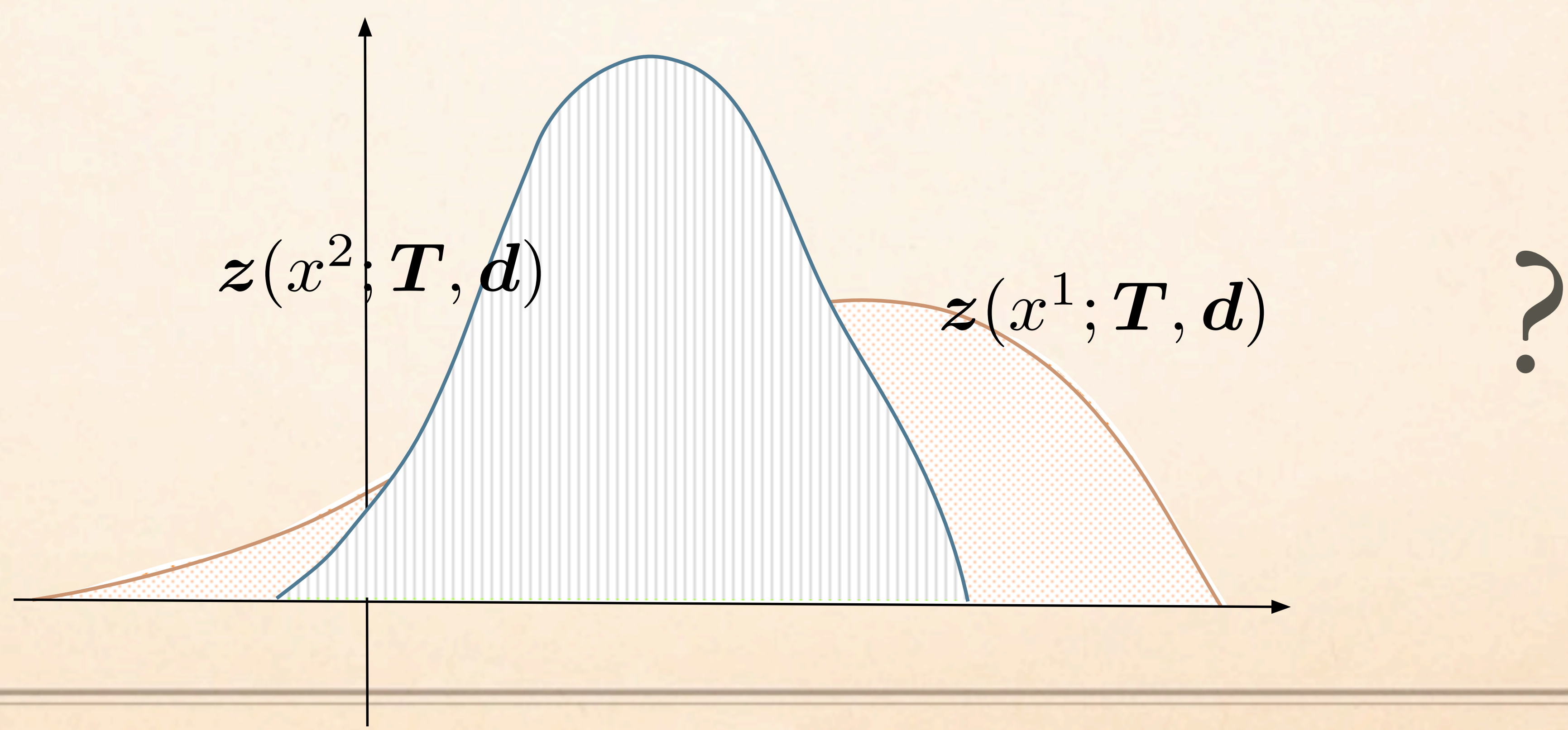
min $z = \langle c, x \rangle$ such that $Tx \le d$, $x \in \mathbb{R}^n_+$, say 4 variables, 2 constraints

$c_j =$ unit cost of activity j, nonegative activities $x_j \ge 0$

$\boldsymbol{t}_{ij} =$ # units of i-resourse consumed by activity j, random

$\boldsymbol{d}_i =$ i-resourse units available, random

Decision problem: choose best x! best returns distribution $\boldsymbol{z}(x; \boldsymbol{T}, \boldsymbol{d})$



A PRODUCT MIX PROBLEM

min $\langle c,x \rangle$ such that $Tx \le d$, $x \in \mathbb{R}^n_+$, say 4 variables, 2 constraints

$c_j = -$ profit of activity x_j

per dresser production profit (manufacturer)

$t_{ij} =$ per unit i-resourse consumed by activity j

time consumed for carpentry and finishing

$d_i =$ i-resourse units available

# of hours availalbe for carpentry and finishing

but actually $\boldsymbol{t}_{ij}$ and $\boldsymbol{d}_i$ are random variables $\Rightarrow$ additional 'overtime'

min $\langle c,x \rangle + \mathbb{E}\{\langle q, \boldsymbol{y} \rangle\}$ such that $\boldsymbol{y} \ge \boldsymbol{T}x - \boldsymbol{d}$, $x \in \mathbb{R}^n_+, \boldsymbol{y} \ge 0$.

$(\boldsymbol{T}, \boldsymbol{d})$ uniformly distributed components $\Rightarrow$ infinite # of variables, constraints

discretized, say each 4-values, $\Rightarrow$ l.p. with $\approx 2*10^6$ variables, constraints

1. consistent approximation? 2. design of solution procedures

VALUATION

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): $S_1^t \equiv 1$

contingent claims: $\left\{G^t(\overset{\rightarrow t}{\xi})\right\}_{t=1}^{T}$; investment strategy: $\left\{X^t(\overset{\rightarrow t}{\xi})\right\}_{t=0}^{T}$

portfolio value at t : $\langle S^t(\overset{\rightarrow t}{\xi}), X^t(\overset{\rightarrow t}{\xi})\rangle$

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

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

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

What if the random vectors are not discrete? What if $t \in [0,T]$?

Associated Risk-Neutral Probabilities: exists?, can be approximated?

HOMOGENIZATION

conductor: $\Omega \subset \mathbb{R}^3$, composite ≥ 2 materials,

$0 \leq a(\xi, x) \leq \kappa_{\text{bdd}}$, stationary process w.r.t. location

heat u: with rapidly varying stochastic coefficients

$\nabla \bullet \big(a(\xi,x) \nabla u(\xi,x) \big) = h(x),\ x \in \Omega$ & bdry conditions

homogenized equation with effective coefficient a

$\nabla \bullet \big(a(x) \nabla u(x) \big) = h(x),\ x \in \Omega$ & brdy cond.

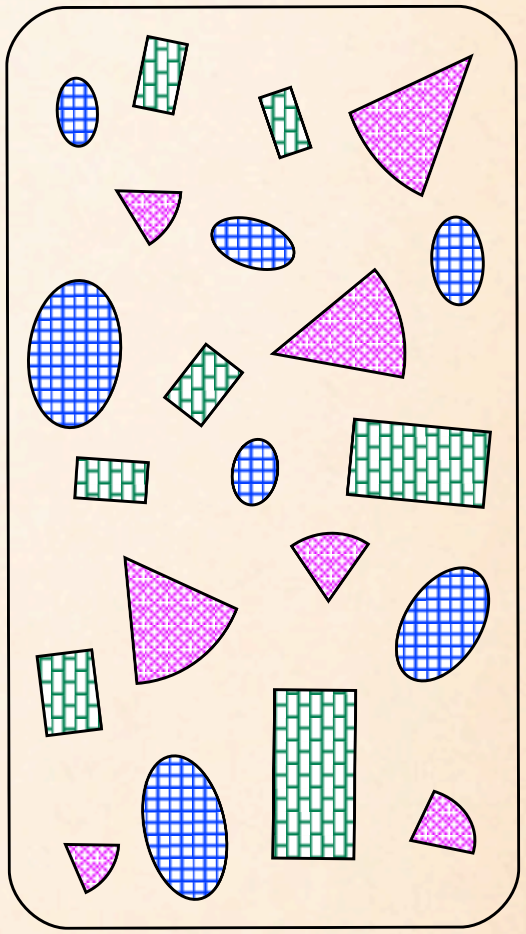
such that $u(x) = \mathbb{E}\{u(\xi,x)\}$. $a(x) \neq \mathbb{E}\{a(\xi,x)\}$

$\min_{u \in H_0^1(\Omega)} g(\xi,u) = \frac{1}{2} \int_\Omega a(\xi,x) |\nabla u|^2 dx - \langle h,u \rangle$

$g : L^2 \to (-\infty]$, to be minimized for all ξ

homogenization: find g^{hom} such that

$\mathbb{E}\{u(\xi,\cdot)\} = \bar{u}(\cdot) \in \arg\min \left[g^{\text{hom}}(u) \,\middle|\, u \in H_0^1(\Omega) \right]$



OPTIMALITY CONDITIONS

min $\mathbb{E}\left\{f_0(\boldsymbol{\xi},x)\right\}$ such that prob$\left\{f_i(\boldsymbol{\xi},x)\leq 0, i=1,\ldots m\right\}\leq\alpha$

simplifying: $\alpha=1$, $f_i(\xi,x)=f_i(x)$, constraint qualification satisfied,

Optimality conditions (KKT) or stationary point

$\bar{x}$ optimal if $\exists\,\bar{y}=(\bar{y}_1,\ldots,\bar{y}_m)$ such that

a) $f_i(\bar{x})\leq 0,\ i=1,\ldots m$

b) $\bar{y}_i\geq 0$ and $\bar{y}_i\perp f_i(\bar{x}),\ i=1,\ldots m$

c) $0\in\nabla\left(Ef_0(x)+\sum_{i=1}^m \bar{y}_i f_i(x)\right)=\mathbb{E}\left\{\nabla f_0(\boldsymbol{\xi},x)\right\}+\sum_{i=1}^m \bar{y}_i\nabla f_i(x)$

OPTIMALITY CONDITIONS

Solving the "generalized equation":

$$C(\boldsymbol{\xi}) = \left\{ (x,y) \,\middle|\, \begin{array}{l} f_i(x) \le 0,\, y_i \ge 0,\, y_i \perp f_i(x),\ i = 1 \to m \\ 0 \in \left[\nabla f_0(\boldsymbol{\xi}, x) + \sum_{i=1}^{m} y_i \nabla f_i(x) \right] \end{array} \right\}$$

$$C : \Xi \rightrightarrows (\mathbb{R}^n \times \mathbb{R}^m), \quad (\bar{x}, \bar{y}) \in \mathbb{E}\{C(\boldsymbol{\xi})\}$$

sample ξ^k, $(x^k, y^k) \in C(\xi^k)$, ? $\frac{1}{\nu} \sum_{k=1}^{\nu} (x^k, y^k) \to_? (\bar{x}, \bar{y})$.

WHAT TO REMEMBER?

- Stochastic problems get quickly unmanageably large
- Approximation (discretization, sampling, ...) is a must
- Approximation "of the classical type" might or might not work, including the standard approx. of stochastic processes
- The presence of constraints, in particular inequality constraints, radically changes the paradigm.
- The search for "averaged solution" doesn't result from straightforward averaging.

VARIATIONAL PROBLEMS

Optimization: $\min f(x)$ sucht that $x \in X \subset \mathcal{X}$

Variational Inequality: $x \in C$ such that $-G(x) \in N_C(x)$

Complementarity Problems: $0 \le x \perp H(x) \ge 0$

Generalized Equations: $S(x) \ni 0, S : X \rightrightarrows U$ (set-valued)

Economic Equilibrium: $\forall a \in \mathcal{A}$, $x \in \text{argmax}_{C_a}\ u_a(p, x)$

market equilibrium: $0 \le p$ such that $D(p, x_A) \in N_C(p)$

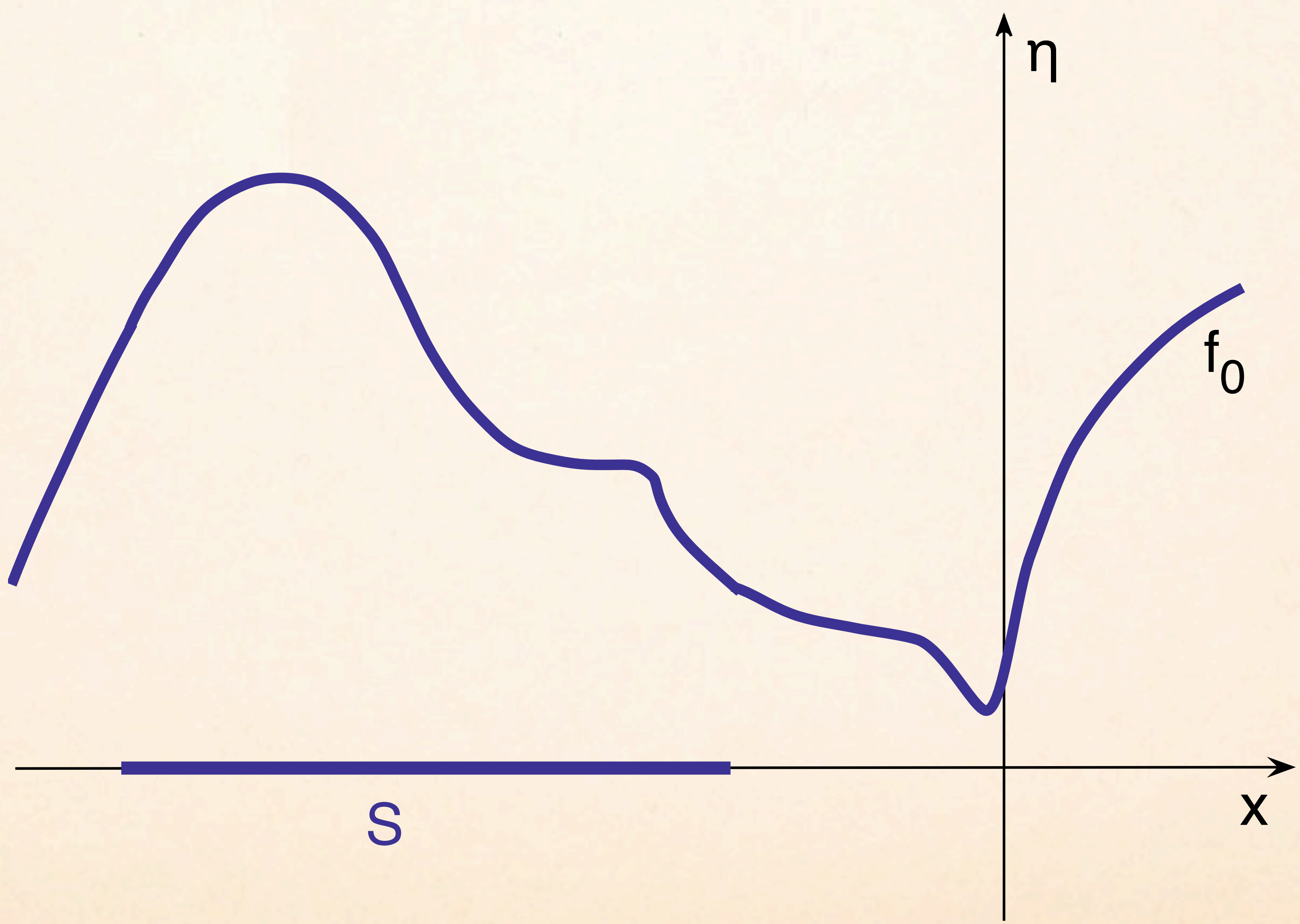
Nash Games: $\bar{x}_a \in \arg\max r_a(x_a, \bar{x}_{-a}), \quad \forall a \in \mathcal{A}$

Each one comes with applications in a stochastic environment

OPTIMIZATION PROBLEM

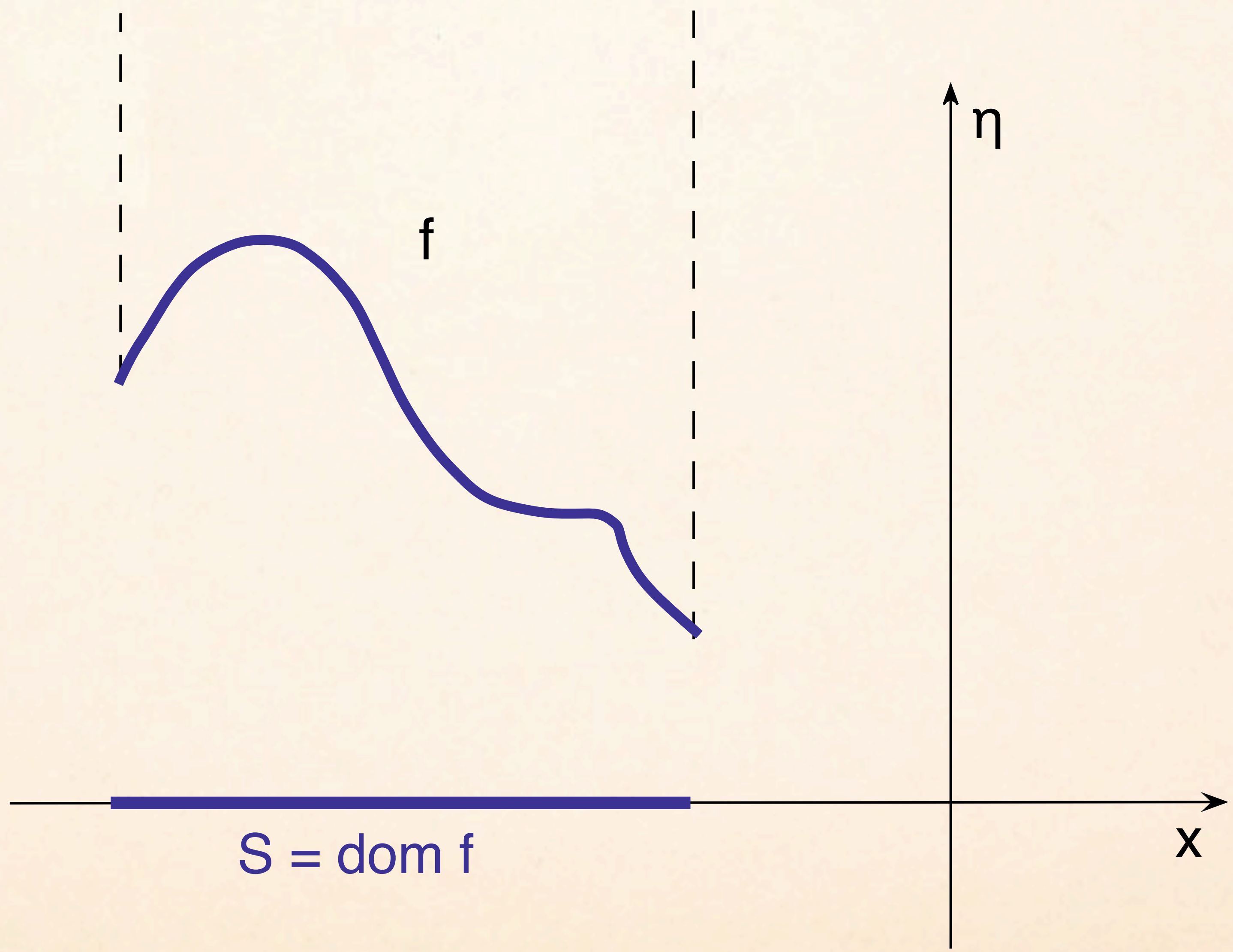
$\min f(x),\ x \in S,$

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



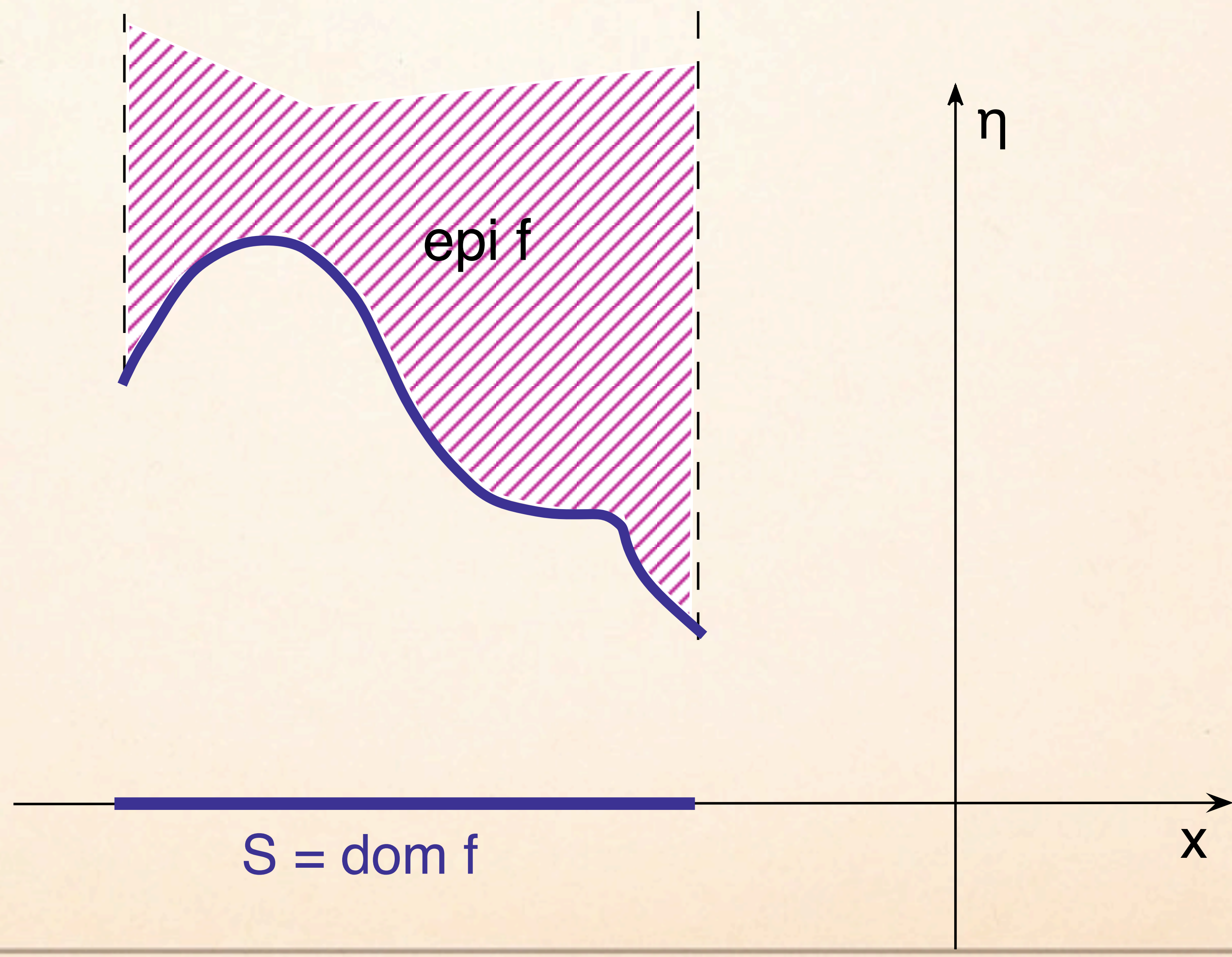
EXTENDED-REAL VALUED FCN

$\min f$ on $\mathbb{R}^n$, $\quad f = f_0 + \iota_S(x), \quad \iota_S$ indicator function of S



FUNCTIONS & EPIGRAPHS

$$\operatorname{epi} f = \left\{ (x,\alpha) \middle| f(x) \le \alpha \right\}$$

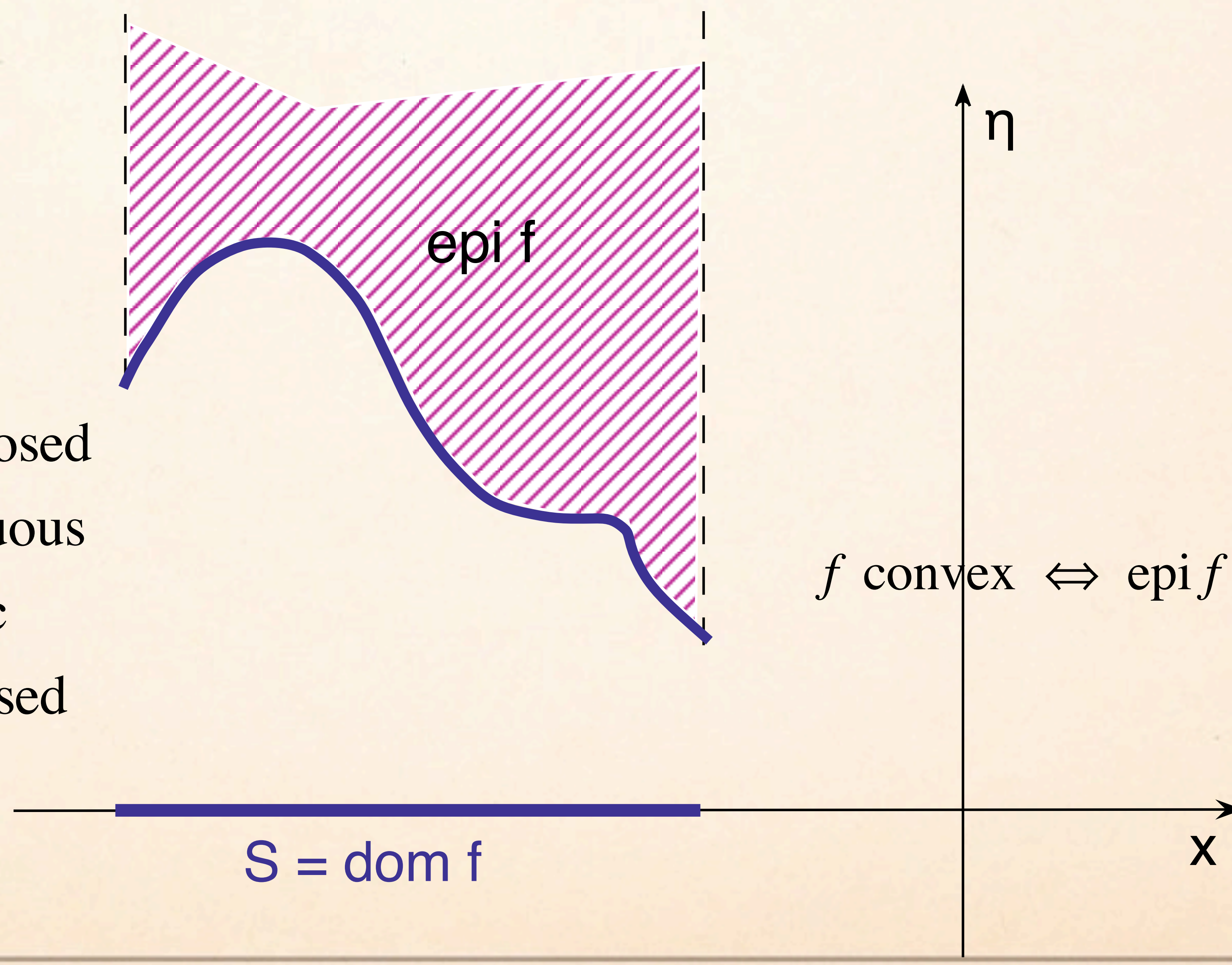


FUNCTIONS & EPIGRAPHS

$$\text{epi} f = \{(x,\alpha) | f(x) \le \alpha\}$$

f lsc at x : $\liminf_{x' \to x} f(x') \ge f(x)$, f usc at x : $\limsup_{x' \to x} f(x') \le f(x)$

f lsc $\Leftrightarrow$ epi f closed f usc $\Leftrightarrow$ hypo f closed



f lsc $\Leftrightarrow$ epi f closed

lower semicontinuous

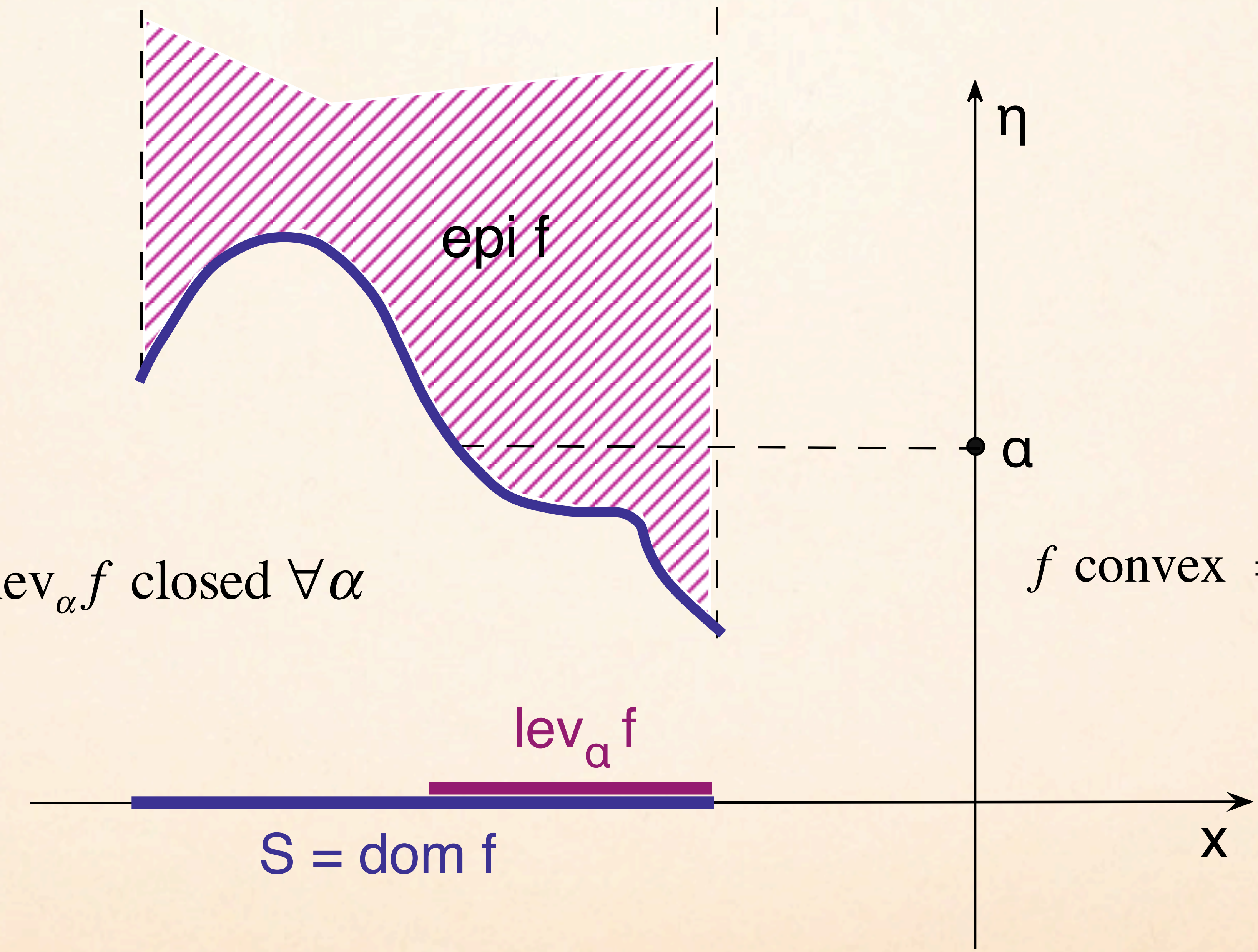
f usc $\Leftrightarrow$ $-f$ lsc

$\Leftrightarrow$ hypo f closed

f convex $\Leftrightarrow$ epi f convex

LEVEL SETS & CONSTRAINTS

$$\text{lev}_\alpha = \left\{ x \in \mathbb{R}^n \,\middle|\, f(x) \leq \alpha \right\}$$

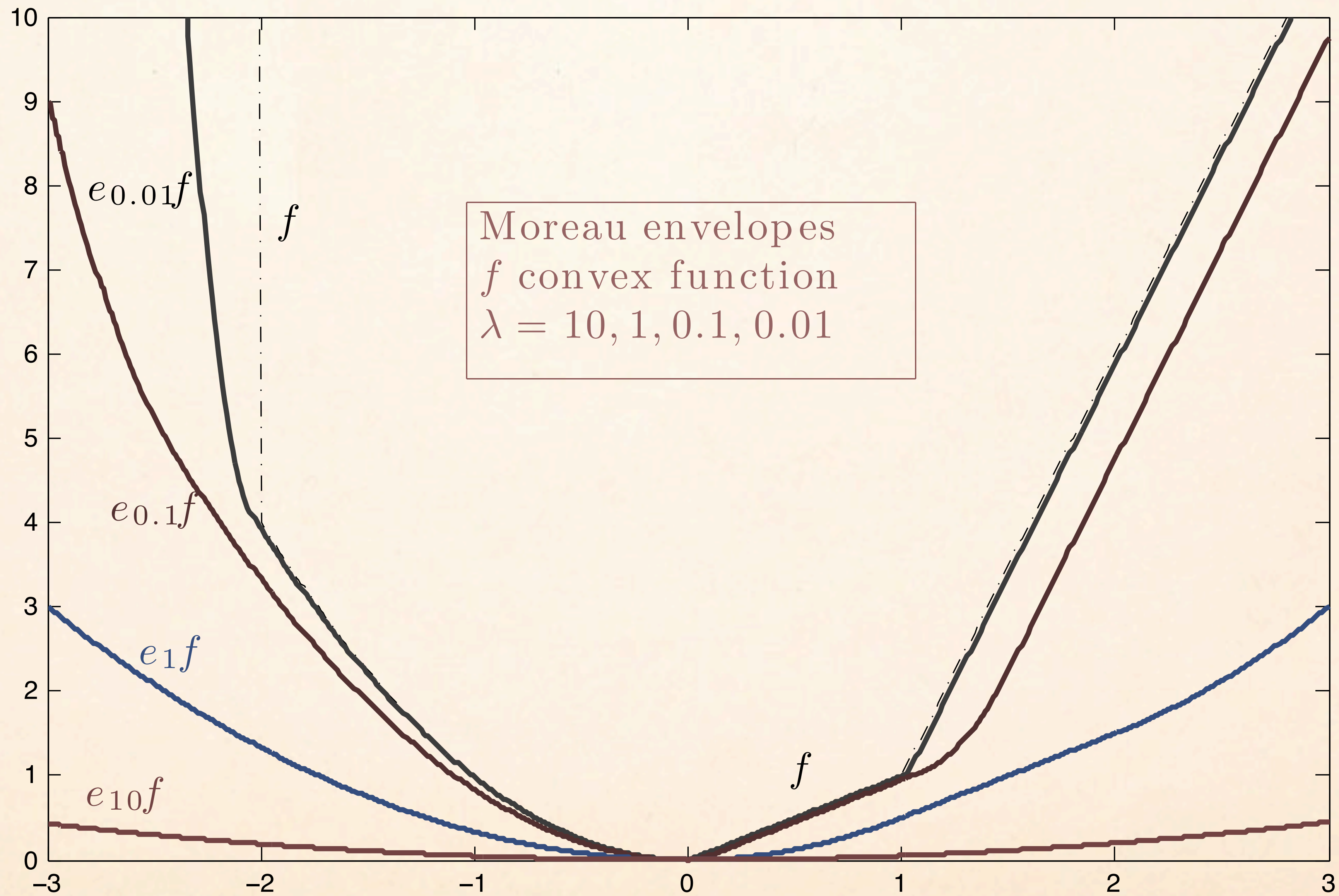


f lsc $\Leftrightarrow$ $\text{lev}_\alpha f$ closed $\forall \alpha$

f convex $\Rightarrow$ $\text{lev}_\alpha f$ convex

EPI-SUMS (INF-CONVOLUTION)

$$\text{epi}\, f \;\#\; \text{epi}\, g = \inf_w \left\{ f(w) + g(w - x) \right\} \qquad e_\lambda f(x) \;\text{ with } g = \frac{1}{2\lambda} |\bullet|$$



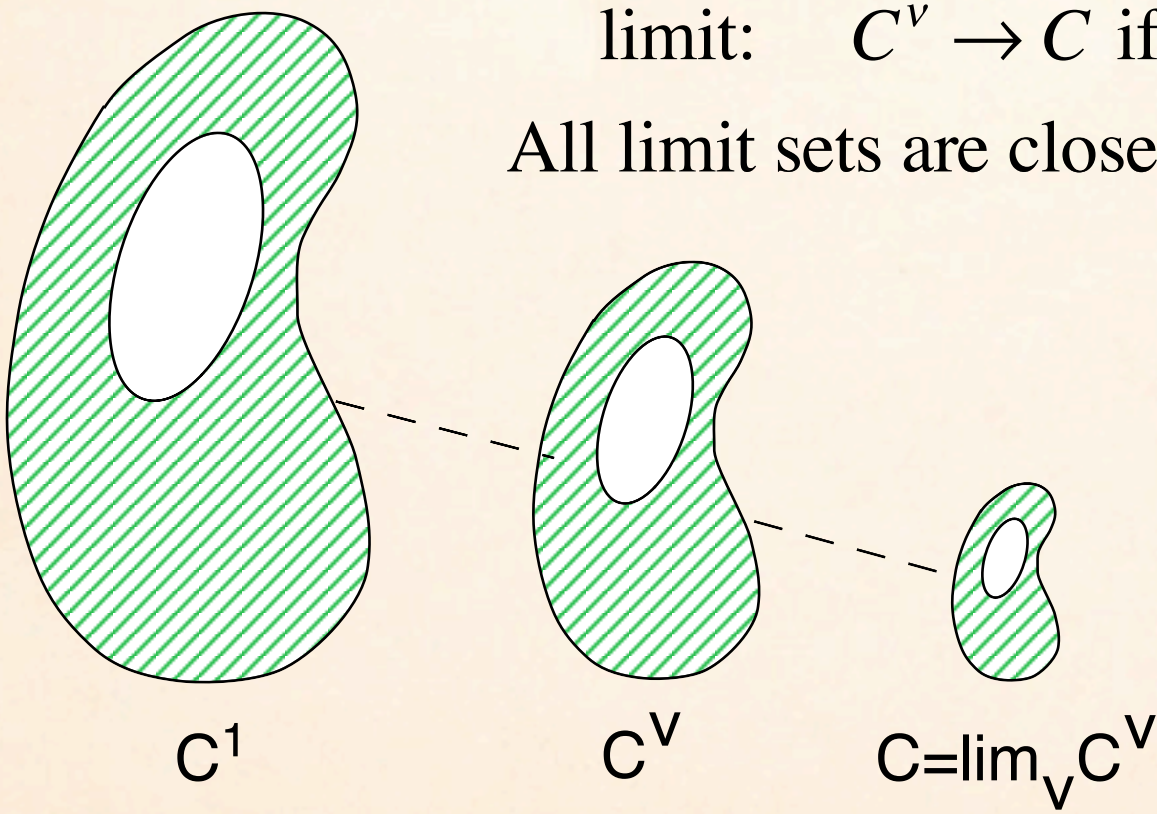
APPROXIMATION: CONVERGENCE

outer limit: $\mathrm{Ls}_v C^v = \left\{ x \in \text{ cluster-points}\{x^v\} , x^v \in C^v \right\}$

inner limit: $\mathrm{Li}_v C^v = \left\{ x = \lim_v x^v , x^v \in C^v \subset \mathbb{R}^n \right\} \subset \mathrm{Ls}_v C^v$

limit: $C^v \to C$ if $C = \mathrm{Li}_v C^v = Ls_v C^v$ (Painlevé)

All limit sets are closed

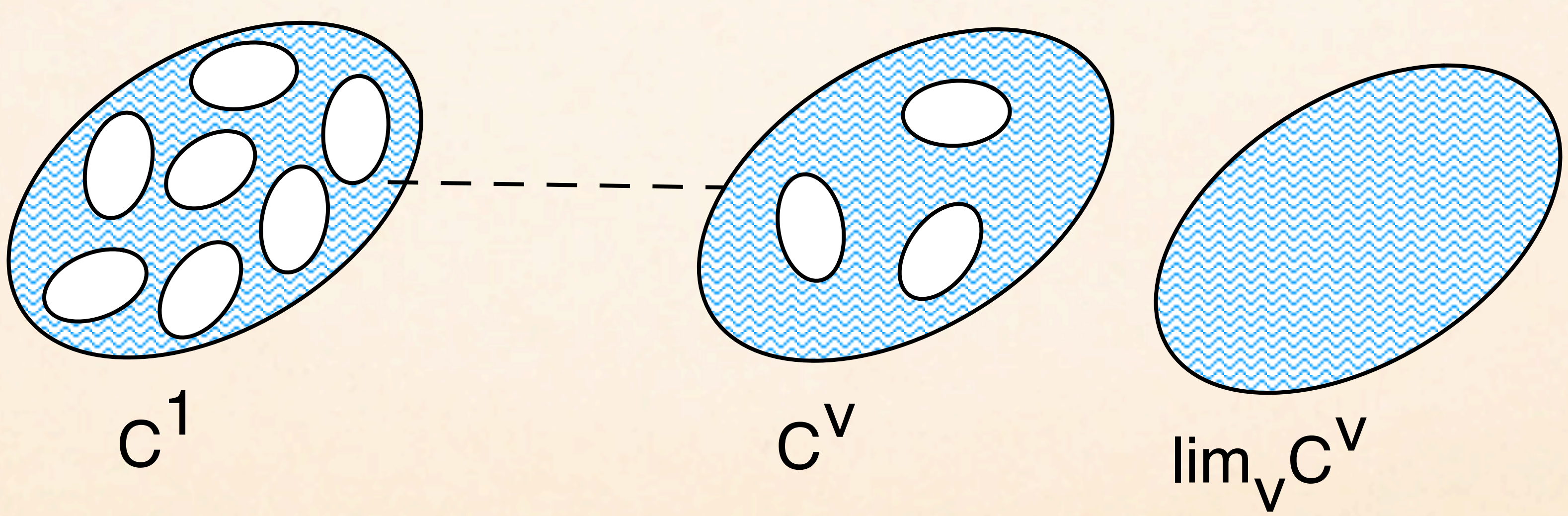


CONVEX LIMIT SETS

C^ν convex $\Rightarrow$ $\mathrm{Li}_\nu C^\nu$ convex $\Rightarrow$ $\mathrm{Lm}_\nu C^\nu$ convex (if it exists)

$\nRightarrow$ $\mathrm{Ls}_\nu C^\nu$ convex

but convexity can result from taking limits



EPI-LIMITS

$\left\{f^\nu : \mathbb{R}^n \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\right\}$

lower epi-limit: $\text{e-li}_\nu f^\nu$ such that $\text{epi}(\text{e-li}_\nu f^\nu) = \text{Ls}_\nu \, \text{epi}\, f^\nu$

upper epi-limit: $\text{e-ls}_\nu f^\nu$ such that $\text{epi}(\text{e-ls}_\nu f^\nu) = \text{Li}_\nu \, \text{epi}\, f^\nu$

epi-limit: $f^\nu \underset{e}{\to} f$ when $f = \text{e-li}_\nu f^\nu = \text{e-ls}_\nu f^\nu$, $f = \text{e-lm}_\nu f^\nu$

all epi-limits are lsc (closed epigraphs), $\text{e-li}_\nu f^\nu \leq \text{e-ls}_\nu f^\nu$

f^ν convex $\Rightarrow$ $\text{e-ls}_\nu f^\nu$ is convex and so is $\text{e-lm}_\nu f^\nu$ (if it exists)

Convergence of level sets / constraint sets:

$f \leq \text{e-li}_\nu f^\nu \Leftrightarrow \text{Ls}_\nu(\text{lev}_{\alpha_\nu} f^\nu) \subset \text{lev}_\alpha f \quad \forall \alpha_\nu \to \alpha$

$f \geq \text{e-ls}_\nu f^\nu \Leftrightarrow \text{Ls}_\nu(\text{lev}_{\alpha_\nu} f^\nu) \subset \text{lev}_\alpha f$ for some $\alpha_\nu \to \alpha$

Operations: sums, scalar multiplication, epi-sums

SV-CONVERGENCE
SOLUTIONS, MINIMIZERS, ...

A^ν solutions of (generalized) equations

minimizers of a sequence of functions

saddle points or min-sup points of bifunctions

$\varepsilon\text{-}A^\nu : \varepsilon > 0$ approximate solutions, minimizers, ...

A solution set, minimizers, ... of corresponding limit

Definition: A^ν sv-converge to A, written $A^\nu \rightrightarrows_\nu A$, if

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

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

CONVERGENCE OF MINIMIZERS
SV-CONVERGENCE OF MINIMIZERS

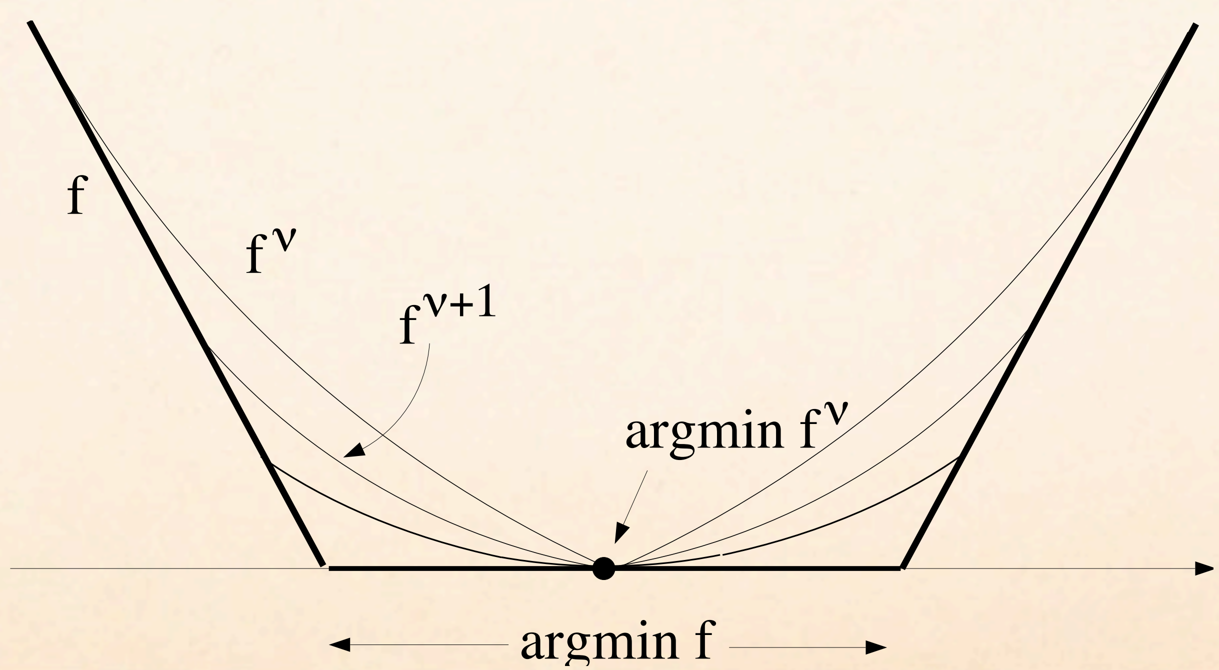
$f^\nu \underset{e}{\to} f, x \in \text{ cluster}\left\{x^\nu \in \arg\min f^\nu\right\} \Rightarrow x \in \arg\min f$

$f^\nu \underset{e}{\to} f, \inf f \in \mathbb{R}, x \in \arg\min f \Rightarrow \exists \varepsilon_\nu \searrow 0, x^\nu \in \varepsilon_\nu\text{-}\arg\min f^\nu \to x$

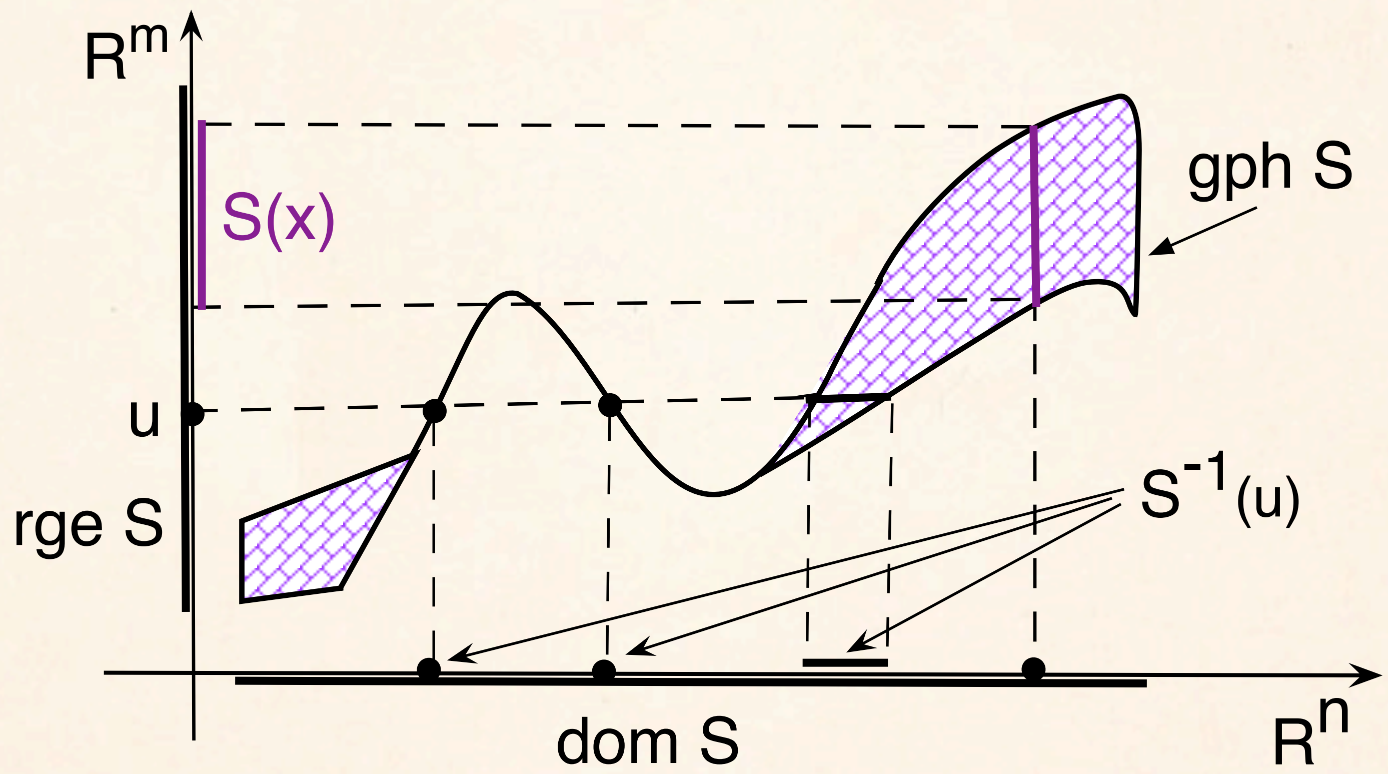
$f^\nu \underset{e}{\to} f \not\Rightarrow \arg\min f^\nu \to \arg\min f$

$f^\nu \underset{e}{\to} f,\ \inf f^\nu \to \inf f \in \mathbb{R} \Leftrightarrow \left\{f^\nu\right\}_{\nu\in\mathbb{N}}$ epi-tight, i.e.

$\forall \varepsilon > 0, \exists B$ compact s.t. $\inf_B f^\nu \leq \inf f^\nu + \varepsilon,\ \forall \nu \geq \nu_\epsilon$



SET-VALUED MAPPINGS



S osc (outer semicontinuous) at $\bar{x}$ if $\mathrm{Ls}_{x \to \bar{x}} S(x) \subset S(\bar{x})$

S osc $\Leftrightarrow$ gph S closed

S isc (inner semicontinuous) at $\bar{x}$ if $\mathrm{Li}_{x \to \bar{x}} S(x) \supset S(\bar{x})$

S continuous if it's isc and osc

GRAPHICAL CONVERGENCE
SV-CONVERGENCE OF SOLUTIONS

$$S^\nu \to_g S \text{ when gph } S^\nu \to \text{gph } S \quad (\text{as subsets of } \mathbb{R}^n \times \mathbb{R}^m)$$

Generalized Equations ~ Inclusions

$S^\nu, S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $S^\nu(x) \ni u^\nu$, $S(x) \ni \bar{u}$ and $S^\nu \to_g S, u^\nu \to \bar{u}$. Then

$$\bar{x} \in \text{ cluster-pts}\left\{x^\nu \,\middle|\, S^\nu(x^\nu) \ni u^\nu\right\} \Rightarrow S(\bar{x}) \ni \bar{u}$$

$$S(\bar{x}) \ni \bar{u} \Rightarrow \exists \hat{u}^\nu \to \bar{u} \text{ with } S^\nu(\hat{x}^\nu) \ni \hat{u}^\nu \text{ and } \hat{x}^\nu \to \bar{x}$$

$S^\nu \to_p S$ pointwise doesn't yield convergence of sol'ns

Applications: $\quad F(x) = b, \quad -G(x) \in N_C(x), \ldots$ variational problems

RATES OF CONVERGENCE

Excess distance function:

$$e_\rho(A,B) = \inf\left\{\eta \geq 0 \middle| A \cap \rho\mathbb{B} \subset B + \eta\mathbb{B}\right\}, \quad \rho > 0$$

Estimate of set distance:

$$\hat{dl}_\rho(A,B) = \max[\, e_\rho(A,B), e_\rho(B,A)\,]$$

Set-distance:

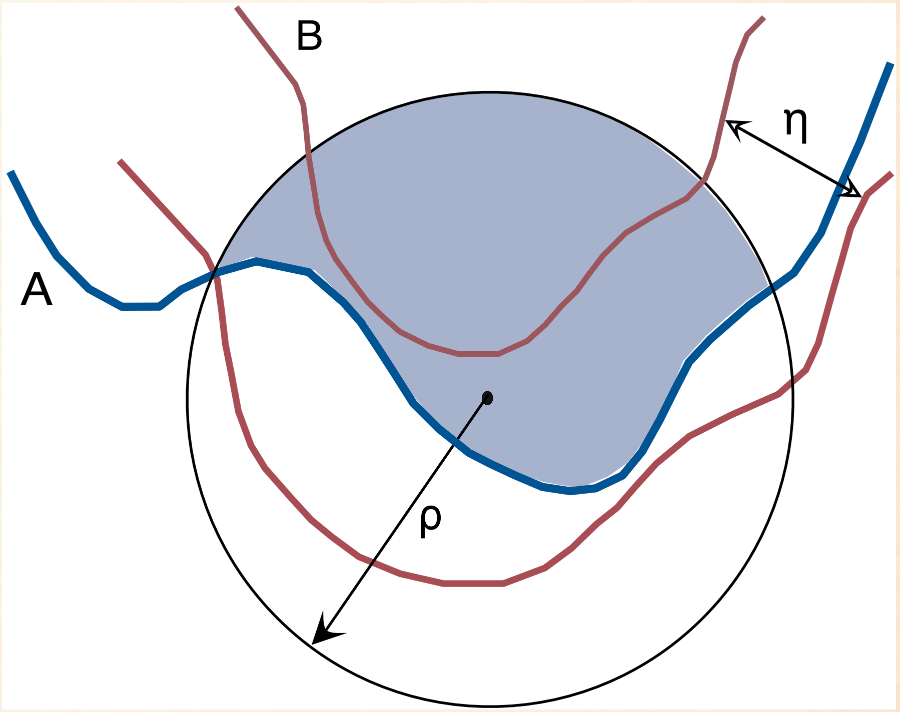
$$dl_\rho(A,B) = \max_{x\in\rho\mathbb{B}} \left| d(x,A) - d(x,B) \right|,$$

$$d(x,C) = \inf_{y\in C} \left| y - x \right|$$

Pompeiu-Hausdroff distance: $\rho = \infty$

$$\hat{dl}_\rho(A,B) \leq dl_\rho(A,B) \leq \hat{dl}_{\rho'}(A,B),$$

$$\rho' \geq 2\rho + \max[\, d(0,A), d(0,B)\,]$$



$$C^\nu \to C \Leftrightarrow dl_\rho(C^\nu,C) \to 0 \Leftrightarrow \hat{dl}_\rho(C^\nu,C) \to 0 \quad \forall \rho \geq 0$$

EPI-DISTANCE

$\text{lsc-fcns}(\mathbb{R}^n) =$ space of all lsc functions from $\mathbb{R}^n \to \overline{\mathbb{R}} = [-\infty, \infty]$

$\hat{dl}_\rho(f,g) = \hat{dl}_\rho(\text{epi}\, f, \text{epi}\, g), \quad dl_\rho(f,g) = dl_\rho(\text{epi}\, f, \text{epi}\, g), \; \rho \geq 0$

$$\mathbb{B}^{n+1} = \mathbb{B}^n \times [-1,1]$$

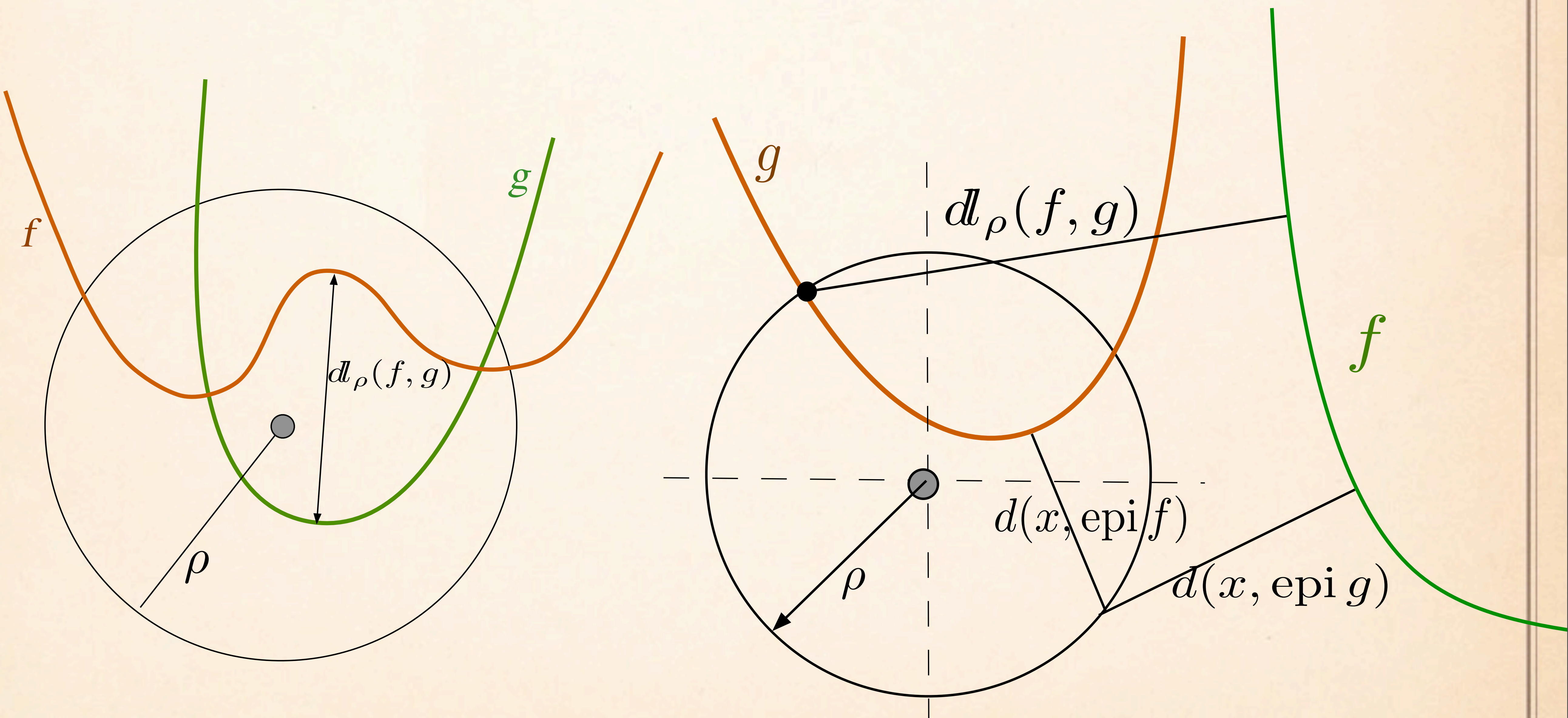
$dl(f,g) = \int_{\rho \geq 0} e^{-\rho} dl_\rho(f,g)\, d\rho,$ epi-distance

$f^\nu, f \in \text{lsc-fcns}(\mathbb{R}^n),\; f^\nu \to_e f \Leftrightarrow dl(f^\nu, f) \to 0$

also $dl_\rho(f^\nu, f) \to 0, \forall \rho \geq \overline{\rho} > 0, \dots$

$\left(\text{lsc-fcns}(\mathbb{R}^n) \setminus \{f \equiv \infty\}, dl\right)$ complete metric space

EPI-DISTANCE

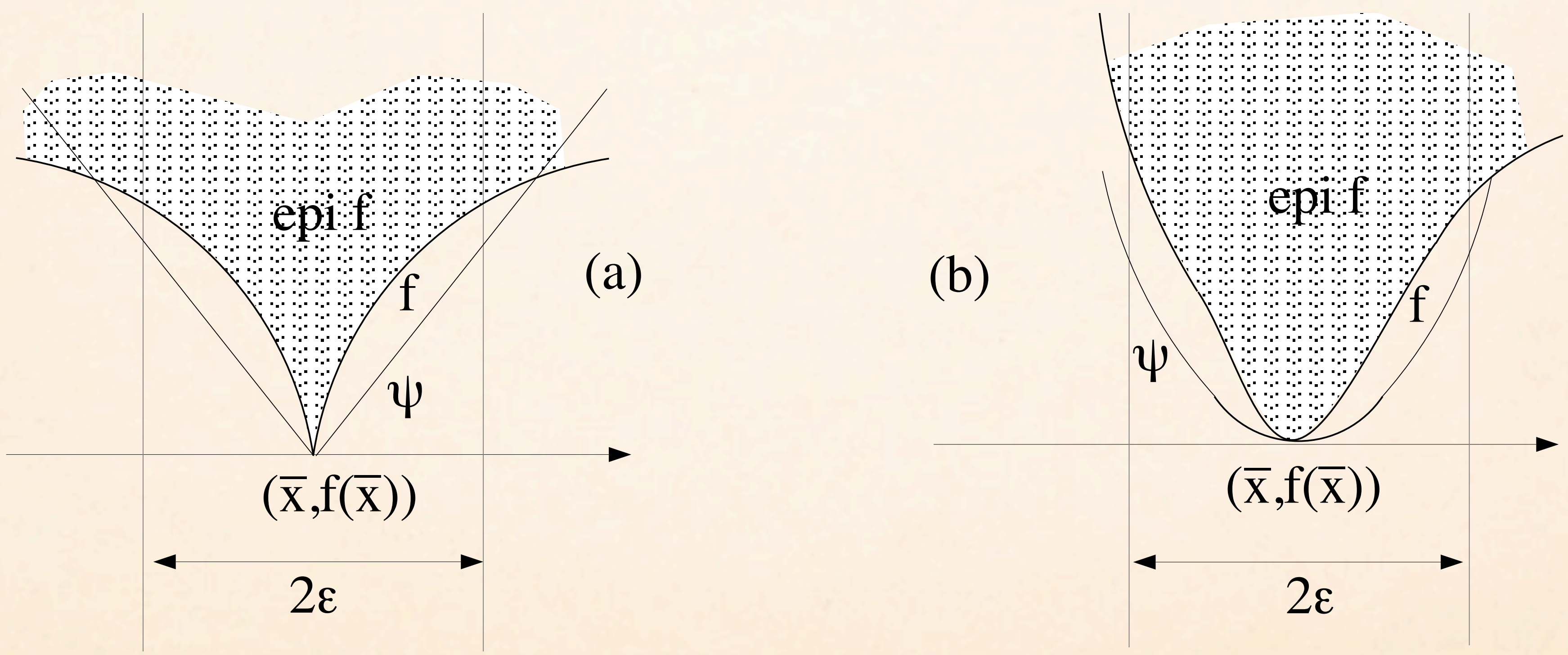


QUANTITATIVE ESTIMATE

under ψ-conditioning for f, $f, g \in \text{lsc-fcns}(\mathbb{R}^n)$, $\inf f, \inf g \in \mathbb{R}$

$$\left|\min_{\rho\mathbb{B}} g - \min f\right| \le dl_\rho(f,g)$$

$$\arg\min_{\rho\mathbb{B}} g \subset \arg\min f + \psi(dl_\rho(f,g))\mathbb{B}$$



APPROXIMATE SOLUTIONS: QUANTITATIVE ESTIMATE

$f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$, proper, lsc, convex functions

$\arg\min f, \arg\min g \neq \emptyset$

ρ_0 large enough so that $\rho_0\mathbb{B}$ meets $\arg\min f$ & $\arg\min g$

$\min f \geq -\rho_0$, $\min g \geq -\rho_0$

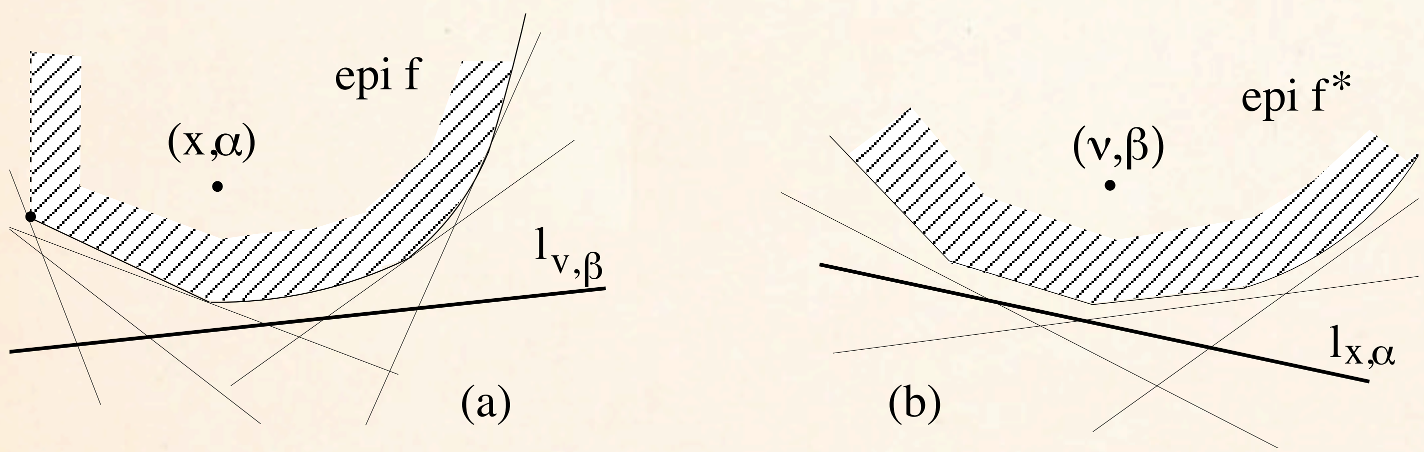
Then, with $\rho > \rho_0$, $\varepsilon > 0$, $\bar{\eta} = dl_\rho(f,g)$

$$d\hat{l}_\rho(\varepsilon\text{-arg min } f, \varepsilon\text{-arg min g}) \leq \bar{\eta}\left(1 + \frac{2\rho}{\bar{\eta} + \epsilon/2}\right)$$

$$\leq (1 + 4\rho/\epsilon)\, d\hat{l}_\rho(f,g)$$

CONVEX FUNCTIONS

(Wijsman) $f^\nu \underset{e}{\to} f \Leftrightarrow (f^\nu)^* \underset{e}{\to} f^* = \sup_x \big(\langle v,x\rangle - f(x)\big),\quad f^\nu$ lsc, convex



conjugate functions

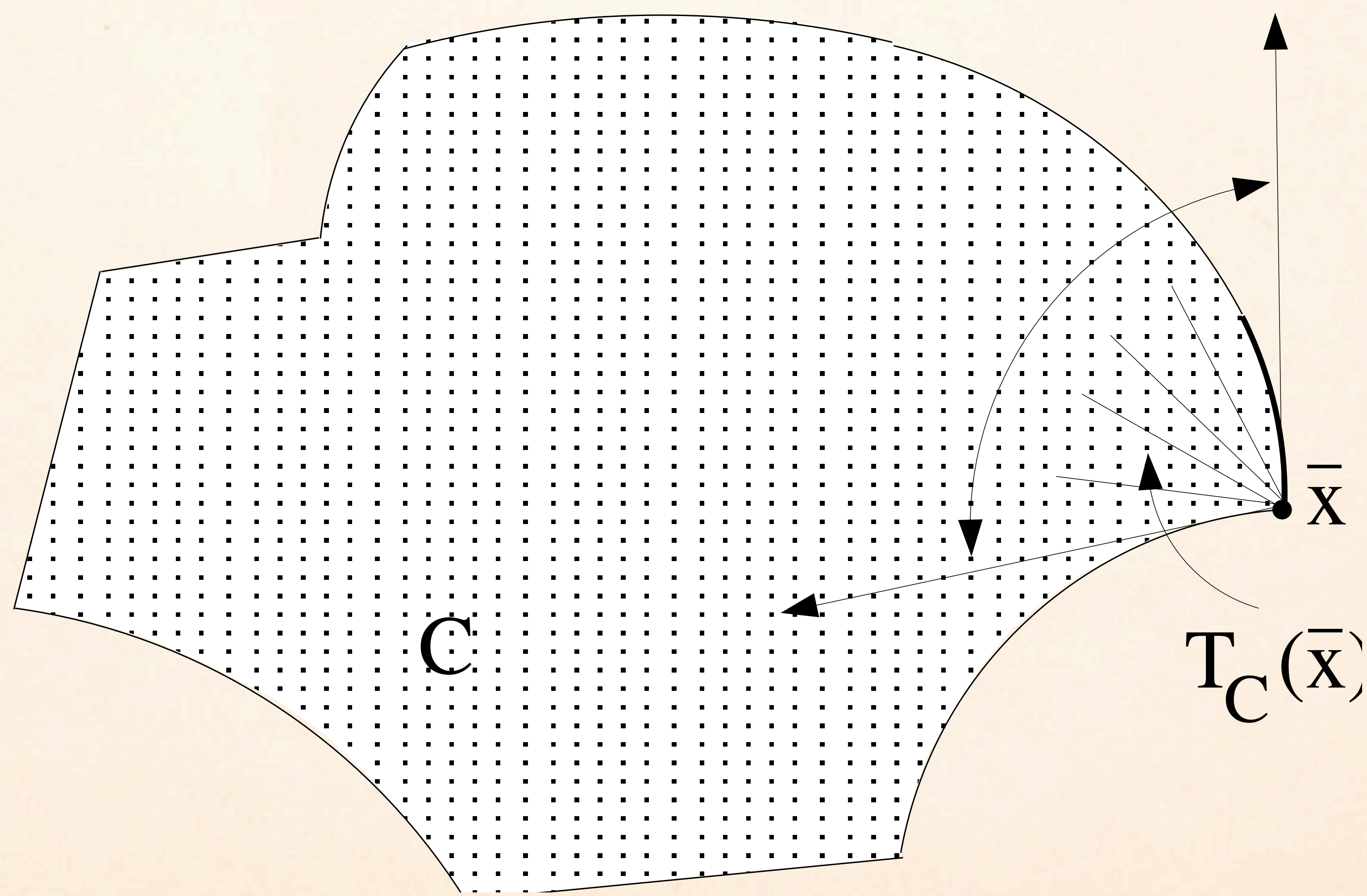
$f^\nu \underset{e}{\to} f \not\Rightarrow f^\nu \underset{p}{\to} f$ (pointwise) & $f^\nu \underset{p}{\to} f \not\Rightarrow f^\nu \underset{e}{\to} f$

$f^\nu \underset{e}{\to} f \equiv f^\nu \underset{p}{\to} f \Leftrightarrow \{f^\nu\}_{\nu\in\mathbb{N}}$ is equi-lsc

(Walkup-Wets) $dl_{csm}(f,g) = dl_{csm}(f^*,g^*)\quad [\approx dl(f,g) = dl(f^*,g^*)]$

VARIATIONAL GEOMETRY
TANGENT CONE

$w \in T_C(x)$, tangent to C at $x \in C$, if $\left| x^\nu - x \right| / \tau_\nu \to w$ for $x^\nu \underset{C}{\to} x, \tau_\nu \searrow 0$

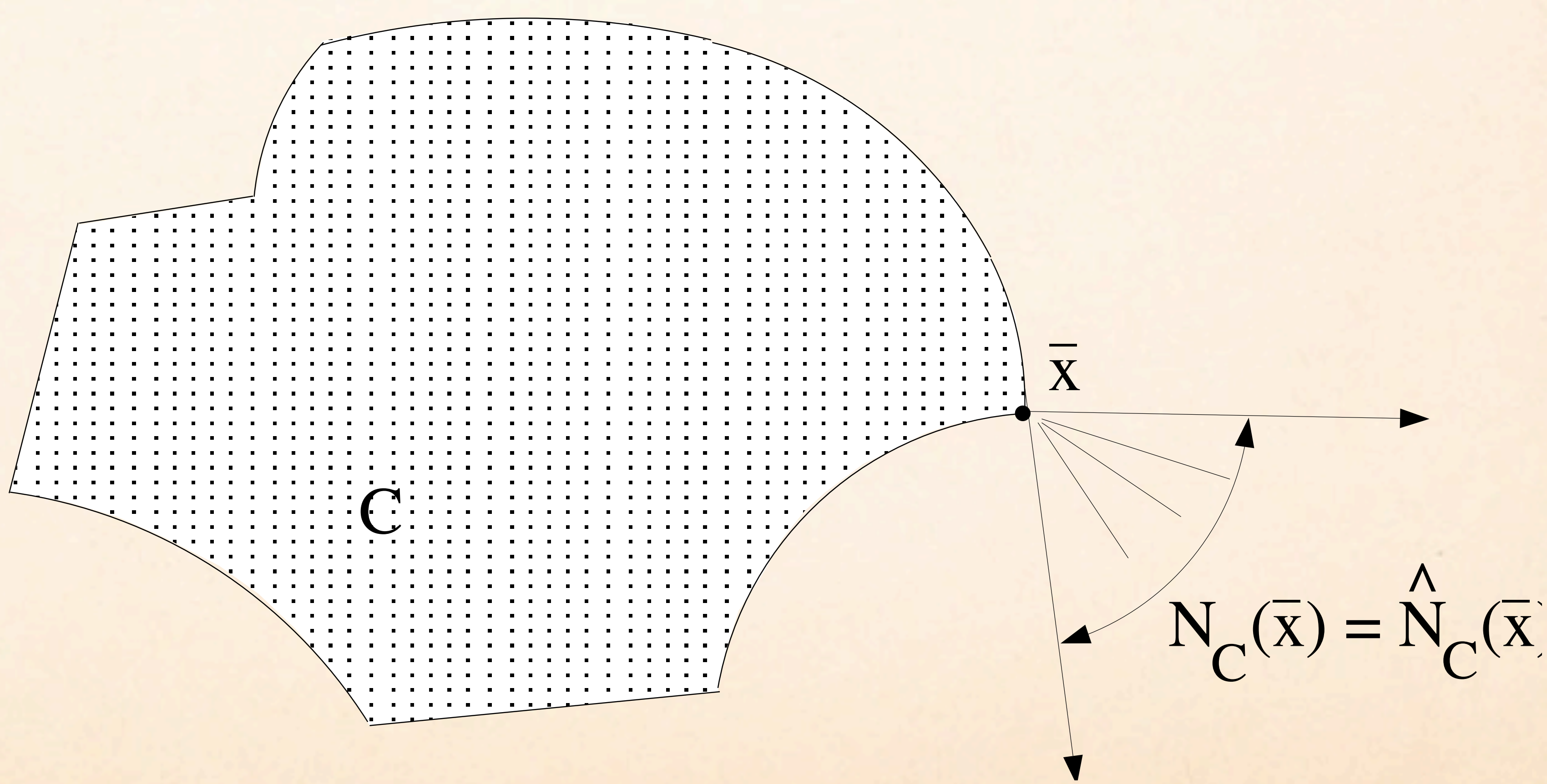


VARIATIONAL GEOMETRY
NORMAL CONE

$v \in \hat{N}_C(\bar{x})$, regular normal at $\bar{x} \in C$, if $\langle v, x - \bar{x} \rangle \leq o(|x - \bar{x}|), \forall x \in C$

$v \in N_C(\bar{x})$, normal at $\bar{x} \in C$, if $\exists x^\nu \underset{C}{\to} x$ and $v^\nu \to v$ with $v^\nu \in \hat{N}_C(x^\nu)$

normal cones: closed cones, $\hat{N}_C(\bar{x})$ convex



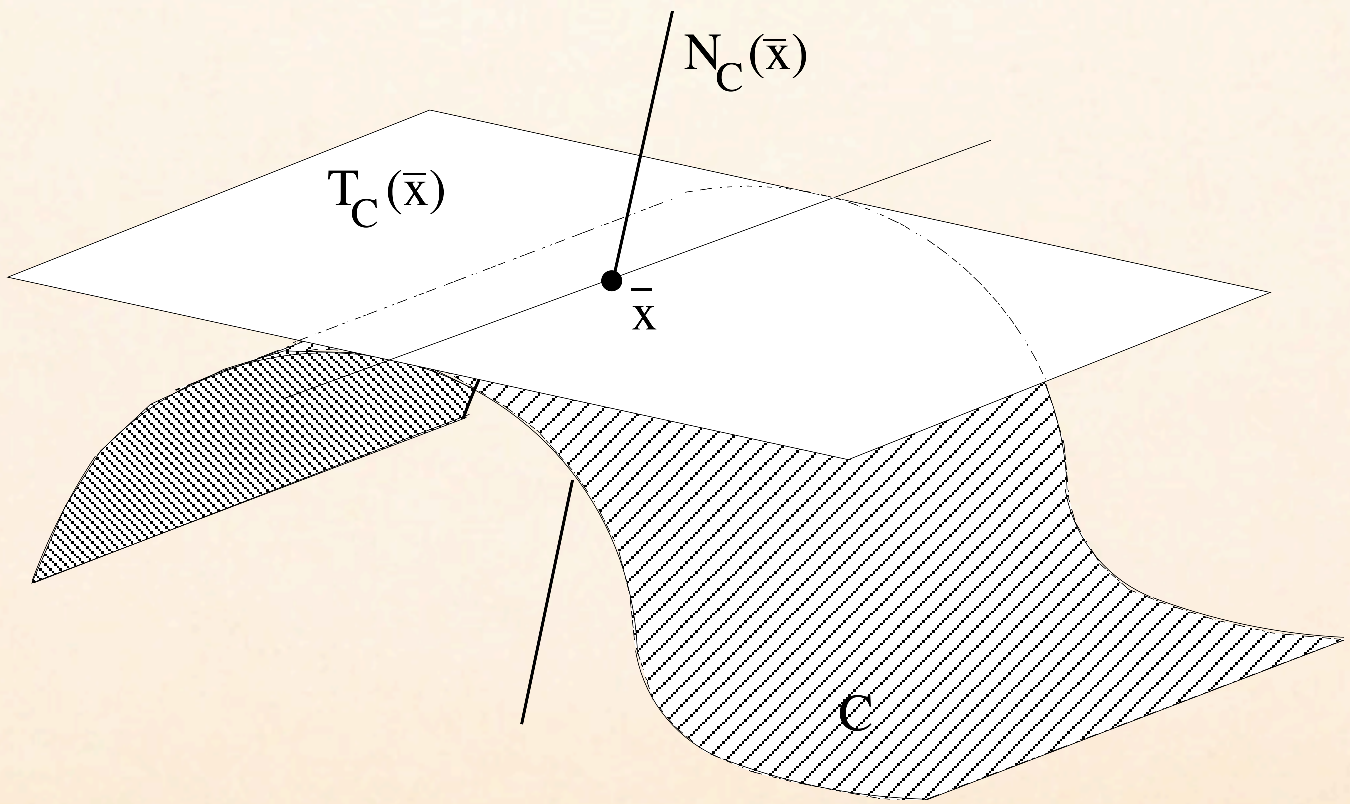
CLARKE REGULARITY

C Clarke regular at $\bar{x}$ if C locally closed & $N_C(x) = \hat{N}_C(\bar{x})$

which implies $N_C(\bar{x})$ is convex if C regular at $\bar{x}$

In general, $N_C(\bar{x}) = Ls_{x \to_C \bar{x}} \ N_C(x) \supset \hat{N}_C(\bar{x})$

Smooth manifolds and closed convex set are regular (also locally)



SUBGRADIENTS

$v \in \hat{\partial} f(\bar{x})$ regular subgradient if $f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle + o(|x - \bar{x}|)$

$$\hat{\partial} f(\bar{x}) = \left\{ v \,\middle|\, (v, -1) \in \hat{N}_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) \right\}, \text{ closed and convex}$$

$v \in \partial f(\bar{x})$ subgradient if $\exists\, x^\nu \to_f \bar{x}, v^\nu \in \hat{\partial} f(x^\nu)$ with $v^\nu \to v$

$$\partial f(\bar{x}) = \left\{ v \,\middle|\, (v, -1) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) \right\}, \text{ closed}$$

$x \mapsto \partial f(x)$ osc f-attentive convergence: $\Rightarrow \operatorname{Ls}_{x \to_f \bar{x}} \partial f(x) \subset \partial f(\bar{x})$

f differentiable at $\bar{x}$: $\hat{\partial} f(\bar{x}) = \nabla f(\bar{x}) = \partial f(\bar{x})$

f regular at $\bar{x}$: f locally lsc with $\partial f(\bar{x}) = \hat{\partial} f(\bar{x})$ (f locally convex, e.g)

$\partial \iota_C(x) = N_C(x)$ when C is convex

OPTIMALITY

$\min f = f_0 + \iota_C$, optimality: ``$0 \in \partial f(\bar{x})$''

generally, $\partial\big(f+g\big) \neq \partial f + \partial g$

$\mathbb{C}.\mathbb{Q}.$ (Constraint Qualification): $-N_C(\bar{x}) \cap \partial^\infty f_0(\bar{x}) = \{0\}$

$v \in \partial^\infty f_0(\bar{x}) =$ horizon subgradient if

$\exists\, x^\nu \to_f \bar{x}, v^\nu \in \hat{\partial} f(x^\nu), \lambda_\nu \searrow 0 \;\&\; \lambda_\nu v^\nu \to v$

with $\mathbb{C}.\mathbb{Q}.$ $\bar{x}$ locally optimal $\Rightarrow \partial f_0(\bar{x}) + N_C(\bar{x}) \ni 0$

f convex ($\Rightarrow$ regular), $\partial f_0(\bar{x}) + N_C(\bar{x}) \ni 0 \;\Rightarrow$

globally optimal (without $\mathbb{C}.\mathbb{Q}$)

ATTOUCH'S THEOREM

(initial proof: via Moreau envelopes)

$f^\nu, f : \mathbb{R}^n \to \overline{\mathbb{R}}$, proper, convex, lsc and $\lambda > 0$

The following are equivalent:

a) $f^\nu \to_e f$

b) the mappings $\partial f^\nu \to_g \partial f$ and

$\exists v^\nu \in \partial f^\nu(x^\nu), \bar{v} \in \partial f(\bar{x}), \left(x^\nu, v^\nu\right) \to (\bar{x}, \bar{v}), f^\nu(x^\nu) \to f(\bar{x})$

(convergence of an integration constant)

c) $P_\lambda f^\nu \to_p P_\lambda f = \arg\min_w \left\{ f(w) + \frac{1}{2\lambda}|w - \bullet|^2 \right\}$ and

$\exists\, \bar{x}, x^\nu \to \bar{x}$ such that $e_\lambda f^\nu(x^\nu) \to e_\lambda f(\bar{x})$

in situation b): also $f^{\nu *}(v^\nu) \to f^*(\bar{v})$

II. MOPEC

"Multi-Optimization Problems with Equilibrium Constraints"

THE MOPEC "FAMILY" ...

- saddle-point problems: Lagrangians, zero-sum games, Hamiltonians
- equilibrium: classical mechanics, Wardrop, economic (Walras, etc.)
- variational inequalities: finance, ecological models, complementarity, PDE
- non-cooperative games: pricing, generalized Nash equilibrium
- finding fixed points: Brouwer-type, Kakutani-type (set-valued), MPEC
- minimal surface problems, ... , mountain pass solutions, ...
- ... and the dynamic versions, and the **stochastic (dynamic)** versions.
- solving inclusions (equivalently, generalized equations): $S(x) \ni 0$

PRIMARY OBJECTIVE: CONSTRUCTIVE THEORY

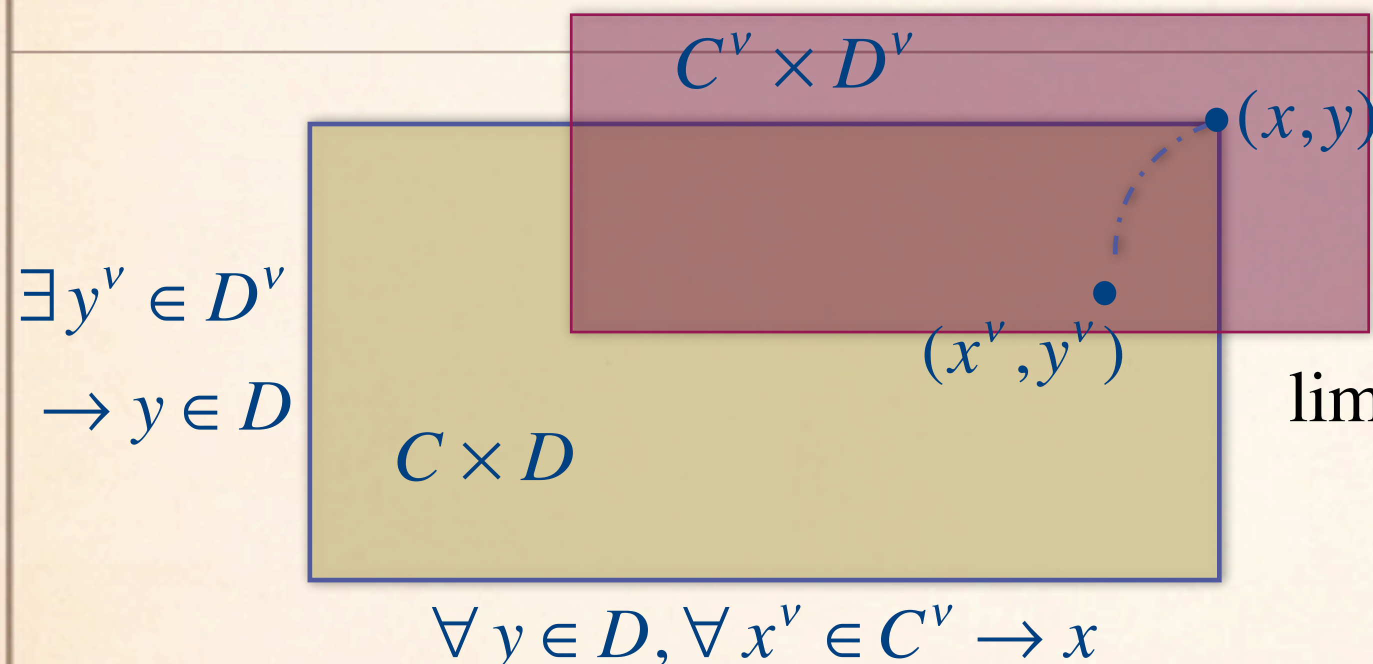
- Exhibits and exploits the interrelation between these problems
- Existence theory: (mostly, not exclusively)
  - Aubin & Ekeland, "Applied Nonlinear Analysis" (Chap. 6), 1984
  - Facchinei & Pang, "Finite Dimensional Variational Inequalities and Complementarity problems" (2003)
  - Iusem & Sosa (+ Kasay), "Existence of solutions to equilibrium problems" (2005-....)
- Approximation theory $\Rightarrow$ algorithmic strategies + existence

SADDLE FUNCTIONS EPI/HYPO CONVERGENCE

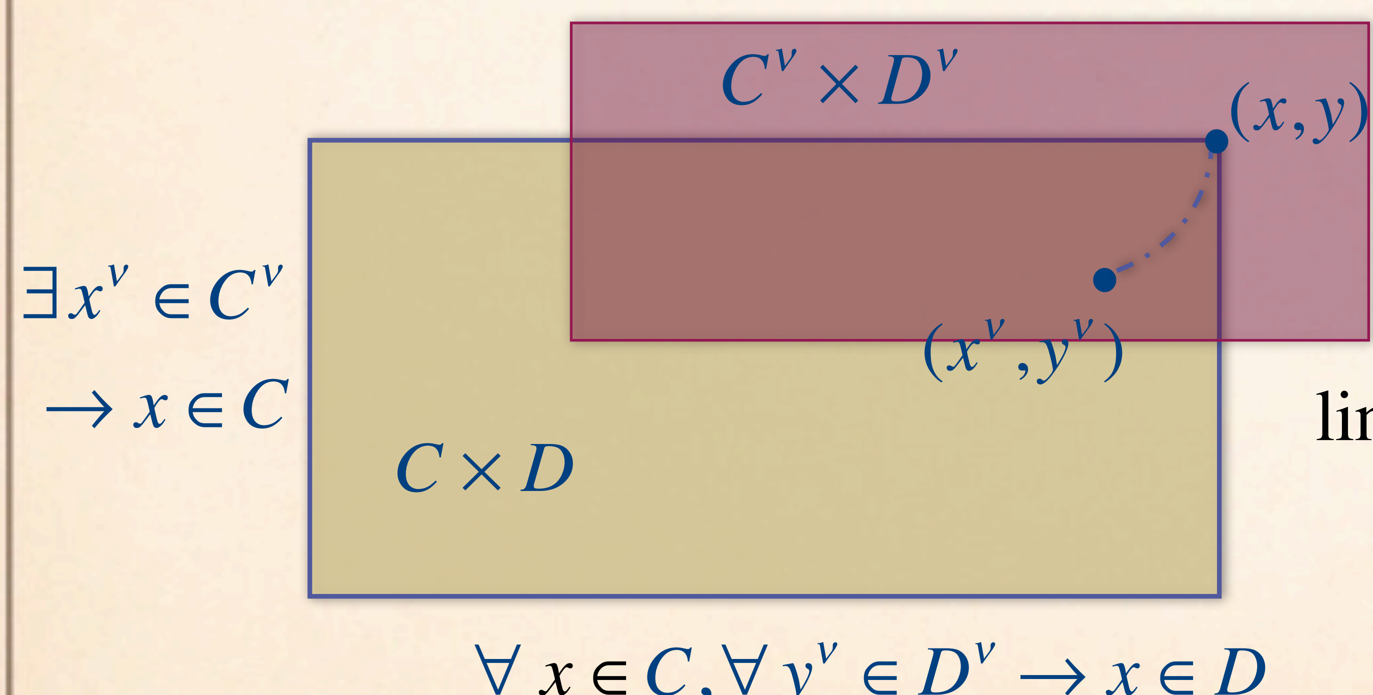
- *Lagrangians (concave/convex)*
- *zero-sum games*
- *Hamiltonians*

EPI/HYPO-CONVERGENCE

finite-valued bifunctions



$\limsup_\nu K^\nu(x^\nu, y^\nu) \le K(x,y)$ when $x \in C$



$\liminf_\nu K^\nu(x^\nu, y^\nu) \ge K(x,y)$ when $y \in D$

CONVERGENCE: SADDLE POINTS

$$K^\nu \xrightarrow{e/h} K : C \times D \to \mathbb{R}, \varepsilon_\nu \searrow 0, \quad \left(x^\nu, y^\nu\right) \in \ \varepsilon_\nu\text{-sdl}(K^\nu)$$

$$\left(\bar{x}, \bar{y}\right) = \lim_{\nu \in N \subset \mathbb{N}} \left(x^\nu, y^\nu\right), \ N \sim \text{ subsequence}$$

$$\Rightarrow \left(\bar{x}, \bar{y}\right) \in \text{sdl}(K) \quad \& \ K\left(\bar{x}, \bar{y}\right) = \lim_{\nu \in N \subset \mathbb{N}} K^\nu\left(x^\nu, y^\nu\right)$$

in the convex/concave case $\Rightarrow$ convergence primal/dual solutions

ancillary tight ($\sim$ y-compact): $\forall \varepsilon > 0, \exists\, B_\varepsilon$ compact, ν_ε

$$\forall \nu \geq \nu_\varepsilon, \sup_{B_\varepsilon \cap D^\nu} K^\nu(x^\nu, \bullet) \geq \sup_{D^\nu} K^\nu(x^\nu, \bullet) - \varepsilon$$

e/h-convergence + ancillary tight $\Rightarrow$ sv-convergence saddle points

ZERO-SUM GAMES

$x^* \in \arg\max_{x\in X} u(x,y^*), \quad y^* \in \arg\min_{y\in Y} u(x^*,y)$

$(x^*,y^*) \in \text{sdl}(u)$

if X,Y convex, compact ($\Rightarrow$ tight)

$\forall y, x \mapsto u(x,y)$ concave, usc, $\quad \forall x, y \mapsto u(x,y)$ convex, lsc

$\Rightarrow$ the zero-sum game $\mathcal{G} = \{(X,u),(Y,-u)\}$ has a solution

moreover, $X^\nu \to X, Y^\nu \to Y$, $u^\nu \underset{e/h}{\to} u$ (with same properties)

$\Rightarrow$ their solutions (x^ν, y^ν) cluster to solution of $\mathcal{G}$

also the case for approximate solutions

VARIATIONAL INEQUALITIES

- $G : C \to \mathbb{R}^n$,$C \subset \mathbb{R}^n$ non-empty, convex set
- find $\bar{u} \in C$ such that $-G(\bar{u}) \in N_C(\bar{u})$

$$v \in N_C(\bar{u}) \Leftrightarrow \langle v, u - \bar{u} \rangle \le 0, \forall u \in C$$

- let $C^v \to C, \quad G^v : C^v \to \mathbb{R}^{\mathrm{n}}$ continuous
- S^v solution set of approximating problems

  S solution of the limit problem. Does $S^v \to S$?

V.I.: THE GAP FUNCTION

- Let $K(u,v) = \langle G(u), v-u \rangle$ on dom $K = C \times C$
  - then $-G(\bar{u}) \in N_C(\bar{u})$ if and only if
  - $\bar{u} \in$ maxinf point of K with $K(\bar{u},\bullet) \geq 0$
- $K^\nu(u,v) := \langle G^\nu(u), v-u \rangle$, dom $K^\nu = C^\nu \times C^\nu$
  - $u^\nu \in \arg\max\text{–}\inf K^\nu$ with $K^\nu(u^\nu,\bullet) \geq 0$
  - $K^\nu \overset{?}{\to} K$ and …
  - $\bar{u} \in$ cluster points $\{u^\nu\} \Rightarrow ?\ \bar{u} \in \arg\min\text{–}\sup K$

NON-COOPERATIVE GAMES

- $a \in \mathcal{A}$, payoff: $u_a(x_a, x_{-a}) : \mathbb{R}^N \to \bar{\mathbb{R}}$, $\therefore$ includes $x_a \in C(x_{-a})$
- Generalized Nash equilibrium: $(\bar{x}_a, a \in \mathcal{A})$ such that
$$\forall a \in \mathcal{A},\ \bar{x}_a \in \arg\max u_a(x_a, \bar{x}_{-a})$$
- Nikaido-Isoda function:
$$N(x,y) = \sum_{a \in \mathcal{A}} u_a(x_a, x_{-a}) - \sum_{a \in \mathcal{A}} u_a(y_a, x_{-a})$$
- $\bar{x} = (\bar{x}_a, a \in \mathcal{A})$ is a Nash equilibrium
$$\Leftrightarrow \bar{x} \in \arg\max\inf N,\ \ N(\bar{x}, \bullet) \geq 0$$

APPROXIMATING GAMES

- Nikaido-Isoda functions of approximating games

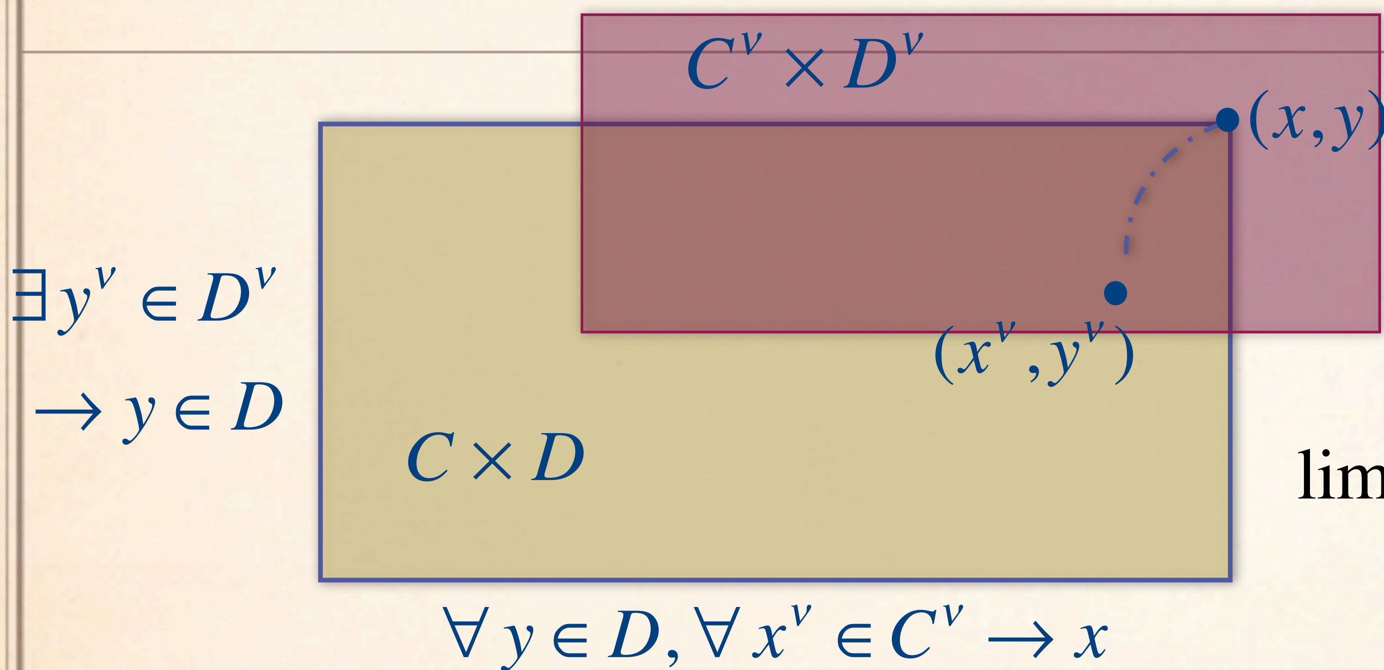
$$N^\nu(x,y) = \sum_{a\in A} u_a^\nu(x_a, x_{-a}) - \sum_{a\in A} u_a^\nu(y_a, x_{-a})$$

- $x^\nu \in \text{arg max–inf } N^\nu, \; \bar{x} \in$ cluster points $\{x^\nu\}$

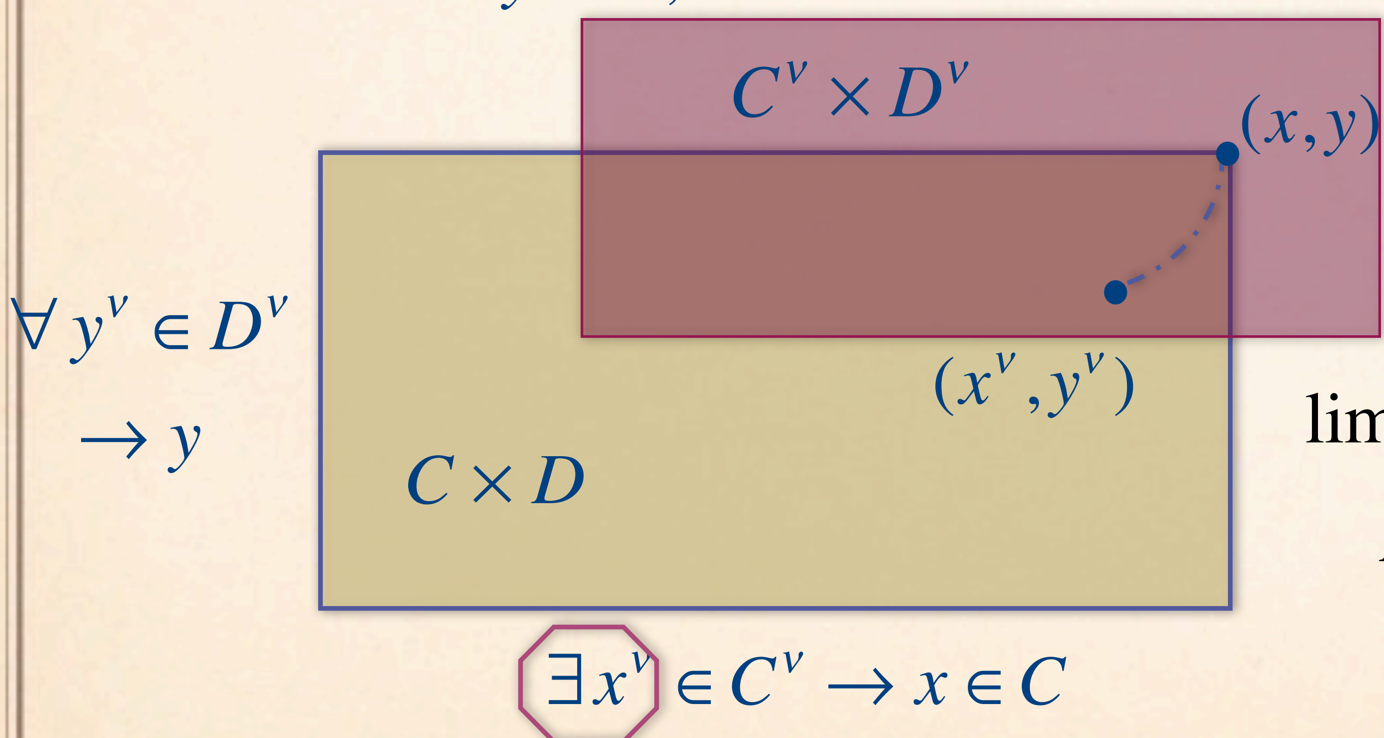
$$N^\nu \overset{?}{\longrightarrow} N \quad \text{and} \quad \ldots$$

- $\Rightarrow ? \; \bar{x} \in \text{arg max–inf } N \sim$ equilibrium point

LOPSIDED CONVERGENCE



$$\limsup_\nu K^\nu(x^\nu, y^\nu) \le K(x, y) \text{ when } x \in C$$



$$\liminf_\nu K^\nu(x^\nu, y^\nu) \ge K(x, y) \text{ when } y \in D$$

$$K^\nu(x^\nu, y^\nu) \to \infty \text{ when y} \notin D$$

ANCILLARY-TIGHTLY ~ 'COMPACT IN Y'

THM. $K^\nu_{C^\nu\times D^\nu} \to_{lop.} K_{C\times D}$ & ancillary-tightly,

$\bar{x} \in$ cluster points of $\{x^\nu \in \text{maxinf}\, K^\nu_{C^\nu\times D^\nu}\}_{\nu\in\mathbb{N}} \Rightarrow \bar{x} \in \text{maxinf}\, K_{C\times D}$

$K^\nu_{C^\nu\times D^\nu} \underset{lop\ ancillary-tight}{\longrightarrow} K_{C\times D}$ if $K^\nu_{C^\nu\times D^\nu} \underset{lop}{\to} K_{C\times D}$ and

(b) $\forall x \in C$, $\exists x^\nu \to x$, $\forall y^\nu \in D^\nu$ and $y^\nu \to y$:

$$\liminf K^\nu(x^\nu, y^\nu) \geq K(x,y) \quad \text{if } y \in D$$

$$K^\nu(x^\nu, y^\nu) \to \infty \quad \text{if } y \notin D$$

but also $\forall \varepsilon>0$, $\exists B_\varepsilon$ compact (depends on $x^\nu \to x$):

$$\inf_{B_\varepsilon \cap D^\nu} K^\nu(x^\nu, \bullet) \leq \inf_{D^\nu} K^\nu(x^\nu, \bullet) + \varepsilon, \ \forall \nu \geq \nu_\varepsilon$$

... EVEN BETTER : CONVERGENCE

$K^\nu_{C^\nu \times D^\nu} \to K_{C\times D}$ lop. ancillary-tightly,

(i) $x^\nu \in \varepsilon\text{-maxinf}\, K^\nu_{C^\nu \times D^\nu}$, $\bar{x}$ cluster point of $\{x^\nu\}_{\nu\in\mathbb{N}}$

$\Rightarrow \bar{x} \in \varepsilon\text{-maxinf}\, K_{C\times D}$

(ii) $x^\nu \in \varepsilon_\nu\text{-maxinf}\, K^\nu_{C^\nu \times D^\nu}$, $\bar{x}$ cluster point of $\{x^\nu\}_{\nu\in\mathbb{N}}$

$\&\ \varepsilon_\nu \searrow 0 \Rightarrow \bar{x} \in \text{maxinf}\, K_{C\times D}$ (special: unique)

(iii) $\bar{x} \in \text{maxinf}\, K_{C\times D} \Rightarrow \exists \varepsilon_\nu \searrow 0\ \&\ x^\nu \in \varepsilon_\nu\text{-maxinf}\, K^\nu_{C^\nu \times D^\nu}$

such that $x^\nu \to \bar{x}$,

Under tight-lop: convergence of the full ε_ν-maxinf sets

and convergence of values

KY FAN FCNS & INEQUALITY

$K: C \times C \to \mathbb{R}$ Ky Fan function if

(a) $\forall\, y \in C: x \mapsto K(x,y)$ usc on C

(b) $\forall\, x \in C: y \mapsto K(x,y)$ convex on C

K Ky Fan fcn, $\operatorname{dom} K = C \times C$ + C compact

$\Rightarrow \arg\max\!-\!\inf K \neq \emptyset$

if $K(x,x) \geq 0$ on $\operatorname{dom} K$, $\bar{x} \in \arg\max\!-\!\inf K$

$\Rightarrow \inf_y K(\bar{x}, y) \geq 0.$

Improvements: Iusem, Kasay, Sosa (locals)

Lignola, Nessah, Tian, X. Yu, ...

KY FAN'S INEQUALITY: AN EXTENSION

$$K^\nu \to K \text{ lopsided tightly with } C^\nu \to C,$$

$$K^\nu \text{ Ky Fan } \Rightarrow K \text{ Ky Fan fcn}$$

$$\& \text{ if } \forall \nu : \arg\max\text{–}\inf K^\nu \neq \emptyset$$

$$\bar{x} \in \text{ cluster-pts } \{\arg\max\text{–}\inf K^\nu\}$$

$$\Rightarrow \bar{x} \in \arg\max\text{–}\inf K \;\&\; K(\bar{x}, \bullet) \geq 0$$

Application: guideline for approximation schemes
truncations, coercivity, ...

LINEAR COMPLEMENTARITY PROBLEMS

LCP: find $z \geq 0$, $Mz + q \geq 0$ and $(Mz+q) \perp z$

$$K(z,v) = \langle Mz + q, v - z \rangle \text{ on } \mathbb{R}^n_+ \times \mathbb{R}^n_+, \text{ Ky Fan fcn}$$

approx. $z \in [0, r^\nu]$, $M^\nu z + q^\nu \geq 0$ and $(M^\nu z + q^\nu) \perp z$

$$K^\nu(z,v) = \langle M^\nu z + q^\nu, v - z \rangle \text{ on } [0, r^\nu] \times \mathbb{R}^n_+$$

- $K^\nu \to_{lop} K$ when $M^\nu \to M$, $q^\nu \to q$, $r^\nu \nearrow \infty$
- $K^\nu \to_{lop} K$ ancillary tightly when also

$$P^\nu = \{ z \in [0, r^\nu] \,|\, M^\nu z + q^\nu \geq 0 \} \to P = \{ z \geq 0 \,|\, Mz + q \geq 0 \}$$

$\Rightarrow$ cluster points of sol'ns of approx. solve LCP

$\left(\textit{note}: \operatorname{int} P \neq \emptyset, \text{ no row of } [M, q] = 0 \Rightarrow P^\nu \to P \right)$

- $K^\nu \to_{lop} K$ tightly (study of quadratic forms)

VARIATIONAL INEQUALITIES

- $-G(u) \in N_C(u)$, G continuous, C convex, compact
- bifunction: $K(u,v) = \langle G(u), v-u \rangle$ on $C \times C$, Ky Fan fcn & $K(u,u) \geq 0$
- THM: $C^\nu \to C \Rightarrow C^\nu$ compact $\nu \geq \bar{\nu}$, G^ν continuous

  $$G^\nu \to_{cont} G: G^\nu(x^\nu) \to G(x),\ \forall x^\nu \in C^\nu \to x$$

  $$K^\nu(u,v) = \left\langle G^\nu(u), v-u \right\rangle \text{ on dom } K^\nu = C^\nu \times C^\nu$$

  lop-converge ancillary tightly to $K \Rightarrow$ sol'ns converge

Continuous convergence (?):

$$\text{sol'ns } S^\nu = G^\nu + N_{C^\nu} \ni 0 \to \text{ sol'ns } S = G + N_C \ni 0$$

FIXED POINTS (SET-VALUED)

find $x \in C$ (convex) : $\quad x \in S(x),\ S : C \rightrightarrows C \subset \mathbb{R}^n$, osc (gph S closed)

$$K(x,v) = \sup\left\{\langle x - v, z - x\rangle \,\middle|\, z \in S(x) \subset C\right\}$$

K a Ky Fan fcn, convex in v, usc in x (sup-projection) $+\ K(x,x) \geq 0$

Approx. bifunctions: $K^v(x,v) = \sup\left\{\langle x - v, z - x\rangle \,\middle|\, z \in S^v(x) \subset C^v\right\}$

THM. $C^v \to C$, gph $S^v \to$ gph S (as sets), C compact. Then,

$\forall \varepsilon_v \searrow 0,\quad \bar{x} \in$ cluster points $\left\{x^v \in \varepsilon_v\text{-maxinf } K^v\right\}$ is a maxinf point of K,

i.e., a fixed point of S. (lop-convergence is tight)

an Application (J.S. Pang) - Cognitive radio multi-user game

$f : C \to C \subset \mathbb{R}^n$ continuous, C compact, convex, $\bar{x}$ fixed point

Pertubation (ε-enlargement): $S(\bullet;\varepsilon) : C \rightrightarrows C$, osc , $S(\bullet;0) = f$

For ε near 0: existence? $\exists\, x^\varepsilon \in S(x^\varepsilon,\varepsilon) = S^\varepsilon(x),\ x^\varepsilon \to \bar{x}$?

LOP- & EPI/HYPO-CONVERGENCE

1. $L^\nu \underset{lop}{\to} L \not\Rightarrow L^\nu \underset{e/h}{\to} L$

2. $L^\nu \underset{e/h}{\to} L$ & convex-concave $\Rightarrow$ $L^\nu \underset{lop}{\to} L$

3. epi/hypo- = hypo/epi-convergence

4. $L^\nu \underset{e/h}{\to} L \Rightarrow$ convergence of saddle points

   $\Rightarrow$ convergence of approximate sadde points

   (without ancillary tightness)

5. Existence requires tightness-conditions (~coercivity, e.g.)

UNIQUENESS OF
LOP- & EPI/HYPO LIMITS

CONVINCING EXAMPLES (?)

- Lagrangians: $L^\nu(x,y) = f_0^\nu(x) + \sum_{i=1}^{m} y_i f_i^\nu(x)$ on $X^\nu \times \left(\mathbb{R}^s \times \mathbb{R}^{m-s}\right)$
- Lopsided convergence (maxmin-framework) - sufficient conditions $f_0^\nu, f_1^\nu, \ldots, f_m^\nu$ hypo-converge to $f_0, f_1, \ldots, f_m$ on $X^\nu \to X$
  - $\left\{ f_i^\nu, \nu \in \mathbb{N} \right\}$ is equi-usc, $i = 0, \ldots, m$
  - Constraint Qualification: $S^\nu = \left\{ x \,\middle|\, f_i^\nu \geq 0, i = 1, \ldots, m \right\} \to S$
  - concave-convex case (epi/hypo): int $S \neq \emptyset$
- lop-limit L is unique

VARIATIONAL INEQUALITIES

$C^\nu \to C, \quad G^\nu : C^\nu \to \mathbb{R}^n$ continuous, C^ν convex

$-G^\nu(x) \in N_{C^\nu}, \ \nu \in \mathbb{N}$

THM: $C^\nu \to C \Rightarrow C$ compact $\nu \geq \bar{\nu}$, G^ν continuous

$G^\nu \to_{cont} G : G^\nu(x^\nu) \to G(x), \ \forall x^\nu \in C^\nu \to x$

$K^\nu(u,v) = \left\langle G^\nu(u), v-u \right\rangle$ on $\operatorname{dom} K^\nu = C^\nu \times C^\nu$

lop-converge ancillary tightly to $K \Rightarrow$ sol'ns converge

lop-limit: $-G(x) \in N_C(x)$ uniquely determined

MPEC (GENERALIZED?)

$\max g(x)$ such that $x \in S(x)$, g continuous, $S : C \rightrightarrows C$ convex

bifunction: $K(x,v) = g(x) + \sup_z \{\langle x - v, z - x \rangle \mid z \in S(x)\}$

V.I.-constraint: $S(x) = N_C(x) + G(x) + Ix$ on C

LCP: $S(x) = \langle Mx + q, v - x \rangle + Ix$ on $\mathbb{R}^n_+$

$\bar{x} \in \arg\max\text{-}\inf K \Rightarrow \bar{x}$ solves MPEC.

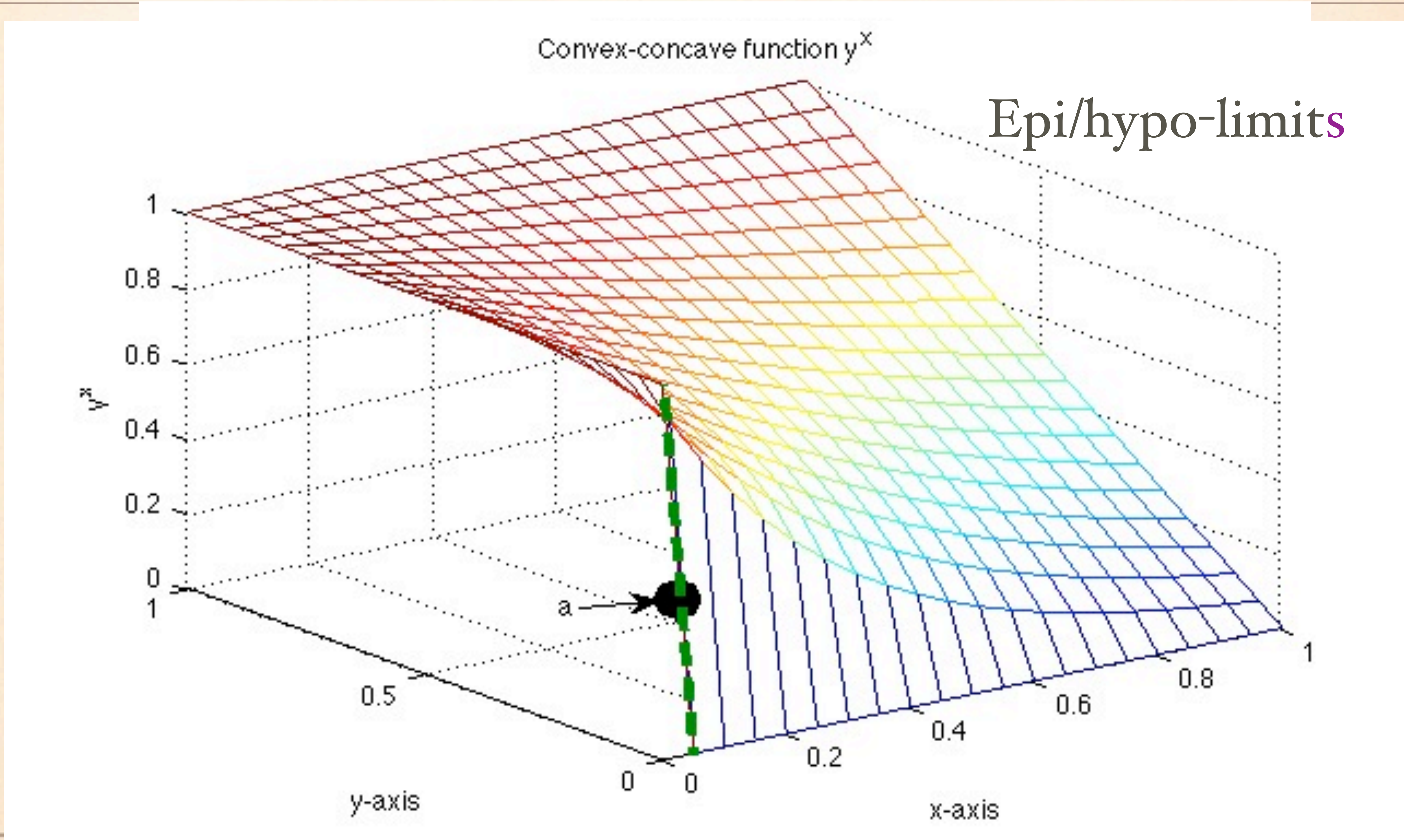
approximating bifunctions: $S^v : C^v \rightrightarrows C$

$K^v(x,v) = g^v(x) + \sup_z \{\langle x - v, z - x \rangle \mid z \in S^v(x)\}$

$C^v \to C$, gph $S^v \to$ gph S, g^v hypo-converges to g

then $K^v \underset{lop}{\to} K$ & K unique

$K^{\nu}(x,y) \equiv y^{x}$ Uniqueness fails!



WALRAS EQUILIBRIUM

- $\forall a \in A$: $d_a(p) \in \arg\max\left\{u_a(x_a) \,\middle|\, \langle p, x_a\rangle \le \langle p, e_a\rangle\right\}$

  $s(p) = \sum_a (e_a - d_a(p))$ excess supply

- find $\bar{p} \in \Delta$ (unit simplex) so that $s(\bar{p}) \ge 0$

- Walrasian: $W(p,q) = \langle q, s(p)\rangle$ Ky Fan fcn

- $\bar{p} \in \operatorname{maxinf} W, W(\bar{p}, \bullet) \ge 0 \Leftrightarrow s(\bar{p}) \ge 0$

- conditions: $u_a^\nu \to_{hypo} u_a,\ e_a^\nu \to e_a \in \text{int dom } u_a \Rightarrow$

- Convergence: W^ν lop-converges ancillary-tight to W

AUGMENTED WALRASIAN

$W(p,q) = \langle q, s(p)\rangle$ on $\Delta \times \Delta$

$\tilde{W}_r(p,q) = \sup_z \left\{ W(p,z) \,\middle|\, \|z-q\|^o \le r \right\}$ ** lop-converges

$$q^{k+1} = \arg\max_{q\in\Delta} \left[\max_z \langle z, s(p^k)\rangle \,\middle|\, \|z-q\|^o \le r_k \right]$$

minimizing a linear form on a ball

reduces to finding the largest element of $s(p^k)$

$$p^{k+1} = \arg\min_{p\in\Delta} \left[\max_z \langle z, s(p)\rangle \,\middle|\, \|z-q^{k+1}\|^o \le r_{k+1} \right]$$

as $r_k \nearrow \infty, p^k \to \bar{p}$ (local quad. approx. Nocedal, Powell)

experiments: 10 agents, 150 goods (easy!)

III. RANDOM SETS AND MAPPINGS

RANDOM CLOSED SETS

$(\Xi, \mathcal{A}, \mu)$, $\Xi \subset \mathbb{R}^N$ $\quad \rightrightarrows$: set-valued mapping,

$C : \Xi \rightrightarrows \mathbb{R}^d, C(\xi) \subset \mathbb{R}^d$ closed set for all $\xi \in \Xi$

& $C^{-1}(O) = \{\xi \,|\, C(\xi) \cap O \neq \emptyset\} \in \mathcal{A}$, $\forall O \subset \mathbb{R}^n$,open (measurability)

$\Rightarrow$ dom $C = C^{-1}(\mathbb{R}^d) \in \mathcal{A}$

==

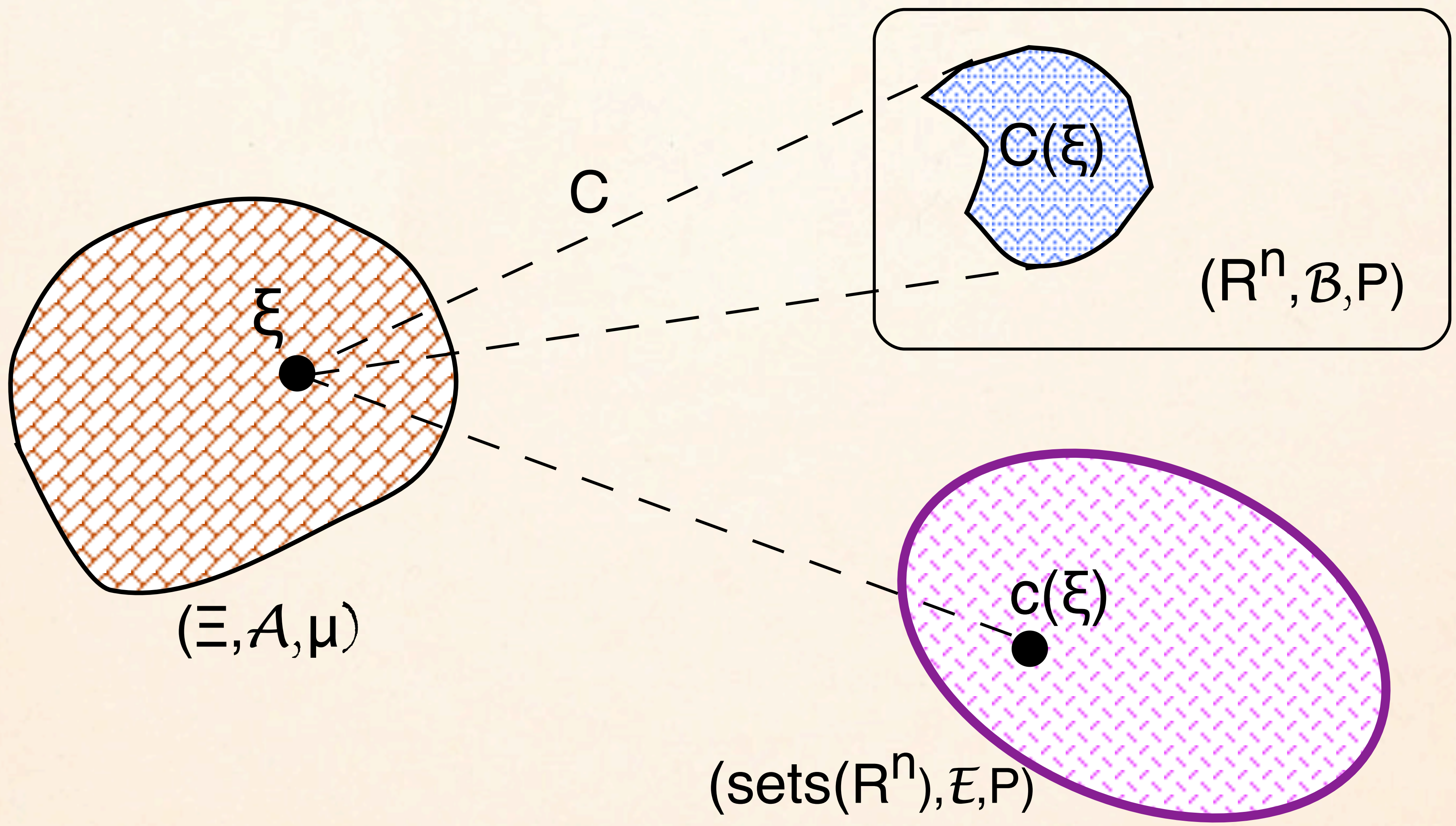
$c : \Xi \to$ sets$(\mathbb{R}^d)$, $c(\xi) \sim C(\xi)$, $\mathcal{F}_O = \{F \subset \mathbb{R}^d \text{ closed} \,|\, F \cap O \neq \emptyset\}$

$\left(\text{sets}(\mathbb{R}^n), \mathcal{E}\right)$, $\mathcal{E}$ Effros field $= \sigma - \{\mathcal{F}_O \in \text{sets}(\mathbb{R}^n), O \text{ open}\}$,

C measurable $\Leftrightarrow c$ measurable $[c^{-1}(\mathcal{F}_O) \in \mathcal{A}]$

$\mathcal{E} = \mathcal{B}$ Borel field ($\mathbb{R}^d$ separable metric space)

SET- & SINGLE-VALUED



CASTAING REPRESENTATION & GRAPH-MEASURABILITY

- a random closed set C always admits a measurable selection!
- (with dom C measurable) C is a random closed set $\Leftrightarrow$ it admits a Castaing representation: $\exists$ a countable family

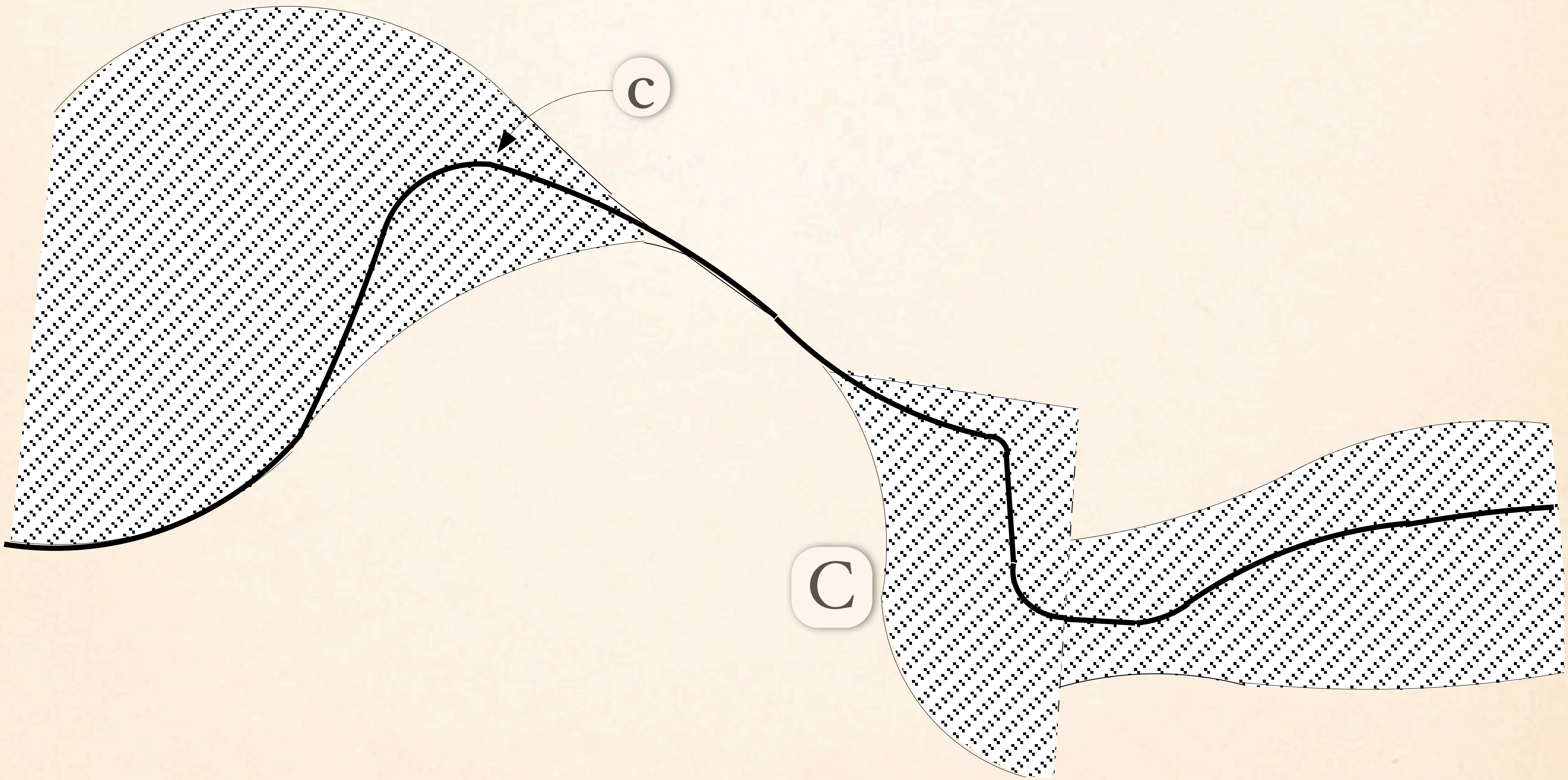
$$\left\{x^\nu : \text{ dom } C \to \mathbb{R}^m, \text{meas.-selections}\right\}$$

$$\text{cl} \bigcup_{\nu \in \mathbb{N}} x^\nu = C(\xi), \forall \xi \in \text{ dom } C \subset \Xi$$

- $(\Xi, \mathcal{A})$ μ-complete for some μ,

  C random set $\Leftrightarrow$ gph C $\mathcal{A} \otimes \mathcal{B}^n$-measurable

MEASURABLE SELECTION



SET-CONVERGENCE TOPOLOGY

$\mathcal{F} =$ cl-sets($\mathbb{R}^d$), all closed subsets of $\mathbb{R}^d$

$\mathcal{F}^D =$ subsets $\mathbb{R}^d$ that miss $D = \{F \cap D = \emptyset\}$

$\mathcal{F}_D =$ subsets $\mathbb{R}^d$ that hit $D = \{F \cap D \neq \emptyset\}$

Hit-and-miss topology (= τ_f Fell topology)

subbase: $\{\mathcal{F}^K \,|\, K \text{ compact}\}$ & $\{\mathcal{F}_O \,|\, O \text{ open}\}$

$\mathbb{B}(x,\rho)$ closed ball, center x radius ρ, $\mathbb{B}^o(x,\rho)$ open

~subbase $\left\{\mathcal{F}^{\mathbb{B}(x,\rho)}, \mathcal{F}_{\mathbb{B}^o(x,\rho)} \;\middle|\; x \in \mathbb{Q}^d, \rho \in \mathbb{Q}_{++}\right\}$

countable base: $\left\{\mathcal{F}^{\mathbb{B}(x^1,\rho_1)\cup\ldots\cup\mathbb{B}(x^r,\rho_r)} \cap \mathcal{F}_{\mathbb{B}^o(x^1,\rho_1)\cup\ldots\cup\mathbb{B}^o(x^s,\rho_s)}\right\}$

$(\mathcal{F} = \text{cl-sets}(\mathbb{R}^d), \tau_f)$ compact, metrizable space

A.S.-CONVERGENCE

- $\{C^\nu : \Xi \rightrightarrows \mathbb{R}^d, \nu \in \mathbb{N}\}$ random closed sets
- a.s. convergence: $\text{Li}_\nu(C^\nu)$ & $\text{Ls}_\nu(C^\nu)$ random closed sets
  $C^\nu \to C$ a.s. $\Rightarrow C$ random closed set on $\Xi_0, \mu(\Xi_0) = 1$
- $C^\nu \to C$ μ-a.s. and dom $C^\nu =$ dom C. Then,
  $\exists$ Castaing representations of $C^\nu \to$ a Castaing representation of C
  If $x : \Xi \to \mathbb{R}^d$ is a measurable selection of C, then
  $\exists x^\nu : \Xi \to \mathbb{R}^d$ selections of C^ν converging μ-a.s. to x
- 'Egorov's Theorem': $C^\nu \to C$ μ-a.s. $\Leftrightarrow C^\nu \to C$ almost uniformly

CONVERGENCE IN PROBABILITY

Let $\varepsilon^o C = \left\{ x \in \mathbb{R}^m \middle| d(x,C) < \epsilon \right\}$, C^ν, C random sets

$\Delta_{\varepsilon,\nu} = \left(C^\nu \setminus \epsilon^o C \right) \cup \left(C \setminus \epsilon^o C^\nu \right)$

μ-a.s. convergence: $\mu\left\{ \xi \middle| C^\nu(\xi) \to C(\xi) \right\} = 1$

in probability: $\mu\left[\Delta^{-1}_{\varepsilon,\nu}(K) \right] \to 0, \forall \epsilon > 0, K \in \mathcal{K}$

C^ν converges to C in probability

$\Leftrightarrow dl(C^\nu, C) \to 0$ in probability

$\Leftrightarrow$ every subsequence of $\{C^\nu\}_{\nu \in \mathbb{N}}$

contains a sub-subsequence converging μ-a.s to C

i.e., in probability $\Rightarrow$ in distribution $\left[\int h(\xi)\, dl(C^\nu(\xi), C(\xi)) \mu(d\xi) \to 0 \right]$

DISTRIBUTION OF A RANDOM SET

Borel σ-field: $\mathcal{B} = \sigma\text{-}\left\{\mathcal{F}^K \,|\, K \text{ compact}\right\}$ or $\sigma\text{-}\left\{\mathcal{F}_O \,|\, O \text{ open}\right\}\ldots$

Distribution $(P, \mathcal{B})$ regular, $\quad \mathcal{K}$ compact subsets $\mathbb{R}^d$

determined by values on $\left\{\mathcal{F}^K \,|\, K \in \mathcal{K}\right\}$ or $\left\{\mathcal{F}_K \,|\, K \in \mathcal{K}\right\}$

Distribution function (Choquet capacity):

$T : \mathcal{K} \to [0,1], T(\emptyset) = 0$ and $\forall \left\{K^\nu, \nu \in \{0\} \cup \mathbb{N}\right\} \subset \mathcal{K}$:

a) $T(K^\nu) \searrow T(K)$ when $K^\nu \searrow K \qquad$ (~ usc on $\mathbb{R}$)

b) $\left\{D_\nu : \mathcal{K} \to [0,1]\right\}_{\nu \in \mathbb{N}}$ where $D_0(K^0) = 1 - T(K^0)$

$D_1(K^0; K^1) = D_0(K^0) - D_0(K^0 \cup K^1)$ and for $\nu = 2, \ldots$

$D_\nu(K^0; K^1, \ldots, K^\nu) = D_{\nu-1}(K^0; K^1, \ldots, K^{\nu-1}) - D_{\nu-1}(K^0 \cup K^\nu; K^1, \ldots, K^{\nu-1})$

(~ monotonicity on $\mathbb{R}$)

EXISTENCE-UNIQUENESS T

P on $\mathcal{B}$ determines a unique distribution function T on $\mathcal{K}$

$$T(K) = P(\mathcal{F}_K)$$

$$D_\nu(K^0; K^1, \ldots, K^\nu) = P(\mathcal{F}^{K^0} \cap \mathcal{F}_{K^1} \cap \cdots \cap \mathcal{F}_{K^\nu})$$

T on $\mathcal{K}$ determines a unique probability measure P.

Proof. via Choquet Capacity Theorem (Matheron)

probabilistic arguments (Salinetti-Wets)

$C : \Xi \rightrightarrows \mathbb{R}^d$ a random closed set

$(P, \mathcal{B})$ induced probability measure:

$$P(\mathcal{F}_G) = \mu\left[C^{-1}(G)\right] \quad \forall G \in \mathcal{B}, \quad T(K) = \mu\left[C^{-1}(K)\right] \quad \forall K \in \mathcal{K}$$

CONVERGENCE IN DISTRIBUTION

random sets C^ν converge in distribution to C when

induced P^ν narrow-converge to $P : P^\nu \to_n P$

$\Leftrightarrow T^\nu \to_p T$ on $\mathcal{K}_{T\text{-cont}}$ (convergence of distribution functions)

what is $\mathcal{K}_{T\text{-cont}}$?

a) $\forall C^\nu, \nu \in N, \exists$ converging subsequence (pre-compact)

b) $K^\nu \nearrow K = \text{cl} \bigcup_\nu K^\nu$ regularly if $\text{int}\, K \subset \bigcup_\nu K^\nu$

c) distribution (fcn) continuity: $\lim_\nu T(K^\nu) = T(\text{cl} \bigcup_\nu K^\nu)$

d) convergence $T^\nu \to_p T$ on C_T continuity set $\Rightarrow P^\nu \to_n P$

e) $P^\nu \to_n P \Leftrightarrow T^\nu \to_p T$ on $C_T^{ub} = C_T \cap \mathcal{K}^{ub}$

$\mathcal{K}^{ub} =$ finite union of rational ball, positive radius

f) $\varepsilon \mapsto T(K + \varepsilon\mathbb{B})$: countable number of distontinuities

A DETOUR ABOUT RATES

$T^\nu \to_p T$ on $C_T \Leftrightarrow P^\nu \to_n P$ (Polish space (E,d))

P^ν, P defined on $\mathcal{B}$

probability sc-measures on cl-sets(E): λ

(i) $\lambda \geq 0$, (ii) $\lambda \nearrow$ $\lambda(C^1) \leq \lambda(C^2)$ if $C^1 \subset C^2$

(iii) λ is τ_f-usc on cl-sets(E), (iv) $\lambda(\emptyset) = 0, \lambda(E) = 1$

(iv) λ modular: $\lambda(C^1) + \lambda(C^2) = \lambda(C^1 \cup C^2) + \lambda(C^1 \cap C^2)$

P and $\lambda = P_{\text{cl-sets}}$ define each other uniquely (E complete $\Rightarrow$ tight)

$\{P^\nu, \nu \in \mathbb{N}\}$ tight: $P^\nu \to_n P \Leftrightarrow \lambda^\nu \to_h \lambda$ $(\sim -\lambda^\nu \to_e -\lambda)$ on cl-sets(E)

tightness ~ equi-usc of $\{\lambda^\nu\}_{\nu \in \mathbb{N}}$ at $\emptyset$

rates: $dl(\lambda^\nu, \lambda) \to 0$ (for $\mathbb{R}$-valued r.v., related to Skorohod distance)

RANDOM SET: EXPECTATION

$$EC = \mathbb{E}\{C(\xi)\} = \left\{ \int_\Xi x(\xi)\,\mu(d\xi) \,\middle|\, x(\bullet)\ \mu\text{-summable selection} \right\}$$

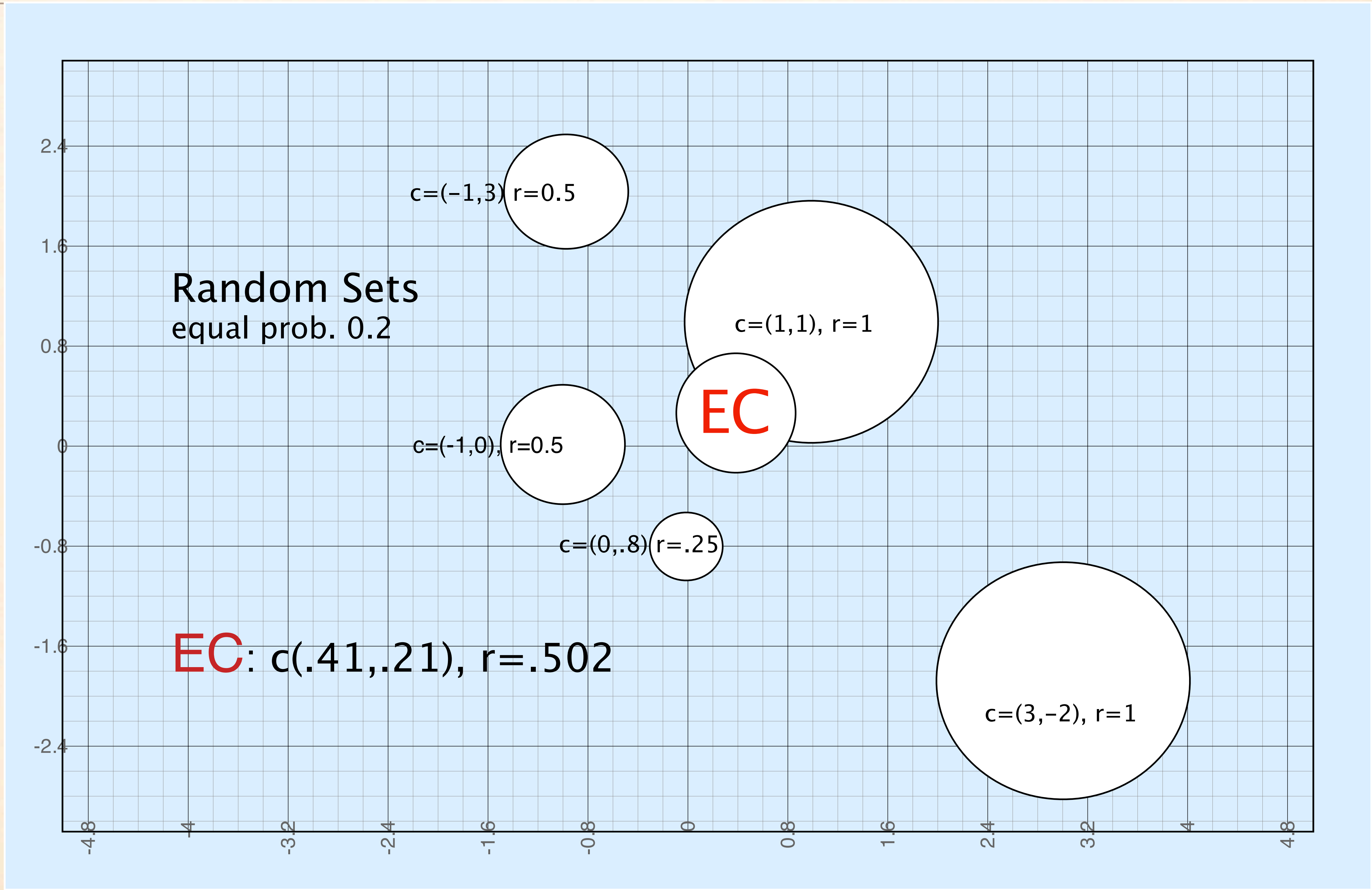
..not necessarily closed even when C is closed-valued

Convexity.

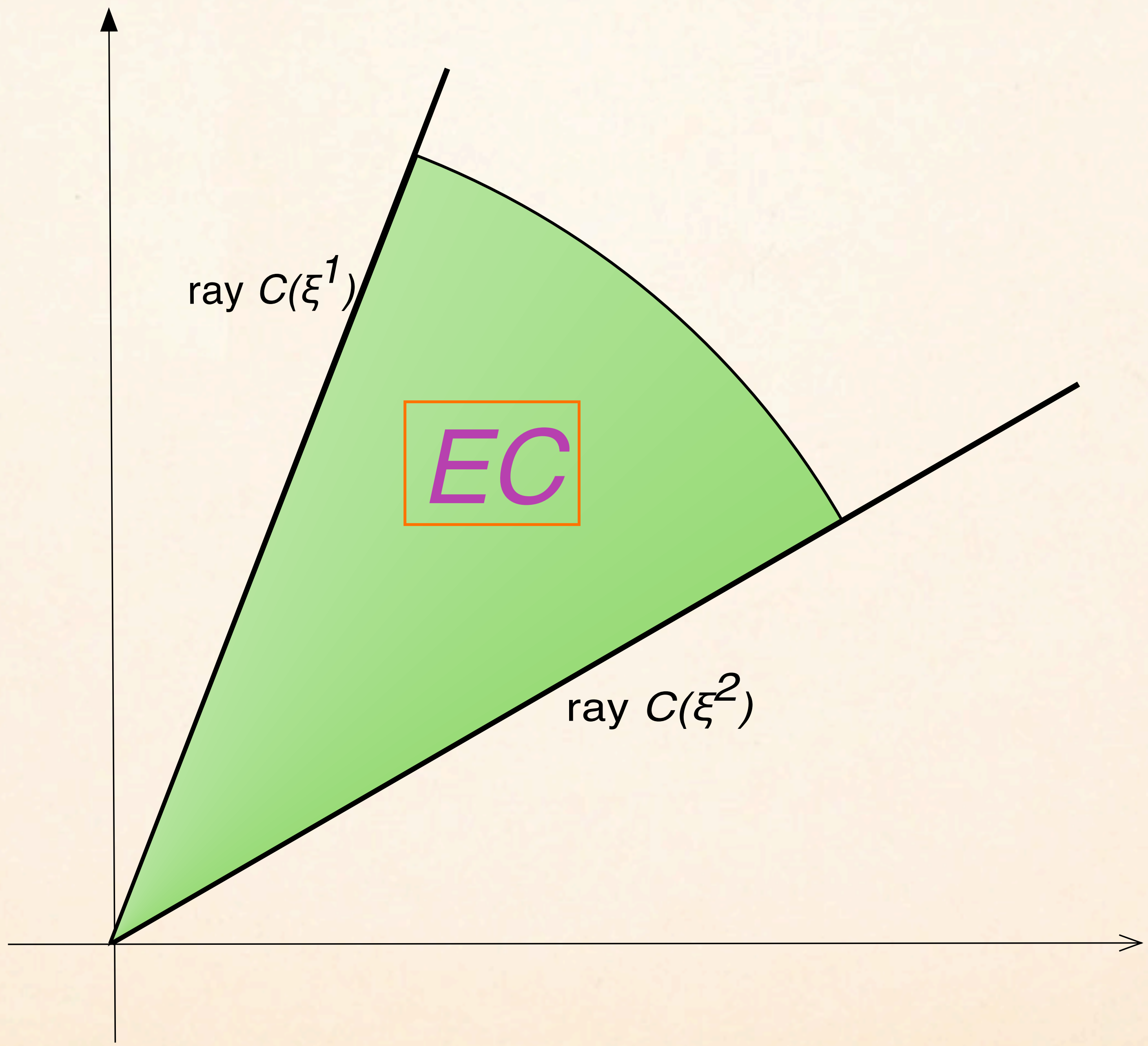
C μ-atom convex $\Rightarrow EC$ is convex

(certainly when P is atomless).

EXPECTATION: BOUNDED RANDOM SETS



EXPECTATION: UNBOUNDED RANDOM SETS



STRONG LAW OF LARGE NUMBERS

ARTSTEIN & HART

$C : \Xi \rightrightarrows \mathbb{R}^m$ measurable, $\{\xi^\nu, \nu \in \mathbb{N}\}$ iid Ξ-valued random variables

$C(\xi^\nu)$ iid random sets (i.e. induced P^ν independent and identical)

$EC = \mathbb{E}\{C(\bullet)\} = \left\{ \int_\Xi x(\xi)\,\mu(d\xi) \,\middle|\, x\ \mu\text{-summable } C(\xi)\text{-selection} \right\}$

independence $\Rightarrow$ all (measurable) selections are independent

$\{C(\xi^\nu) : \Xi \rightrightarrows \mathbb{R}^m\ \nu \in \mathbb{N}\}$ iid with $EC \neq \emptyset$. Then, with

$$C^\nu(\xi^\infty) = \nu^{-1}\left(\sum_{k=1}^{\nu} C(\xi^k)\right) \to \bar{C} = \text{cl con } EC \quad \mu^\infty\text{-a.s.}$$

$\text{Ls}_\nu C^\nu(\xi^\infty) \subset \bar{C} \Leftrightarrow \limsup_\nu \sigma_{C^\nu} \le \sigma_{\bar{C}}$ support functions

$\text{Li}_\nu C^\nu(\xi^\infty) \supset \bar{C}$ relies on LLN for (vector-valued) selections

RESOURCES ALLOCATIONS
AVERAGE OF EPI-SUMS

$q \in \mathbb{R}^n_{++}$, q central resources allocated to ν firms

Optimal allocation: p_i production functions

suppose k large, $p_i = p_i(\xi, x)$ with $\xi \in \Xi$,

$$\forall \xi: \quad z_\nu(\xi, q) = \max \; \nu^{-1} \sum_{i=1}^{\nu} p_i(\xi, x^i) \;\; \text{s.t.} \;\; \nu^{-1} \sum_{i=1}^{k} x^i \le q$$

$(\Xi, \mathcal{A}, \mu)$, p_i : usc in x, jointly measurable $\mathcal{A} \otimes \mathcal{B}$

"Limit" Problem:

$$z(q) = \max \int p(\xi, x(\xi))\, d\mu \;\; \text{s.t.} \;\; \int x(\xi)\, d\mu \le q$$

Suppose $\left\{ p_i(\xi, \cdot) \in \text{lsc-fcns}(\mathbb{R}^n) \right\}$ are iid $\Leftrightarrow$ epi p_i iid

Then, $z_\nu(\xi, q) \to z(q)$ μ-a.s. where $p = p_1$ if μ nonatomic

or $-p =$ con $-p_1$ (must not depend on ξ)

Argument: set-LLN on hypographs ($\sim$ epi $-p_i$)

SAMPLE AVERAGE APPROXIMATION

stochastic variational problem: $\bar{S}(x) = \mathbb{E}\{S(\boldsymbol{\xi},x)\} \ni 0$

$S : \Xi \times \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ random set-valued mapping

$\boldsymbol{\xi}$ random vector with values $\xi \in \Xi \subset \mathbb{R}^N$

solution (a 'stationary point') $\bar{x} \in \bar{S}^{-1}(0)$

—o —o —o —o —o —o —o —o —o —o —o —o —o —o —o —o —o —o

sample $\overset{\rightarrow\nu}{\xi} = (\xi^1,\ldots,\xi^\nu)$ of $\boldsymbol{\xi}$

$\frac{1}{\nu}\left(\sum_{k=1}^{\nu} S(\xi^k,x)\right) = S^\nu(\overset{\rightarrow\nu}{\xi},x) \ni 0,$ approximating system?

i.e., $\left(S^\nu\right)^{-1}(0) \overset{?}{\rightarrow} \bar{S}^{-1}(0)$ a.s.

STOCHASTIC OPTIMIZATION

$\min Ef(x) = \mathbb{E}\{f(\boldsymbol{\xi},x)\}$ --stationary point-- $\partial Ef(x) \ni 0$

assuming $\mathbb{E}\{\partial f(\boldsymbol{\xi},x)\} = \partial Ef(x)$ (not generally correct)

could $\partial Ef(x) \ni 0$ get replaced (?) by

$$\nu^{-1}\left(\sum_{k=1}^{\nu} \partial f(\xi^k, x)\right) \ni 0 \text{ from sample } \overset{\to\nu}{\xi}$$

$\mathrm{dom}\, Ef \approx \bigcap_{\xi\in\Xi} \mathrm{dom}\, f(\xi,\cdot)$,

unless $\xi \mapsto \mathrm{dom}\, f(\xi,\cdot)$ constant,

interchanging $\mathbb{E}$ & ∂ is only exceptionally valid

STOCHASTIC V.I.
(VARIATIONAL INEQUALITY)

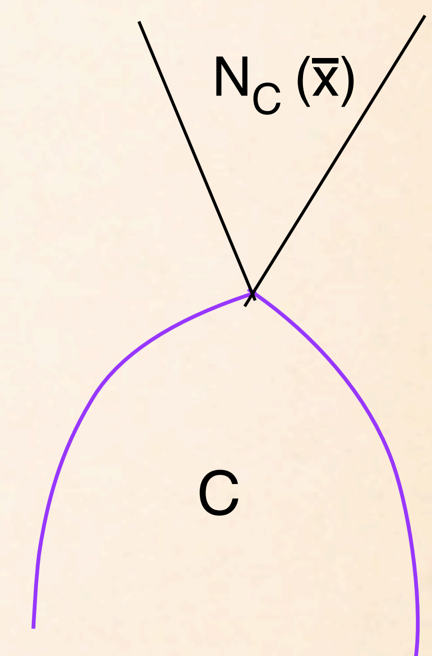
$\boldsymbol{\xi} = (\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \ldots)$, $G^\nu(\cdot, x)$ σ-$(\boldsymbol{\xi}^1, \ldots \boldsymbol{\xi}^\nu)$ measurable

$-G^\nu(\boldsymbol{\xi}, x) \in N_C(x)$, C compact, convex

$$G^\nu(\boldsymbol{\xi}, \cdot) \to^? G(\boldsymbol{\xi}, \cdot)$$

$x^\nu(\xi)$ solution of $-G^\nu(\xi, x) \in N_C(x)$ for sample $\xi \approx \overset{\to \nu}{\xi}$

does $x^\nu(\boldsymbol{\xi}) \to$ a solution of $-G(\boldsymbol{\xi}, x) \in N_C(x)$? a.s.



what if C depends on ν, ξ : sequence of random sets $C^\nu(\boldsymbol{\xi})$

RANDOM MAPPINGS

$S : \Xi \times E \rightrightarrows \mathbb{R}^m, \;\; E \subset \mathbb{R}^n$

$\mathcal{A} \otimes \mathcal{B}^n$-jointly measurable: $S^{-1}(O) \in \mathcal{A} \otimes \mathcal{B}^n$, O open

$\Rightarrow \forall\, x : \xi \mapsto S(\xi, x)$ a random set

random closed set when S is closed-valued

$ES : E \rightrightarrows \mathbb{R}^m$ with $ES(x) = \mathbb{E}\{S(\boldsymbol{\xi}, x)\}$ expected mapping

ES convex-valued when $\xi \mapsto S(\xi, \cdot)$ μ-atom convex

Law of Large Numbers for random sets applies

SAMPLE AVERAGE APPROXIMATIONS

$\boldsymbol{\xi} = (\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \ldots)$ iid, sample $\vec{\xi}^{\,\nu} = (\xi^1, \ldots, \xi^\nu)$

SAA-mapping: given $S : \Xi \times E \rightrightarrows \mathbb{R}^m$ random mapping

$S^\nu : \Xi^\infty \times E \rightrightarrows \mathbb{R}^m$ with

$$\forall \xi \in \Xi^\infty, x \in E : S^\nu(\xi, x) = \frac{1}{\nu} \sum_{k=1}^{\nu} S(\xi^k, x) = S^\nu(\vec{\xi}^{\,\nu}, x)$$

S^ν depends only on $\vec{\xi}^{\,\nu}$

SAA-mappings S^ν are random mappings

not necessarily closed-valued

(the sum of closed sets is not necessarily closed)

POINTWISE LIMITS: SAA-MAPPINGS

$ES(x) = \mathbb{E}\{S(\boldsymbol{\xi},x)\} \neq \emptyset$, then

$\forall x \in X: \; S^\nu(\xi,x) \to \mathrm{cl\,con}\, ES(x) =: \bar{S}(x) \;\; \mu^\infty$-a.s.

If $S(\cdot,x)$ is P-atom convex, $S^\nu(\xi,\cdot) \to \mathrm{cl}\, ES(x) =: \bar{S}(x) \;\; \mu^\infty$-a.s.

Proof: LLN for random sets. □

CONSISTENT APPROXIMATIONS?

$$S^{\nu}(\xi,\cdot) \underset{\text{point}}{\longrightarrow} \bar{S} \;\; \mu^{\infty}\text{-a.s.} \Rightarrow ? \;\; S^{\nu}(\xi,\cdot)^{-1}(0) \rightrightarrows_{a} \bar{S}^{-1}(0)$$

sometimes!

graphical rather than pointwise convergence is required

$$S^{\nu}(\xi,\cdot) \underset{\text{gph}}{\longrightarrow} \bar{S} \;\; \mu^{\infty}\text{-a.s. is needed}$$

relationship between graphical and pointwise convergence?

GRAPHICAL & POINTWISE

$D, D^\nu : X \rightrightarrows \mathbb{R}^m$. Then, $D^\nu \underset{\text{point}}{\longrightarrow} D$ and $D^\nu \underset{\text{gph}}{\longrightarrow} D$ (at x)

$\Leftrightarrow \left\{ D^\nu, \nu \in \mathbb{N} \right\}$ are equi-osc (asymptotically) (at x)

~ Arzela-Ascoli Theorem for set-valued mappings

S random mapping, μ^∞-a.s., $S^\nu(\xi, \cdot) \underset{\text{point}}{\longrightarrow} \text{cl con}\, ES = \bar{S}$

then $S^\nu \underset{\text{gph}}{\longrightarrow} \bar{S} \Leftrightarrow \left\{ S^\nu, \nu \in \mathbb{N} \right\}$ are equi-osc (asymptotically)

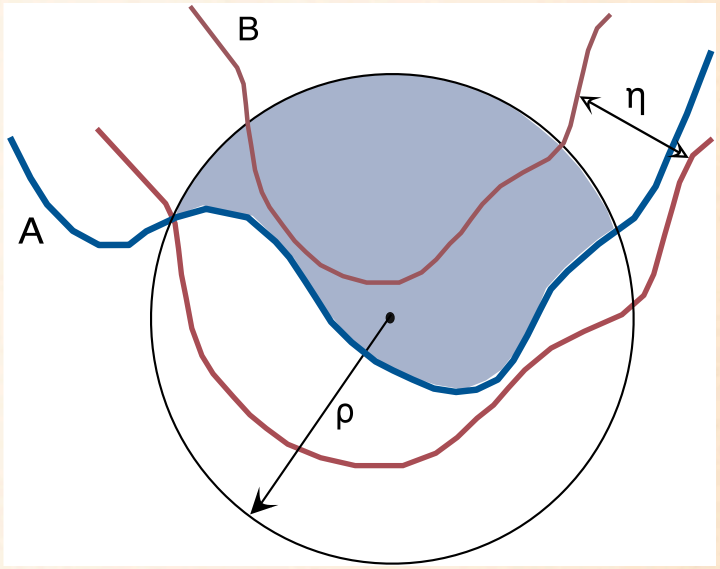
EQUI-OSC MAPPINGS

$D: X \rightrightarrows \mathbb{R}^m, X \subset \mathbb{R}^n$ is osc if gph S is closed

osc at $\bar{x}$ if $D(\bar{x}) \supset \operatorname{Ls}_{x^\nu \to \bar{x}} D(x^\nu)$

$\sim$ given any $\rho > 0, \varepsilon > 0$

$\exists V \in \mathcal{N}(\bar{x}): e_\rho\big(D(x), D(\bar{x})\big) < \varepsilon, \ \forall\, x \in V$



$\{D^\nu : X \rightrightarrows \mathbb{R}^m\}$ are equi-osc at $\bar{x}$

$\sim$ given any $\rho > 0, \varepsilon > 0$

$\exists V \in \mathcal{N}(\bar{x}): e_\rho\big(D^\nu(x), D^\nu(\bar{x})\big) < \varepsilon, \ \forall\, x \in V$

$V = V(\rho, \varepsilon)$ doesn't depend on ν.

GRAPHICAL CONVERGENCE OF SAA-MAPPINGS

$S\ \Xi\times X \rightrightarrows \mathbb{R}^m$ random mapping, $(\Xi,\mathcal{A},\mu)$

μ^∞-a.s.: $S^\nu(\xi,\cdot)\underset{\text{gph}}{\longrightarrow}\bar{S}$ at $\bar{x} \Leftrightarrow$ SAA-mappings $\{S^\nu(\xi,\cdot)\}$ equi-osc at $\bar{x}$

$\Rightarrow$ sol'ns of $S^\nu(\xi,\cdot)\ni 0 \rightrightarrows_\nu$ sol'ns of $\bar{S}(\cdot)\ni 0$

Sufficient conditions: μ^∞-a.s.

$S(\xi,\cdot)$ stably osc & steady under averaging $\Rightarrow$ $\{S^\nu(\xi,\cdot)\}$ equi-osc

Law of large Numbers for Random Mappings

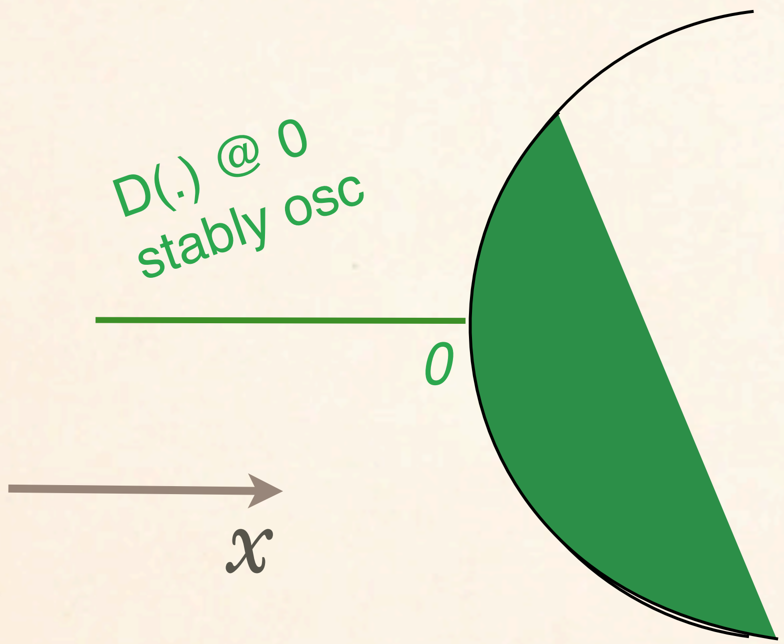
S random osc mapping: $\Xi\times\mathbb{R}^n \rightrightarrows \mathbb{R}^m$

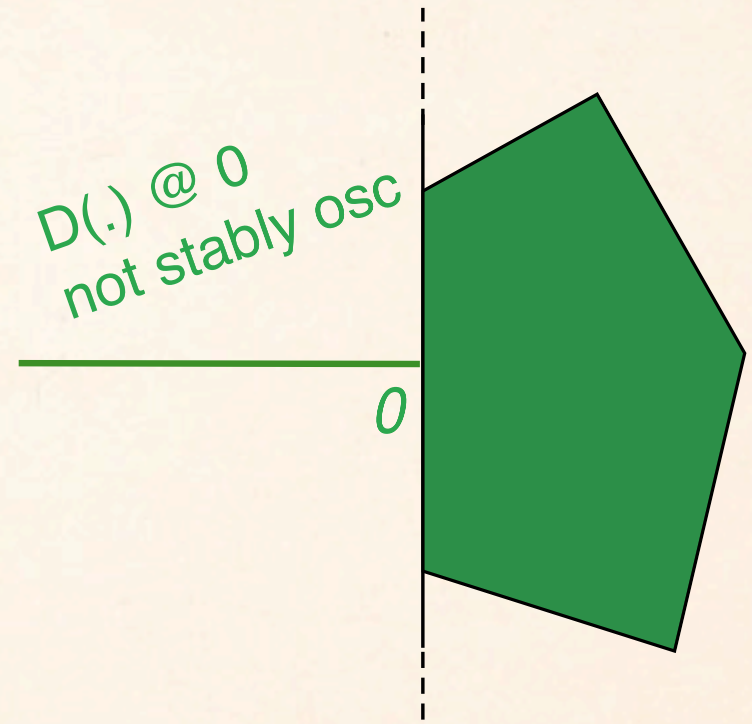
stably osc & steady under averaging

$\xi^1,\xi^2,\ldots,$ iid random variables (values in Ξ), distribution μ

Then, $\nu^{-1}\sum_{k=1}^{\nu} S(\xi^k,\cdot)\to_{\text{gph}} \bar{S} = \text{cl con}\, E\{S(\xi^0,\cdot)\}$ μ^∞-a.s.

STABLY OSC



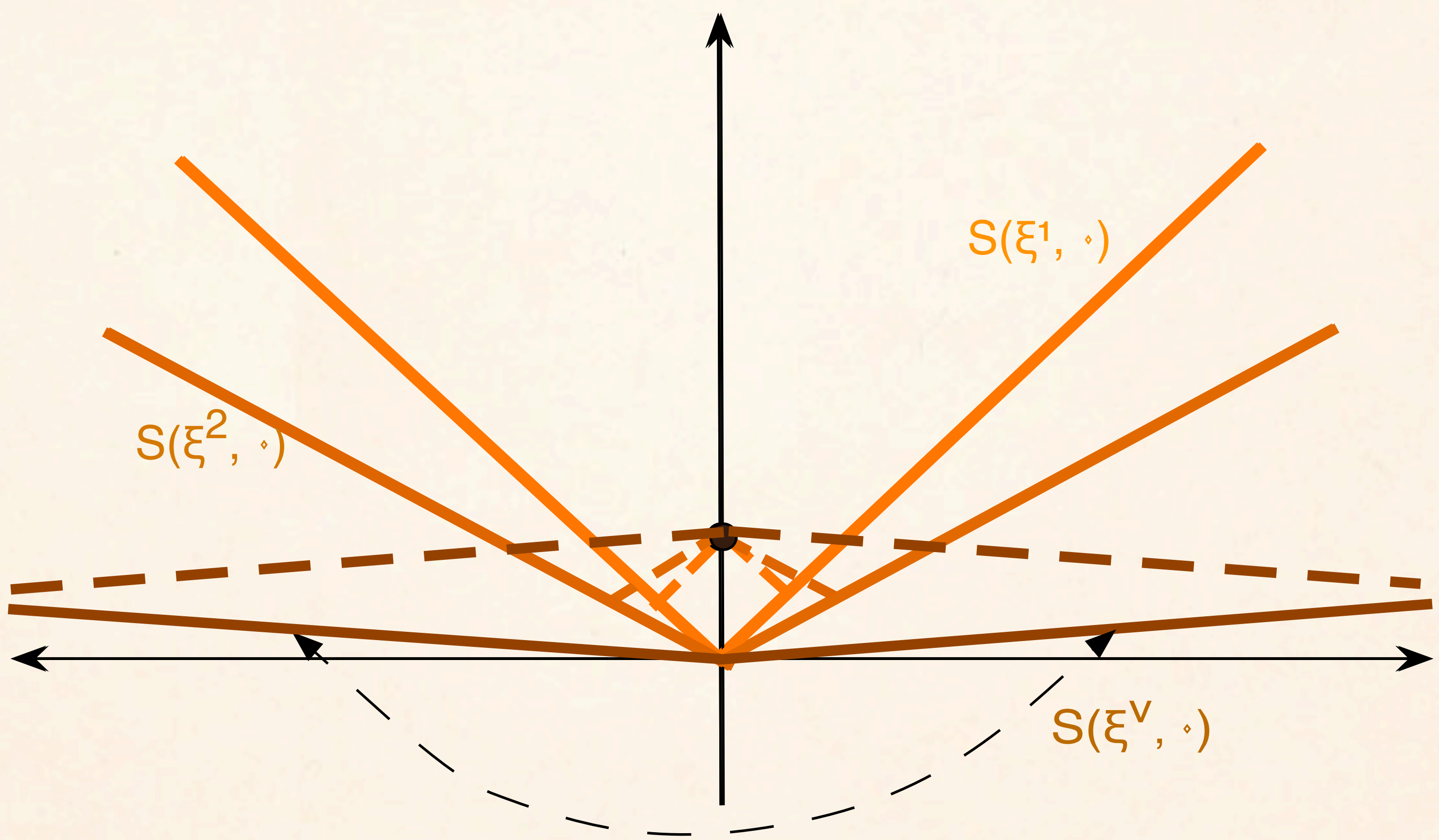


S stably osc near $\bar{x}$ if μ-a.s.,

$$\forall \rho > 0, \varepsilon > 0,\ \exists W \in \mathcal{N}(\bar{x})\ \&\ \eta\mathbb{B}\ (\eta > 0):$$

$$e_\rho\big(S(\xi, x'), S(\xi, x)\big) < \varepsilon, \forall\, x' \in x + \eta\mathbb{B},\ x \in W$$

STEADY UNDER AVERAGING



$$u \in S^\nu(\vec{\xi}^\nu, x) \cap \rho\mathbb{B} \Rightarrow \exists\, \hat{\rho} \geq \rho,\ u^k \in S(\xi^k, x) \cap \hat{\rho}\mathbb{B} \text{ such that}$$

$$u = \nu^{-1}(u^1 + \cdots + u^\nu);\ S^\nu(\vec{\xi}^\nu, x) \cap \rho\mathbb{B} \subset \frac{1}{\nu}\left[\sum_{k=1}^{\nu} S(\xi^k, x) \cap \hat{\rho}\mathbb{B}\right]$$

STEADY UNDER AVERAGING & STABLY OSC

rge $S \subset B$ bounded $\Rightarrow$ steady under averaging

S cone-valued and rge $S \subset$ pointed cone K. Then,

$\bar{S} = ES$ and $\Rightarrow$ steady under averaging.

S, R steady under averaging $\Rightarrow$ so is $S + R$

$R(\xi, x) = R(x) \Rightarrow R$ steady under averaging

rge S bounded + R constant $\Rightarrow$ steady under averaging

$G(\xi, x) + N_C(x) \Rightarrow$ steady under averaging (V.I.)

$G : \Xi \times X \to \mathbb{R}^n$ is bounded

S, R stably osc $\Rightarrow S + R$ stably osc

although D^1, D^2 osc $\not\Rightarrow D^1 + D^2$ osc

$\mathbb{B}$ closed, convex $x \mapsto N_{\mathbb{B}}(x)$ osc

but not stably osc ($x^\nu \in$ int $\mathbb{B} \to \bar{x} \in$ bdry $\mathbb{B}$)

IMPLEMENTING SAA ** LOCALLY

$EG(x) = \mathbb{E}\{G(\boldsymbol{\xi},x)\} \in S(x)$

(V.I.: $S = N_C$, applied to option pricing, ...)

$G^\nu(\overset{\to\nu}{\xi},\cdot) = \nu^{-1}\sum_{k=1}^{\nu} G(\xi^k,x)$. Assume $G^\nu(\overset{\to\nu}{\xi},\cdot)$, $EG \in C^1(\mathbb{R}^n;\mathbb{R}^n)$,

$\bar{x}$ strongly regular solution [Robinson] of $EG(x) \in S(x)$,

$\exists V \in \mathcal{N}(\bar{x}), \rho > 0$ such that $\forall z \in \rho\mathbb{B}$:

$$z + EG(\bar{x}) + \nabla EG(\bar{x})\big(x - \bar{x}\big) \in S(x)$$

has a unique solution $\bar{x}(z) \in V$, Lipschitz continuous on $\rho\mathbb{B}$, and

$\left\| G^\nu(\overset{\to\nu}{\xi},\cdot) - EG \right\| \to 0$ μ-a.s. Then, for ν sufficiently large

on a neighborhood of $\bar{x}$, $G^\nu(\overset{\to\nu}{\xi},\cdot) \in S(x)$ has a unique solution

$$\bar{x}(\overset{\to\nu}{\xi}) \to \bar{x} \quad \mu\text{-a.s.}$$

IMPLEMENTING SAA ** EXAMPLE

stochastic program with recourse (simple): $\boldsymbol{\xi}$ uniform on [1,2]

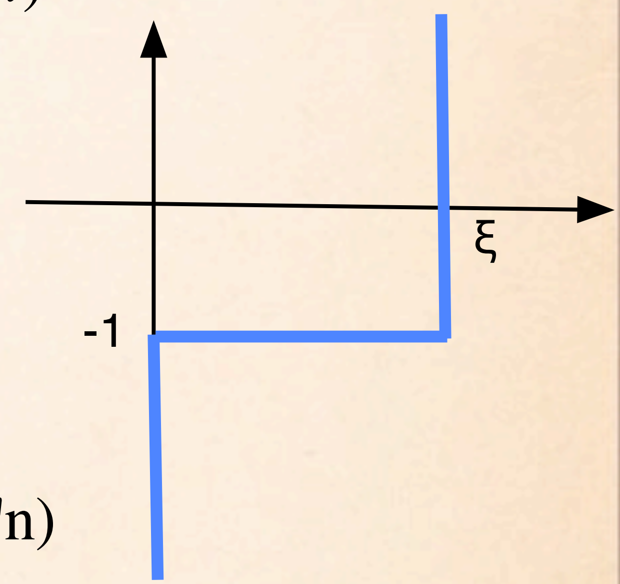
$$\min_{x,y_\xi} \mathbb{E}\left\{-x \,\middle|\, x + y_\xi \le \xi, x \in [0,2], y_\xi \ge 0\right\} = \min\left(Ef(x) = \mathbb{E}\{f(\boldsymbol{\xi},x)\}\right)$$

$$f(\xi,x) = -x + \iota_{[0,2]} + \iota_{(-\infty,\xi]} = -x + \iota_{[0,\xi]}$$

to solve $0 \in \partial Ef(x)$ gets replaced by $0 \in \nu^{-1}\sum_{k=1}^{\nu} S(\xi^k,x) = S^\nu(\vec{\xi}^{\,\nu},x)$

$$S(\xi,x) = \partial f(\xi,x) = -1 + N_{[0,\xi]}(x), \quad \text{dom } S(\xi,\cdot) = [0,\xi]$$

$$= \begin{cases} (-\infty,-1] & \text{when } x = 0, \\ -1 & \text{for } x \in (0,\xi), \\ [1,\infty) & \text{when } x = \xi \end{cases}$$



Solution of $0 \in S^\nu(\vec{\xi}^{\,\nu},x)$: $x^\nu = \min\left\{\xi^1,\ldots,\xi^\nu\right\} \to_{\text{a.s.}} \bar{x} = 1$ (opt. sol'n)

but x^ν is never a feasible solution,

$\nexists y_\xi \ge 0$ such that $x^\nu + y_\xi \le \xi$ when $\xi \in [1,x^\nu)$

Problem: $\partial Ef(x) \ne \mathbb{E}\{\partial f(\boldsymbol{\xi},x)\}$ *** interchange is not valid.

IV. RANDOM LSC FUNCTIONS

STOCHASTIC PROGRAM WITH RECOURSE

$$f(\xi,x) = f_{10}(x) + \inf_{y\in Y}\left\{ f_{20}\left(\xi;x,y\right) \middle| f_{2i}(\xi;x,y) \le 0,\, i=1,..,m_2 \right\}$$

$$\text{when } f_{1i}(x) \le 0,\, i = 1,\dots,m_1,$$

$$= \infty \text{ otherwise}$$

2-stage stochastic program with recourse: $\min_{x\in\mathbb{R}^n} E\{f(\boldsymbol{\xi},x)\}$

$$f\left(\xi,x(\bullet)\right) = \begin{cases} f_0\left(\xi,x(\bullet)\right) \text{ if } x(\xi)\in C(\xi,x(\xi))\; a.s \\ \infty \text{ otherwise} \end{cases}.$$

with $\xi \mapsto x(\xi)$ & $\xi \mapsto C(\xi,x(\xi)) \in \mathcal{N}_\infty$ (non-anticipative)

(dynamic) stochastic programs with recourse

$$\boxed{\min_{x\in\mathcal{N}^a} E\{f(\boldsymbol{\xi},x(\boldsymbol{\xi}))\}}$$

SOLVING VIA APPROXIMATION

$P^\nu \to_n P$ usually a discretization (P^ν)

say, generated via taking conditional expectation, …

$\min_{x\in\mathbb{R}^n} E^\nu f = \mathbb{E}^\nu\{f(\boldsymbol{\xi},x)\} = \int_\Xi f(\xi,x)P^\nu(d\xi)$

approximates $\min_{x\in\mathbb{R}^n} Ef = \int_\Xi f(\xi,x)P(d\xi)$?

If $Ef^\nu \to_e Ef$ then "$\arg\min Ef^\nu \to \arg\min Ef$"

& "ε-$\arg\min Ef^\nu \to \varepsilon$-$\arg\min Ef$" (confidence intervals)

holds if $\{P^\nu, \nu\in\mathbb{N}\}$ is f-tight: for all $x\in$ dom Ef,

$\forall\varepsilon>0, \exists$ compact K_ε such that $\int_{\Xi\setminus K_\varepsilon} |f(\xi,x)| P^\nu(d\xi) < \varepsilon$

certainly the case when supp P^ν is bounded

SOLVING VIA SAMPLING

$\xi^1, \xi^2, \ldots$ iid samples of $\boldsymbol{\xi}$ (or p.iid)

$$\arg\min_{x\in X} \frac{1}{\nu}\sum_{k=1}^{\nu} f(\xi^k, x) \xrightarrow{?} \arg\min_{x\in X} E\{f(\boldsymbol{\xi}, x)\}$$

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

(random) empirical measure P^ν, prob$\left[\boldsymbol{\xi} = \xi^k\right] = \nu^{-1}$

$$\boldsymbol{\xi}^\infty = (\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \ldots), \text{"}E^\nu f(\boldsymbol{\xi}^\infty, x)\text{"} = E^\nu f(x) = \int f(\xi, x)P^\nu(d\xi)$$

$$E^\nu f \xrightarrow{?} Ef \quad \mu^\infty\text{-}a.s. \text{ (arg min}^\nu \ \rightrightarrows_a \text{arg min)}$$

STATISTICAL ESTIMATION
FUSION OF HARD & SOFT INFORMATION

Observation (hard data): $\xi^1, \xi^2, \ldots, \xi^\nu$

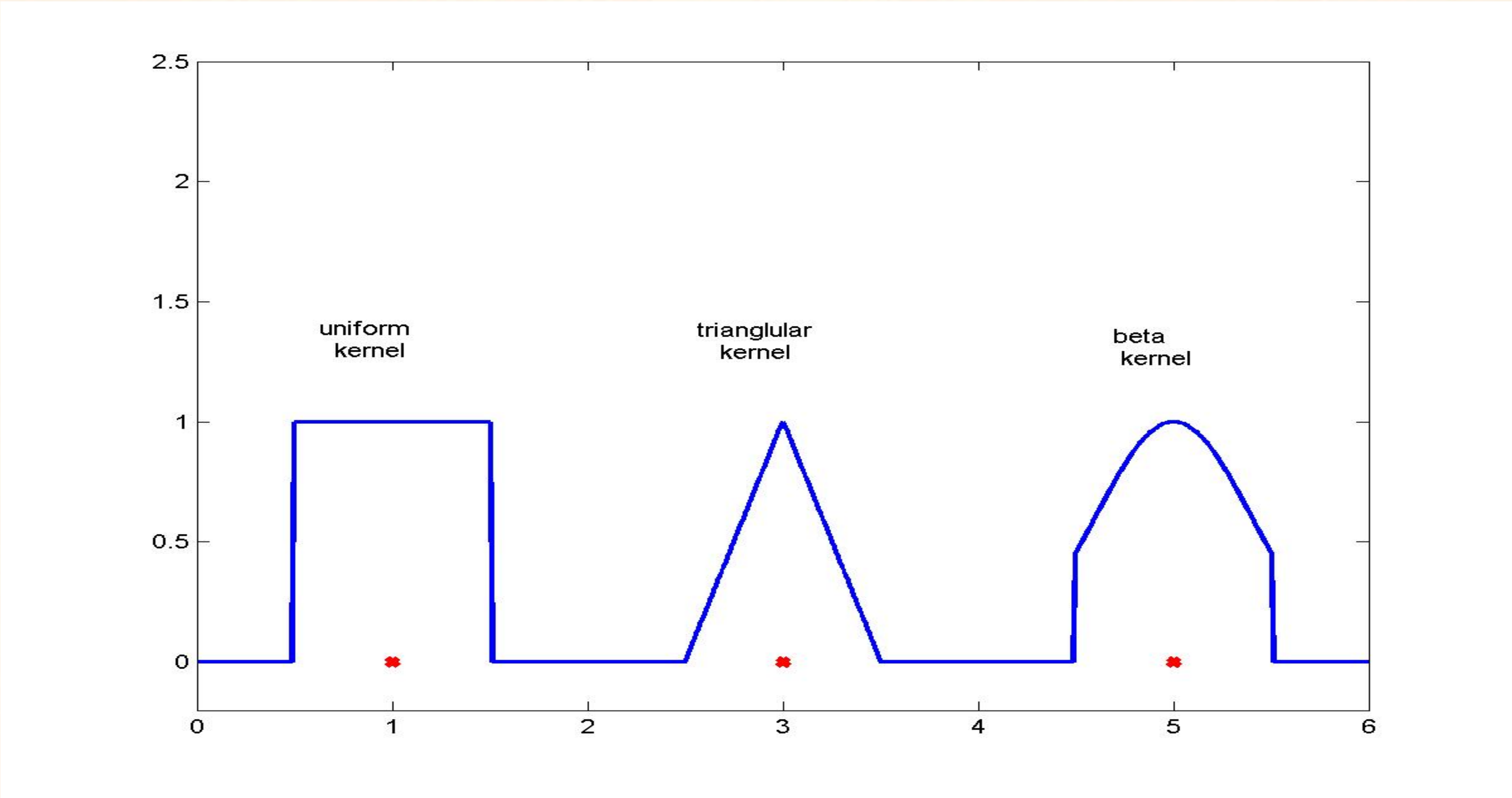
Soft data (non-data knowledge):
- Support: (un)bounded
- density or discrete distribution,
- bounds on expectation, moments,
- heavy tails
- shape: unimodal, decreasing, parametric class

Softer data (modeling assumptions):
- see above + …
- level of smoothness, `Bayesian' neighborhood, ..

POTENTIAL APPLICATIONS

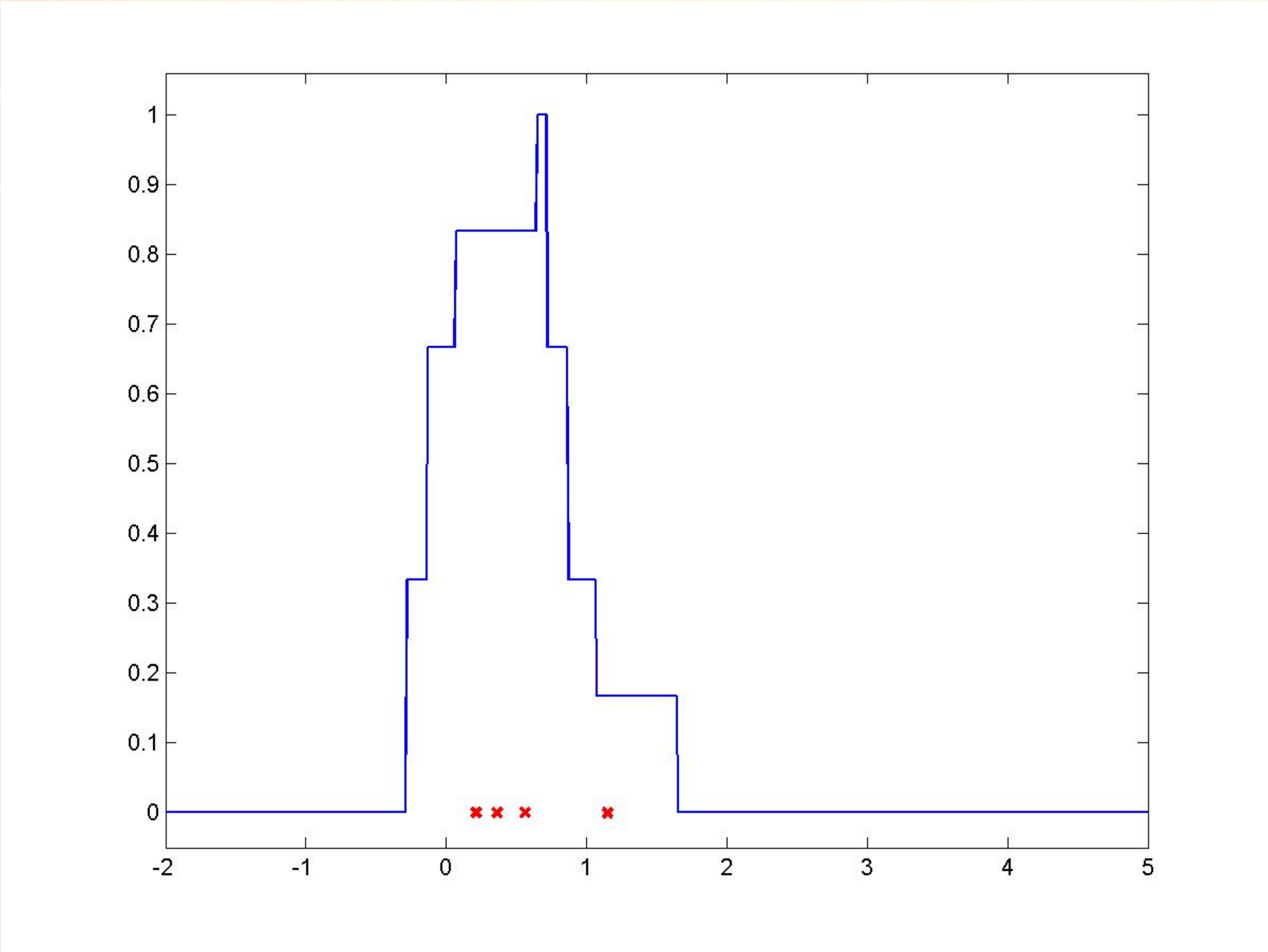
- Estimating (cum.) distribution functions
- Estimating coefficient of time series
- Estimating coefficients of stochastic differential equation (SDE)
- Estimating financial curves (zero-curves)
- Kalman filtering
- Dealing with lack of data: ***few** observations*
- Estimating density functions h^{est} (non-parametric)

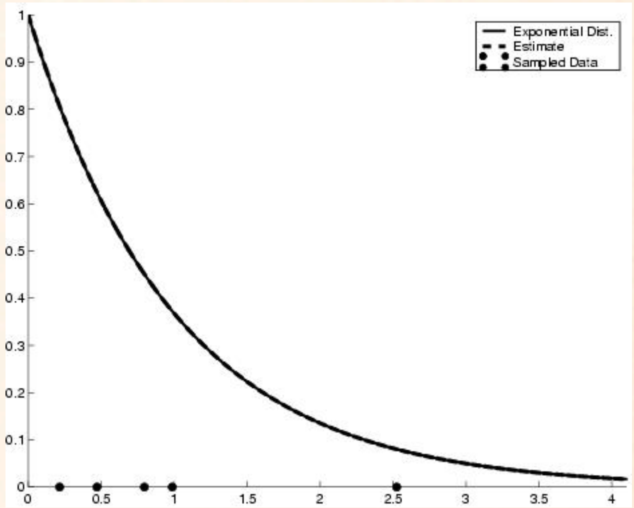
KERNEL ESTIMATES



"frequentist" viewpoint: observations
optimal bandwith: kernel support

KERNEL ESTIMATES: LOW DATA





ν=5 samples
exponential distribution

... AS STOCHASTIC OPTIMIZATION

$$\max E^{\nu}\left\{\ln h(x)\right\} = \frac{1}{\nu}\sum_{l=1}^{\nu} \ln h(x_l)$$

$$\text{such that } \int h(x)dx = 1,$$

$$h(x) \geq 0,\ \forall x \in \mathbb{R}^n$$

$$h \in A^{\nu} \subset H$$

$E^{\nu}\left\{\ln h(x)\right\} \sim \max \prod_{l=1}^{\nu} h(x_l)$ maximum likelihood

A^{ν} soft information (constraint set)

H: density functions space, $C^2(\mathbb{R}^n)$, $HRKS(\text{supp } h)$, $H^1(\text{supp } h)$

NUMERICAL PROCEDURES

1. $h \simeq \sum_{k=1}^{q} u_k \varphi_k(\bullet)$ φ_k basis-functions

   Fourier coefficients, wavelets, kernel-like functions

2. $h(x) = e^{-s(x)}$, s epi-spline of order n (cubic, quadratic, ...)

$$s(x) = s_0 + v_0 x + \int_0^x dr \int_0^r dt\, z(t),\ z(t) \equiv z_k \text{ on } \left(x_k, x_{k+1}\right]$$

$$= s_0 + v_0 x + \sum_{j=1}^{k} a_{kj} z_k \text{ when } x \in \left(x_k, x_{k+1}\right]$$

ESTIMATION PROBLEM $X = \mathbb{R}$

$$\max E^\nu\left\{\ln h(x)\right\} \sim \min \frac{1}{\nu}\sum_{l=1}^{\nu} s(x_l)$$

$$\text{such that } \int_{\text{"supp" } h} e^{-s(x)}\,dx \le 1, \qquad (h \ge 0)$$

$$z_k \in \left[-\kappa_l, \kappa_u\right] \text{ 'constrained'-spline}$$

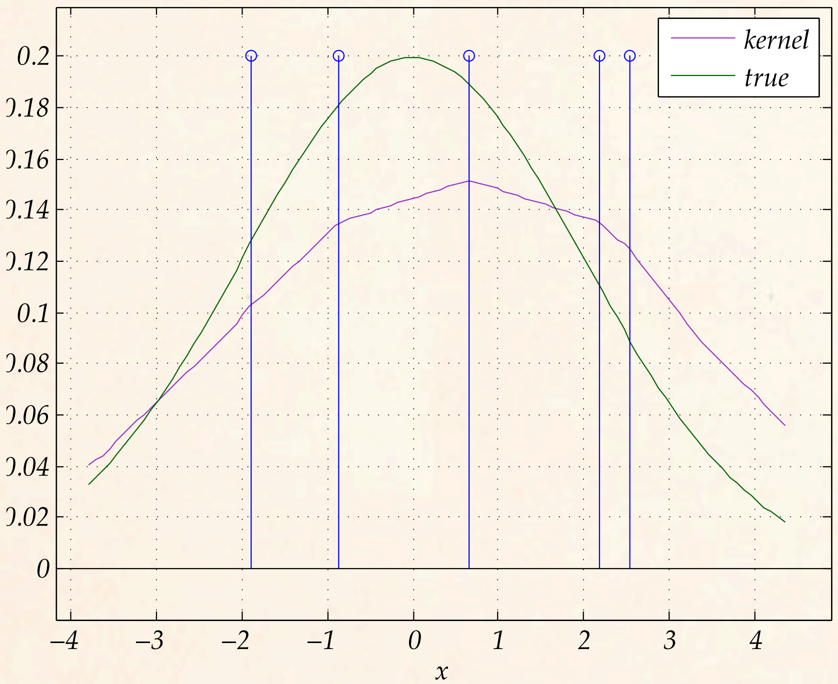
unimodal: $\kappa_l = 0$

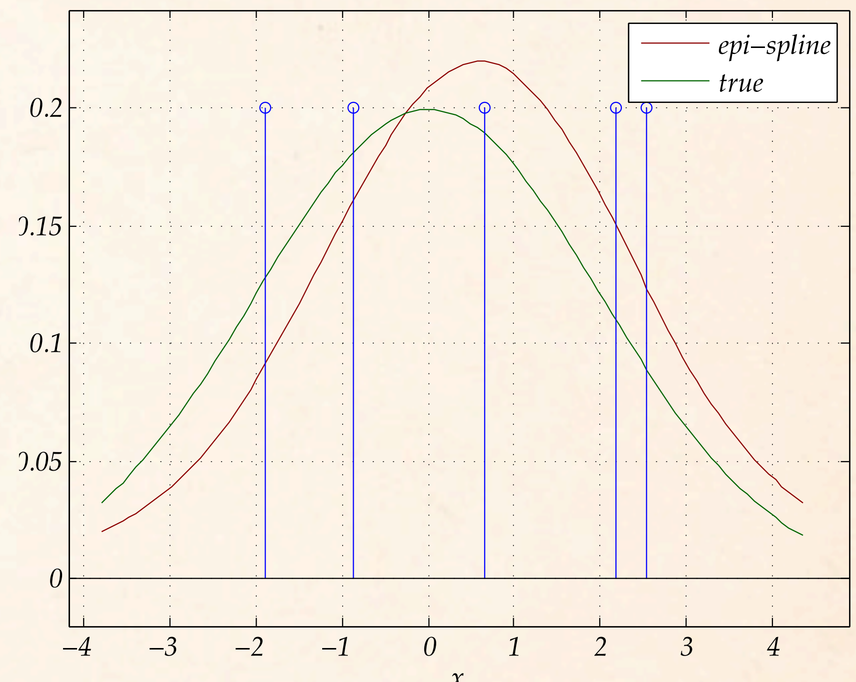
$$s(x) = s_0 + v_0 x + \sum\nolimits_{j=1}^{k} a_{kj} z_k \text{ when } x \in \left(x_k, x_{k+1}\right]$$

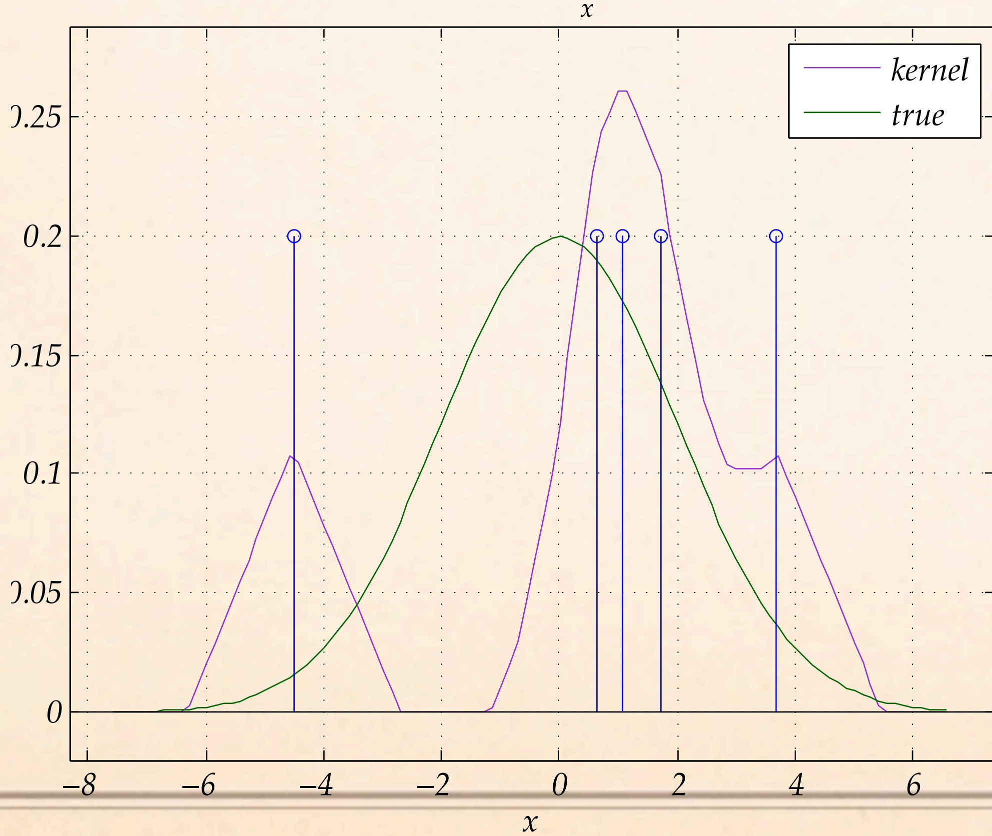
constraints on z_k : on curvature of s

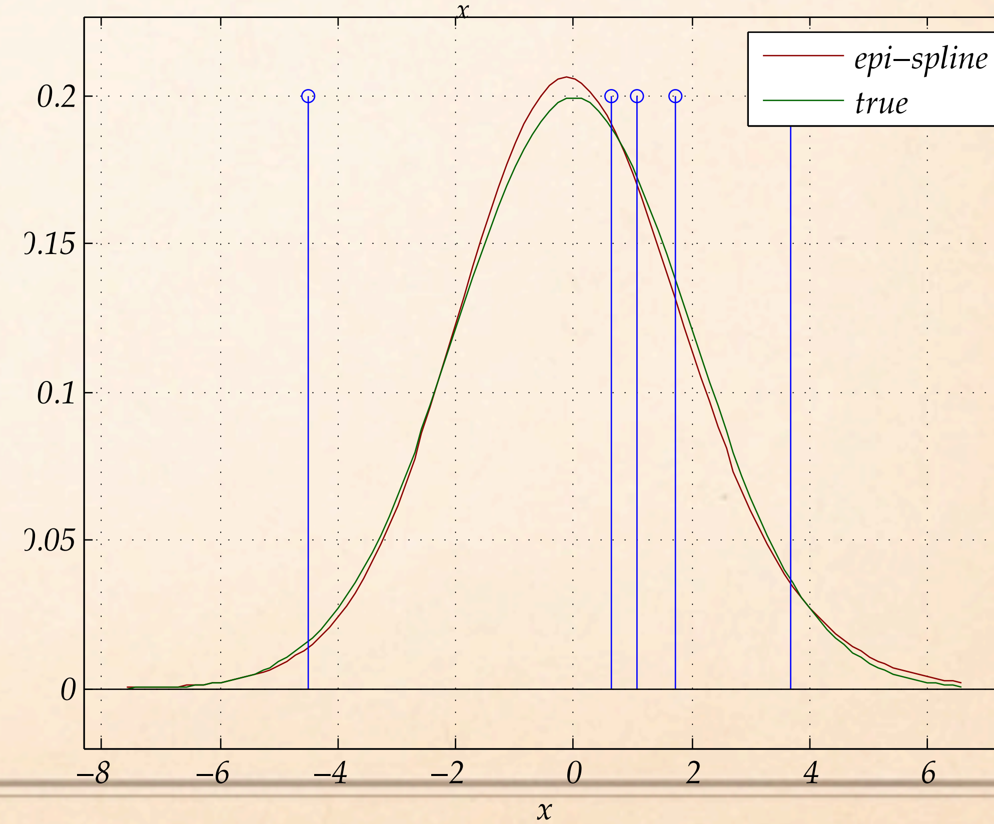
on "supp h" : bounds on support of h

NORMAL DISTRIBUTION
KERNEL & EPI-SPLINE ESTIMATES

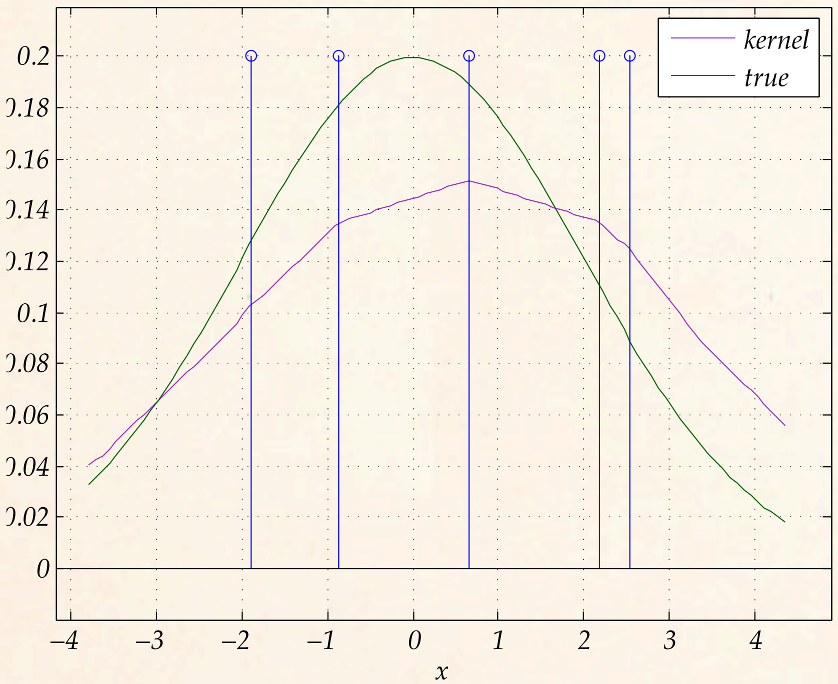


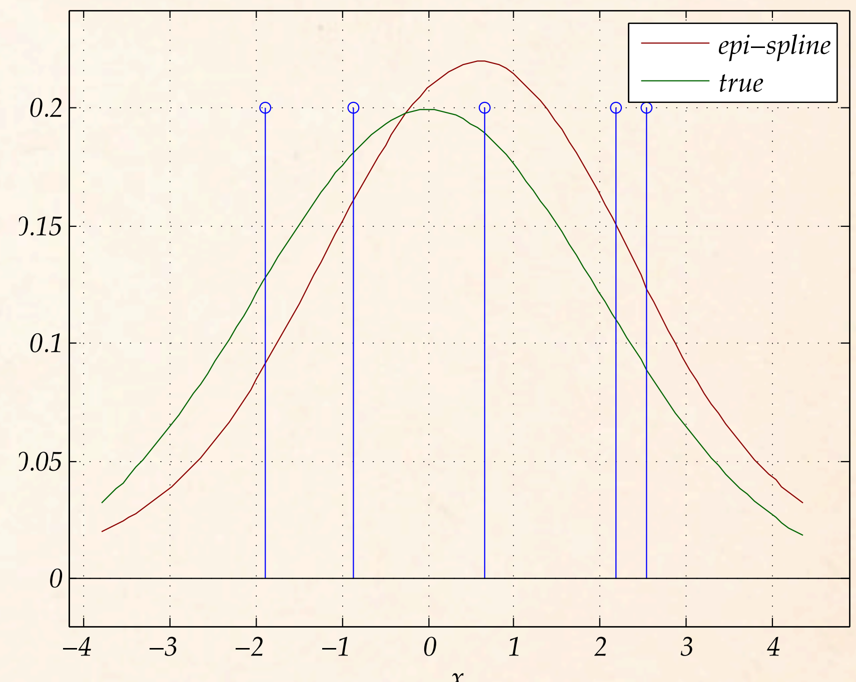


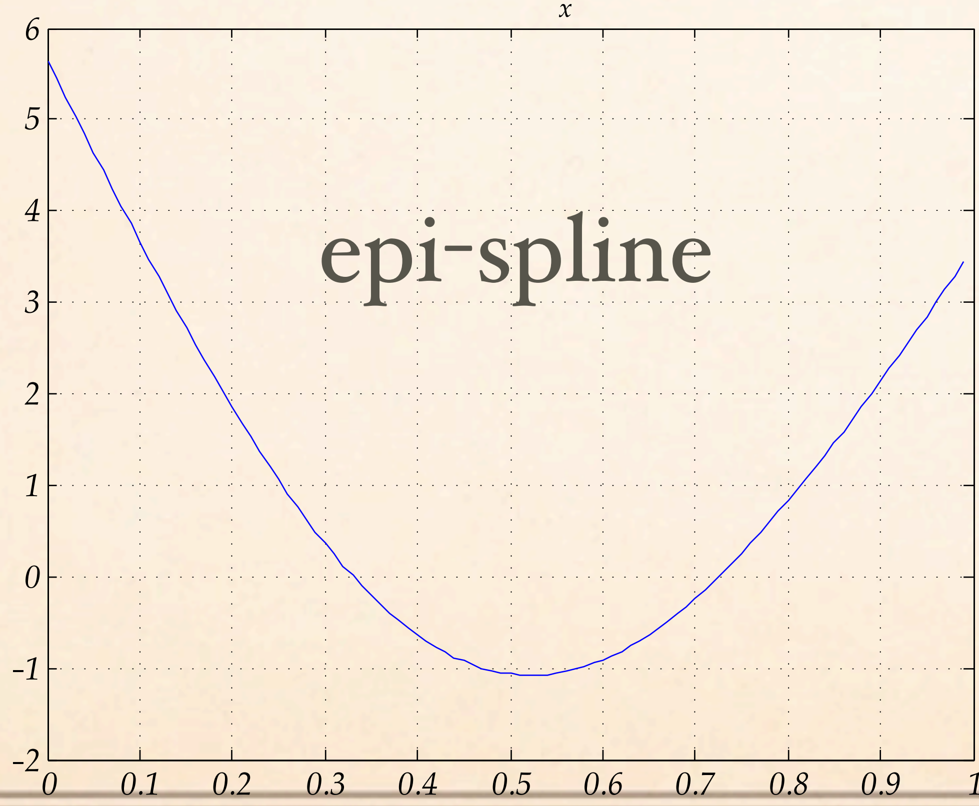


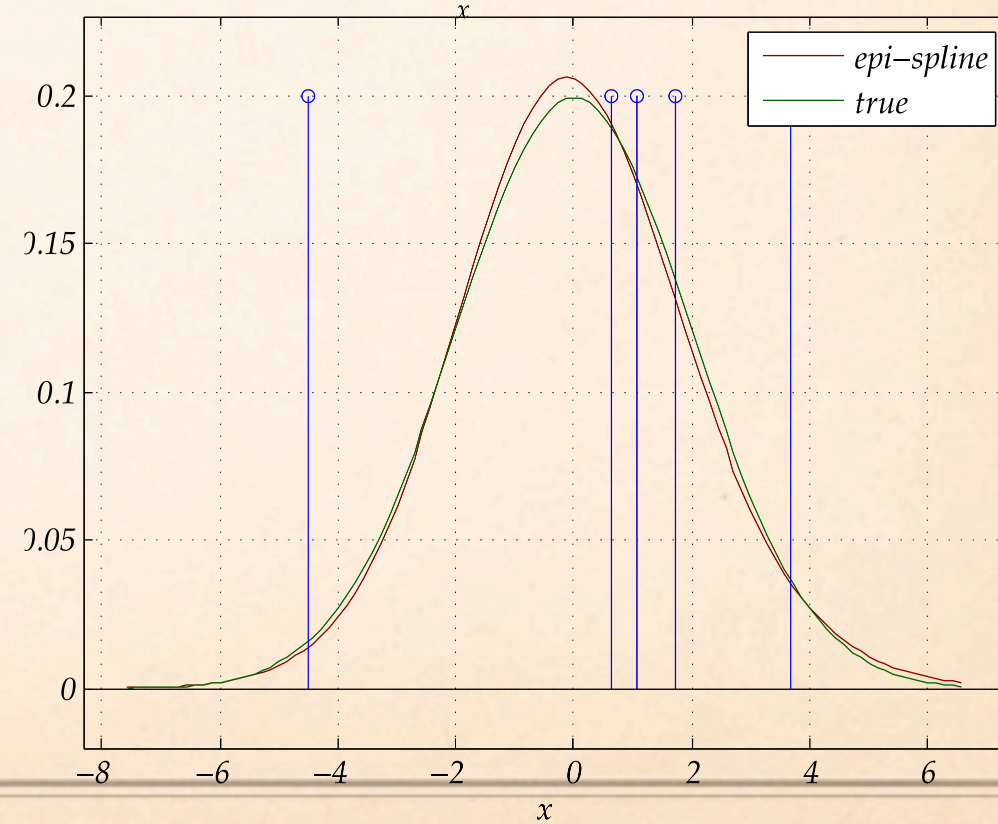


NORMAL DISTRIBUTION
KERNEL & EPI-SPLINE ESTIMATES

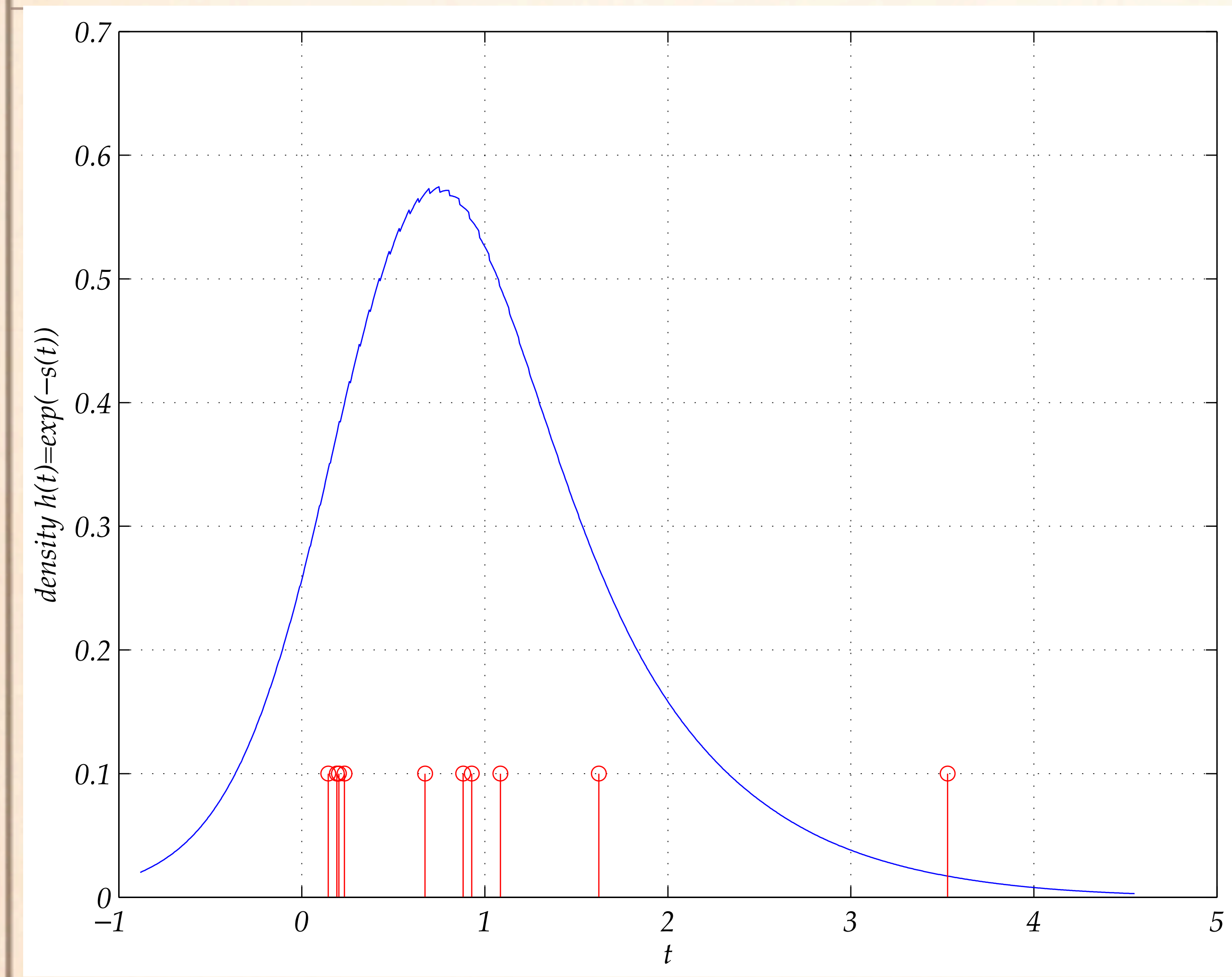


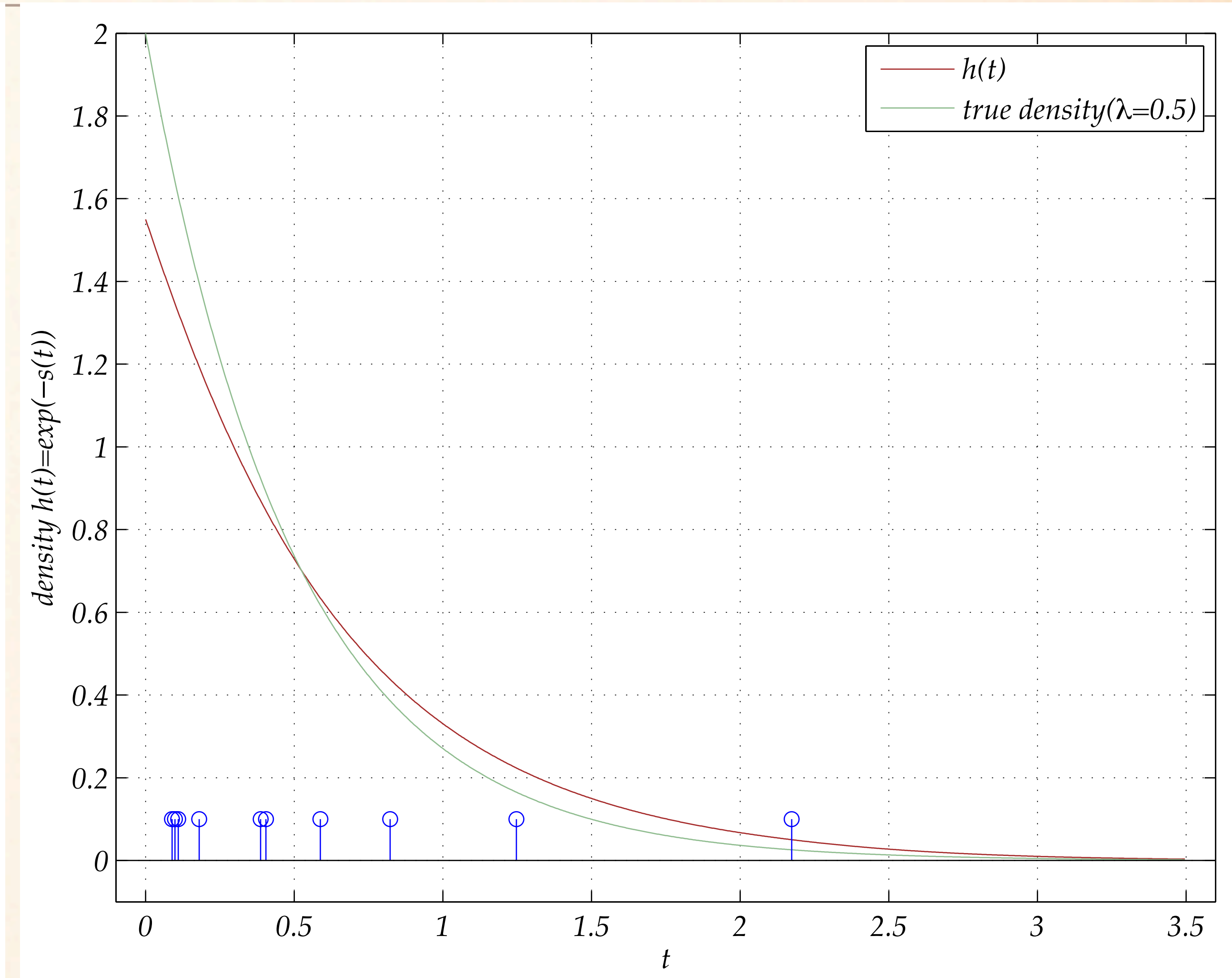






KERNEL & EPI-SPLINE ESTIMATED SAMPLES FROM EXPONENTIAL DISTRIBUTION





kernel-based estimate
R-stat

epi-spline based estimate
soft info: support R_+
h decreasing

FUNCTIONAL LLN

[Kolmogorov, Mourier ('53)] X separable Banach space

$\{f^\nu : \Xi \times X \to \mathbb{R}, \nu \in \mathbb{N}\}$, iid and $\mathbb{E}\{|f(\boldsymbol{\xi},\cdot)|\} < \infty$. Then,

$$\forall x \in X,\ \frac{1}{\nu}\sum_{k=1}^{\nu} f^k(\boldsymbol{\xi},x) \to Ef^k(x) = \mathbb{E}\{f^k(\xi,x)\}\ \ \mu\text{-a.s.}$$

For our purposes:

a) $\mathbb{E}\{|f(\boldsymbol{\xi},\cdot)|\} < \infty$ but f is $\overline{\mathbb{R}}$-valued!

b) convergence is pointwise $\not\Rightarrow$ $\arg\min$ convergence

LLN FOR RANDOM LSC FUNCTIONS

$\xi^1, \xi^2, \ldots$ iid samples of $\boldsymbol{\xi}$ (or p.iid)

(random) empirical measure P^ν, $\text{prob}\left[\boldsymbol{\xi} = \xi^k\right] = \nu^{-1}$

$$\boldsymbol{\xi}^\infty = (\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \ldots), \text{"}\mathbb{E}^\nu\{f(\boldsymbol{\xi}^\infty, x)\}\text{"} = E^\nu f(x)$$

$$= \int f(\xi, x) P^\nu(d\xi) = \frac{1}{\nu} \sum_{k=1}^{\nu} f(\xi^k, x)$$

LLN (# 1): $E^\nu f \underset{\text{epi}}{\longrightarrow} Ef \quad \mu^\infty\text{-}a.s.$ $\quad \xi \mapsto \inf_{x \in X} f(\xi, x)$ summable

$$\text{"}\arg\min_{x \in X} E^\nu f \to \arg\min_{x \in X} Ef\text{"} \quad \mu^\infty\text{-a.s.}$$

f random lsc *convex* function, $\quad f(\xi, \cdot)$ convex

$\Rightarrow \partial E^\nu f \underset{\text{gph}}{\longrightarrow} \partial Ef \quad \mu^\infty$-a.s. (Attouch's Theorem)

solutions of $\partial E^\nu f \ni 0 \rightrightarrows_a$ solutions of $\partial E f \ni 0 \quad \mu^\infty$-a.s.

$\partial E^\nu f = E^\nu \partial f$? sometimes but not always

LLN (#2) RANDOM LSC FUNCTIONS

(X,d) Polish (also, a linear space, for convenience)

$\forall \lambda > 0, \rho \geq 0 : d_{\lambda,\rho}(f,g) = \sup_{x\in\rho\mathbb{B}} \left| e_\lambda f(x) - e_\lambda g(x) \right|,\ f,g$ proper lsc-fcns(X)

metric: $m_\lambda(f,g) = \int_0^\infty e^{-\rho} d_{\lambda,\rho}(f,g)\, d\rho$, topologically $\equiv dl$, τ_{aw} topology

$\lambda > 0$ sufficiently small, $\quad m_\lambda d(f^\nu, f) \to 0 \Leftrightarrow f^\nu \underset{\text{epi}}{\longrightarrow} f$

$\left(\xi^0, \xi^1, \ldots\right)$ iid values in $\Xi \subset \mathbb{R}^N$, support of μ

$f : \Xi \times X \to \overline{\mathbb{R}}$ random lsc function such that

$S = \left\{ f(\xi,\cdot) \,\middle|\, \xi \in \Xi \right\}$ separable subspace of (proper lsc-fcns(X), τ_{aw})

$d_{\lambda,\rho}\left(E^\nu f, e_\eta E^\nu f\right) \searrow 0$ μ^∞-a.s. as $\eta \searrow 0$, $\forall$ samples $\vec{\xi}^{\,\nu}$

$d_{\lambda,\rho}\left(E f, e_\eta E f\right) \searrow 0$ μ^∞-a.s. as $\eta \searrow 0$.

f random lsc convex function

Then, $E^\nu f \underset{aw\text{-epi}}{\longrightarrow} E f$ μ^∞-a.s

AUTO-REGRESSIVE TIME SERIES

$$Y_t = a_0 + a_1 Y_{t-1} + \cdots + a_p Y_{t-p} + \zeta_t,\ t = \ldots,0,1,\ldots$$

data: $\eta_{1-p},\ldots,\eta_\nu$, non-data info: $a_1 \geq a_2 \geq \cdots \geq a_p$,

$$f(\xi^t, a) = \begin{cases} \left|\eta_t - a_0 - \eta_{t-1}a_1 - \cdots - \eta_{t-p}a_p\right|^2 & \text{if } a_1 \geq \cdots \geq a_p, \\ \infty & \text{otherwise} \end{cases}$$

$f : \Xi \times \mathbb{R}^{p+1} \to \mathbb{R}$, $\{\xi^t\}$ stationary

$$\left(\bar{a}_0^\nu, \bar{a}_1^\nu, \ldots, \bar{a}_p^\nu\right) \in \arg\min \frac{1}{\nu}\sum_{t=1}^{p} f(\xi^t, a)$$

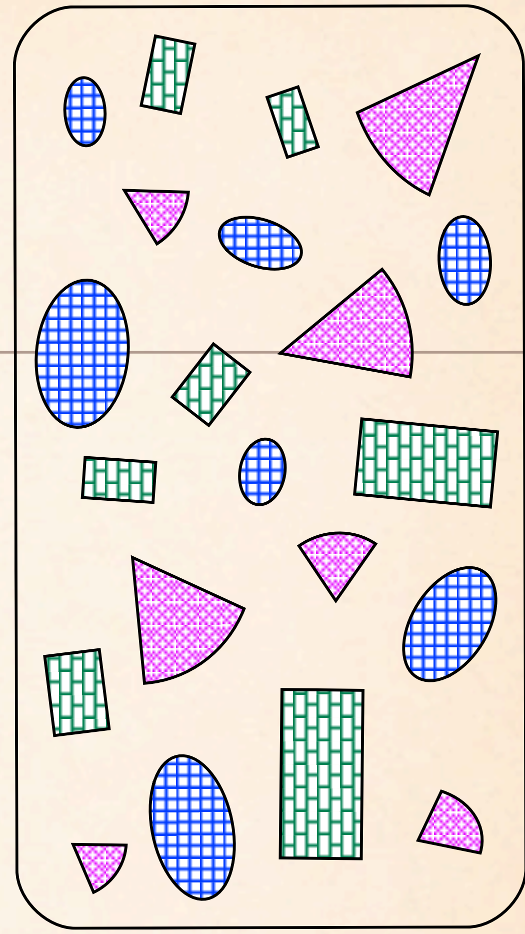
$$\overset{?}{\to} \left(\bar{a}_0, \bar{a}_1, \ldots, \bar{a}_p\right) \in \arg\min \mathbb{E}\left\{ f(\xi, a) \,\middle|\, \mathcal{I} \right\} \text{ (invariant field)}$$

Ergodic Theorem: $f : \Xi \times X \to \overline{\mathbb{R}}$ random lsc fcn,

$\vartheta : \Xi \to \Xi$ ergodic measure preserving transformation

$$\frac{1}{\nu}\sum_{k=1}^{\nu} f(\vartheta^k(\xi), \bullet) \xrightarrow{epi} Ef \ a.s., \ \inf_x f(\bullet, x) \text{ summable}$$

HOMOGENIZATION



conductor: $\Omega \subset \mathbb{R}^3$, composite ≥ 2 materials,

$\neq$ conductivity spatial location: $a(\boldsymbol{\xi},x)$ dependend

$0 \leq a(\boldsymbol{\xi},x) \leq \kappa_{\text{bdd}}$, stationary process w.r.t. location

heat u: with rapidly varying stochastic coefficients

$$\nabla \bullet \big(a(\boldsymbol{\xi},x)\nabla u(\boldsymbol{\xi},x)\big) = h(x), \; x \in \Omega$$

$$\text{u}(\boldsymbol{\xi},x) = 0, \; x \in \text{ bdry } \Omega$$

homogenized equation with effective coefficient a

$$\nabla \bullet \big(a(x)\nabla u(x)\big) = h(x), \; x \in \Omega$$

$$\text{u}(\xi) = 0, \; x \in \text{ bdry } \Omega$$

such that $u(x) = \mathbb{E}\{u(\boldsymbol{\xi},x)\}$. $\qquad a(x) \neq \mathbb{E}\{a(\boldsymbol{\xi},x)\}$

HOMOGENIZATION
A "NUMERICAL" METHODOLOGY

$$\min_{u\in H_0^1(\Omega)} g(\xi,u) = \frac{1}{2}\int_\Omega a(\xi,x)|\nabla u|^2 dx - \langle h,u\rangle$$

$g : L^2 \to (-\infty]$, to be minimized for all ξ

homogenization: find g^{hom} such that

$$\mathbb{E}\{u(\xi,\cdot)\} = \bar{u}(\cdot) \in \arg\min\left[g^{\text{hom}}(u) \,\middle|\, u \in H_0^1(\Omega)\right]$$

—∘ —∘ —∘ —∘ conjugate duality —∘ —∘ —∘ —∘

$$g^{\text{hom}}(u) = \left(\text{epi-}\int_\Xi g(\xi,\cdot)\star P(d\xi)\right)(u) \text{ on } H_0^1(\Omega)$$

$$\left(\text{epi-}\int_\Xi g(\xi,\cdot)\star P(d\xi)\right)(x) =$$ epi-integral

$$\inf_{z(\cdot)}\left\{\int_\Xi g(\xi,z(\xi))P(d\xi) \,\middle|\, \int_\Xi z(\xi)P(\xi) = x\right\}$$

approximation: $\theta\star f(\cdot) = \theta f(\theta^{-1}\cdot)$

$$\bar{u}^\nu \in \arg\min\left\{\nu^{-1}\star\left(g(\xi^1,\cdot)\#\cdots\# g(\xi^\nu,\cdot)\right) \,\middle|\, u\in H_0^1(\Omega)\right\}$$

Ergodic Theorem

RANDOM LSC FUNCTIONS
[ROCKAFELLAR: NORMAL INTEGRANDS]

$f : \Xi \times X \to \overline{\mathbb{R}}$ a random lsc function

(X, d) Polish space, Borel field $\mathcal{B}$

$(\Xi, \mathcal{A}, \mathcal{P})$ probability space

(a) $f(\xi, \bullet)$ lsc $\forall \xi \in \Xi$

(b) $(\xi, x) \mapsto f(\xi, x)$ $\mathcal{A} \times \mathcal{B}^n$-measurable (jointly)

(a & b) imply $\xi \mapsto \text{epi}\, f(\xi, \bullet) = \text{epi}\, f(\xi)$ is a closed random set

$\text{epi}\, f : \Xi \rightrightarrows X \times \mathbb{R}$ ($\Rightarrow$ all properties can be trasferred)

$\text{epi}\, f(\xi) \subset X \times \mathbb{R}$ closed epigraph for all $\xi \in \Xi$

& $(\text{epi}\, f)^{-1}(O) = \{\xi \,|\, \text{epi}\, f(\xi) \in O\} \in \mathcal{A}$, $\forall O \subset X \times \mathbb{R}$, open

$\Rightarrow$ dom epi $f = (\text{epi}\, f)^{-1}(X \times \mathbb{R}) \in \mathcal{A}$ (measurable set)

EXAMPLES: RANDOM LSC FCNS

$f:\Xi\times X\to\mathbb{R}$, $\mathcal{A}$-measurable in ξ, continuous in x

f and $-f$ are random lsc functions

f is a *Carathéodory random lsc function*

$f(\xi,x)\equiv g(x),\quad g$ lsc

Ξ Borel subset of $\mathbb{R}^d, f$ lsc $\Rightarrow$ random lsc function

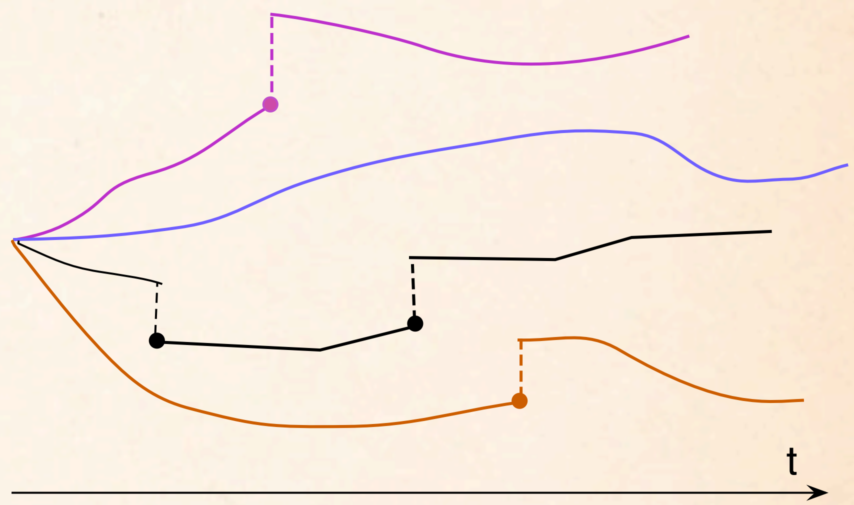
$f(\xi,x)=\iota_{C(\xi)}(x)$ with C random closed set

$f(\xi,x)=f_0(\xi,x)+\iota_{C(\xi)}(x)$ is a random lsc function

when f_0 random lsc function, C random closed set

Proof. $\quad \operatorname{epi} f(\xi)=\operatorname{epi} f_0(\xi)\cup\left[C(\xi)\times\mathbb{R}\right]$

RANDOM LSC FCNS: PROPERTIES



f random lsc function

$\Rightarrow$ $\text{lev}_\alpha f(\xi,\bullet)$ random closet set

$\Rightarrow \xi \mapsto p(\xi) = \inf_X f(\xi,\bullet)$ is $\mathcal{A}$-measurable

$p(\xi) = \inf_\nu \alpha_\nu(\xi)$ with

$\{(x^\nu,\alpha_\nu)\}_{\nu\in\mathbb{N}}$ Castaing representation of epi f

$\Rightarrow \xi \mapsto \arg\min f(\xi,\bullet) = \left\{x \,\middle|\, f(\xi,x) \le p(\xi)\right\}$ is $\mathcal{A}$-measurable

$\Rightarrow \exists$ $\mathcal{A}$-measurable selections: $f(\xi,\bar{x}(\xi)) = \min_X f(\xi,\bullet)$

Moreau envelopes: $e_{\lambda(\xi)} f(\xi,\cdot)$ are random lsc functions

$\lambda(\xi) > 0$ sufficiently small

$e_{\lambda(\xi)} f$ Carathéodory random lsc function

RANDOM CONSTRAINT SYSTEMS

$C : \Xi \rightrightarrows \mathbb{R}^n$ random closed set

$f_i : \Xi \times \mathbb{R}^n \to \overline{\mathbb{R}}$ random lsc function, $\mathrm{i} \in I_1$

$f_i : \Xi \times \mathbb{R}^n \to \mathbb{R}$ Carathéodory random lsc fcn, $\mathrm{i} \in I_2$

$\alpha_i : \Xi \to \mathbb{R}$ random variable $i \in I_1 \cup I_2$ (countable index)

Then, $S : \Xi \rightrightarrows \mathbb{R}^n$ is a random closed set where

$$S(\xi) = \left\{ x \in C(\xi) \middle| \begin{array}{l} f_i(\xi, x) \le \alpha_i(\xi),\ i \in I_1 \\ f_i(\xi, x) = \alpha_i(\xi),\ i \in I_2 \end{array} \right\}$$

EPI-TOPOLOGY: REVIEW

$\{f^\nu : \mathbb{R}^n \to \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$

$\text{epi}(\text{e-li}_\nu f^\nu) = \text{Ls}_\nu \,\text{epi}\, f^\nu, \quad \text{epi}(\text{e-ls}_\nu f^\nu) = \text{Li}_\nu \,\text{epi}\, f^\nu$

epi-limit: $f^\nu \underset{e}{\to} f$ when $f = \text{e-li}_\nu f^\nu = \text{e-ls}_\nu f^\nu$, $f = \text{e-lm}_\nu f^\nu$

Hit-and-Miss topology translated to lsc-fcns($\mathbb{R}^n$) : τ_{epi}

subbase: hit open sets, miss compact sets

$\{g \in \text{lsc-fcns}(\mathbb{R}^n) \mid \inf_O g < \alpha\} \quad \{g \in \text{lsc-fcns}(\mathbb{R}^n) \mid \inf_K g > \alpha\}$

(lsc-fcns($\mathbb{R}^n$), τ_{epi}) compact, metrizable space -- *dl* a metric

$f \geq \text{e-ls}_\nu f^\nu \Leftrightarrow \limsup_\nu (\inf_O f^\nu) \geq \inf_O f, \ \forall O$ open ($\inf_o$ usc)

$f \leq \text{e-li}_\nu f^\nu \Leftrightarrow \liminf_\nu (\inf_K f^\nu) \leq \inf_K f, \ \forall K$ compact ($\inf_K$ lsc)

SCALARIZATION

Effrös field for lsc-fcns($\mathbb{R}^n$): $\mathcal{E}$ ($= \mathcal{B}_X, X$ Polish)

generated by $\mathcal{A}_{D,\alpha} = \left\{ f \in \middle| \inf_D f \le \alpha \right\}$, D closed or open

distribution of a random lsc function $f : \Xi \times \mathbb{R}^n \to \overline{\mathbb{R}}$

$P_f(\mathcal{A}) = \mu\left\{ \xi \in \Xi \middle| f(\xi, \bullet) \in \mathcal{A} \right\}$, $\mathcal{A} \in \mathcal{E}$

$\pi_D(\xi) = \inf_{x \in D} f(\xi, x)$ with $f : \Xi \to$ lsc-fcns(X), X Polish

then f random lsc fcn $\Leftrightarrow \xi \mapsto \pi_D(\xi)$ measurable for all $D \in \mathcal{D}$

where $\mathcal{D}$ is anyone of the following:

a) all closed (or open) sets

b) all open rational balls, centers at R dense subset of X

c) if X is σ-compact, all closed rational balls, …

$\left\{ \pi_{x,\rho},\ x \in R, \rho \in \mathbb{Q}_+ \right\}$ countable collection $\overline{\mathbb{R}}$-valued r.v.

inherit independence, identically distributed

ERGODICITY

$\left\{ f^\nu : \Xi \times X \to \bar{\mathbb{R}},\ \nu \in \mathbb{N} \right\}$ random lsc functions

$\left\{ \pi^\nu_{x,\rho} = \inf_{\mathbb{B}(x,\rho)},\ x \in R, \rho \in \mathbb{Q}_+ \right\}$ countable collection $\bar{\mathbb{R}}$-valued r.v.

f^ν independent: $\left\{ \nu_1, \nu_2, \ldots, \nu_k \right\} \subset \mathbb{N},\ \alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R} \cup \{\infty\}$

$$x^1, x^2, \ldots, x^k \in R,\ \rho_1, \rho_2, \ldots, \rho_k \in \mathbb{Q}_{++}$$

$$P\left\{ \xi \in \Xi \,\middle|\, \pi^{\nu_i}_{\mathbb{B}(x^i,\rho_i)}(\xi) < \alpha_i, i = 1 \to k \right\} = \prod_{i=1}^k P\left\{ \xi \in \Xi \,\middle|\, \pi^{\nu_i}_{\mathbb{B}(x^i,\rho_i)}(\xi) < \alpha_i \right\}$$

stationarity: joint distributions of $f^{\nu_1}, \ldots, f^{\nu_k}$ invariant under shift

$\varphi : \Xi \to \Xi$ meas. preserving transformation: $P(\varphi^{-1}(A)) = P(A), \forall A \in \mathcal{A}$

$\mathcal{I}$ invararian σ-field: $P(\varphi^{-1}(A) \vartriangle A) = 0$

φ ergodic if $\mathcal{I}$ is trivial $P(A) \in \{0,1\} \forall A \in \mathcal{I}$

$\left\{ f^\nu,\ \nu \in \mathbb{N} \right\}$ ergodic $\Leftrightarrow \left\{ f \circ \varphi^\nu, \nu \in \mathbb{N} \right\}$ for φ associated meas. p. transform.

$\Rightarrow \forall O$ open $\left\{ \pi_O \circ \varphi^\nu, \nu \in \mathbb{N} \right\}$ ergodic sequence of $\bar{\mathbb{R}}$-valued r.v.

ERGODIC THEOREM

(X,d) Polish, $(\Xi,\mathcal{A},\mu)$ probability space

$\varphi : \Xi \to \Xi$ a measure preserving transformation

$\mathcal{I}$ its invariant σ-field

$f : \xi \times X \to \overline{\mathbb{R}}$ random lsc function, inf-locally summable:

$\forall x \in X, \exists V \in \mathcal{N}_{\text{closed}}(x) :\ \mathbb{E}\left\{\pi_X(\xi) = \inf_V f(\xi,x)\right\} > -\infty$

$\Rightarrow \mathcal{R} \subset \mathcal{A},\ \exists\, E^{\mathcal{R}} f(\xi,\cdot)$ random lsc fcn

$\exists$ countable dense subset of epi $E^{\mathcal{I}} f(\xi,\cdot)$ μ-a.s.

epi $f(\xi,\cdot)$ solid set, cl(int(epi g)) = epi g (cont. on dom g)

Then,

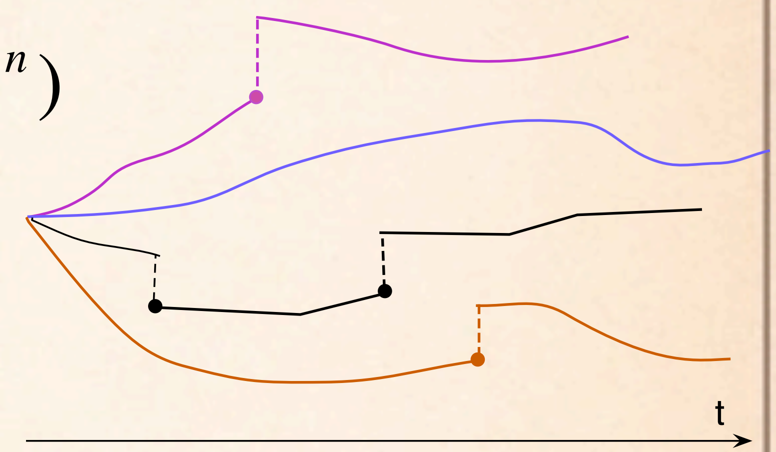
$$\frac{1}{\nu}\sum_{k=1}^{\nu} f(\varphi^{\nu}(\xi),\cdot) \underset{\text{epi}}{\to} E^{\mathcal{I}} f(\xi,\cdot)\ \ \mu\text{-a.s.}$$

$\underset{\text{epi}}{\to} Ef$ μ-a.s. when φ is ergodic

LSC STOCHASTIC PROCESSES

$\left\{ f^{\nu} : \mathbb{R}^n \to \overline{\mathbb{R}}, \nu \in \mathbb{N} \right\}$ stochastic process with lsc-paths ($\mathbb{R}^n$)

(lsc-fcns($\mathbb{R}^n$), τ_{epi}) compact metrizable space -- dl a metric



τ_{epi} can be generated by $\quad x \in \mathbb{Q}^n, \rho \in \mathbb{Q}_{++}, \alpha \in \mathbb{Q}$

$$\left\{ g \in \text{ lsc-fcns}(\mathbb{R}^n) \,\middle|\, \inf_{\mathbb{B}^o(x,\rho)} g < \alpha \right\} \quad \left\{ g \in \text{ lsc-fcns}(\mathbb{R}^n) \,\middle|\, \inf_{\mathbb{B}(x,\rho)} g > \alpha \right\}$$

a.s.-, in probability, in distribution convergence ~ as for their epigraphs

process $\left\{ f^{\nu} : \mathbb{R}^n \to \overline{\mathbb{R}}, \nu \in \mathbb{N} \right\}$ converges in distribution in

'classical' sense (i.e., pointwise) $\Rightarrow$ epi-converge in distribution

if the paths $x \mapsto f^{\nu}(\xi, x)$ are equi-lsc μ-a.s.:

$$\forall x, \varepsilon > 0,\ \exists V \in \mathcal{N}(x) : \inf \left\{ f^{\nu}(\xi, x) \,\middle|\, x \in V \right\} > \min \left[\varepsilon^{-}1, \inf f^{\nu}(\xi, \cdot) - \varepsilon \right]$$

DISTANCE & INDICATOR FUNCTIONS

$\left\{C^\nu : \Xi \Rightarrow \mathbb{R}^m, \nu \in \mathbb{N}\right\}$ random closed sets

(set-)converge in distribution to $C \Leftrightarrow$

processes $\left\{d(C^\nu(x,\cdot), \nu \in \mathbb{N}\right\}$ convergerve

in distribution to $d(C(x,\cdot)$ for all x $\in \mathbb{R}^m$

$C^\nu(\xi) =$ ray through $(1, \nu^{-}1)$

$C(\xi) = \{(0,0)\}, \ \tilde{C}(\xi) =$ ray through $(1,0)$

C^ν converges in distribution to $\tilde{C}$ (clearly)

but ι_{C^ν} converges in distribution to ι_C (exercise)

... IN DISTRIBUTION OF SELECTIONS

$\{C^\nu : \Xi \rightrightarrows \mathbb{R}^m, \nu \in \mathbb{N}\}$ random closed sets

converge in distribution to C. Then,

$\exists$ measurable selections x^ν of C^ν

converging in distribution to a seclection of C.

(also holds for Castaing representations)

$\{f^\nu : \Xi \times X \to \overline{\mathbb{R}},\ \nu \in \mathbb{N}\}$ random lsc functions

epi-converge in distribution to f

$\xi \mapsto \arg\min f^\nu(\xi, \cdot)$ are random sets $\Rightarrow$

selections (minimizers) converge in distribution if

$(\arg\min f^\nu)$ converge in distribution

1. $f^\nu \to f$ + epi-tightness $\Rightarrow \inf f^\nu \to \inf f$ in distribution
2. + under μ-a.s. convergence, $\text{argmin}\, f^\nu(\xi, \cdot) \rightrightarrows_\nu \arg\min f(\xi, \cdot)$

V. EXPECTATION FUNCTIONALS ⊡⊡ CALCULUS

EXPECTATION FUNCTIONALS

$f : \Xi \times \mathcal{X} \to \bar{\mathbb{R}}$, random lsc function,

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

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

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

$$= \infty \text{ whenever } \int_\Xi f_+(\xi, x(\xi))\,\mu(d\xi) = \infty$$

$Ef : \mathcal{X} \to \bar{\mathbb{R}}$ always defined

Regression: ($\mathcal{X}$ is not a linear space)

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

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

DECOMPOSABILITY

$\mathcal{X} \subset \mathcal{M}$ decomposable (w.r.t. μ) when

$\forall x^0 \in \mathcal{X}, A \in \mathcal{A}$ and $x^1 : A \to \mathbb{R}^n \in \mathcal{M},$ bounded

$$x(\xi) = \begin{cases} x^0(\xi) & \text{for } \xi \in \Xi \setminus A \\ x^1(\xi) & \text{for } \xi \in A \end{cases}$$

$\Rightarrow \mathcal{X}$ is a linear space $(0 \in \mathcal{X})$

$\mathcal{L}^p(\Xi, \mathcal{A}, \mu; \mathbb{R}^n)$, $\mathcal{M}$ are decomposable

$C(\Xi, \mathcal{A}; \mathbb{R}^n)$, constant-fcns($\Xi$) not decomposable

f random lsc function, $Ef \not\equiv \infty$ on $\mathcal{X}$. Then,

$$\inf_{x \in \mathcal{X}} \int_\Xi f(\xi, x(\xi))\,\mu(d\xi) = \int_\Xi \left[\inf_{x \in \mathbb{R}^n} f(\xi, x) \right] \mu(d\xi)$$

$\bar{x} \in \arg\min_{x \in \mathcal{X}} Ef(x) \Leftrightarrow \bar{x}(\xi) \in \arg\min_{x \in \mathbb{R}^n} f(\xi, x)$ μ-a.s. $(\inf ES > -\infty)$

INTERCHANGE OF SUBDIFFERENTIATION AND INTEGRATION

$f : \Xi \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, random convex lsc function,

$Ef : \mathcal{X} \to \bar{\mathbb{R}}$, here $\mathcal{X} = \mathcal{L}^\infty(\Xi, \mathcal{A}, \mu; \mathbb{R}^n)$

$\mathcal{G}$ subfield of $\mathcal{A}$ (possibly the trivial field $= \{\emptyset, \Xi\}$)

$f^{\mathcal{G}} : \Xi \times \mathbb{R}^n \to \bar{\mathbb{R}}$ with $f^{\mathcal{G}}(\xi, x) = \int_\Xi f(\zeta, x) P^{\mathcal{G}}(d\zeta \,|\, \xi)$

$E f^{\mathcal{G}} : \tilde{\mathcal{X}} \to \bar{\mathbb{R}}$, here $\tilde{\mathcal{X}} = \mathcal{L}^\infty(\Xi, \mathcal{G}, \mu; \mathbb{R}^n)$

assume Ef & Ef^* finite for some $\mathcal{G}$-measurable functions $x(\bullet)$

$\partial^* Ef \subset (\mathcal{L}^\infty_n)^* = \mathcal{L}^1(\Xi, \mathcal{A}, \mu; \mathbb{R}^n) \oplus S(\Xi, \mathcal{A}, \mu; \mathbb{R}^n)$

$\partial^* Ef(x) = \partial Ef(x) + N^S_{\text{dom } Ef}(x)$, $\partial Ef(x) = \left\{ v \in \mathcal{L}^1 \,\middle|\, v(\xi) \in \partial f(\xi, x(\xi)) \ \mu\text{-a.s.} \right\}$

$\partial^* Ef^{\mathcal{G}}(x) = \partial Ef^{\mathcal{G}}(x) + N^S_{\text{dom } Ef^{\mathcal{G}}}(x)$, $\partial Ef^{\mathcal{G}}(x) = \ldots$ in $\mathcal{L}^1_n(\mathcal{G})$

$x \in \mathcal{L}^\infty_n$, $\partial Ef^{\mathcal{G}}(x) = \mathbb{E}^{\mathcal{G}}\{\partial f(\xi, x(\xi))\}$ μ-a.s. ?

VALIDATING THE INTERCHANGE

$f : \Xi \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, random convex lsc fcn, $Ef : \mathcal{L}_n^\infty \to \bar{\mathbb{R}}$, $\mathcal{G} \subset \mathcal{A}$,

suppose $\xi \mapsto D(\xi) = \text{cl dom} f(\xi,\cdot)$ $\mathcal{G}$-measurable

$Ef(x) < \infty : \forall$ $\mathcal{G}$-measurable selection of D (+ summ. condition)

Then, $\forall x \in \mathcal{L}_n^\infty(\mathcal{G}) : \partial Ef^{\mathcal{G}}(\boldsymbol{\xi}, x(\boldsymbol{\xi})) = \mathbb{E}^{\mathcal{G}}\{\partial f(\boldsymbol{\xi}, x(\boldsymbol{\xi}))\}$ μ-a.s.

i.e., the closed-valued $\mathcal{G}$-measurable mappings $\partial Ef^{\mathcal{G}} = \mathbb{E}^{\mathcal{G}} \partial f$ μ-a.s.

$\inf \ \mathbb{E}\left\{f(\boldsymbol{\xi}, x^1(\boldsymbol{\xi}), x^2(\boldsymbol{\xi}))\right\}$, $x^1 \in \mathcal{L}_n^\infty(\mathcal{G})$, $x^2 \in \mathcal{L}_n^\infty(\mathcal{A})$

$\exists\, h \in \mathcal{L}^1(\mathcal{A})$ such that $(x^1, x^2) \in \text{dom} f(\xi,\cdot,\cdot) \Rightarrow \left| f(\xi, x^1, x^2) \right| < \infty$

$\xi \mapsto D_1(\xi) = \text{cl}\left\{ x^1 \in \mathbb{R}^n \,\middle|\, \exists x^2 : f(\xi, x^1, x^2) < \infty \right\}$ $\mathcal{G}$-measurale

$\inf Eg(x^1)$ on $x^1 \in \mathcal{L}_n^\infty(\mathcal{G})$ with $g(\xi, x^1) = \mathbb{E}^{\mathcal{G}}\left[\inf_{x^2 \in \mathbb{R}^n} f(\cdot, x^1, x^2)\right](\xi)$

is "equivalent" to given problem. [Consider $\mathcal{G} = \left\{\varnothing, \mathbb{R}^n\right\}$]

$\min_{x, y_\xi} \left\{ -x^1 \,\middle|\, x^1 + x_\xi^2 \le \xi, x^1 \in [0,2], x_\xi^2 \ge 0 \right\} = \min\left(Ef(x) = \mathbb{E}\{f(\boldsymbol{\xi}, x)\}\right)$

$f(\xi, x) = -x^1 + \iota_{[0,2]} + \iota_{(-\infty,\xi]} = -x^1 + \iota_{[0,\xi]}$

$\mathcal{G} = \{\varnothing, [1,2]\}$, $D(\xi) = [0,\xi]$ is not $\mathcal{G}$-measurable!

PRICING A CONTINGENT CLAIM

environment process: $\left\{\boldsymbol{\xi}^t \in \mathbb{R}^d\right\}_{t=0}^T$ history: $\vec{\xi}^t$, $\xi = \xi^T$

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

contingent claims: $\left\{G^t(\vec{\xi}^t)\right\}_{t=1}^T$; investment strategy: $\left\{X^t(\vec{\xi}^t)\right\}_{t=0}^T$

portfolio value at t : $\langle S^t(\vec{\xi}^t), X^t(\vec{\xi}^t)\rangle$

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

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

(T+1)-stage linear stoch. opt. $\langle S^0, X^0\rangle \le G^0,\ \langle S^T, X^T\rangle \le G^T$ a.s.

feasible if $G^0 + \cdots + G^T \ge 0\ \ \forall \xi$, with arbitrage when unbounded

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

RISK NEUTRAL PROBABILITIES: DUALITY

pricing via risk-neutral probabilities (obtained from dual variables)

$$f(\xi, x(\xi)) = \begin{cases} -\langle S^T(\xi), X^T(\xi)\rangle & \text{when } x(\xi) \in C(\xi) \\ \infty & \text{otherwise} \end{cases}$$

$$x(\xi) = \left(X^0(\xi^0), \ldots, X^T(\overset{\to T}{\xi}) \right), \quad C(\xi) = \left\{ x(\xi) \,\middle|\, \text{satisfies the constraints a.s.} \right\}$$

$$\min_{x \in \mathcal{N} \subset \mathcal{M}} \mathbb{E}\left\{ f(\xi, x(\xi)) \right\}, \quad f \text{ random convex lsc function}$$

$\mathcal{L}^1$-"Perfect" duality: (1) $\mathbb{C}.\mathbb{Q}.$ $(\mathcal{M} = \mathcal{L}_n^\infty)$, (2) $\xi \mapsto C(\xi)$ nonanticipative

$$\forall\, t : \mathbb{E}\left\{ C(\boldsymbol{\xi}) \,\middle|\, \overset{\to t}{\boldsymbol{\xi}} \right\} \ \overset{\to t}{\boldsymbol{\xi}}\text{-measurable (depend only on past history)}$$

Pricing a contingent claim doesn't satisfy (2) $\Rightarrow$ no "perfect" duality

Full duality requires: dual variables $\in \mathcal{L}_n^1 \oplus S_n$, but ...

i.e.,the risk-neutral probabilities are in $\mathcal{L}_n^1 \oplus S_n$!

SOLUTION PROCEDURES
FOR STOCHASTIC VARIATIONAL PROBLEMS

INFORMATION-DECISION PROCESS

$$\xi^0 \to \; x^1(\xi^0) = x^1 \;\to\; \xi^1 \;\to\; x^2(\xi^0, \xi^1) \;\to \xi^2$$

More specifically,

(dynamic) Stochastic Programs with Recourse:

$$\min_{x \in \mathcal{N}^a} \mathbb{E}\{ f(\boldsymbol{\xi}, x(\boldsymbol{\xi})) \}$$

time scale: $t = 0, 1, 2, \ldots, T$, $\quad x(\xi) = \left(x^1(\xi), \ldots, x^T(\xi) \right)$

$$\xi = (\xi^0, \xi^1, \ldots, \xi^T)$$

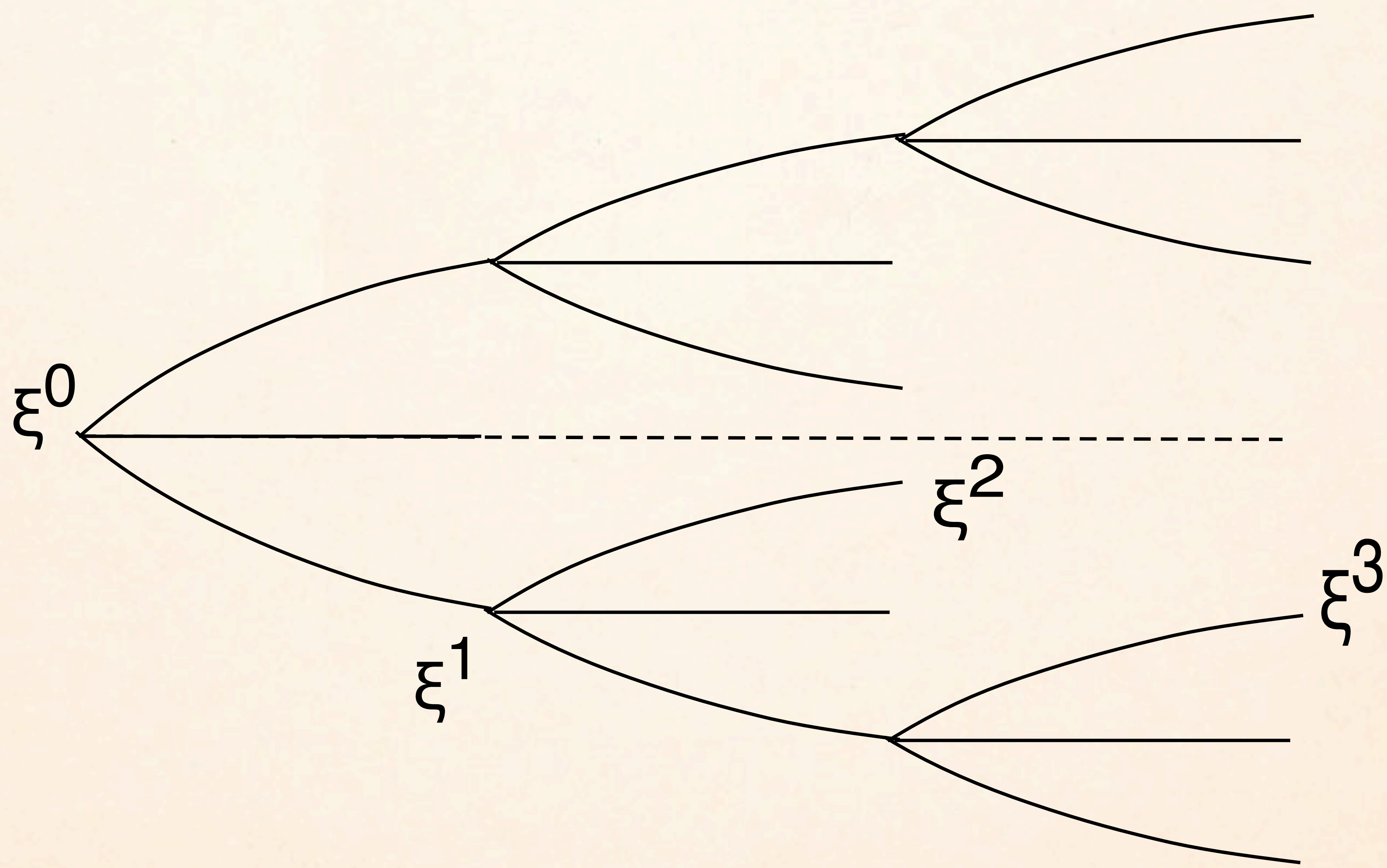
information (observation) available at time t: $\mathcal{A}_{t-1}$

filtration : $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_T = \mathcal{A}$, $\mathcal{A}_0$ trivial

$x \in \mathcal{N}^a$ if x^t $\mathcal{A}_{t-1}$-measurable $\approx \sigma$-field($\overset{\to \nu - 1}{\boldsymbol{\xi}}$)

here $\boldsymbol{\xi}^0$ deterministic, $x^1(\boldsymbol{\xi}) \equiv x^1$

DISCRETE SCENARIO TREE



DETERMINISTIC EQUIVALENT

$$\min_{x\in\mathcal{N}^a} \mathbb{E}\{f(\boldsymbol{\xi},x(\boldsymbol{\xi}))\} = \mathbb{E}\left\{\mathbb{E}\left\{\mathbb{E}\left\{f(\boldsymbol{\xi},x(\boldsymbol{\xi}))\middle|\mathcal{A}_T\middle|\cdots\middle|\mathcal{A}_1\middle|\mathcal{A}_0\right\}\right\}\right\}$$

"time-staged objective":

$$= f_1(x^1) + \mathbb{E}\left\{f_2(\xi;x^1,x^2(\xi)) + \mathbb{E}\left\{f_3(\xi;x^1,x^2(\xi),x^3(\xi)\middle|\mathcal{A}_2\right\}\middle|\mathcal{A}_1\right\} \quad \cdots$$

$$= f_1(x^1) + \mathbb{E}\left\{f_2(\xi;x^1,x^2(\xi)) + EQ_2(\xi;x^1,x^2(\xi))\middle|\mathcal{A}_1\right\}$$

$$EQ_2(\xi;x^1,x^2(\xi)) = \mathbb{E}\left\{\inf_{x^3\in\mathbb{R}^{n3}} f_3(\xi;x^1,x^2(\xi),x^3)\middle|\mathcal{A}_2\right\}$$

$$= f_1(x^1) + \mathbb{E}\left\{EQ_1(\xi;x^1,x)\middle|\mathcal{A}_1\right\}$$

$$EQ_1(\xi;x^1) = \mathbb{E}\left\{\inf_{x^2\in\mathbb{R}^{n2}} f_2(\xi;x^1,x^2) + EQ_2(\xi;x^1,x^2)\middle|\mathcal{A}_1\right\}$$

$$= f_1(x^1) + EQ_1(x^1)$$

SOLUTION PROCEDURES

$$\min_{x\in\mathcal{N}^a} \mathbb{E}\left\{f(\boldsymbol{\xi},x(\boldsymbol{\xi}))\right\} = \min_{x^1\in\mathbb{R}^{n_1}} f_1(x^1) + EQ_1(x^1)$$

$$EQ_1(\xi;x^1) = \mathbb{E}\left\{\inf_{x^2\in\mathbb{R}^{n_2}} f_2(\xi;x^1,x^2) + EQ_2(\xi;x^1,x^2)\Big|\mathcal{A}_1\right\}$$

$$EQ_2(\xi;x^1,x^2(\xi)) = \mathbb{E}\left\{\inf_{x^3\in\mathbb{R}^{n_3}} f_3(\xi;x^1,x^2(\xi),x^3)\Big|\mathcal{A}_2\right\}$$

deterministic optimization! convex when f random lsc convex function

in theory: any algorithmic procedure!

hurdles: values, (sub)gradients, "Hessians" of $f_1(x^1) + EQ_1(x^1)$

are either not acessible or at best, computationally EXPENSIVE

Approaches: $\mu^\nu \sim \mu \Rightarrow$ approximating stochastic process $\left\{\boldsymbol{\xi}_t, t \le T\right\}$

sampling: a) same as approximation except μ^s random measure

b) SAA-strategy for $\partial\Big(\mathbb{E}\left\{f(\boldsymbol{\xi},x(\boldsymbol{\xi}))\right\} + N_{\mathcal{N}^a}(x(\boldsymbol{\xi}))\Big)$

SEQUENTIAL L.P. STRATEGY

$\min f_0(x), \quad x \in X \in \mathbb{R}^n, \quad f_0$ linear (not essential)

$\quad f_i(x) \le 0, \;\; i = 1,\dots,s, \quad f_i(s) = 0, i = s+1,\dots,m$ (affine)

in the $s+1$ first constraints: $f_i(x) = \sup_{t\in T} f_{i,t}(x), \quad f_i \ge f_{i,t}$ affine

0. $\nu = 0$, pick polytope (box) $K^0 \ni x^{opt}$

1. $x^\nu \in \arg\min f_0$ on K^ν, set $\mathrm{i}_\nu : f_{i_\nu}(x^\nu) = \max_{1\le i\le s} f_i(x^\nu)$

   if $f_{i_\nu}(x^\nu) \le 0, x^\nu$ optimal, otherwise go to 2.

2. return to 1. with $K^{\nu+1} = K^\nu \cap \left\{ \left\langle \nabla f_{i_\nu}(x^\nu), x - x^\nu \right\rangle + f_{i_\nu}(x^\nu) \le 0 \right\}$

when f_0 is not linear (but convex): $\min \theta$ such that $f_0(x) - \theta \le 0$

convergence: finite # of steps or iterates cluster to optimal sol'n

SLP FOR STOCHASTIC PROGRAMS

$\min f_1(x) + EQ_1(x)$ s.t. $Ax = b,\ x \geq 0 \quad (x = x^1)$

$$EQ_1(x) = \sum_{l=1}^{L} p_l Q_1(\xi^l, x) \quad L \text{ large}$$

$$Q_1(\xi^l, x) = \inf_{x^2 \in X_2} \left\{ f_2(\xi^l; x, x^2) + (EQ_2(\cdots) \right\}$$

$$\text{dom } EQ_1 = \bigcap_{l=1}^{L} \text{dom } Q_1(\xi^l, \cdot) = \bigcap_{l=1}^{L} \left\{ x \middle| \exists x^2 \in X_2, f_2(\xi^l; x, x^2) < \infty \right\}$$

0. $\nu = r = s = 0$

1. $\nu = \nu + 1$, solve: $\min f_1(x) + \theta,\ Ax = b,\ x \geq 0$ such that

(feasibility cuts) $\quad \langle E_k, x \rangle \geq e_k,\ k = 1 \to r$

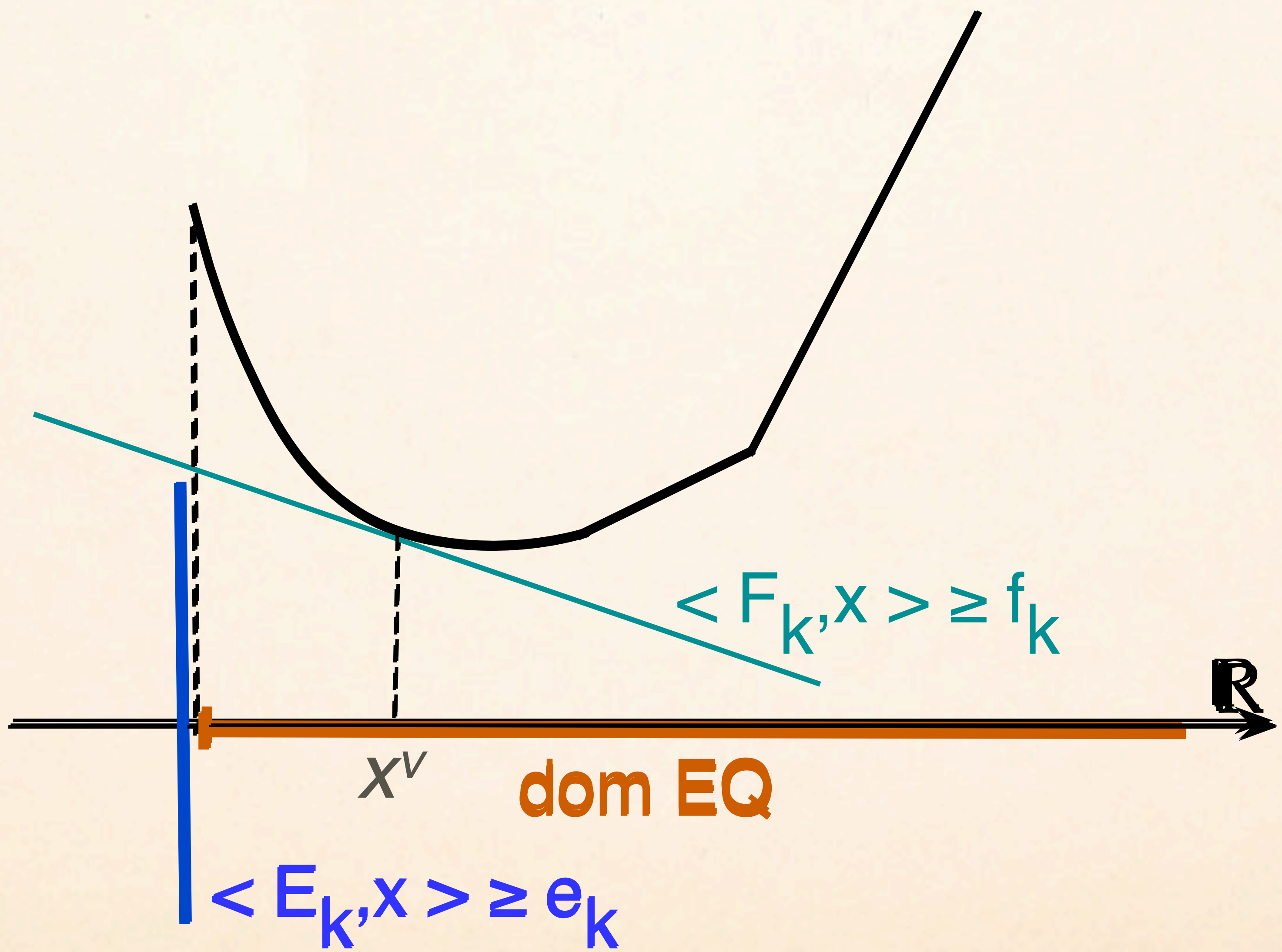
(optimality cuts) $\quad \langle F_k, x \rangle + \theta \geq f_k,\ k = 1 \to s$

2. generate feasibility cuts: check if $x \in \text{dom } EQ_1$.

   No: E_k separates x from dom EQ_1, go to 1. Yes, go to 3.

3. generate optimality cuts: $F_k \in \partial EQ_1(x^k)$, go to 1.

GENERATING CUTTING HYPERPLANES



STOCHASTIC QUASI-GRADIENTS (~ SAA-APPROACH)

$\min Ef(x) = \mathbb{E}\{f(\boldsymbol{\xi},x)\}$ on $X \subset \mathbb{R}^n$,

X convex (compact), $f : \Xi \times \mathbb{R}^n \to \mathbb{R}$

$f(\xi,\cdot)$ convex (gen. semi-smooth)

$x^{\nu+1} = \mathrm{prj}_X(x^\nu - \lambda_\nu d^\nu)$, descent direction, step size

d^ν : stochastic quasi-gradient

$$\mathbb{E}\left\{d^\nu \middle| x^0,\ldots,x^\nu\right\} \in \partial Ef(x^\nu) + \eta_\nu$$

for example: $d^\nu \in \partial f(\xi^\nu, x^\nu)$ sample ξ^ν

or $d^\nu \in \partial\left(\sum_{l=1}^{L} f(\xi^l, x^\nu)\right)$ sample $\xi^1,\ldots,\xi^L$

convergence: $\rho_\nu \geq 0, \sum_{\nu=0}^{\infty} \rho_\nu = \infty, \sum_{\nu=0}^{\infty} \rho_\nu^2 < \infty$

HERE-&-NOW VS. WAIT-&-SEE

- Basic Process: decision --> observation --> decision

  $$x^1 \rightsquigarrow \xi \rightsquigarrow x^2_\xi$$

- Here-&-now problem!

  not all contingencies available at time 0

  x^1 can't depend on ξ!

- Wait-&-see problem

  implicitly all contingencies available at time 0

  choose (x^1_ξ, x^2_ξ) after observing ξ

- incomplete information to anticipative information ?

Fundamental Theorem of Stochastic Optimization

A here-and-now problem can be "reduced" to a wait-and-see problem by introducing the

appropriate 'information' costs
(price of non-anticipativity)

PRICE OF NON-ANTICIPATIVITY

Here-&-now

$$\min \ \mathbb{E}\left\{ f(\xi, x^1, x^2_\xi) \right\}$$

$$x^1 \in C^1 \subset \mathbb{R}^n,$$

$$x^2_\xi \in C^2(\xi, x^1), \forall \xi.$$

Explicit non-anticipativity

$$\min \mathbb{E}\left\{ f(\xi, x^1_\xi, x^2_\xi) \right\}$$

$$x^1_\xi \in C^1 \subset \mathbb{R}^n,$$

$$x^2_\xi \in C^2(\xi, x^1_\xi), \forall \xi.$$

$$x^1_\xi = \mathbb{E}\left\{ x^1_\xi \right\} \quad \forall \xi$$

$$w_\xi \perp \mathrm{c}^{\mathrm{ste}} \text{ functions}$$

$$\Rightarrow \mathbb{E}\left\{ w_\xi \right\} = 0$$

ADJUSTED HERE-&-NOW

min $\mathbb{E}\left\{f(\boldsymbol{\xi},x^1,x^2_\xi)\right\}$ such that $x^1 \in C^1 \subset \mathbb{R}^n,\ x^2_\xi \in C^2(\xi,x^1), \forall\xi$

x^1 must be $\mathcal{G}$-measurable, $\mathcal{G} = \sigma\text{-}\{\emptyset,\Xi\}$

x^2 is $\mathcal{A}$-measurable, $\mathcal{A} \supset \mathcal{G}$,

in general, interchange $\mathbb{E}$ & ∂ is not valid

required: $\forall\ \xi, x^1 \in C^1, C^2(\xi,x^1) \neq \emptyset$ $\mathcal{G}$-measurability of constraints

Now, suppose w_ξ are the (optimal) non-anticipativity multipliers (prices)

min $\mathbb{E}\left\{f(\boldsymbol{\xi},x^1_\xi,x^2_\xi) - \langle w_\xi, x^1_\xi\rangle + \langle w_\xi, \mathbb{E}\{x^1_\xi\}\rangle\right\}$

such that $x^1_\xi \in C^1 \subset \mathbb{R}^n,\ x^2_\xi \in C^2(\xi,x^1_\xi), \forall\xi$

Interchange is now O.K. , $\mathbb{E}\left\{\langle w_\xi, \mathbb{E}\{x^1_\xi\}\rangle\right\} = \langle \mathbb{E}\{w_\xi\}, \mathbb{E}\{x^1_\xi\}\rangle = 0$, yields

$\forall\ \xi$, solve: min $f(\boldsymbol{\xi},x^1,x^2) - \langle w_\xi, x^1\rangle$ s.t. $x^1 \in C^1,\ x^2 \in C^2(\xi,x^1)$

a collection of deterministic optimization problems in $\mathbb{R}^{n_1+n_2}$

FINDING w_ξ

Progressive Hedging Algorithm

0. $w^0(\cdot)$ such that $\mathbb{E}\{w^0(\xi)\} = 0,\ \nu = 0.$ Pick $\rho > 0$

1. for all ξ :

$$(x_\xi^{1,\nu}, x_\xi^{2,\nu}) \in \arg\min f(\xi; x^1, x^2) - \langle w_\xi^\nu, x^1 \rangle$$

$$x^1 \in C^1 \subset \mathbb{R}^{n_1},\ x^2 \in C^2(\xi, x^1) \subset \mathbb{R}^{n_2}$$

2. $\bar{x}^{1,\nu} = \mathbb{E}\{x_\xi^{1,\nu}\}$. Stop if $\left| x_\xi^{1,\nu} - \bar{x}^{1,\nu} \right| = 0$ (approx.)

othersie $w_\xi^{\nu+1} = w_\xi^\nu + \rho\left[x_\xi^{1,\nu} - \bar{x}^{1,\nu} \right]$, return to 1. with $\nu = \nu + 1$

Convergence: add a proximal term

$$f(\xi; x^1, x^2) - \langle w_\xi^\nu, x^1 \rangle - \frac{\rho}{2}\left| x^1 - \bar{x}^{1,\nu} \right|^2$$

linear rate in $(\mathrm{x}^{1,\nu}, w^\nu)$ ... eminently parallelizable

PH: IMPLEMENTATION

implementation: choice of ρ ... scenario ($\times$), decision (+) dependent

(heuristic) extension to problems with integer variables

non-convexities: e.g. ground-water remediation with non-linear PDE recourse

asynchronous

partitioning (= different information feeds)

$$\min \mathbb{E}\{f(\boldsymbol{\xi},x)\} \text{ such that x} \in C = \bigcap_{\xi\in\Xi} C_\xi$$

$S = \{\Xi_1, \Xi_2, \ldots, \Xi_N\}$ a partitioning of Ξ, $p_n = \mu(\Xi_n)$

$\mathbb{E}\{f(\boldsymbol{\xi},x)\} = \sum_n p_n \mathbb{E}\{f(\boldsymbol{\xi},x) \mid \Xi_n\}$ (Bayes)

defining $g(n,x) = \mathbb{E}\{f(\boldsymbol{\xi},x) \mid \Xi_n\}$ if $x \in C_n = \bigcap_{\xi\in\Xi_n} C_\xi$

solve the problem as: $\min \sum_{n=1}^{N} p_n g(n,x)$

MULTISTAGE STOCHASTIC PROGRAMS

$\min_{x \in \mathcal{N}^a} \mathbb{E}\{f(\boldsymbol{\xi}, x(\boldsymbol{\xi}))\}, \quad x(\xi) = \left(x^1(\xi), \ldots, x^T(\xi)\right)$

filtration : $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_T = \mathcal{A}, \;\; \mathcal{A}_0$ trivial

$x \in \mathcal{N}^a$ if x^t $\mathcal{A}_{t-1}$-measurable $\approx \sigma$-field($\overset{\rightarrow \nu-1}{\boldsymbol{\xi}}$)

(here $\boldsymbol{\xi}^0$ deterministic, $x^1(\boldsymbol{\xi}) \equiv x^1$)

under usual $\mathbb{C}.\mathbb{Q}.$ (convex case): $\bar{x} \in \mathcal{X}$ optimal if

$\exists\, \bar{w} \perp \mathcal{N}^a, \bar{w} \in \mathcal{X}^*$ such that $\bar{x} \in \arg\min_{x \in \mathcal{X}} Ef(x) - \mathbb{E}\{\langle \bar{\boldsymbol{w}}, x \rangle\}$

$\bar{w} \perp \mathcal{N}^a \Leftrightarrow \mathbb{E}\left\{\bar{w}(\boldsymbol{\xi}) \,\middle|\, \mathcal{A}_{t-1}\right\} = 0, \forall t = 1, \ldots, T$

$\bar{w}$ non-anticipativity prices

at which to buy the right to adjust decision (after observation)

can be viewed as insurance premiums, ...

PROGRESSIVE HEDGING ALGO.

0. initialize: pick $\mathrm{w}^0(\xi) \in (\mathcal{N}^a)^\perp$, $\hat{x}_0, \rho > 0, \nu = 1$

1. $\forall \xi \in \Xi$, solve (approximately): min $\mathrm{f}^\nu(\xi, x), x \in$ dom $\mathrm{f}(\xi, \cdot)$

$$f^\nu(x,\xi) = f(x,\xi) + \sum_{t=1}^T \left[\langle w_{\nu-1}^t(\xi), x^t \rangle + \frac{\rho}{2} \left| x^t - \hat{x}^{t,\nu-1} \right|^2 \right]$$

minimizer: $x_\nu(\xi) = \left(x_\nu^1(\xi), \ldots, x_\nu^T(\xi) \right),\ \xi \in \Xi$

2. $w_\nu^t(\xi) = w_{\nu-1}^t(\xi) + \rho\left(x_\nu^t(\xi) - \hat{x}_\nu^t(\xi) \right)$ where $\hat{x}_\nu^t(\xi) =$ "averaged" solution

$\hat{x}_\nu^t(\xi) = \mathbb{E}\left\{ \boldsymbol{x}_\nu^t(\cdot) \middle| A \right\}(\xi)$ for each $A \in \mathcal{A}_t$

go to 1. with $\nu = \nu + 1$

convergence: linear in (x, w)

WALRAS EQUILIBRIUM

agent's problem: Agents: a $\in \mathcal{A}$, $|\mathcal{A}|$ finite, possibly "large"

$\bar{x}_a \in \arg\max u_a(x_a)$ so that $\langle p, x_a \rangle \le \langle p, e_a \rangle$, $x_a \in X_a$

e_a: endowment of agent a, $e_a \in \operatorname{int} X_a$

u_a: utility of agent a, concave, usc

$u_a : X_a \to \mathbb{R}$, $X_a \subset \mathbb{R}^n$ (survival set) convex

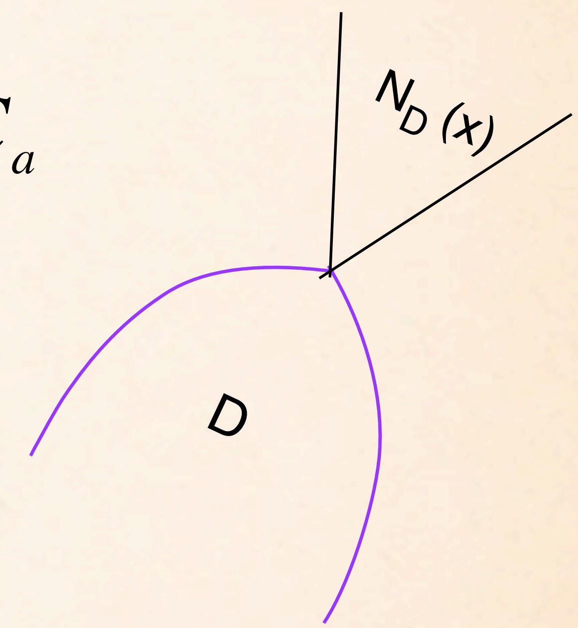
market clearing: $s(p) = \sum_{a \in \mathcal{A}} (e_a - \bar{x}_a)$ excess supply

equilibirum price: $\bar{p} \in \Delta$ such that $s(\bar{p}) \ge 0$, Δ unit simplex

VARIATIONAL INEQUALITY

$$c_a = \arg\max_x u_a(x) \text{ so that } \langle p,x\rangle \le \langle p,e\rangle,\ x \in C_a$$

$$\sum_a (e_a - c_a) = s(p) \ge 0.$$



$$N_D(\bar{z}) = \{v \,|\, \langle v, z-\bar{z}\rangle \le 0, \forall z \in D\}$$

$$G(p,(x_a),(\lambda_a)) = \left[\sum_a (e_a - x_a); (\lambda_a p - \nabla u_a(x_a)); \langle p, e_a - x_a\rangle\right]$$

$$D = \Delta \times \left(\prod_a C_a\right) \times \left(\prod_a \mathbb{R}_+\right)$$

$$-G(\bar{p},(\bar{x}_a),(\bar{\lambda}_a)) \in N_D(\bar{p},(\bar{x}_a),(\bar{\lambda}_a))$$

D unbounded $\to \hat{D}$ bounded

EQUILIBRIUM: STOCHASTIC ENVIRONMENT

$$(c_a^1, y_a, c_{a,\xi}^2) = \arg\max_{x^1, y \in \mathbb{R}^L, x_\bullet^2 \in \mathcal{M}} u_a^1(x^1) + \mathbb{E}^a\left\{u_a^2(\boldsymbol{\xi}, x^2(\boldsymbol{\xi}))\right\}$$

$$\text{such that } \left\langle p^1, x_a^1 + T_a^1 y\right\rangle \leq \left\langle p^1, e_a^1\right\rangle$$

$$\left\langle p_\xi^2, x_{a,\xi}^2\right\rangle \leq \left\langle p_\xi^2, e_{a,\xi}^2 + T_{a,\xi}^2 y\right\rangle, \ \forall \xi \in \Xi$$

$$x_a^1 \in X_a^1, \ x_{a,\xi}^2 \in X_{a,\xi}^2, \quad \forall \xi \in \Xi$$

$\mathbb{E}^a\{\bullet\}$ rational expecttion with respect to a-beliefs, Ξ finite support

2-stage stochastic programs with recourse

solution procedures & approximation theory "well-estblished"

$T_a^1, T_{a,\xi}^2$: input-output matrices (production, investments)

$e_a^1 \in \text{int}\, X_a^1, \ e_{a,\xi}^2 \in \text{int}\, X_{a,\xi}^2$ for all ξ

MARKET CLEARING

excess supply:

agent-a: $\left(c_a^1, y_a^1, \left\{ c_{a,\xi}^2 \right\}_{\xi\in\Xi} \right)$

$$\sum_{a\in\mathcal{A}} \left(e_a^1 - (c_a^1 + T_a^1 y_a) \right) = s^1\left(p^1, \{p_\xi^2\}_{\xi\in\Xi} \right) \geq 0$$

and for all $\xi \in \Xi$:

$$\sum_{a\in\mathcal{A}} \left((e_{a,\xi}^2 + T_{a,\xi}^2) - c_{a,\xi}^2 \right) = s_\xi^2\left(p^1, \{p_\xi^2\}_{\xi\in\Xi} \right) \geq 0$$

Walras' auctioneer:

$$\max_{p=\left(p^1, \{p_\xi^2\}_{\xi\in\Xi}\right)} \inf_{q=\left(q^1, \{q_\xi^2\}_{\xi\in\Xi}\right)} \mathbb{E}\left\{ \langle q, s \rangle \right\}, \quad s = \left(s^1, \{s_\xi^2\}_{\xi\in\Xi} \right)$$

AGENT'S PROBLEM: DISAGGREGATION

with $p_{\blacklozenge} = \left(p^0, \left\{p_\xi^1\right\}_{\xi\in\Xi}\right)$

$(c_{a,\xi}^1, y_a, c_{a,\xi}^2) \in$

$$\arg\max_{x^1\in\mathbb{R}^l, y\in\mathbb{R}^L, x^2\in\mathbb{R}^L} \left\{u_a^1(x^1) - \left\langle \bar{w}_{a,\xi}, (x^1, y)\right\rangle + u_a^2(\xi, x^2)\right\}$$

$$\left\langle p^1, x^1\right\rangle \le \left\langle p^1, e_a^1 - T_a^1 y\right\rangle$$

$$\left\langle p_\xi^2, x^2\right\rangle \le \left\langle p_\xi^2, e_{a,\xi}^2 + T_{a,\xi}^2 y\right\rangle,$$

$$x^1 \in X_a^1,\ x^2 \in X_{a,\xi}^2.$$

solved for each ξ **separately**

INDIVIDUALLY COMPLETED MARKET

$\forall \xi \in \Xi$ (separately),

agent's problem (individually completed market):

$$\left(c_a^1, y_a, c_{a,\xi}^2\right) \in \arg\max\left\{u_a^{w_{a,\xi}}\left(x^1, y, x^2\right) \text{ on } \widehat{X}_{a,\xi}(p^1, p_\xi^2)\right\}$$

for $\{w_{a,\xi}\}_{\xi\in\Xi}$ associated with (p^0, p_ξ^1)

clear market:

$$s^1(p^1, p_\xi^2) \geq 0, \quad s_\xi^2(p^1, p_\xi^2) \geq 0$$

- Arrow-Debreu 'un-stochastic' equilibrium problem

EXAMPLE USING PATH-SOLVER

- Economy: (5 goods)
  - Skilled & unskilled workers
  - Businesses: Basic goods & leisure
  - Banker: bonds (riskless), 2 stocks
- 2-stages,   280 scenarios, 2776 scenarios
- utilities: CES-functions (gen. Cobb-Douglas)
  - Utility in stage 2 assigned to financial instruments
  - only used for transfer in stage 1
- on laptop: ~4 min, ~14 min, but
  extremely parallelizable algorithm

PATH-SOLVER: CONVERGENCE

objective: $u_a^1(x^1) + u_a^2(x^2) \quad \Rightarrow$

$$u_a^1(x^1) - \langle w_{a,\xi}^v, (x^1, y) \rangle - \frac{\rho}{2} \left| (x^1, y) - (\hat{x}_a^{1,v}, \hat{y}_a^v) \right|^2 + u_a^2(x^2)$$

updating:

$(\hat{x}_a^{1,v}, \hat{y}_a^v) = \mathbb{E}^a \left\{ (c_{a,\xi}^{0,v}, y_{a,\xi}^v) \right\}$ projection on non-anticipative subspace

$$w_{a,\xi}^{v+1} = w_{a,\xi}^v + \rho_a \left((c_{a,\xi}^{0,v}, y_{a,\xi}^v) - (\hat{x}_a^{1,v}, \hat{y}_a^v) \right)$$

convergence: $\rho_a > 0$

also requires a proximal term to support

the convergence of the equilibrium prices $p_{\blacklozenge} = \left(p^0, \{ p_\xi^1 \}_{\xi \in \Xi} \right)$