Stochastic Variational Problems

Roger J-B Wets
University of California, Davis

Abgekürzte Vorlesung
@ Universität Wien, Juni 2013

“Optimization & Equilibrium”

“Nothing at all takes place in the universe in which some rule of maximum or minimum does not appear” L. Euler 1744

“Optimization & Equilibrium”

“Nothing at all takes place in the universe in which some rule of maximum or minimum does not appear” L. Euler 1744

“Il n’est pas certain que tout soit incertain” B. Pascal   ±1645

Prelude: 14th-Century ~ 1950

~ 1350 Oresme's rule: $f:\mathbb{R}\to\mathbb{R}$

$$x^* \in \arg\min f \Rightarrow df(x^*;w) = \lim_{\tau\to 0}\frac{f(x^*+\tau w)-f(x^*)}{\tau} = 0,\ \forall w\in\mathbb{R}$$

~ 1650 Fermat's rule: $f:\mathbb{R}\to\mathbb{R},\ f$ smooth

$$x^* \in \arg\min f \Rightarrow f'(x^*) = \nabla f(x^*) = 0$$

$$df(x;w) = \langle \nabla f(x0,w\rangle \Rightarrow \text{(Oresme rule} \Leftrightarrow \text{Fermat rule)}$$

(~ 1950 Dantzig simplex method for linear programming)

Oresme's rule: $f:\mathbb{R}^n\to\mathbb{R}$

$$x^* \in \arg\min f \Rightarrow df(x^*;w) = 0,\ \forall w\in\mathbb{R}^n$$

Fermat's rule: $f:\mathbb{R}^n\to\mathbb{R},\ f$ smooth

$$x^* \in \arg\min f \Rightarrow \nabla f(x^*) = 0, \qquad df(x;w) = \langle \nabla f(x), w\rangle$$

Curve Fitting

find $h:[0,1] \to \mathbb{R}$, given $h(z_1),\ldots,h(z_L)$

approximate by $p(x) = a_n x^n + \cdots + a_1 x + a_0$

$$\min_{a\in\mathbb{R}^{n+1}} \sum_{l=1}^{L} \Big(\sum_{j=0}^{n} a_j z_l^j - h(z_l)\Big)^2 = \min_{a\in\mathbb{R}^{n+1}} \langle Za - y, Za - y\rangle$$

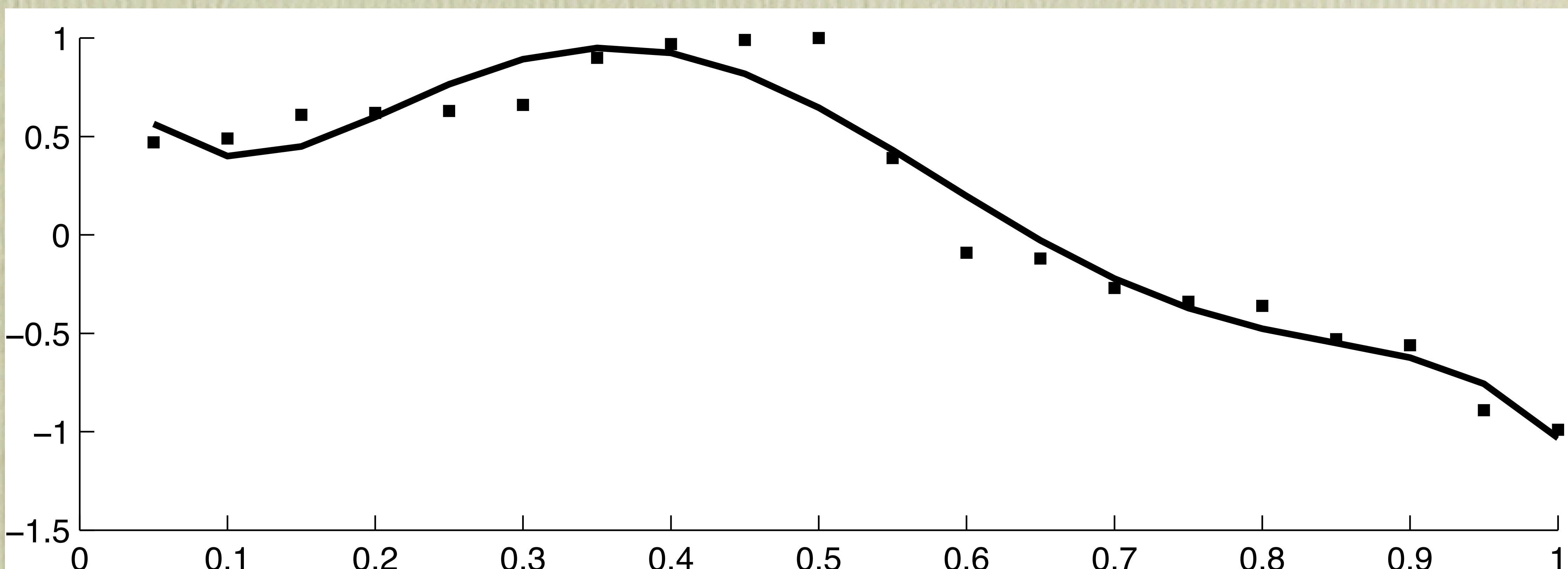
$$Z = \begin{bmatrix} z_1^n & \ldots & z_1 & 1 \\ z_2^n & \ldots & z_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ z_L^n & \ldots & z_L & 1 \end{bmatrix}, \quad (y_1,\ldots,y_L) = \big(h(z_1),\ldots,h(z_l)\big)$$

Solution: applying Fermat's rule,

$a^* = \big(Z^\top Z)^{-1} Z^\top y$; assuming col. Z linearly independent

Curve fitting: polynomial-approx.

with best degree $n = 5$



Steepest descent - Newton Methods

direction of descent $\exists \tau_\star > 0, \forall\, \tau \in (0, \tau_\star): \quad f(\bar{x} + \tau d) < f(\bar{x})$

$df(\bar{x}; d) < 0$ smooth: $\langle \nabla f(\bar{x}), d \rangle \le 0 \Rightarrow d$ direction of descent

Steepest descent: x^0; $\nabla f(x^\nu) = 0$ stop;

$$d^\nu = -\nabla f(x^\nu), x^{\nu+1} = \arg\min \left\{ f(x) \middle| x \in [x^\nu, x^\nu + \lambda d^\nu), \lambda \ge 0 \right\}$$

Newton direction: $d^\nu = -\left(\nabla^2 f(x^\nu)\right)^{-1} \nabla f(x^\nu) \Rightarrow$ Newton Method

quadratic convergence locally: (f $C^2 \ldots$), with local sol'n @ $\bar{x}$

$\exists \rho > 0, \kappa \ge 0: \quad \| \nabla^2 f(x) - \nabla^2 f(x') \| \le \kappa \,|\, x - x' \,|, \ \forall x, x' \in \mathbb{B}(\bar{x}, \rho)$

Quasi-Newton Method(s)

0. $x^{\nu}, B^{0}(=I), \nu = 0$

1. $\nabla f(x^{\nu}) = 0$ stop or $B^{\nu} d^{\nu} = -\nabla f(x^{\nu})$

2. $x^{\nu+1} = \arg\min\left\{ f(x) \middle| x \in [x^{\nu}, x^{\nu} + \lambda d^{\nu}), \lambda \geq 0 \right\}$

$B^{\nu} \approx$ Hessian $\nabla^2 f(x^{\nu})$ in Newton (curvature)

$B^{\nu+1}(x^{\nu+1} - x^{\nu}) = (B^{\nu} + U)(x^{\nu+1} - x^{\nu}) = \nabla f(x^{\nu+1}) - \nabla f(x^{\nu})$, i.e.,

$U^{\nu} s^{\nu} = c^{\nu} - B^{\nu} s^{\nu}, s^{\nu} = x^{\nu+1} - x^{\nu}, c^{\nu} = \nabla f(x^{\nu+1}) - \nabla f(x^{\nu})$

Quasi-Newton condition

Integral Functionals (Calculus of Variations)

$$\min\left\{f(x)=\int_0^1 L(t,x(t),\dot{x}(t))\,dt \,\middle|\, x\in \text{ fcns}\big([0,1],\mathbb{R}\big),\; x(0)=\alpha, x(1)=\beta\right\}$$

Oresme rule : $\quad x^* \in \arg\min f \Rightarrow df(x^*;w)=0, \forall w\in W \subset X,$

W admissible variations: $=\left\{w\in \text{ fcns}\big([0,1],\mathbb{R}\big)\,\middle|\, w(0)=w(1)=0\right\}$

Bernoulli (Jacob, Johan), Newton (~ 1700) $\Rightarrow$ Euler equation

$$\boxed{L_x\big(t,x^*(t),\dot{x}^*(t)\big)=\frac{d}{dt}L_{\dot{x}}\big(t,x^*(t),\dot{x}^*(t)\big) \quad \text{for } t\in[0,1]}$$

Mathematical Shift (a paradigm change)

differentiability $\rightarrow$ non-smooth

typical example: $f(x) = \min\left\{ g(x,y) \middle| y \in S(x) \right\}$

dom f $\rightarrow$ open to closed

INEQUALITIES!

A formulation -- Product mix problem

A furniture manufacturer must choose $x_j \geq 0$, how many dressers of type $j = 1, \ldots, 4$ to manufacture so as to maximize profit

$$\sum_{j=1}^{4} c_j x_j = 12x_1 + 25x_2 + 21x_3 + 40x_4$$

The constraints:

$$t_{c1}x_1 + t_{c2}x_2 + t_{c3}x_3 + t_{c4}x_4 \leq d_c$$
$$t_{f1}x_1 + t_{f2}x_1 + t_{f3}x_1 + t_{f4}x_1 \leq d_f$$

t_{cj} (t_{fj}) carpentry (finishing) man-hours: dresser type j

d_c (d_f) = total time available for carpentry (finishing)

Product mix problem (2)

Solution via linear programming:

$$\max\langle c, x\rangle \text{ so that } Tx \le d,\ x \in I\!R^n_+.$$

With

$$T = \begin{bmatrix} t_{c1} & t_{c2} & t_{c3} & t_{c4} \\ t_{f1} & t_{f2} & t_{f3} & t_{f4} \end{bmatrix} = \begin{bmatrix} 4 & 9 & 7 & 10 \\ 1 & 1 & 3 & 40 \end{bmatrix}, \quad \begin{bmatrix} d_c \\ d_f \end{bmatrix} = \begin{bmatrix} 6000 \\ 4000 \end{bmatrix}$$

Optimal: $\boxed{x^d = (4000/3,\ 0,\ 0,\ 200/3)}$

Value: \$ 18,667.

Product mix problem (3)

But . . . "reality" can't be ignored!

$$\boldsymbol{t}_{cj} = t_{cj} + \boldsymbol{\eta}_{cj}, \quad \boldsymbol{t}_{fj} = t_{fj} + \boldsymbol{\eta}_{fj}$$

entry	possible values			
$d_c + \zeta_c$:	5,873	5,967	6,033	6,127
$d_f + \zeta_f$:	3,936	3,984	4,016	4,064

10 random variables, say, 4 possible values each

$$\boxed{L = 1,048,576 \text{ possible pairs } (T^l, d^l)}$$

Product mix problem (4)

What if $\sum_{j=1}^{4}(t_{cj}+\eta_{cj})x_j > d_c + \zeta_c$? $\implies$ overtime

With $\xi = (\eta_{\{\cdot,\cdot\}}, \zeta_{\{\cdot\}})$, **recourse**: $(y_c(\xi), y_f(\xi))$ @ cost (q_c, q_f).

$$
\begin{array}{lllllll}
\max & \langle c, x\rangle & -p_1\langle q, y^1\rangle & -p_2\langle q, y^2\rangle & \cdots - & p_L\langle q, y^L\rangle & & \\
\text{s.t.} & T^1 x & -y^1 & & & & \leq & d^1 \\
 & T^2 x & & -y^2 & & & \leq & d^2 \\
 & \vdots & & & \ddots & & & \vdots \\
 & T^L x & & & & -y^L & \leq & d^L \\
 & x \geq 0, & y^1 \geq 0, & y^2 \geq 0, & \cdots & y^L \geq 0. & &
\end{array}
$$

Structured large scale l.p. ($L \approx 10^6$)

Product mix problem (5)

Define $\Xi = \{\xi = (\eta, \zeta)\}$, $p_\xi =$ prob $[\boldsymbol{\xi} = \xi]$

$$Q(\xi, x) = \max\left\{\langle -q, y\rangle \,\middle|\, T_\xi x - y \geq d_\xi,\ y \geq 0\right\}$$

$$EQ(x) = E\{Q(\boldsymbol{\xi}, x)\} = \sum_{\xi \in \Xi} p_\xi Q(\xi, x)$$

the equivalent deterministic program (DEP):

$$\max\langle c, x\rangle + EQ(x) \ \text{ so that } \ x \in I\!R^n_+.$$

a *non-smooth* convex optimization problem: EQ concave.

Product mix problem (6)

Solution of DEP, or large scale l.p.,:

Optimal: $\boxed{x^* = (257,\ 0,\ 665.2,\ 33.8)}$

expected Profit: \$ 18,051

The solution x^* is *robust* : it considered all $\approx 10^6$ possibilities.

$$\text{Recall: } x^d = (1{,}333.33,\ 0,\ 0,\ 66.67)$$

$$\text{expected "profit" relying on } x^d = \$\,16{,}942$$

- x^d is not close to optimal
- x^d isn't pointing in the right direction

Curve Fitting (2)

find $h:[0,1] \to \mathbb{R}$, given $h(t_1),\ldots,h(t_L)$ & h smooth

approximate by z, C^{1+}-curve

mesh $=\{0,\delta,2\delta,\ldots,N\delta=1\}$, on $((k-1)\delta,k\delta]$ $\quad z''(t)=a_k$

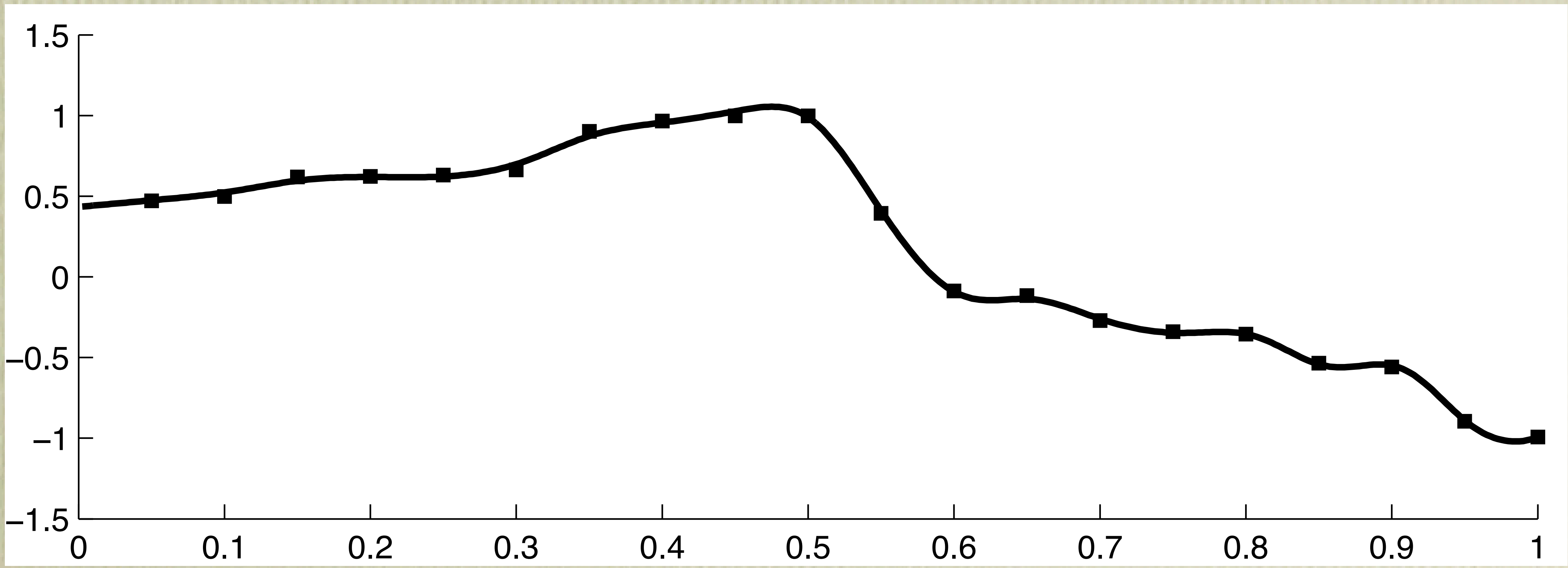
$$z(t)=z_0+v_0t+\delta\sum_{j=1}^{k-1}(t-t_j+\delta/2)a_j+\tfrac{1}{2}(t-t_{k-1})^2a_k,\quad t\in((k-1)\delta,k\delta]$$

$$\boxed{\begin{array}{l}\text{find } z_0,v_0,a_1,\ldots,a_N \quad \text{all } \in\mathbb{R}\\ \quad \min\left\|\big(z(t_l)-h(t_l)\big),l=1,\ldots,L\right\|_{\square} \quad -\kappa\le a_k\le\kappa,\ k=1,\ldots,N\end{array}}$$

Epi-spline fit

z is an epi-spline of order 2,

on each (open) sub-interval a polynomial of order 2.



Preliminaries -- Convexity
"minimization framework"

$$f : \mathbb{R}^n \to \bar{\mathbb{R}} = [-\infty, \infty], \ \ \text{dom} f = \{x \mid f(x) < \infty\}$$

$$\text{proper: dom} f \neq \emptyset, f > -\infty$$

$$f(x) = f_0(x) \text{ when } f_i(x) \leq 0, \ f_i(x) = 0, \ x \in X \subset \mathbb{R}^n$$

$$i = 1,..,s \qquad i = s+1,..,m$$

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

$$\text{epi} f = \{(x,\alpha) \mid f(x) \geq \alpha\} \subset \mathbb{R}^{n+1}$$

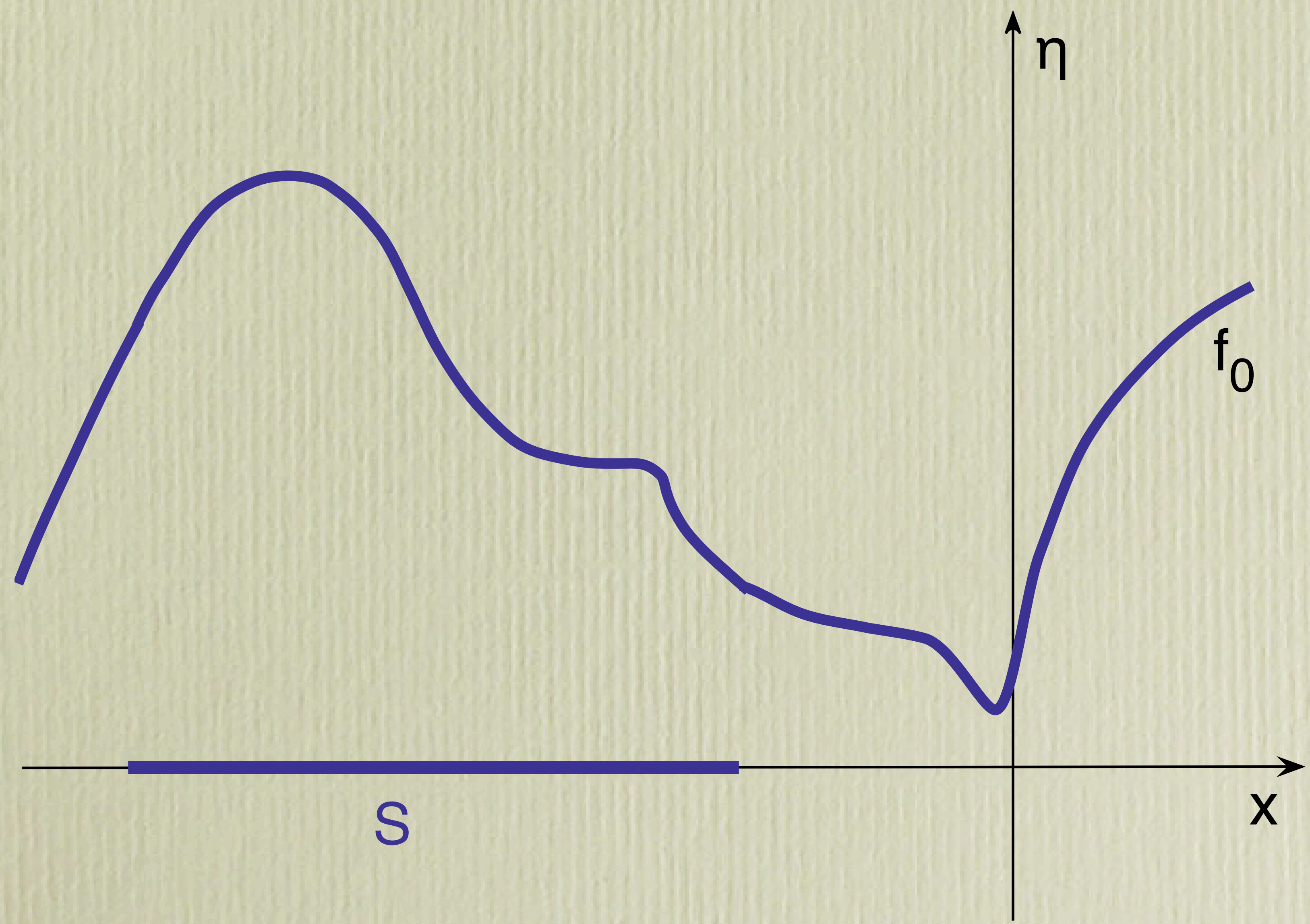
$$f \text{ lsc } \Leftrightarrow \text{epi} f \text{ closed (lower semicontinuous)}$$

$$f \text{ convex } \Leftrightarrow \text{epi} f \text{ convex}$$

Optimization problem

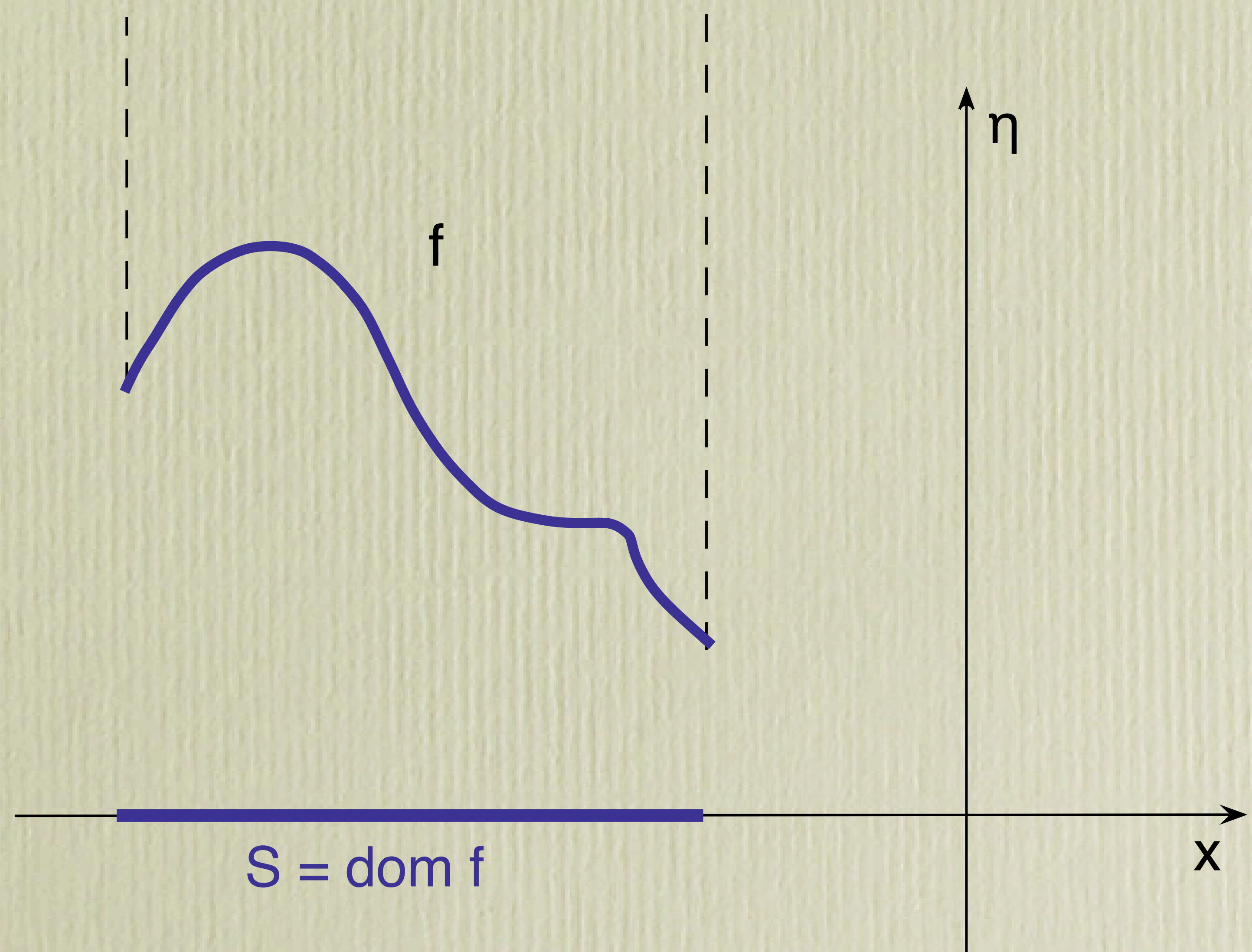
$$\min f(x),\ x \in S,$$

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



Extended-real valued fcn

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

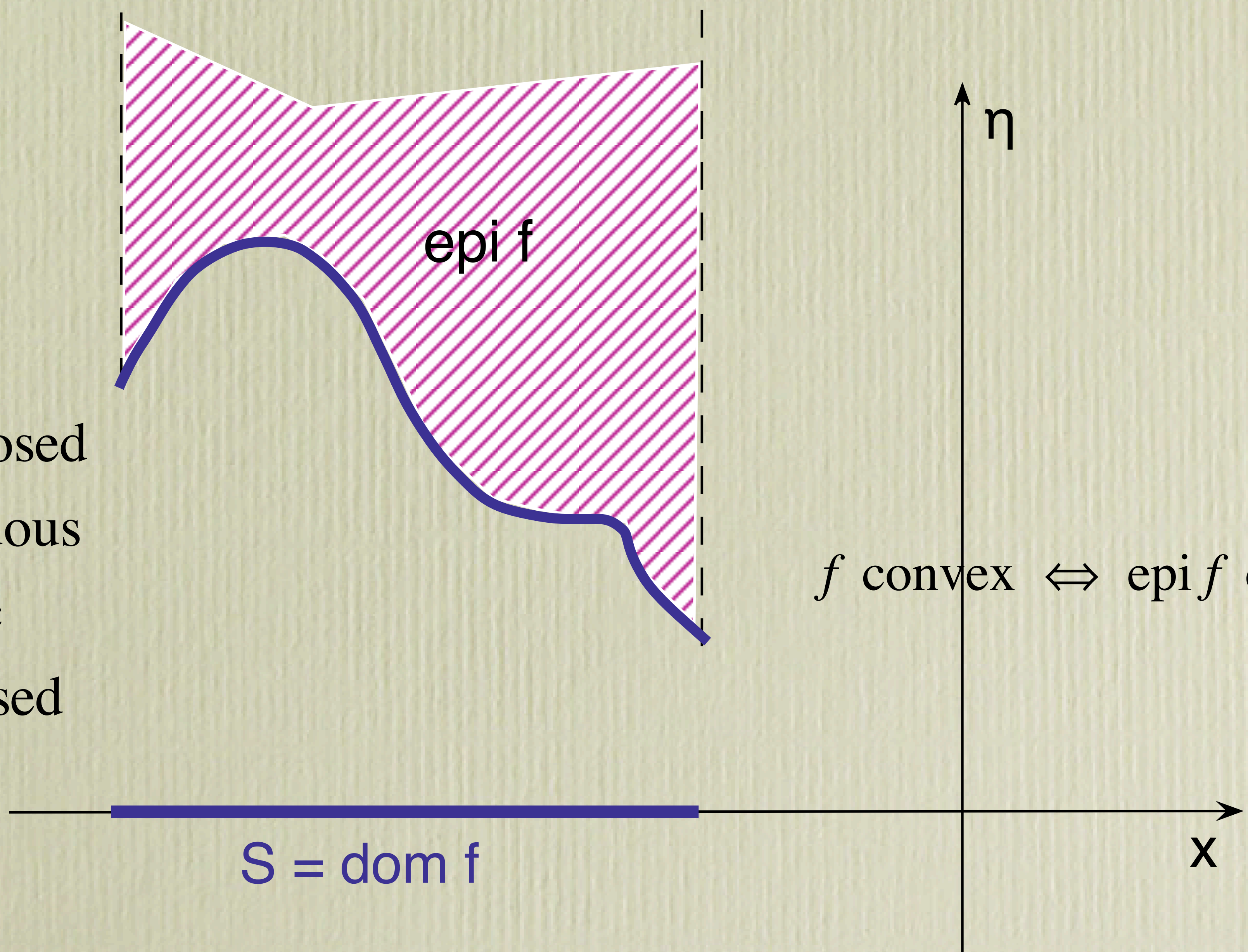


Functions & Epigraphs

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

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

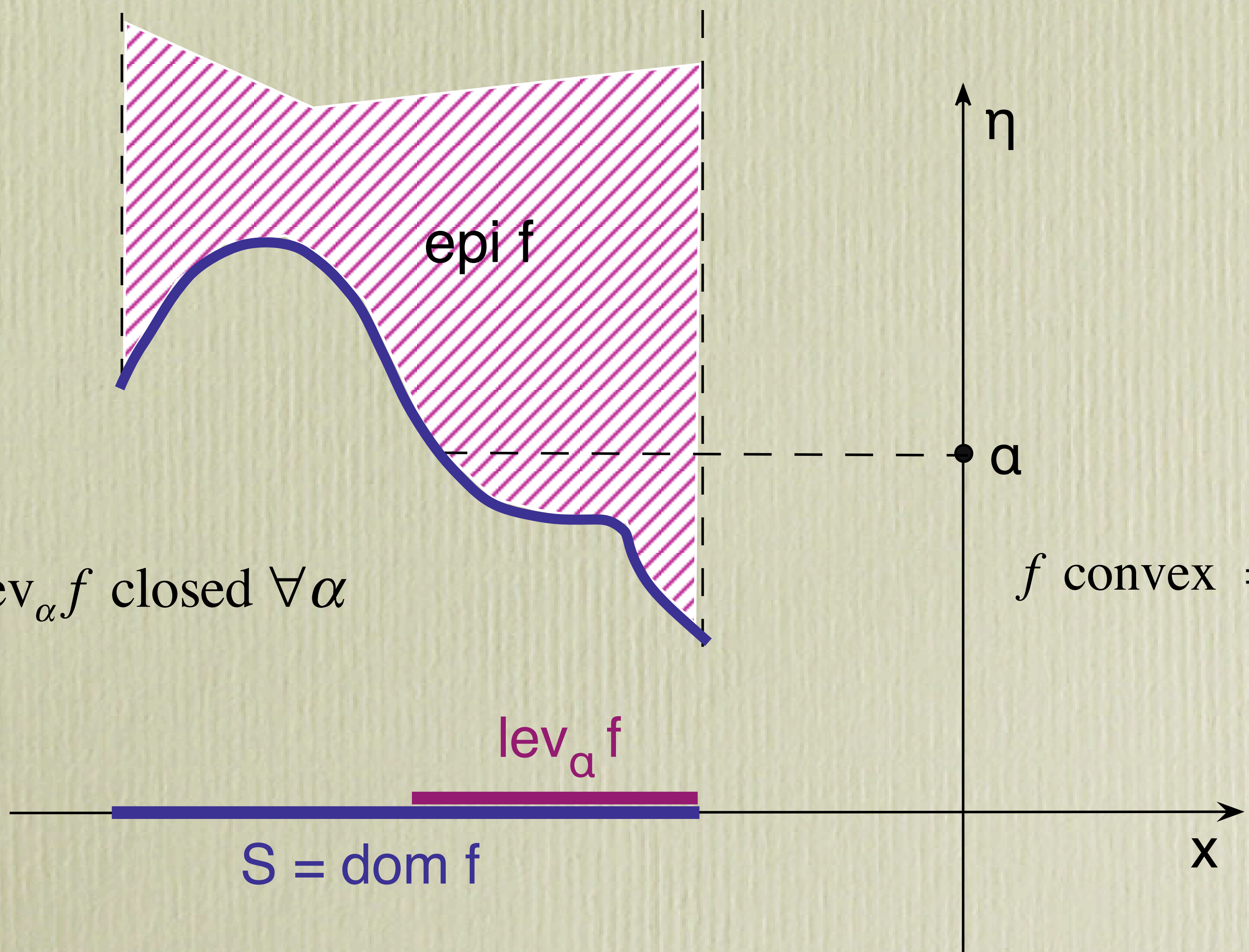
f usc $\Leftrightarrow$ $-f$ lsc

$\Leftrightarrow$ hypo f closed

f convex $\Leftrightarrow$ epi f convex

Level sets & Constraints

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

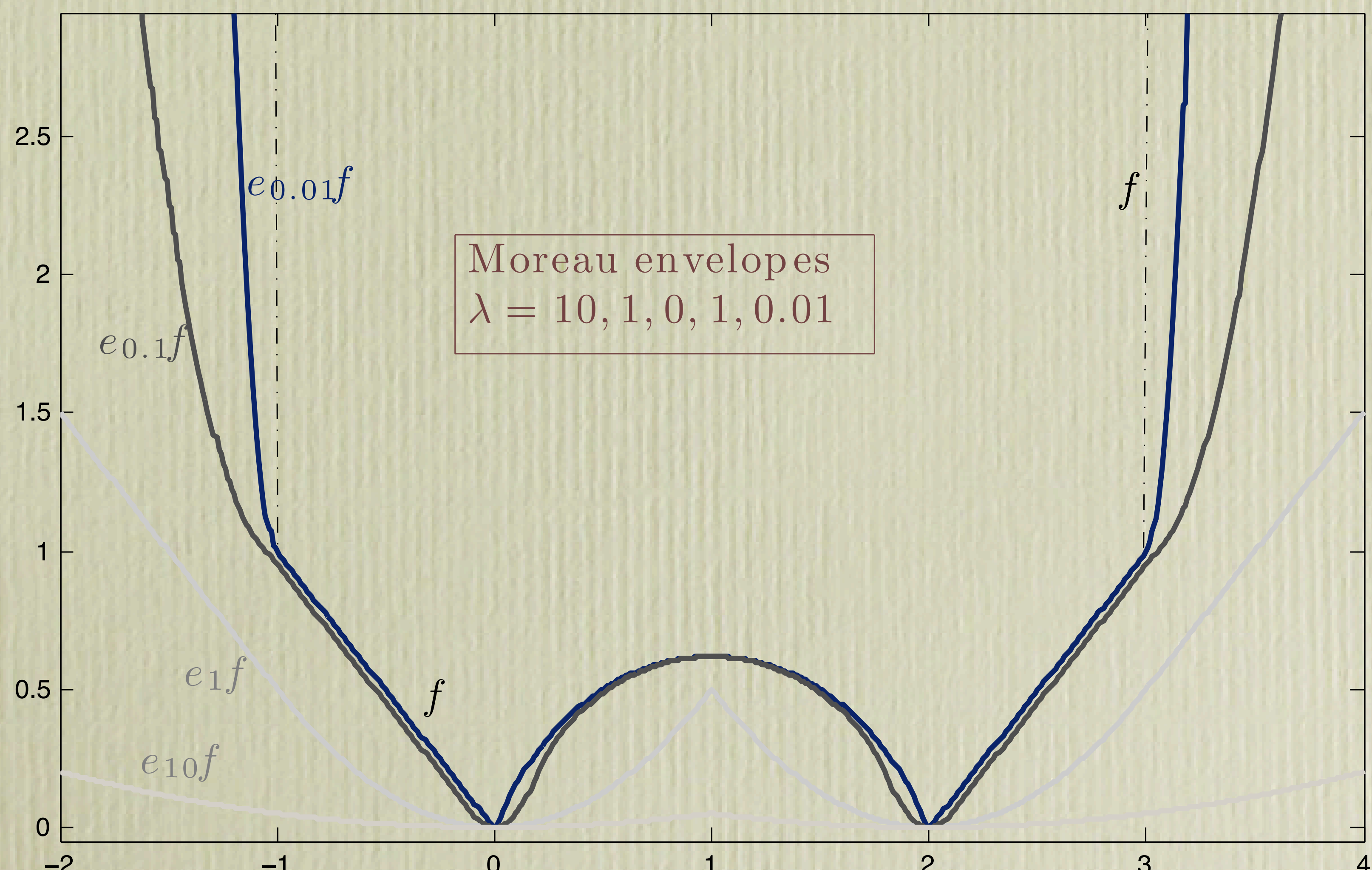


f lsc $\Leftrightarrow$ $\text{lev}_\alpha f$ closed $\forall \alpha$

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

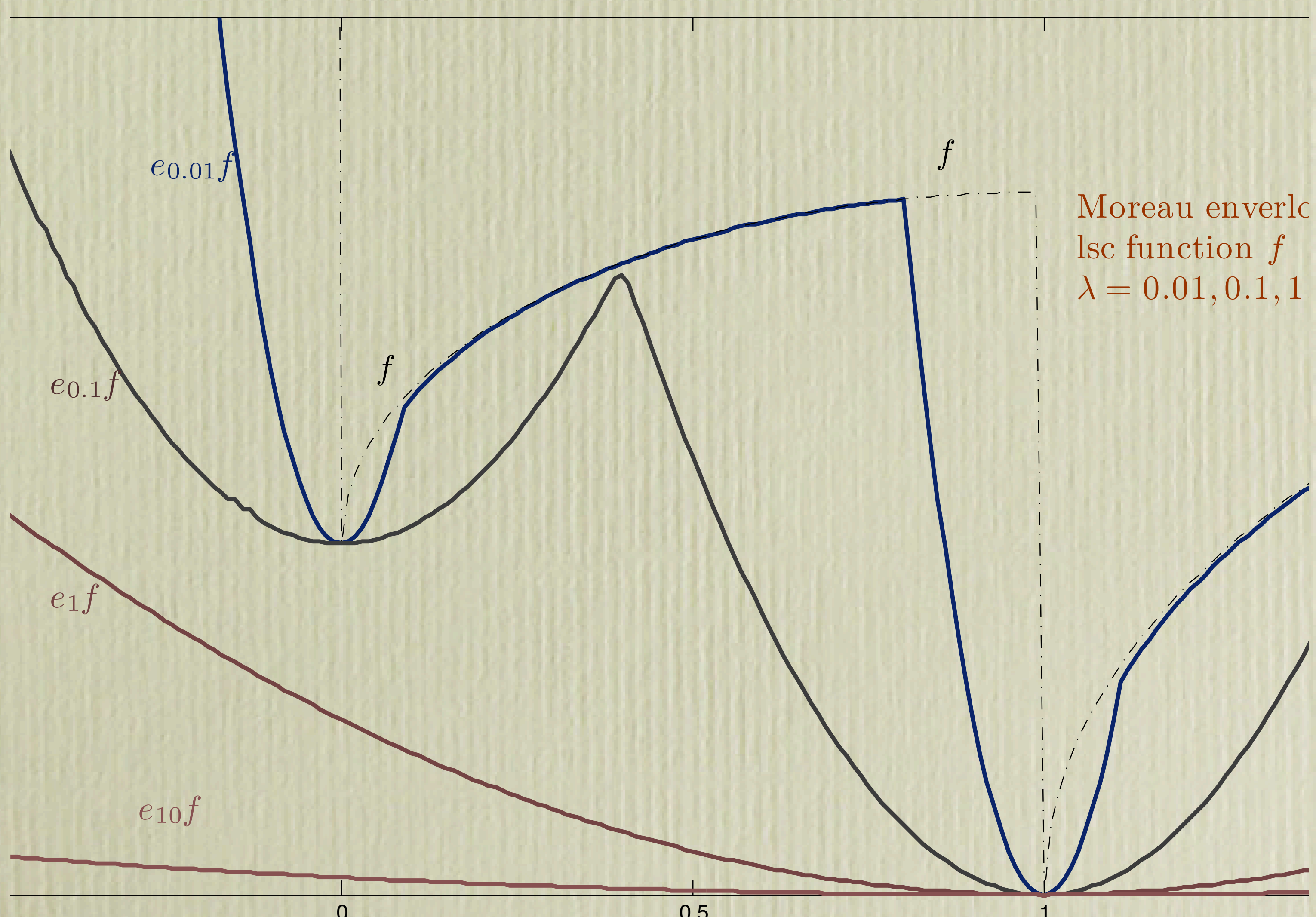
MOREAU ENVELOPES

$$e_\lambda f = \inf_w \left\{ f(w) + \frac{1}{2\lambda} | w - x |^2 \right\} \qquad \text{epi } e_\lambda f \approx \text{epi } f + \text{epi} \frac{1}{2\lambda} \left| \bullet \right|^2$$



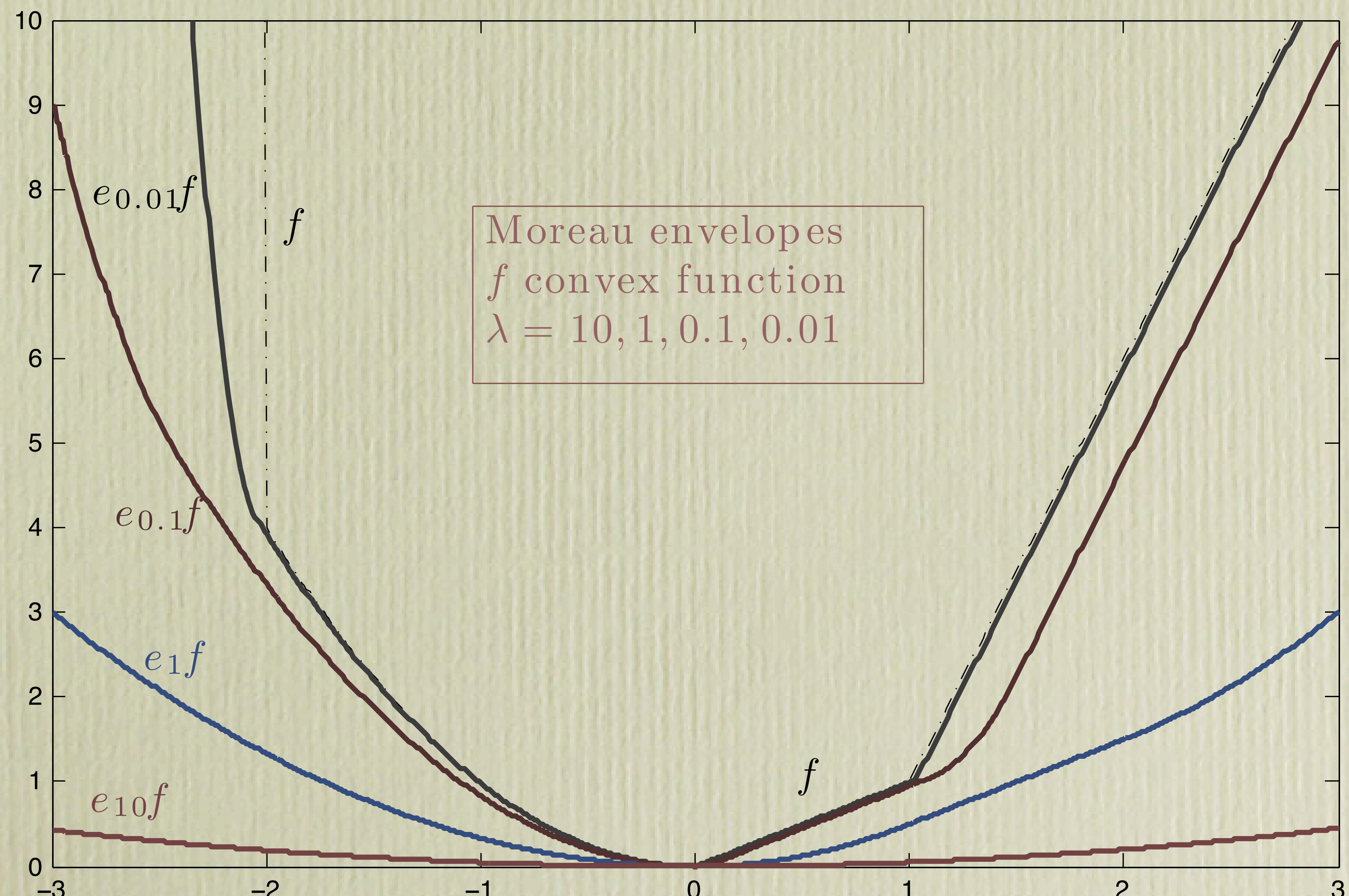
MOREAU ENVELOPE: APPROXIMATIONS

$$e_\lambda f = \inf_w \left\{ f(w) + \frac{1}{2\lambda} |w - x|^2 \right\} \qquad \text{epi } e_\lambda f \approx \text{epi } f + \text{epi} \frac{1}{2\lambda} |\bullet|^2$$



EPI-SUMS (inf-convolution)

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



Convex functions: Properties

- $f_1 + f_2,\ \mathbb{E}\{f(\xi,\cdot)\} = \int_\Xi f(\xi,\bullet)P(d\xi) = Ef$ convex $\quad (\Xi,\mathcal{A},P)$

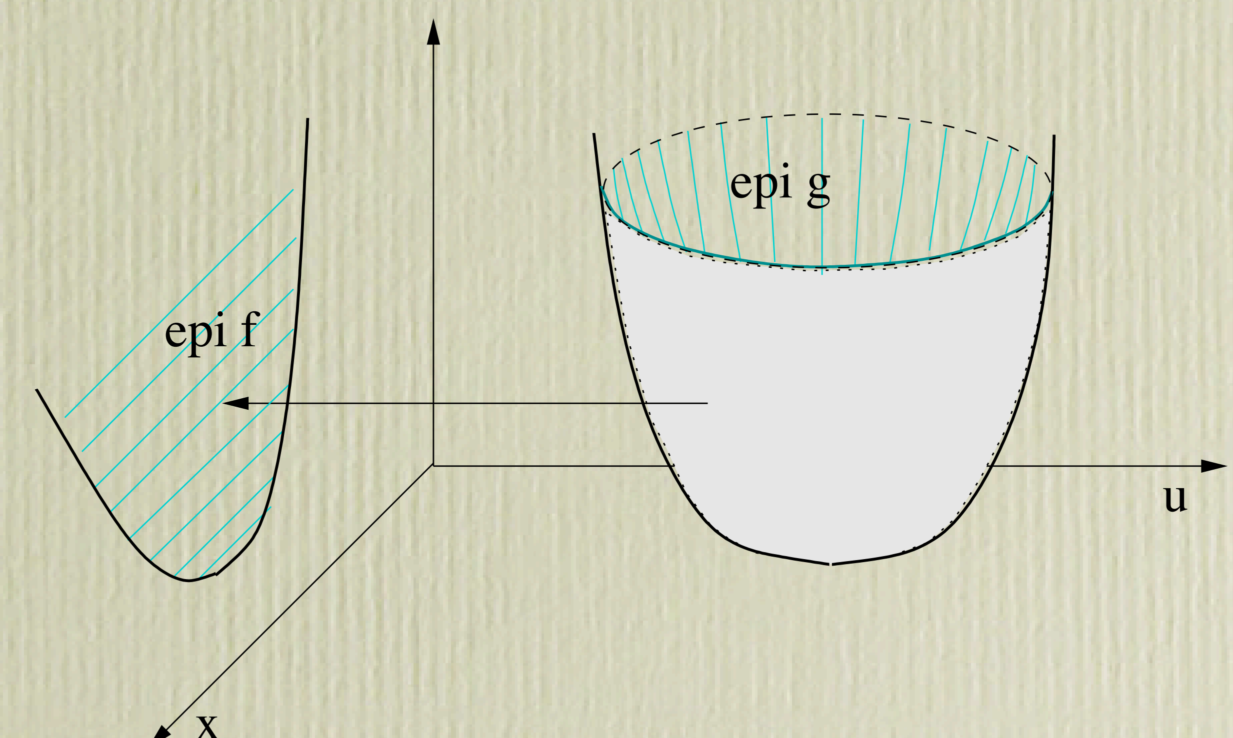
  $f : \Xi \times \mathbb{R}^n \to \overline{\mathbb{R}} = [-\infty,\infty]$ random lsc function (normal integrand)

  $x \mapsto f(\xi,x)$ lsc $\forall \xi \in \Xi$

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

- inf-projection: $f(x) = \inf_{u\in\mathbb{R}^m} g(u,x)$ convex when g is convex

Proof:



- f convex, $\boxed{\text{local minimum} \iff \text{global minimum}}$

Convex fcns: differentiability

- $f : O_{\text{open}} \subset \mathbb{R}^n$ differentiable, convex $\Leftrightarrow$
  - (a) $\langle x^1 - x^0, \nabla f(x^1) - \nabla f(x^0)\rangle \geq 0 \;\; \forall x^0, x^1$; montonicity
  - (b) $f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle, \forall\, x, y \in \mathbb{R}^n$; affine support
  - (c) $\nabla^2 f(x)$ positive semi-definite $\forall x$; f twice differentiable
- $f(x) = \frac{1}{2}\langle x, Qx\rangle + \langle c, x\rangle$ quadratic form, convex $\Leftrightarrow$ Q psd

Convex fcns: subdifferentiability

- $df(\bar{x};w) = \lim_{\tau \searrow 0} \dfrac{f(\bar{x}+\tau w) - f(\bar{x})}{\tau}$ subderivative (fcn) direction w
- $\partial f(\bar{x}) = \left\{ v \in \mathbb{R}^n \,\middle|\, f(x) \geq f(\bar{x}) + \langle v, x-\bar{x}\rangle,\ \forall\, x \in \mathbb{R}^n \right\}$ subgradients

  $= \left\{ v \in \mathbb{R}^n \,\middle|\, \langle v, w\rangle \leq df(\bar{x};w), \forall\, w \in \mathbb{R}^n \right\}$

  $= \nabla f(\bar{x}),\ f$ diferentiable
- $\forall\, w : df(\bar{x};w) = \max_v \left\{ \langle v,w\rangle \,\middle|\, v \in \partial f(\bar{x}) \right\}$

- $f(x) = \max_{i \in I} f_i(x)$ of convex fcns is convex, ∂f(x) = ??
- $f =$ inf-projection of $g(u,\bullet)$, ∂f(x) = ??

Generalized Oresme, Fermat Rules

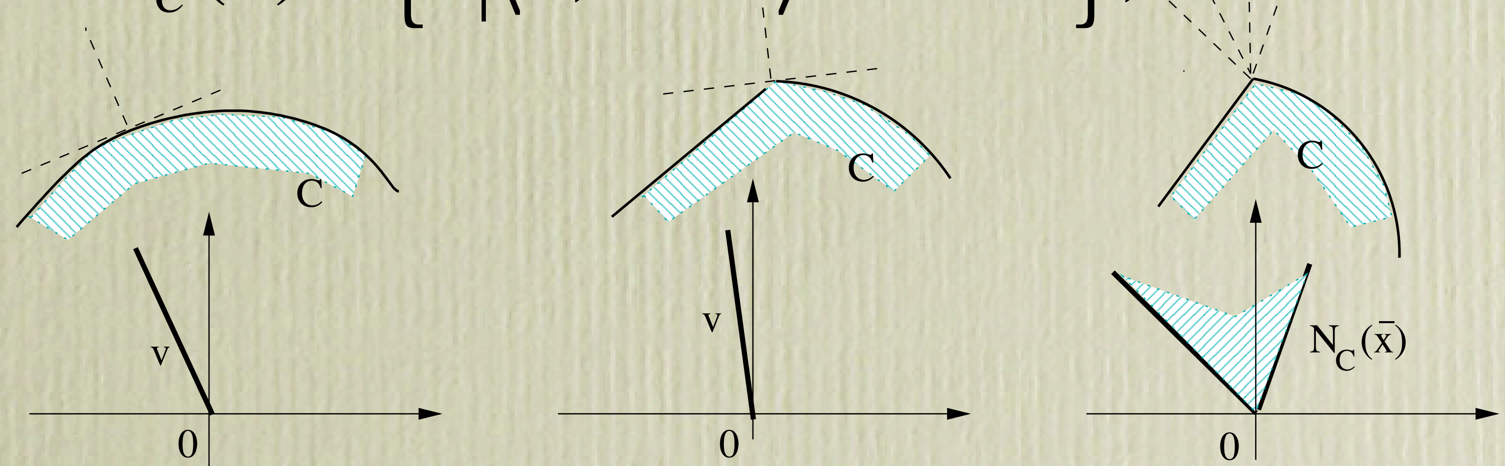
$f:\mathbb{R}^n \to \bar{\mathbb{R}}$, convex

- Oresme rule: $x^* \in \arg\min f \iff df(x^*;\cdot) \geq 0$
- Fermat rule: $x^* \in \arg\min f \iff 0 \in \partial f(x^*)$

$g = \iota_C$, C convex -- indicator function (of constraints, e.g.)

$f = f_0 + \iota_C$, $\quad 0 \in \partial f(x^\star) \iff ?\ 0 \in \partial \mathrm{f}_0(x^\star) + \partial \iota_C(x^\star)$

- $\partial \iota_C(\bar{x}) = N_C(x) = \left\{ v \middle| \langle v, x - \bar{x} \rangle \ \forall x \in C \right\}$, normal cone to C @ $\bar{x}$



Subdifferential Calculus

proper fcn: $\operatorname{dom} f \neq \emptyset,\ f > -\infty.\quad f, g$ proper, convex

$$df(x;\cdot) + dg(x;\cdot) = d(f+g)(x;\cdot),$$

$$\partial f(x) + \partial g(x) \subset \partial(f+g)(x)$$

when $\operatorname{int}\operatorname{dom} f \cap \operatorname{dom} g = \emptyset$, then $\boxed{\partial f(x) + \partial g(x) = \partial(f+g)(x)}$

$$f = f_0 + \iota_S,\ f_0(x) \in \mathbb{R} \Rightarrow x^\star \in \arg\min f \Leftrightarrow 0 \in \partial f(x^\star) + N_C(x^\star)$$

convex programs:

$$\min f = f_0 + \iota_S,\ S = \left\{ x \in X \,\middle|\, f_i \le 0, i:1 \to s,\quad f_i = 0,\ i: s+1 \to m \right\}$$

X closed, convex   convex   affine

linearly constrained: (linear, quadratic, ... programs)

X polyhedral (box)   affine   affine

Convex, Linearly Constrained

$C = \{x | \langle A_i, x\rangle \le b_i, i = 1,\ldots,s,\ \langle A_i, x\rangle = b_i, i = s+1,\ldots,m\}$ polyhedral

with $\bar{x} \in C$,

$v \in N_C(\bar{x}) \Leftrightarrow \exists \lambda_1 \ge 0,\ldots,\lambda_s \ge 0,\ \lambda_{s+1}\ldots,\lambda_m \in \mathbb{R}$ such that

$$v = \sum_{i=1}^{m} \lambda_i A_i \quad i = 1,\ldots,s, \quad \lambda_i(\langle A_i, \bar{x}\rangle - b_i) = 0 \qquad \left(\lambda \perp (A\bar{x} - b)\right)$$

$x^\star$ sol'n of linearly constrained convex program if and only if

one can find KKT-multipliers $y \in \mathbb{R}^m$ such that

(a) $\left\langle A_i, x^\star \right\rangle \ge b_i, i = 1,\ldots,s, \quad \left\langle A_i, x^\star \right\rangle = b_i, i = s+1,\ldots,m,$

(b) $i = 1,\ldots,m, \quad y_i \ge 0,\ y_i(\langle A_i, x^\star\rangle - b_i) = 0,$

(c) $x^\star \in \arg\min f_0(x) - \left\langle A^\top y, x\right\rangle, x \in X$ (box)

(~c) -$v \in N_X(x^\star)$ such that $\partial f_0(x^\star) \ni v + A^\top y$

Expectation Functionals

plain probability measures: for $A \in \mathcal{A}$, $P(A) = P_d(A) + P_c(A)$

$$\Rightarrow \mathbb{E}\{h(\xi)\} = \sum_{\xi \in \Xi_d} h(\xi) p_\xi + \int_{\xi \in \Xi_c \subset \mathbb{R}^N} h(\xi)\, p(\xi)\, d\xi)$$

density

$$\mathbb{E}\{S(\xi)\} = \{\mathbb{E}\{s(\xi)\} \,|\, s(\xi) \in_{a.s.} S(\xi), s \text{ summable}\} \subset \mathbb{R}^N$$

properties of E :

linearity, order preserving, dominated convergence

▵ f random convex lsc functions $\Rightarrow x \mapsto Ef(x)$ convex

when Ef finite-valued:

$$dEf(x;w) = \mathbb{E}\{df(\xi,x;w)\} \quad \& \quad \partial Ef(x) = \mathbb{E}\{\partial f(\xi,x)\}$$

▵ $x^* \in \arg\min Ef \Leftrightarrow \exists\, v \in \partial f(\bullet, x^*)$, $\mathbb{E}\{v(\xi)\} = 0$ &

$$\boxed{\forall \xi \in \Xi : \; x^* \in \arg\min_x \left[f(\xi,x) - \langle v(\xi), x \rangle \right]}$$

!! Remark: to solve $0 \in \partial Ef(x^*)$ it suffices to know $v(\xi)$ for just one ξ and solve one problem of the same size as the deterministic version

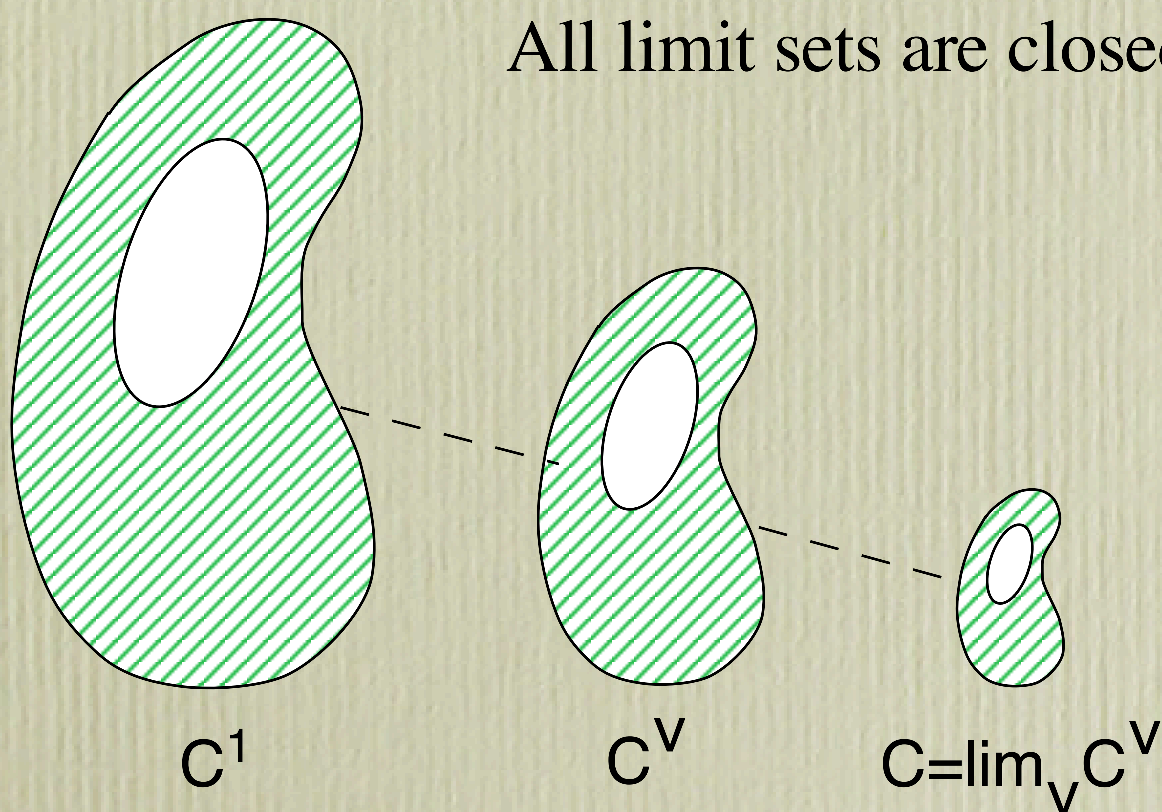
Approximation: Convergence

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

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

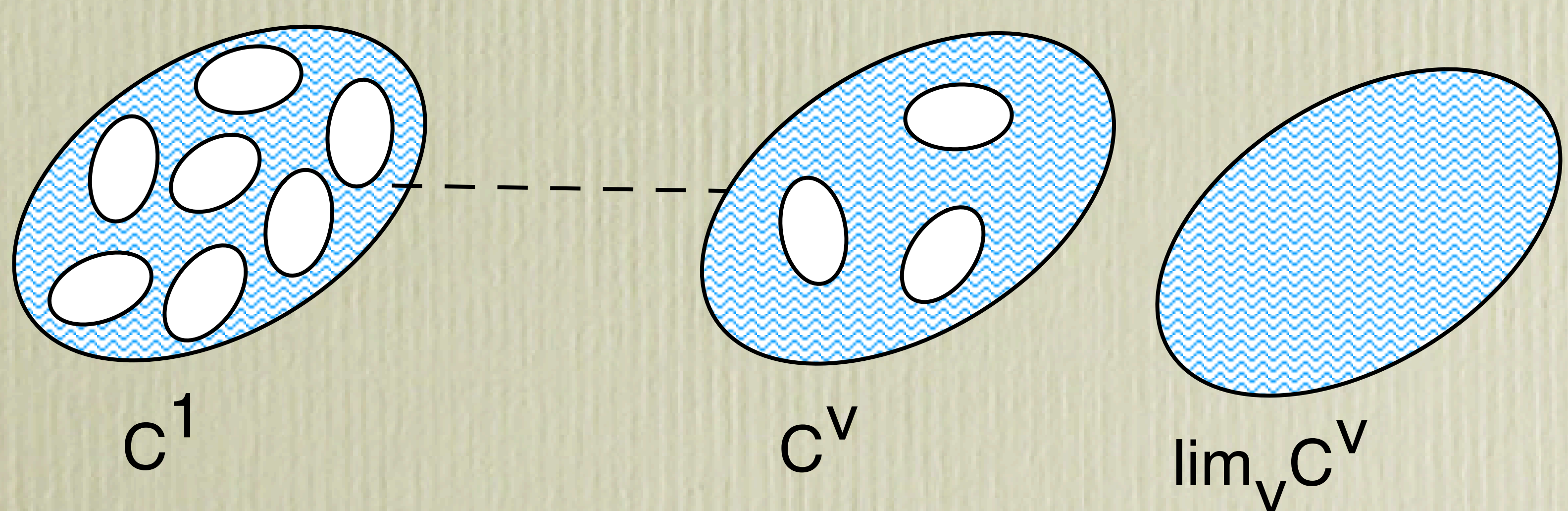


Convex limit sets

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

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

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), $\text{e-li}_\nu f^\nu \le \text{e-ls}_\nu f^\nu$

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:

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

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

ε-$A^\nu : \varepsilon > 0$ 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) $\bar{x} \in A \Rightarrow \exists\ \varepsilon_\nu \searrow 0, \mathrm{x}^\nu \in \varepsilon_\nu\text{-}A^\nu \to \bar{x}$

Convergence of Minimizers
Sv-convergence of Minimizers

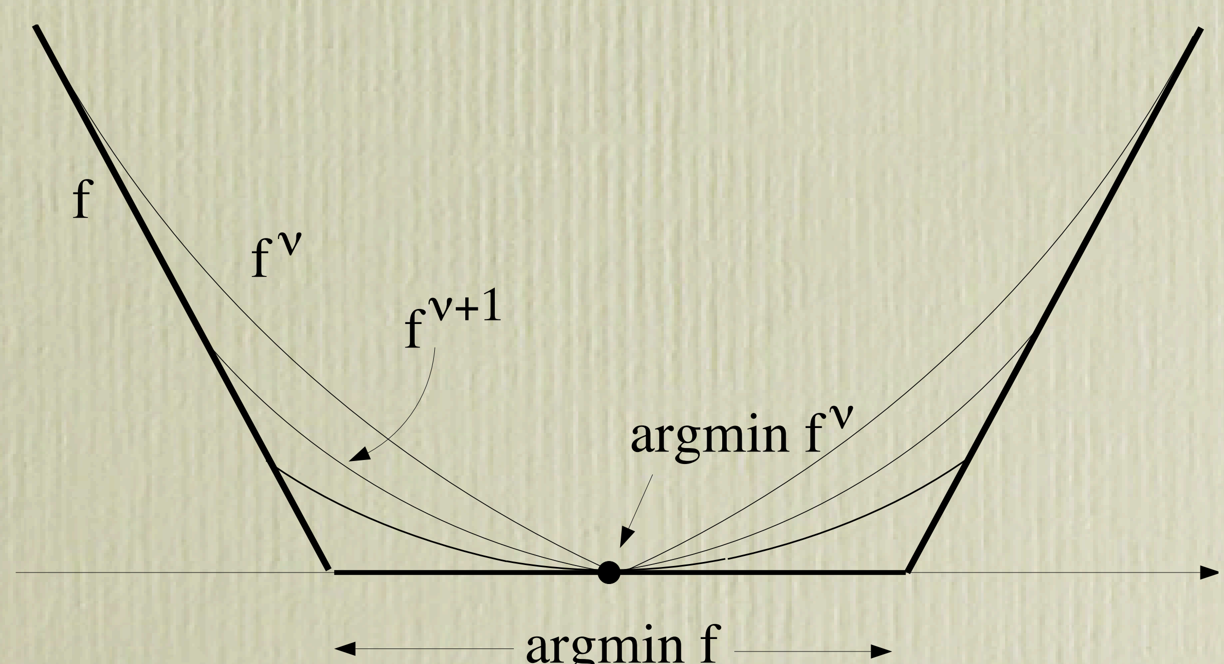
$$f^\nu \underset{e}{\to} f, x \in \text{ cluster}\left\{x^\nu \in \arg\min f^\nu\right\} \Rightarrow x \in \arg\min f$$

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

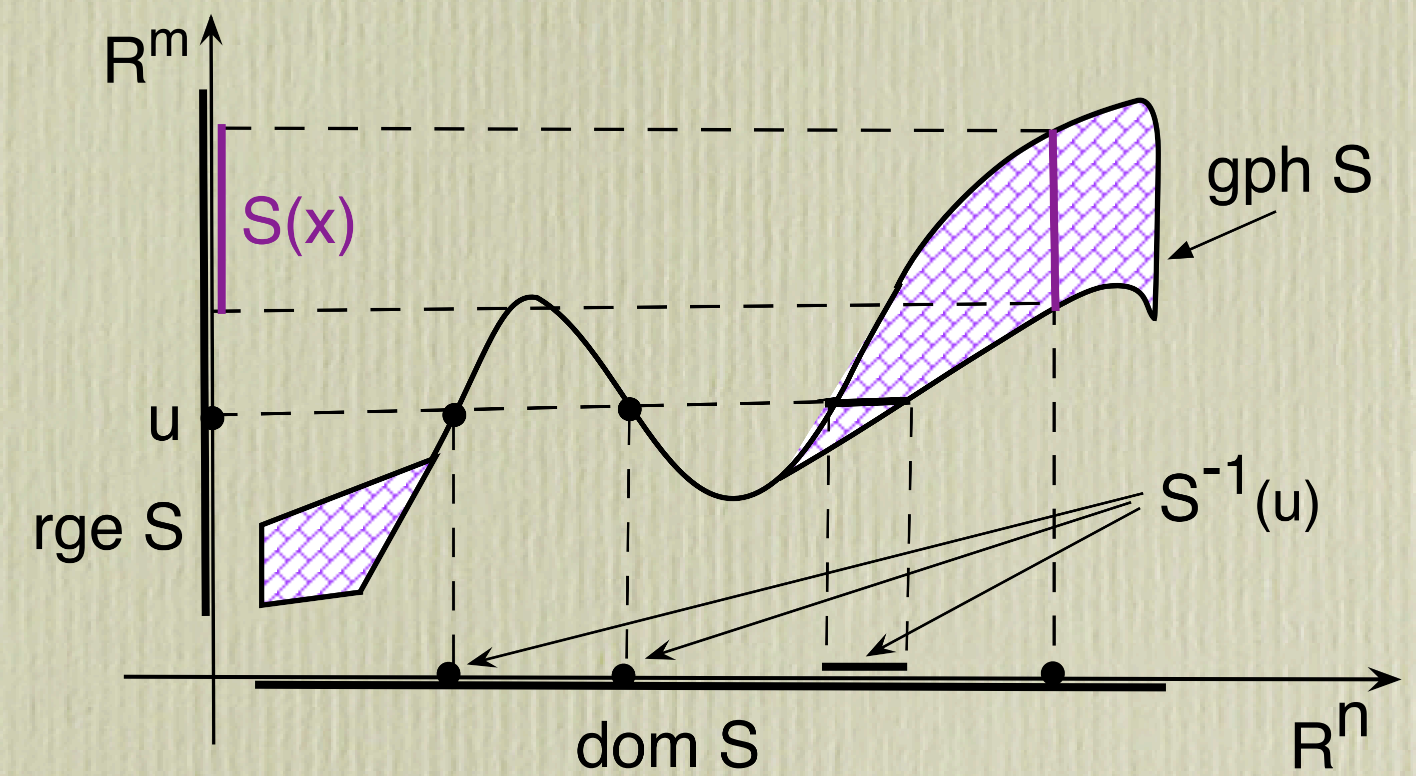
$$f^\nu \underset{e}{\to} f \not\Rightarrow \arg\min f^\nu \to \arg\min f$$

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

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



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

$S^\nu \to_g S$ when gph $S^\nu \to$ gph S (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

CONVERGENCE RATES

Excess distance function:

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

Estimate of set distance:

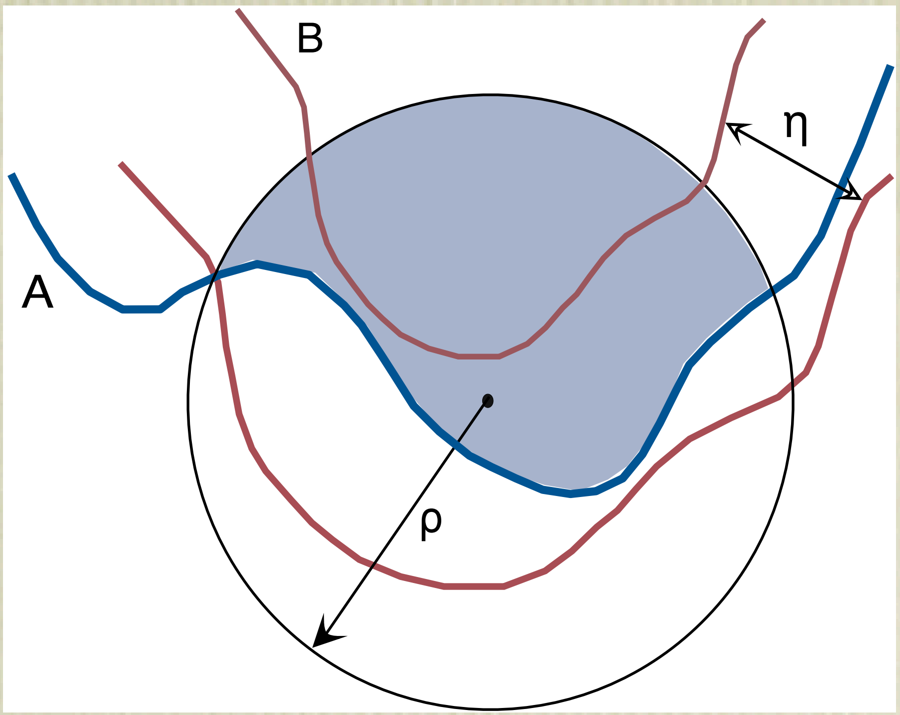
$$d\hat{l}_\rho(A,B) = \max[\,e_\rho(A,B), e_\rho(B,A)\,]$$

Set-distance:

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

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

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



$$d\hat{l}_\rho(A,B) \le dl_\rho(A,B) \le d\hat{l}_{\rho'}(A,B),$$

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

$$C^\nu \to C \Leftrightarrow dl_\rho(C^\nu, C) \to 0 \Leftrightarrow d\hat{l}_\rho(C^\nu, C) \to 0 \quad \forall \rho \ge 0$$

EPI-DISTANCE

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

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

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

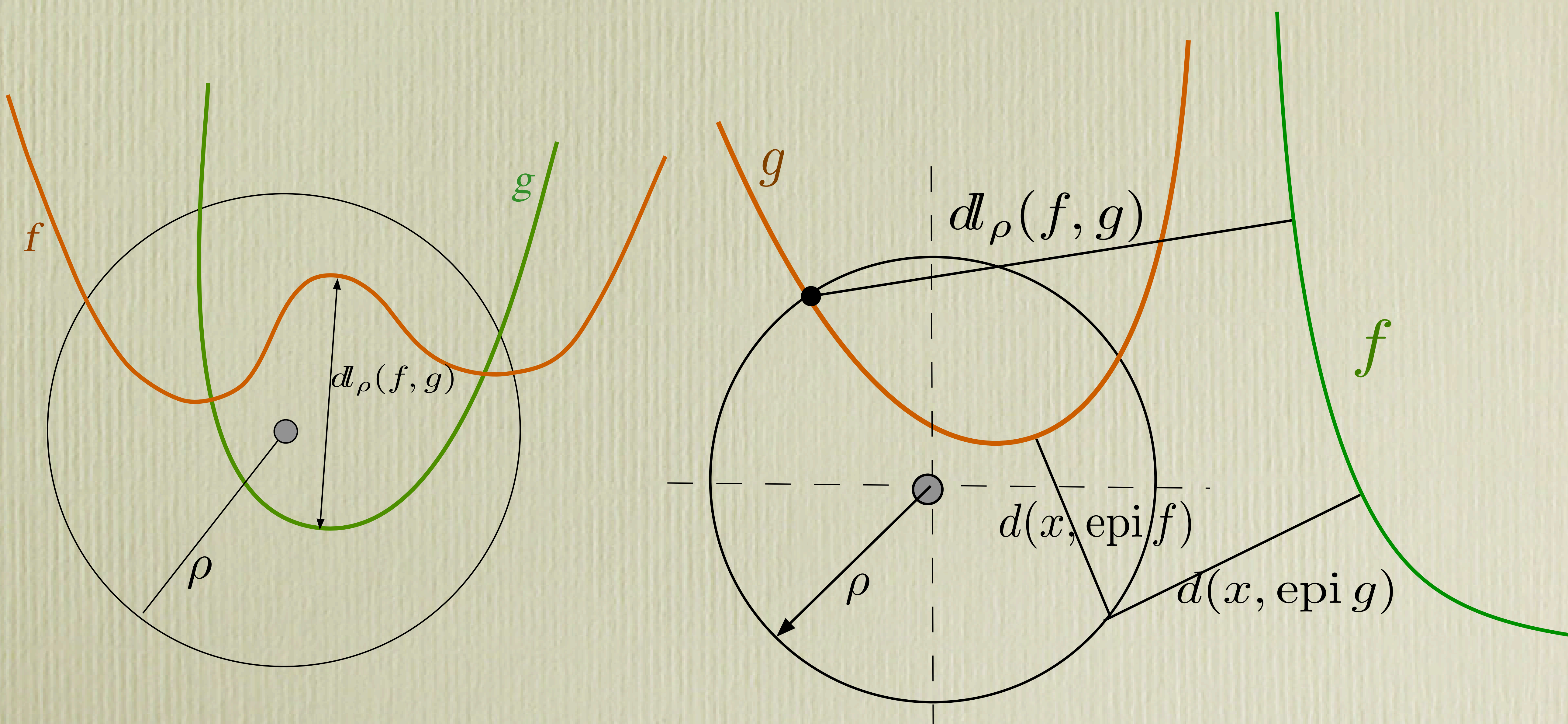
$dl(f,g) = \int_{\rho \geq 0} e^{-\rho} dl_\rho(f,g)\, d\rho$, epi-distance, Attouch-Wets topology

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

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

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

Epi-distance

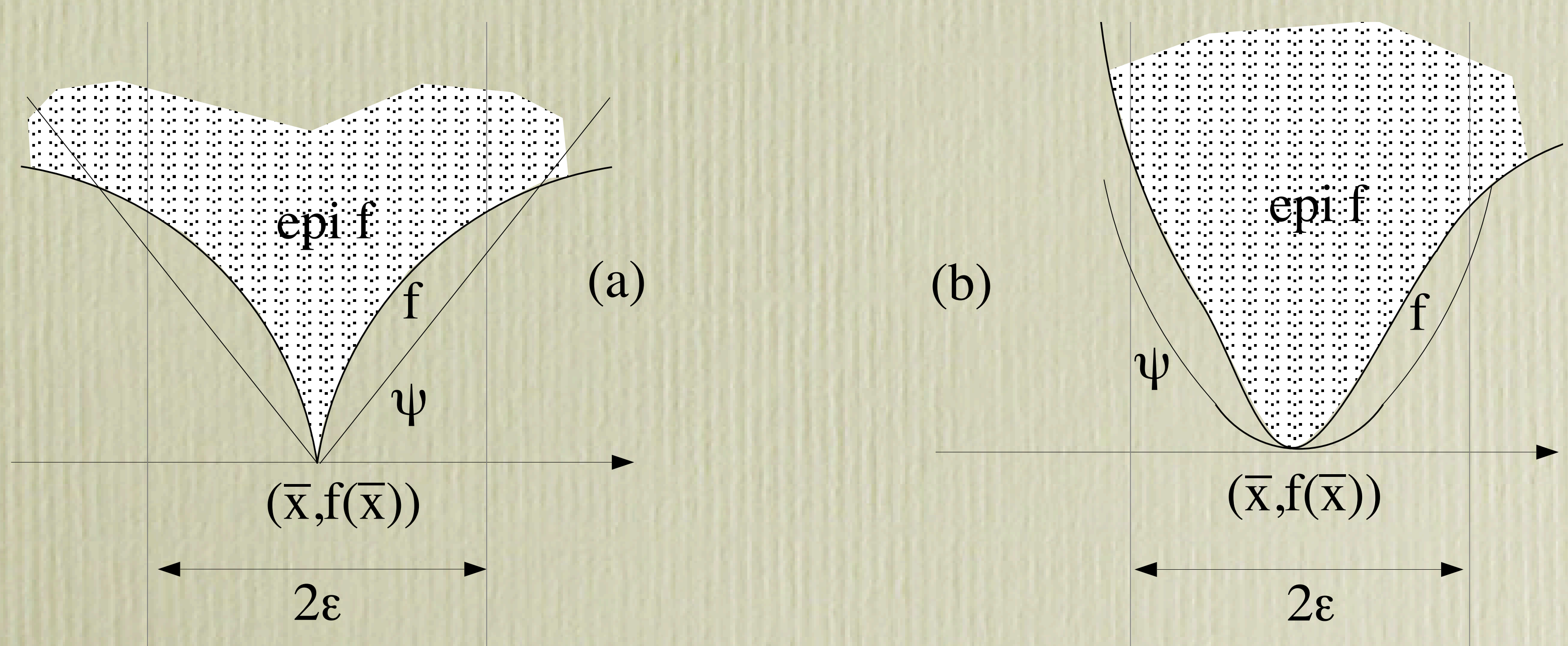


QUANTITATIVE ESTIMATES

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

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

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



QUANTITATIVE ESTIMATE
convex functions

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

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

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

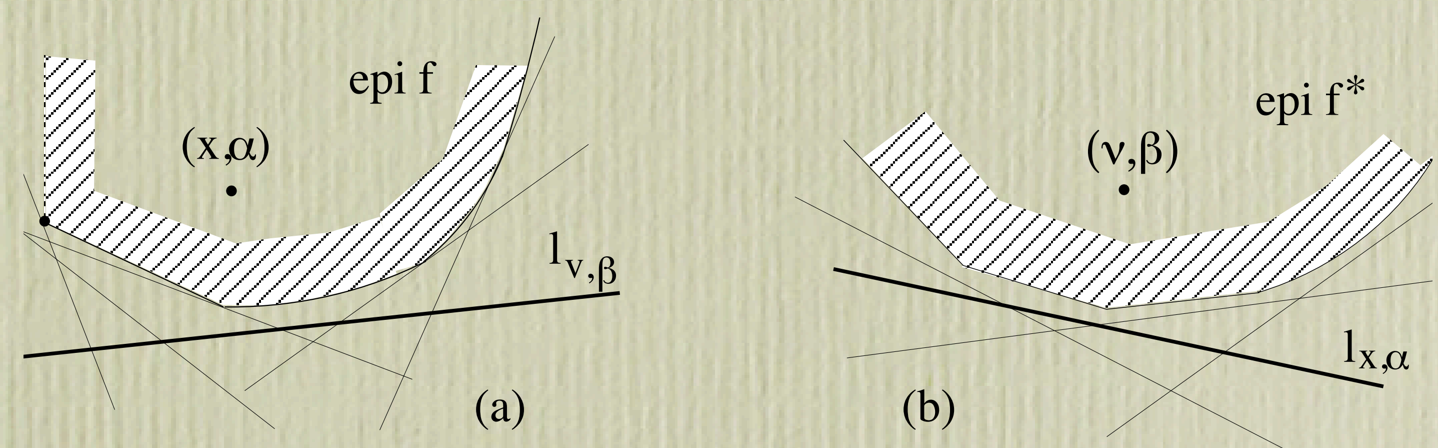
$\min f \geq -\rho_0, \ \min g \geq -\rho_0$

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

$$\boxed{\begin{aligned} d\hat{l}_\rho(\varepsilon\text{-arg min} f, \varepsilon\text{-arg min g}) &\leq \bar{\eta}\left(1 + \frac{2\rho}{\bar{\eta} + \epsilon/2}\right) \\ &\leq (1 + 4\rho/\epsilon) d\hat{l}_\rho(f, g) \end{aligned}}$$

Convex functions

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



conjugate functions

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

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

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

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 $f^{\nu*}(v^\nu) \to f^*(\bar{v})$

Epi-Convergence

review

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

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

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

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

Pointwise:

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

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

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

Epi-Convergence $\Rightarrow$ convergence of minimizers

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

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

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

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

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

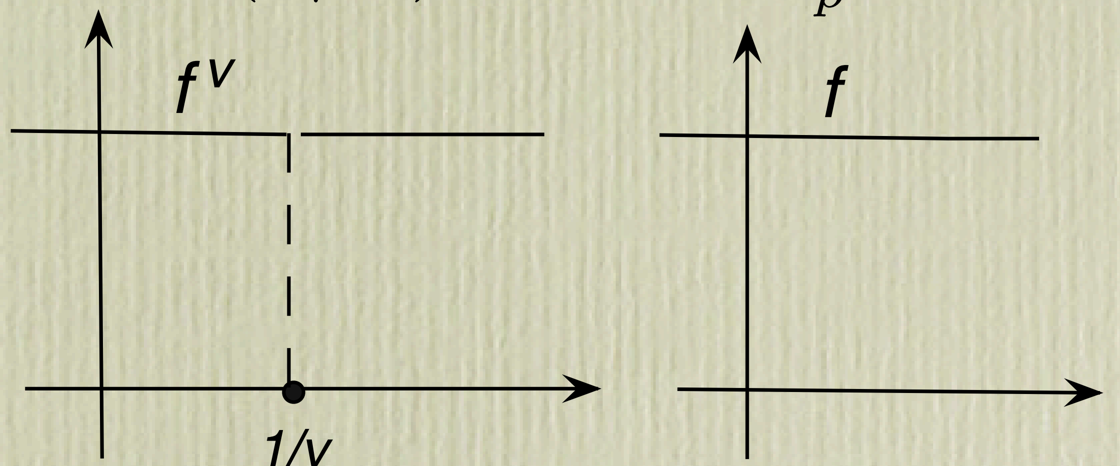
$f^\nu \xrightarrow{e} f$ implies ε-$A^\nu \Rightarrow_v \varepsilon$-$A$, $\forall\, \varepsilon \geq 0$
A unique minimizer, ε^ν-$A^\nu \rightrightarrows A$ as $\varepsilon^\nu \searrow 0$.

(inf f > -∞)

Why epi-convergence?

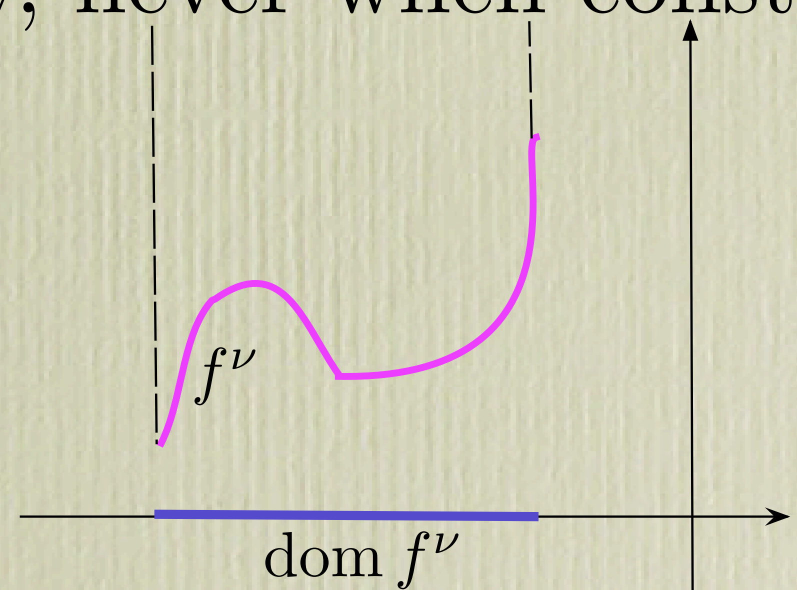
1. pointwise convergence $\not\Rightarrow$ convergence of minimizers

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



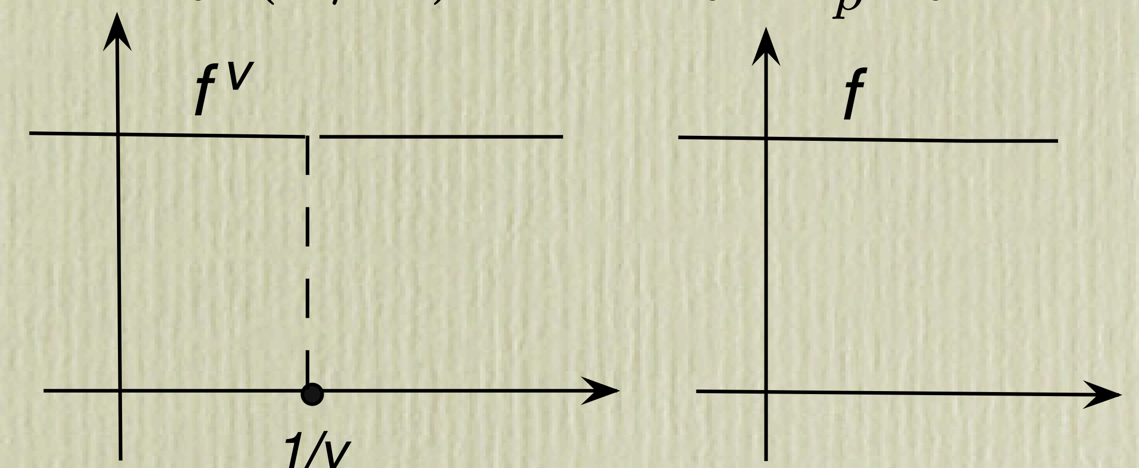
2. uniform convergence implies convergence of minimizers

   but applies rarely, never when constraints depend on ν

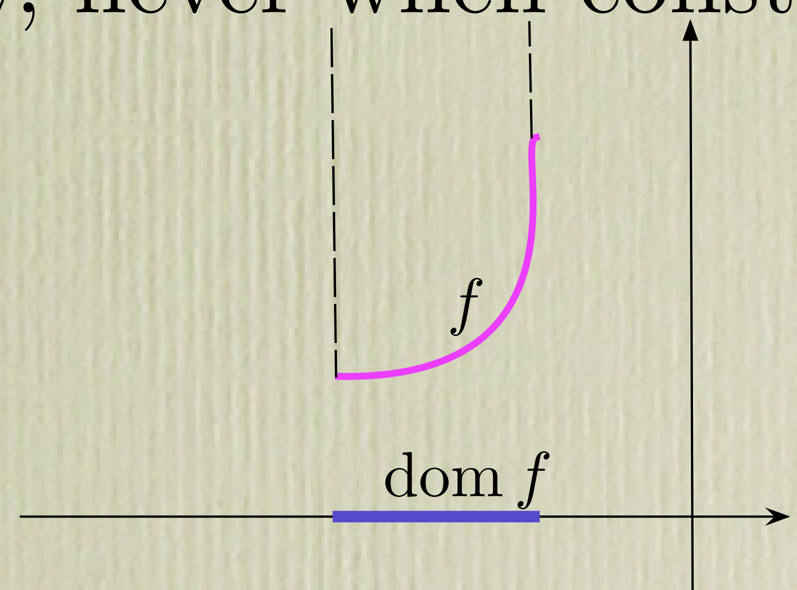


Why epi-convergence?

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



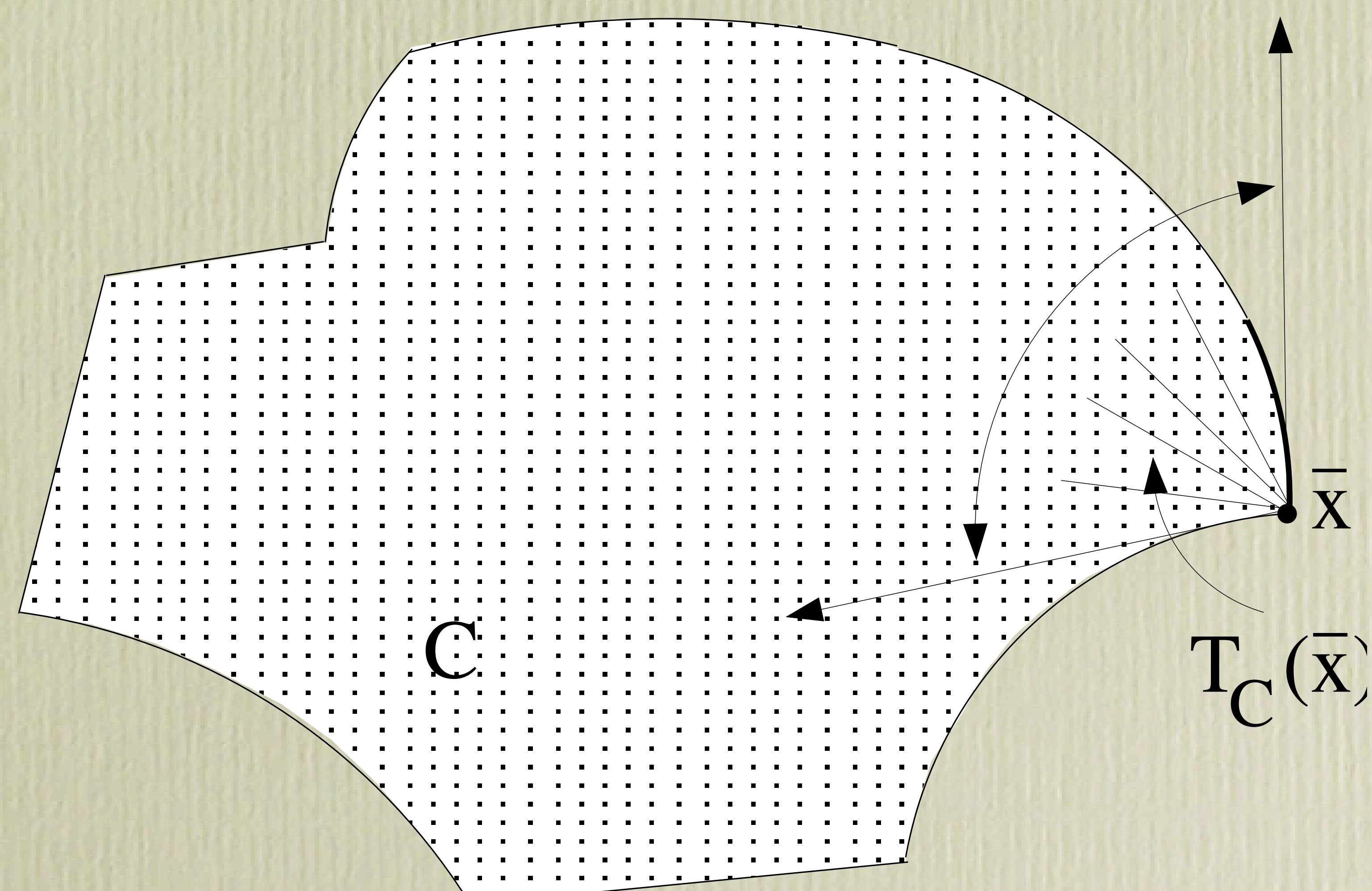
2. uniform convergence implies convergence of minimizers
   but applies rarely, never when constraints depend on ν



Variational epi-convergence

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 $x^\nu \underset{C}{\to} x, \tau_\nu \searrow 0$

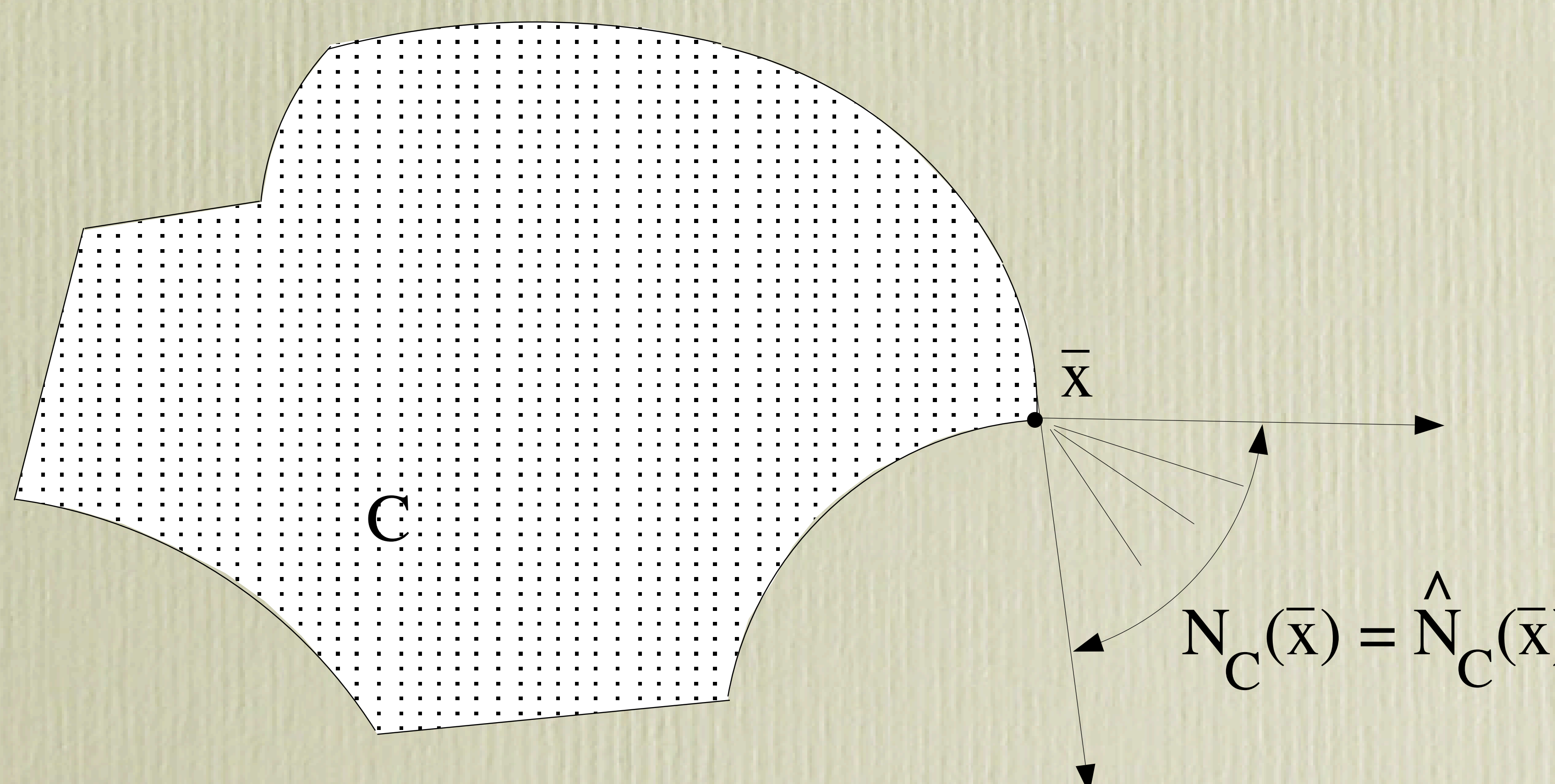


Variational Geometry
Normal Cone

$v \in \hat{N}_C(\bar{x})$, 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 $\exists x^\nu \underset{C}{\rightarrow} x$ and $v^\nu \rightarrow v$ with $v^\nu \in \hat{N}_C(x^\nu)$

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



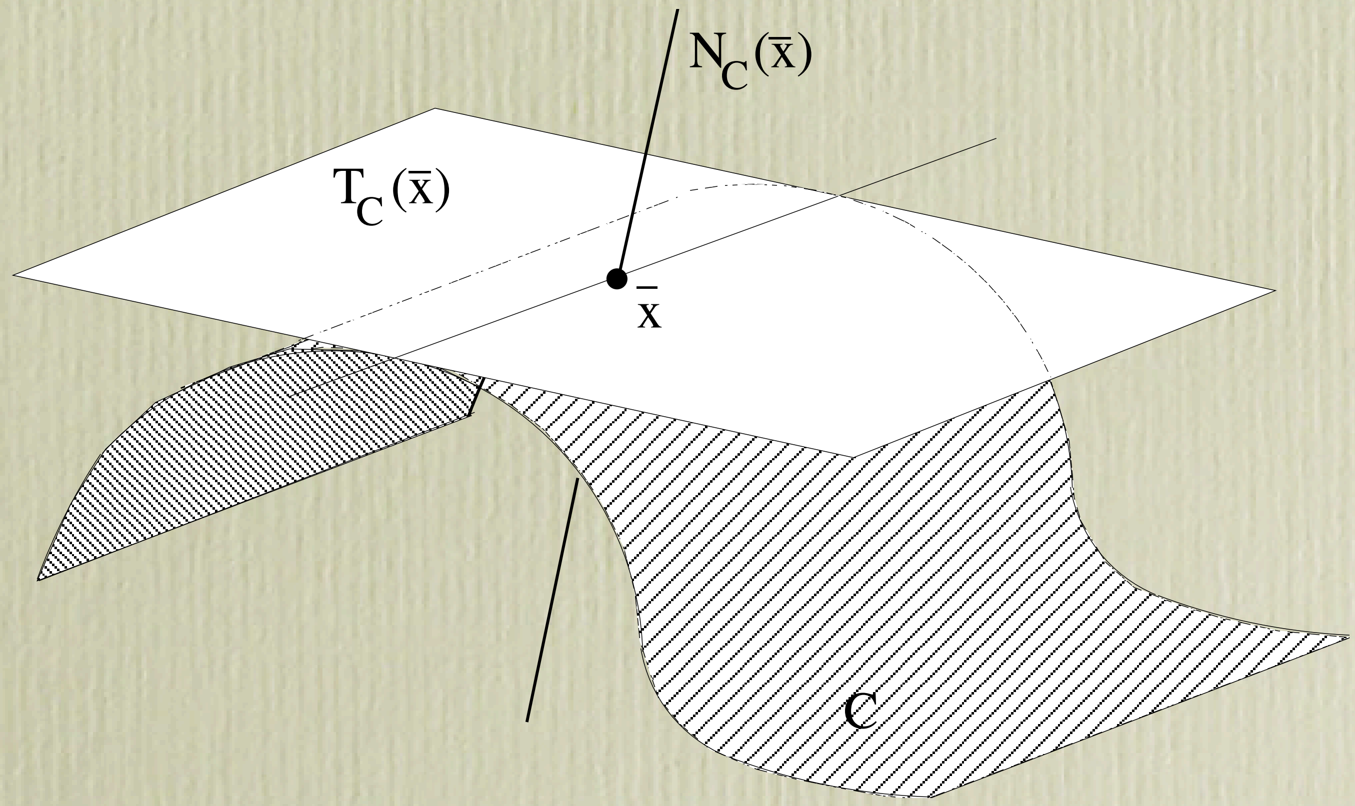
Clarke regularity

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

(never supergradients)

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

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

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

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

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

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, $\partial\left(f+g\right) \neq \partial f + \partial g$

$\mathbb{C}.\mathbb{Q}.$ (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

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

Fermat's Rule (quite a bit generalized):

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 (no $\mathbb{C}.\mathbb{Q}.$ in this form)

Stochastic Variational Analysis

Why?

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

$EG(x) = \mathbb{E}\{G(\boldsymbol{\xi}, x)\} = 0$ "approximated" by $G^\nu(x) = 0$

$\xi^1, \dots, \xi^\nu$ sample, $G^\nu(x) = \frac{1}{\nu}\sum_{l=1}^{\nu} G(\xi^l, x)$

$G : \Xi \times D \rightrightarrows E$, set-valued $G(\xi, x) \subset E$, inclusion $\mathbb{E}\{G(\boldsymbol{\xi}, x)\} \ni 0$

$\xi^1, \dots, \xi^\nu$ sample, approximation $\frac{1}{\nu}\sum_{l=1}^{\nu} G(\xi^l, x) \ni 0$

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

$\xi^1, \dots \xi^\nu$ sample P^ν (random) empirical measure

approx.: $\min \mathbb{E}^\nu\{f(\boldsymbol{\xi}, x)\} = \frac{1}{\nu}\sum_{l=1}^{\nu} f(\xi^l, x),\ x \in C$

Some Examples:

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

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

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

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

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

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

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

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

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

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

Pricing financial instruments

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

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

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

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

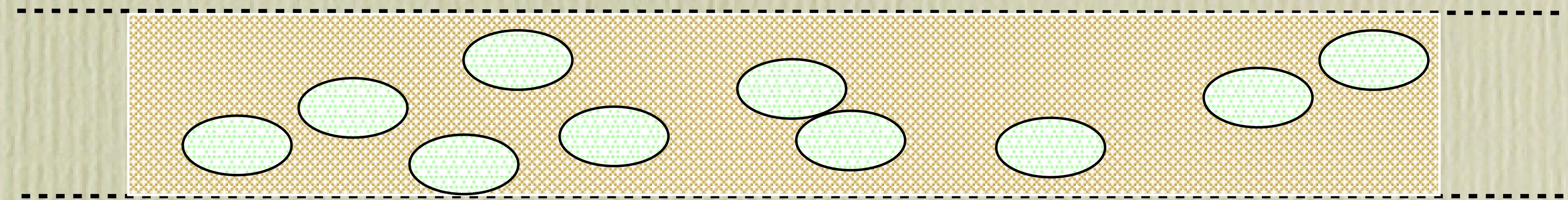
$$\max \mathbb{E}\left\{\langle S^T, X^T\rangle\right\} \text{ such that } \langle S^t, X^t\rangle \le G^t + \langle S^t, X^{t-1}\rangle,\ t = 1 \to T$$

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

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

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

4. Stochastic homogenization, ...



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

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

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

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

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

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

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

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

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

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

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

Regression: (E is not a linear space)

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

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

Random lsc functions

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

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

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

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

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

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

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

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

Random lsc functions

(via inf-projections)

D countable dense subset of E

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

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

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

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

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

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

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

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

Epi-convergence: Characterization

(via inf-projections)

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

$$\boxed{\operatorname{limsup}_\nu c^\nu_{x\delta} \le c_{x\delta}, \quad \operatorname{liminf}_\nu o^\nu_{x\delta} \ge o_{x\delta}}$$

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

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

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

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

Scalarization of random lsc fcns

$f : \Xi \times E \to \overline{\mathbb{R}}$, random lsc fcn, completely identified by
$\left\{ o_{x\delta}(\xi) = \inf_{\mathbb{B}^o(x,\delta)} f(\xi,\cdot) \,\middle|\, x \in D, \delta \in \mathbb{Q}_+ \right\}$, countable
or $\left\{ c_{x\delta}(\xi) = \inf_{\mathbb{B}(x,\delta)} f(\xi,\cdot) \,\middle|\, x \in D, \delta \in \mathbb{Q}_+ \right\}$, $\forall \xi \in \Xi$

COUNTABLE

$\forall\, x \in \mathbb{R}^n, \delta > 0$
$\xi \mapsto o_{x\delta} : \Xi \to \overline{\mathbb{R}}$ are measurable,
$o_{x\delta}(\boldsymbol{\xi})$ extended real-valued random variable
$\xi \mapsto c_{x\delta} : \Xi \to \overline{\mathbb{R}}$ are measurable,
$c_{x\delta}(\boldsymbol{\xi})$ extended real-valued random variable.

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

Probabilistic properties
iid-properties (pairwise "i")

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

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

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

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

Countable $\Rightarrow$ a.s.

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

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

Proof. $f \le g$ on D only if $\{(x,\alpha) \,|\, \alpha \ge g(x), x \in R\} \subset \operatorname{epi} f$.
Taking closure on both sides $\implies$ $\operatorname{epi} g \subset \operatorname{epi} f$. $\square$

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

LLN: random lsc functions?

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

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

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

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

in general

Law of Large Numbers: Random lsc functions

LLN: Proof

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

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

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

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

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

Law of Large Numbers
(random lsc fcns)

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

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

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

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

SAA-applies without 'any' restrictions

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

Ergodic Theorem

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

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

allows for stationary rather than iid.

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

Approximation: Probability Measures

1. $\{P^\nu,\ v = 1, \ldots\} \underset{\mathcal{D}}{\to} P$, $f : \Xi \times \mathbb{R}^n \to \overline{\mathbb{R}}$ random lsc fcn function and for given x, $\xi \mapsto f(\xi, x)$ continuous on Ξ, $\forall\, \varepsilon > 0$,

(a) there exists a neighborhood V of x such that

$$|f(\xi, y) - f(\xi, x)| < \varepsilon, \quad \text{for all } y \in V.$$

(b) there exists a subset $\Xi_\varepsilon \in \mathcal{A}$ such that

$$\int_{\Xi \setminus \Xi_\varepsilon} |f(\xi, x)| P^\nu(d\xi) < \varepsilon, \quad \text{for all } \nu \in \mathbb{N}.$$

Then, $E^\nu f = \xrightarrow{c} Ef = \int f(\xi, \cdot)\, dP(\xi)$ at x, i.e.,

$$\forall\, \{x^\nu,\ v = 1, \ldots\} \to x, \quad Ef(x) = \lim_{\nu \to \infty} E^\nu f(x^\nu).$$

Approx.: Probability Measures 2

2. $\{P^\nu\}_{\nu\in\mathbb{N}} \underset{\mathcal{D}}{\to} P$

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

$\forall\, \varepsilon > 0, \exists\, \Xi_\varepsilon \in \mathcal{A}$ such that $\forall\, x \in \operatorname{dom} f$,

$$\int_{\Xi\setminus\Xi_\varepsilon} |f(\xi,x) P^\nu(d\xi) < \varepsilon, \quad \forall\, \nu \quad (f\text{-tight})$$

Then, $E^\nu f \xrightarrow{e} EF = \int_\Xi f(\xi,\cdot)\, P(d\xi)$.

Quantitative Approximation

3. f random convex lsc fcn, $\operatorname{dom} f(\xi,\cdot) = X$ (constant)

$$\mathcal{F} = \left\{ \xi \mapsto f(\xi,x) \,\middle|\, x \in X \right\}$$

$$\mathcal{P}_{\mathcal{F}} = \left\{ Q \in \mathcal{P}(\Xi) \,\middle|\, \begin{array}{c} \forall \rho > 0,\ \sup_{\rho\mathbb{B}} \int f(\xi,\cdot) Q(d\xi) < \infty \\ \int \inf_{\rho\mathbb{B}} f(\xi,\cdot) Q(d\xi) > -\infty \end{array} \right\}$$

$$d_{\mathcal{F},\rho}(P,Q) = \sup_{x \in \rho\mathbb{B}} \left| E^P f(x) - E^Q f(x) \right|, \text{ pseudo-metric on } \mathcal{P}_{\mathcal{F}}$$

$$P \in \mathcal{P}_{\mathcal{F}}, \emptyset \neq \operatorname{arg\,min} Ef \text{ bounded } \Rightarrow \exists \hat{\rho} > 0, \hat{\varepsilon} > 0:$$

$$\forall \varepsilon \in (0,\hat{\varepsilon}),\ Q \in \mathcal{P}_{\mathcal{F}} \text{ such that } d_{\mathcal{F},\hat{\rho}+\varepsilon}(P,Q) < \varepsilon:$$

$$\boxed{dl_\infty(\varepsilon\text{-}\operatorname{arg\,min} E^P f, \varepsilon\text{-}\operatorname{arg\,min} E^Q f) \leq \frac{4\hat{\rho}}{\varepsilon} d_{\mathcal{F},\hat{\rho}+\varepsilon}(P,Q)}$$

Proof: via epi distance between $E^P f$ and $E^Q f$

Solving Stochastic Programs

L-Shaped Strategy

Benders, Dantzig-Wolfe dual

taking advantage of structure

Two-Stage Recourse Model

With (arbitrary linear) recourse:

$\min_x \langle c, x\rangle + \mathbb{E}\big\{Q(\boldsymbol{\xi}, x)\big\},\ Ax = b,\ \ x \geq 0$

where $\boldsymbol{\xi} = (q(\boldsymbol{\xi}, \cdot\,), \boldsymbol{T}, \boldsymbol{d}, \boldsymbol{W})$

$Q(\xi, x) = \inf\big\{\langle q(\xi, y)\rangle \,\big|\, W_\xi y = d_\xi - T_\xi x,\ y \geq 0\big\},\ \ q(\xi, \cdot\,)$ convex

The *deterministic equivalent problem*, a convex program:

$\min_x \langle c, x\rangle + EQ(x)$ such that $Ax = b,\ x \geq 0$

with $EQ(x) = E\{Q(\boldsymbol{\xi}, x)\} = \int_\Xi Q(\xi, x)\, P(d\xi)$

but, generally, EQ is not finite valued.

Q finite-valued implies for all decision x,

for all events ξ a recourse is available

.

Multi-Stage: Deterministic Equivalent

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

"time-staged objective":

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

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

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

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

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

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

Solution procedures

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

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

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

deterministic optimization! convex when f convex random lsc 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, prohibitively EXPENSIVE

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

sampling: a) same as approximation except P^ν random measure

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

Sequential l.p. Strategy

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

$\quad f_i(x) \le 0,\ \ i = 1,\ldots,s,\ \ f_i(s) = 0, i = s+1,\ldots,m$ (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 $\mathrm{i}_\nu : f_{i_\nu}(x^\nu) = \max\left\{ {}_{1\le i\le s} f_i(x^\nu),\ {}_{s+1<i<m} \left| f_i(x^\nu) \right| \right\}$

   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

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

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

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

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

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.

Cut Generation: Fixed Recourse

(l.p.)-solution: (x^ν, θ^ν)

x^ν feasible?

$\forall \xi \in \Xi : \; z_\xi = \operatorname{argmax}_z \left\{ \langle d_\xi - T_\xi x^\nu, z \rangle \,\middle|\, W^\top z \leq 0, \; -1 \leq z_j \leq 1 \right\}$

if $\eta_\xi = \langle d_\xi - T_\xi x^\nu, z \rangle = 0$, x^ν feasible

for some $\xi, \eta_\xi > 0$, then $E_{k+1} = (T_\xi)^\top z_\xi, \; e_{k+1} = \langle d_\xi, z_\xi \rangle$

.

x^ν optimal?

$\forall \xi \in \Xi : \; v_\xi = \operatorname{argmax}_v \left\{ \langle d_\xi - T_\xi x^\nu, v \rangle \,\middle|\, W^\top z \leq q_\xi \right\}$

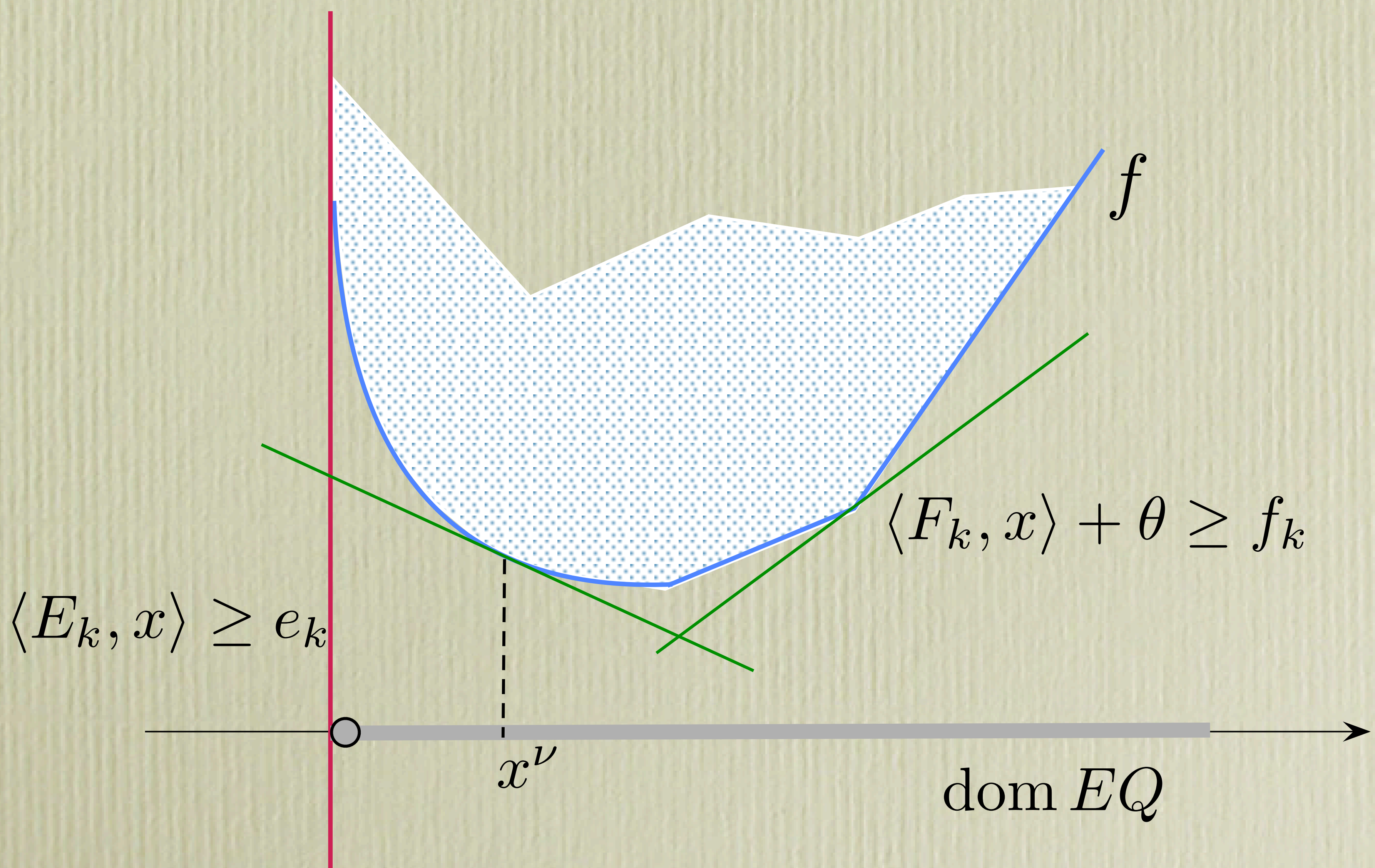
if infeasible for some $\xi \quad \Longrightarrow \quad$ unbounded problem

otherwise $F_{k+1} = \mathbb{E}\{\boldsymbol{T v}\}, \quad f_{k+1} = \mathbb{E}\{\langle \boldsymbol{d}, \boldsymbol{v} \rangle\}$

if $\theta^\nu \geq fk + 1 - \langle F_{k+1}, x^\nu \rangle \quad \Longrightarrow \quad x^\nu$ optimal

add optimality cut

Generating cutting hyperplanes



Aggregation Principle

in

Stochastic

Optimization

Interchanging: $\boxed{\mathbb{E} \;\&\; \min}$

Evident: with $E = \{x : \Xi \to \mathbb{R}^N \mid \text{ measurable, ... }\}$
$$\min \mathbb{E}\left\{ f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big) \,\middle|\, x \in E \right\} = \mathbb{E}\left\{ \min f(\boldsymbol{\xi}, x) \,\middle|\, x \in \mathbb{R}^N \right\}$$
when $\exists\, x(\,\cdot\,) \in E$ such that P-$a.s.$ $x(\xi) \in \operatorname{argmin} f(\xi, \cdot\,)$
x is measurable, ...

But our problem is: $\min \mathbb{E}\{f(\boldsymbol{\xi}, x)\}$, equivalently,

$$\min Ef(x) = \mathbb{E}\{f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big)\}$$
$$\text{such that } \; x(\boldsymbol{\xi}) = \mathbb{E}\{x(\boldsymbol{\xi})\} \; P\text{–}a.s.$$

x can not depend on 'anticipated' (future) information

Dynamic Information Process

So far, x mostly restricted to $\{\emptyset, \Xi\}$-measurable, i.e., constant on Ξ

Generally, as $t \nearrow T$ (possibly ∞) additional information is acquired

$$\mathcal{A}_0 = \{\emptyset, \Xi\} \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}_T = \mathcal{A}, \text{ a filtration}$$

with x_t decision @ time t depend on available information, i.e. $\mathcal{A}_t$-measurable

<u>Reformulation</u>

Let $x(\xi) = \big(x_0(\xi), x_1(\xi), \dots, x_T(\xi)\big) : \Xi \to \mathbb{R}^N,\ N = \sum_{t=0}^T n_t$

$$\mathcal{N}_a = \left\{x \in E \,\middle|\, x_t\ \mathcal{A}_t\text{-measurable},\ t = 0, \dots T\right\}$$

$$\boxed{\text{find } x \in \mathcal{N}_a \text{ such that } Ef(x) = \mathbb{E}\{f(\boldsymbol{\xi}, x(\boldsymbol{\xi}))\} \text{ is minimized}}$$

Nonanticipativity <u>constraints</u>: $x \in \mathcal{N}_a$ (linear subspace)

Here-&-Now vs. Wait-&-See

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

$$x^1 \quad \to \quad \xi \quad \to \quad x^2_\xi$$

- Here-&-now problem! x^1

  not all contingencies available at time 0

  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 ?

Stochastic Optimization: Fundamental Theorem

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 Nonanticipativity

Here-&-now

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

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

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

Explicit non-anticipativity

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

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

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

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

$w_\xi \perp$ subspace of constant fcns

multipliers

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

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

such that $x^1_\xi \in C_1, \quad x^2_\xi \in C_2(\boldsymbol{\xi}, x^1_\xi)$

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

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

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

Progressive Hedging Algorithm

0. w^0_ξ such that $\mathbb{E}\{w^0_\xi\} = 0,\ \nu = 0.$ Pick $\rho > 0$

1. for all ξ :

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

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

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

otherwise $w^{\nu+1}_\xi = w^\nu_\xi + \rho\left[x^{1,\nu}_\xi - \bar{x}^{1,\nu} \right]$, return to 1. with $\nu = \nu + 1$

Convergence: add a proximal term

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

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

Nonanticipativity

Recall $\quad \min Ef(x) = \mathbb{E}\{f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big)\}$ such that $x(\xi) = \mathbb{E}\{x(\boldsymbol{\xi})\}$ P–$a.s.$

Nonanticipativity <u>constraints</u>:

$\mathcal{N}_a = \big\{x : \Xi \to \mathbb{R}^n\big\} \subset$ linear subspace of constant fcns

$\Longrightarrow\ \exists\, w : \Xi \to \mathbb{R}$ "multipliers" $\perp \mathcal{N}_a$ ($\Rightarrow \mathbb{E}\{w(\boldsymbol{\xi})\} = 0$) such that

$$x^* \in \operatorname{argmin} Ef \quad \Longrightarrow \quad x^* \in \operatorname{argmin} \big\{\mathbb{E}\{f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big) + \langle w(\boldsymbol{\xi}), (x(\boldsymbol{\xi}) - \mathbb{E}\{x(\boldsymbol{\xi})\})\rangle\big\}$$

$$\Longrightarrow \quad x^* \in \operatorname{argmin} \big\{\mathbb{E}\{f\big(\boldsymbol{\xi}, x(\boldsymbol{\xi})\big) + \langle w(\boldsymbol{\xi}), x(\boldsymbol{\xi})\rangle\}\big\}$$

$$P\text{–}a.s. \quad \Longrightarrow \quad x^* \in \operatorname*{argmin}_{x \in E} \{f\big(\xi, x\big) + \langle w(\xi), x\rangle\}\big\}, \quad \xi \in \Xi$$

$w(.)$: contingencies equilibrium prices, $\sim$ 'insurance' prices

PH: Implementation issues

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)

$$\min \mathbb{E}\{f(\boldsymbol{\xi},x)\}\ , \quad f(\boldsymbol{\xi},x) = f_0(x) + \iota_{C(\xi,x)}(x)$$

$S = \{\Xi_1, \Xi_2, \ldots, \Xi_K\}$ a partitioning of Ξ, $\quad p_k = P(\Xi_k)$

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

defining $g(k,x) = \mathbb{E}\{f_0(\boldsymbol{\xi},x) \mid \Xi_n\}$ if $x \in C_k = \bigcap_{\xi\in\Xi_k} C_\xi$

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

Multistage Stochastic Programs

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

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($\overset{\rightarrow v-1}{\boldsymbol{\xi}}$)

(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

$$\boxed{\begin{array}{l}\exists\,\bar{w}\perp\mathcal{N}^a, \bar{w}\in\mathcal{X}^* \text{ such that } \bar{x}\in\arg\min_{x\in\mathcal{X}} Ef(x)-\mathbb{E}\{\langle\bar{\boldsymbol{w}},x\rangle\}\\ \bar{w}\perp\mathcal{N}^a \Leftrightarrow \mathbb{E}\left\{\bar{w}(\boldsymbol{\xi})\,\middle|\,\mathcal{A}_{t-1}\right\}=0, \forall t=1,\dots,T\end{array}}$$

$\bar{w}$ non-anticipativity prices

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

can be viewed as insurance premiums, ...

A dual PH-strategy

single-stage case

minimize $Ef(x) := \sum_{\xi \in \Xi} p_\xi f(\xi, x)$ over all $x \in \mathbb{R}^n$

Strategy: better estimates of the w-variables
"aggregation" of the solutions" to

minimize $f^a(\xi, x)$ over all $\xi \in \mathbb{R}^n$ for fixed $\xi \in \Xi$

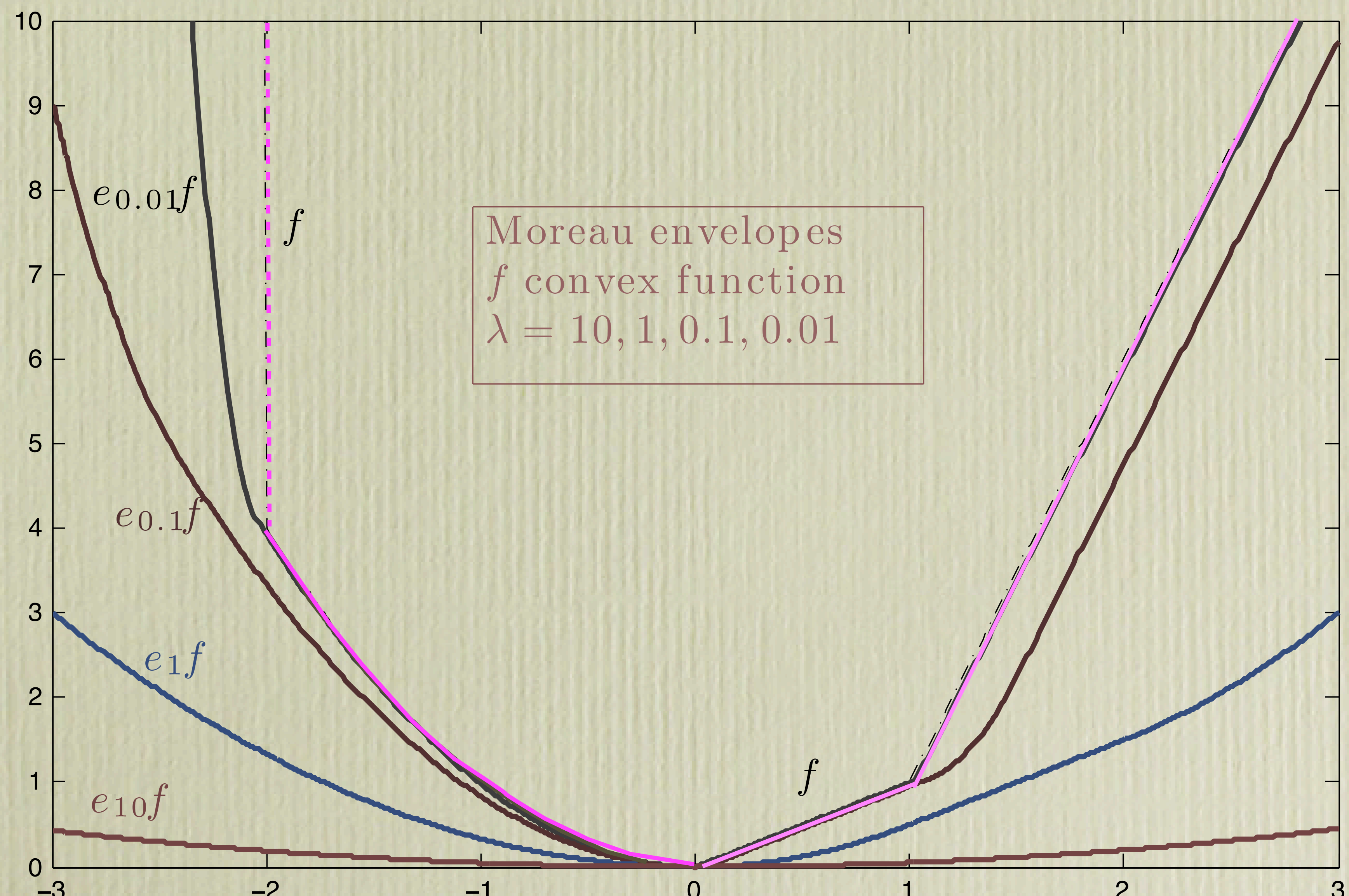
where f^a approximates f.

Moreau envelopes

a.k.a. Moreau-Yosida approximations

$$\operatorname{epi} f \# \operatorname{epi} g = \inf_u \{ f(u) + g(u-x) \} \quad e_\lambda f(x) \text{ with } g = \frac{1}{2\lambda} \left| \cdot \right|^2$$

epi-sums



Approximating problem

$\min F_\lambda(x) := \sum_{\xi\in\Xi} p_\xi f_\lambda(\xi, x), x \in \mathbb{R}^n$
$f_\lambda(\xi, x) = \inf_u \left\{ f(\xi, u) + \frac{1}{2\lambda} |u - x|^2 \right\}$
$F_\lambda \neq (F)_\lambda$ but $F_\lambda \xrightarrow{p} F, F_\lambda \xrightarrow{e} F$, finite-valued

Dual: $\max G_\lambda(w) = -\sum_{\xi\in\Xi} p_\xi f_\lambda^*(\xi, w_\xi),$
such that $\sum_{\xi\in\Xi} p_\xi w_\xi = 0$

Solution strategy:
$\min \sum_{k=0}^{\nu-1} \beta_k \alpha_k$ such that $\sum_{k=0}^{\nu-1} \beta_k \bar{w}^k = 0$
$\sum_{k=0}^{\nu-1} \beta_k = 1,\ \beta_k \geq 0,\ k = 0, \dots, \nu - 1$
for a 'desirable' collection of $\{\bar{w}^k\}$
Check for optimality, if not, generate $\bar{w}^\nu$

Algorithmic procedure

Step 0. Initialize by setting $\nu = 1$ picking $\{\widehat{w}^0_\xi\}_{\xi\in\Xi}$ in such a way that
$\bar{w}^0 := \sum p_\xi \widehat{w}^0_\xi = 0$ & $\alpha_0 = \sum_{\xi\in\Xi} p_\xi[\lambda/2|\widehat{w}^0_\xi|^2 + \sup_u(\widehat{w}^0_\xi u - f(\xi,u))]$

Step 1. $\min_\beta \sum_{k=0}^{\nu-1} \beta_k\alpha_k$, such that $\sum_{k=0}^{\nu-1} \beta_k \bar{w}^k = 0$,
$\sum_{k=0}^{\nu-1} \beta_k = 1,\ \beta_k \geq 0,\ k = 0,\dots,\nu-1$
Let $(z^\nu, \theta^\nu) \in \mathbb{R}^{n+1}$ be the associated multipliers

Step 2. For each $\xi \in \Xi$, let

$$u^\nu_\xi \in \operatorname*{argmin}_u \left\{ f(\xi,u) + \tfrac{1}{2\lambda}|z^\nu - u|^2 \right\}$$

$$\widehat{w}^\nu_\xi = \lambda^{-1}(z^\nu - u^\nu_\xi), \quad \bar{w}^\nu = \lambda^{-1}\textstyle\sum_{\xi\in\Xi} p_\xi(z^\nu - u^\nu_\xi)$$

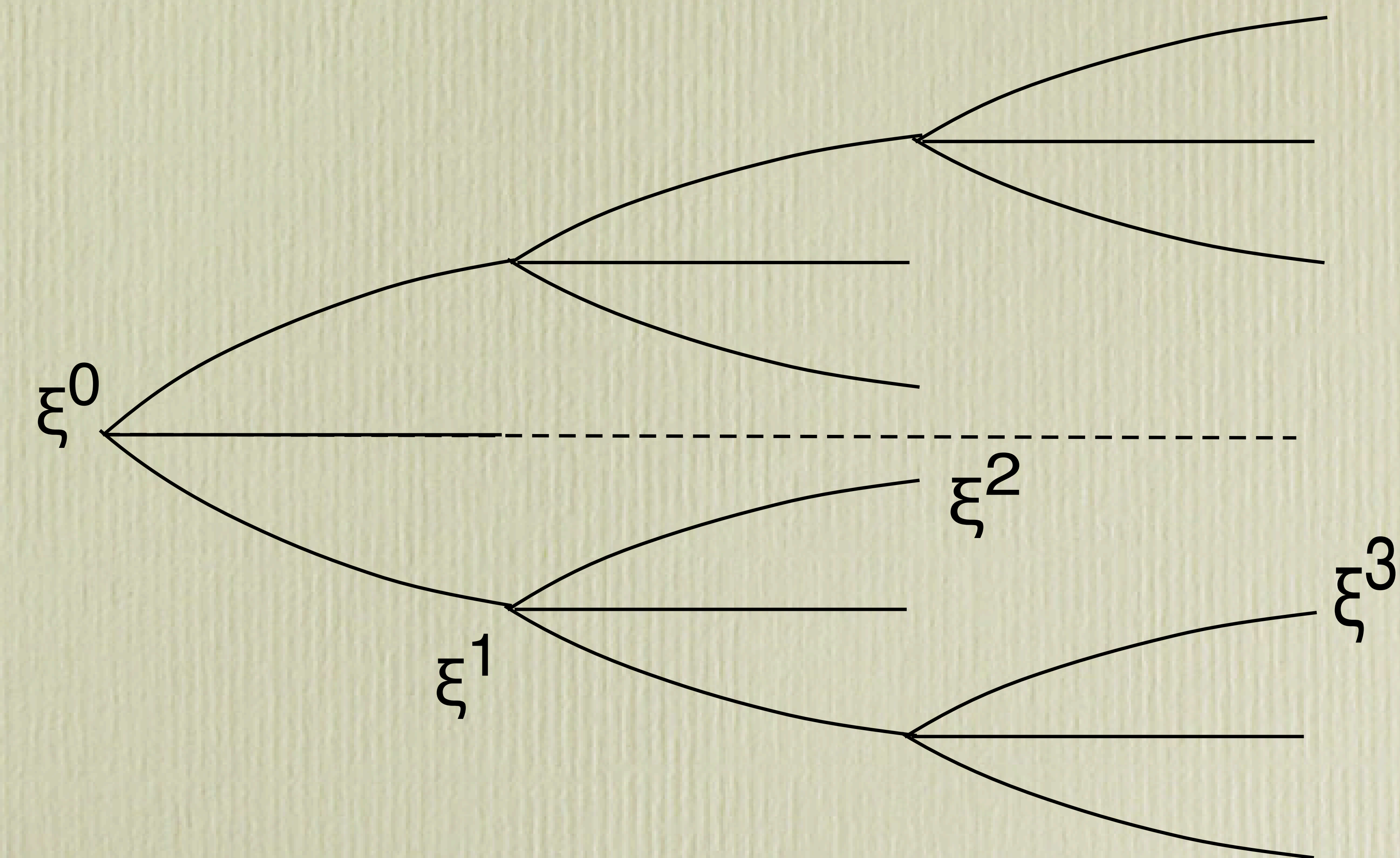
$$\alpha^\nu_\xi = z^\nu \widehat{w}^\nu_\xi - f(\xi, z^\nu - \lambda\widehat{w}^\nu_\xi) - \tfrac{\lambda}{2}|\widehat{w}^\nu_\xi|^2, \quad \alpha_\nu = \textstyle\sum_{\xi\in\Xi} p_\xi \alpha^\nu_\xi$$

Step 3. If $\alpha_\nu < \bar{w}^\nu z^\nu + \theta^\nu$ return to Step 1 with $\nu + 1 \to \nu$
If $\alpha_\nu \geq \bar{w}^\nu z^\nu + \theta^\nu$; z^ν is optimal

Adjust λ if appropriate; always generates bounds for original problem.

Bundling

Discrete Scenario Tree



Algorithm: (1) Nested Sequential SLP
(2) Progressive Hedging + Bundling

Bundling Decomposition

$\min \mathbb{E}\{f(\boldsymbol{\xi}, x)\}, \qquad f + \iota_{C(\xi)}, \quad \xi \in \Xi$

$\iota_C(x) = 0$ when $x \in C$, $= \infty$ otherwise

Ξ discrete or discretization (not based on best approximation of Ξ)

$\{\Xi_k, k = 1, \ldots, K\}$ a partition of Ξ

$p_k = \int_{\Xi_k} P(d\xi),\ k = 1, \ldots, K$

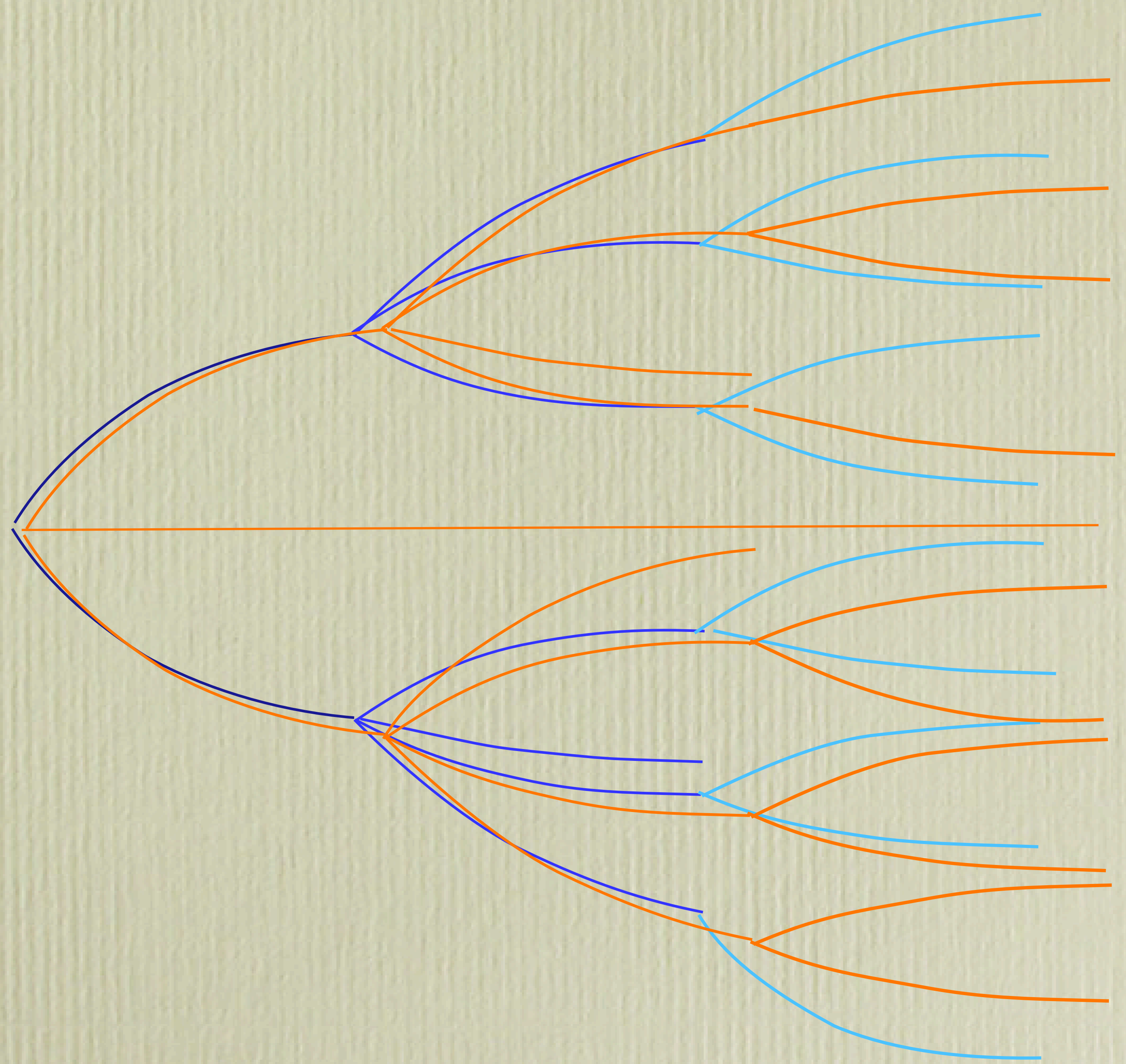
$$\mathbb{E}\{f(\boldsymbol{\xi}, x) = \sum_{k=1}^{K} p_k \mathbb{E}\{f(\boldsymbol{\xi}, x) \,\big|\, \Xi_k\}$$

$$g(k, x) = \mathbb{E}\{f(\boldsymbol{\xi}, x) \,\big|\, \Xi_k\}$$

$$\min \sum_{k=1}^{K} p_k g(k, x)$$

Bayes' Rule

Scenario Tree Decomposition



Ξ_1, Ξ_2

⭐ PH with Bundling ⭐

0. w_k^0 such that $\mathbb{E}\left\{w_k^0\right\} = 0,\ \nu = 0.$ Pick $\rho > 0$

1. for all k : $\widehat{x}^1,\ \widehat{x}^2 = (x_\xi^2, \xi \in \Xi_k)$

$$(\widehat{x}_k^{1,\nu}, x_k^{2,\nu}) \in \arg\min g(k; \widehat{x}^1, \widehat{x}^2) - \langle w_k^\nu, \widehat{x}^1 \rangle$$

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

2. $\overline{x}^{1,\nu} = \sum_{k=1}^K p_k \widehat{x}_k^{1,\nu}$. Stop if $\left|\widehat{x}_k^{1,\nu} - \overline{x}^{1,\nu}\right| = 0$ (approx.)

otherwise $w_k^{\nu+1} = w_k^\nu + \rho\left[\widehat{x}_k^{1,\nu} - \overline{x}^{1,\nu}\right]$, return to 1. with $\nu = \nu + 1$

Convergence: add a proximal term

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

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

Arrow-Debreu model

pure-exchange economy: goods $\in \mathbb{R}^L$, prices $p = \left(p_1, \ldots, p_L\right)$, free disposal

agents: $i \in I$, $|I|$ finite ---- initial holdings: $(e_i, i \in I)$

demand functions: $x_i(p) \in \arg\max \left\{ u_i(x) \,\middle|\, \langle p, x \rangle \leq \langle p, e_i \rangle \right\}$

utility fcn: u_i : dom $u_i = X_i \to \mathbb{R}$, usc, concave $\not\Rightarrow X_i$ closed (but convex)

excess supply function: $s(p) = \sum_{i \in I} \left(e_i - x_i(p)\right)$, market clearing: $s(p) \geq 0$

$$\bar{p} \geq 0 \text{ equilibrium } \Leftrightarrow s(\bar{p}) \geq 0$$

Existence: $x = (x_m, x_g)$, $x_m =$ 'money' allows $p = (1, p_g)$, under

ample survivability: $\left(e_{im}, e_{ig}\right) \Rightarrow \exists\, (\hat{x}_{im}, \hat{x}_{ig}) \in X_i$

such that $\hat{x}_{ig} \leq e_{ig}$, $\hat{x}_{im} < e_{im}$ and $\sum_{i \in I} \hat{x}_{ig} < \sum_{i \in I} e_{ig}$

+ indispensability & unactractiveness

Solution Procedures

Walras' law: $\bar{p} \perp s(\bar{p}) \sim \bar{p}_l s_l(\bar{p}) = 0,\ l = 1,\ldots,L, \qquad s(p) = s(\alpha p)$ for $\alpha > 0$

scaling: $\bar{p} \in \Delta = \left\{ p \in \mathbb{R}^L_+ \middle| \sum_l p_l = 1 \right\}$ since $\forall\, \alpha > 0 : \langle \alpha p, x \rangle \leq \langle \alpha p, e_i \rangle$

one possible way

find $\bar{p}$ $(\in \Delta)$ such that $\quad 0 \leq \bar{p} \perp s(\bar{p}) \geq 0,$

0. (very) special instances: via convex programming

1. tâtonnement, : $\dot{p} = -s(p),\ p(0) = p^0$ (Adam Smith, Léon Walras)

variant: 'Global Newton' (S. Smale) :

$$\nabla s(p)\dot{p} = \lambda s(p),\ \operatorname{sgn}(\lambda) = (-1)^L \operatorname{sgn} \det\big(\nabla s(p)\big)$$

requires s single-valued and differentiable,

$e_i \in \operatorname{int} X_i$ or bdry conditions on s

fails, “in general”

source of doubts about economic equilibrium theory

Solution Procedures

2. simplicial methods (based on "pivoting")

   - Scarf (& Hansen) '73: $\Rightarrow$ find fixed point of $p \mapsto s(p) - p$ in Δ

     partitioning Δ in a simplicial complex, pivoting á la Lemke-Howson

   - piece-wise linear homotopy methods: Eaves '74, Saigal, ...

3. homotopy continuation methods

   - homotopy methods $G(x) = 0$, Yorke *et al*. ('72, '78)

     variants: Kojima, Meggido and Noma for NCP ('89)

     Newton homotopy: Wu ('05), ...

   - 'interior point' homotopy method: Dang and Ye ('11)

a maxinf approach

recall: $s(p) = \sum_{i \in I} \big(e_i - x_i(p)\big)$, market clearing: $s(p) \geq 0$

Walrasian: $W(p,q) = \langle q, s(p) \rangle$, $W : \Delta \times \Delta \to \mathbb{R}$ (a bifunction)

Key observation:

$\bar{p} \in$ maxinf W, $W(\bar{p}, \bullet) \geq 0$ on $\Delta \Rightarrow \bar{p}$ is an equilibrium point.

under insatiability, $\bar{p}$ an equilibrium $\Rightarrow \bar{p} \in$ maxinf W, $W(\bar{p}, \bullet) \geq 0$ on Δ

Moreover:

p_ε : ε-equilibrium point if $\forall l$ (good), $s_l(p_\varepsilon) \geq -\varepsilon$

$p_\varepsilon \in$ ε-maxinf W, $W(p_\varepsilon, \bullet) \geq -\varepsilon$ on $\Delta \Rightarrow p_\varepsilon$ is an ε-equilibrium point.

with insat., p_ε an ε-equilibrium $\Rightarrow p_\varepsilon \in$ ε-maxinf W, $W(p_\varepsilon, \bullet) \geq -\varepsilon$ on Δ

... even more moreover

recall: $\max\left\{u_i(x) \,\middle|\, \langle p,x\rangle \le \langle p,e_i\rangle\right\}$ i-agent problem, $i \in I$

$u_i^\nu \to_{hypo} u_i$, $e_i^\nu \to e_i$ & p_ε^ν equilibrium points and $\varepsilon \searrow 0$

$\Rightarrow$ every cluster point of $\left\{p_\varepsilon^\nu\right\}_{(\nu,\varepsilon)}$ is an equilibrium point!

Consequence of $W_\varepsilon^\nu \to_{\mathrm{lop}} W$ lopsided-convergence

$\left\{C, C^\nu \subset \mathbb{R}^n\right\}$, $\left\{D, D^\nu \subset \mathbb{R}^m\right\}$ $K : C \times D \to \mathbb{R}$, $K^\nu : C^\nu \times D^\nu \to \mathbb{R}$

$K^\nu \to_{\mathrm{lop}} K$ (lopsided convergence) if

multi-hypoconvergence
$\Rightarrow$ lopsided (Gürkan & Pang)

(a) $\forall \left(y \in D, (x^\nu \in C^\nu) \to x \in C\right)$,

$\limsup_\nu K^\nu(x^\nu, y^\nu) \le K(x,y)$ for some $(y^\nu \in D^\nu) \to y \in D$

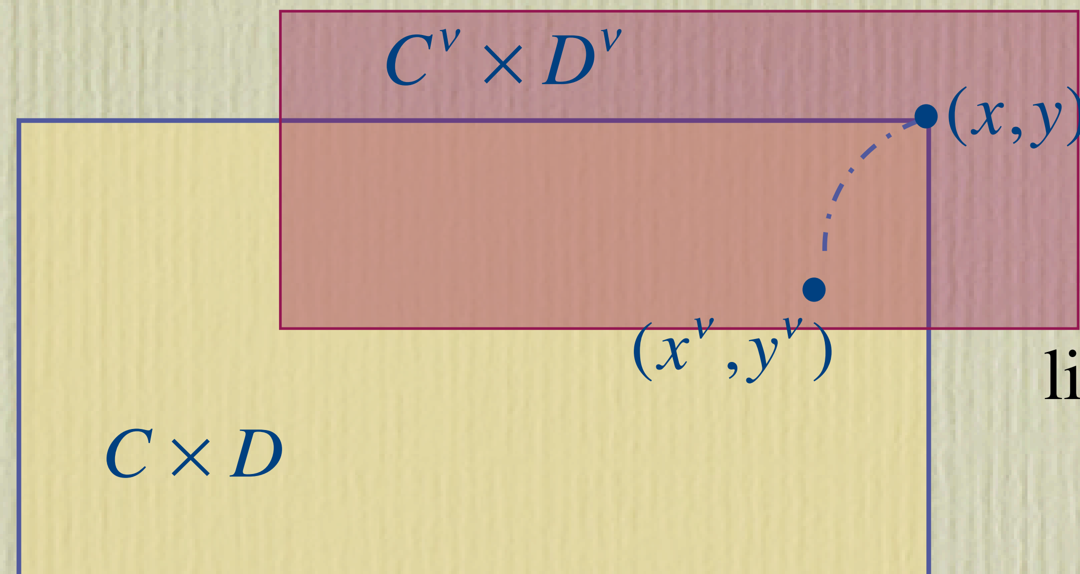
(b) $\forall\, x \in C$, $\exists (x^\nu \in C^\nu) \to x$ such that for any $(y^\nu \in D^\nu) \to y$

$\liminf_\nu K^\nu(x^\nu, y^\nu) \le K(x,y)$ when $y \in D$, $K^\nu(x^\nu, y^\nu) \to \infty$ $y \notin D$

Attouch & Wets '83, Jofré & Wets '09

Lopsided Convergence

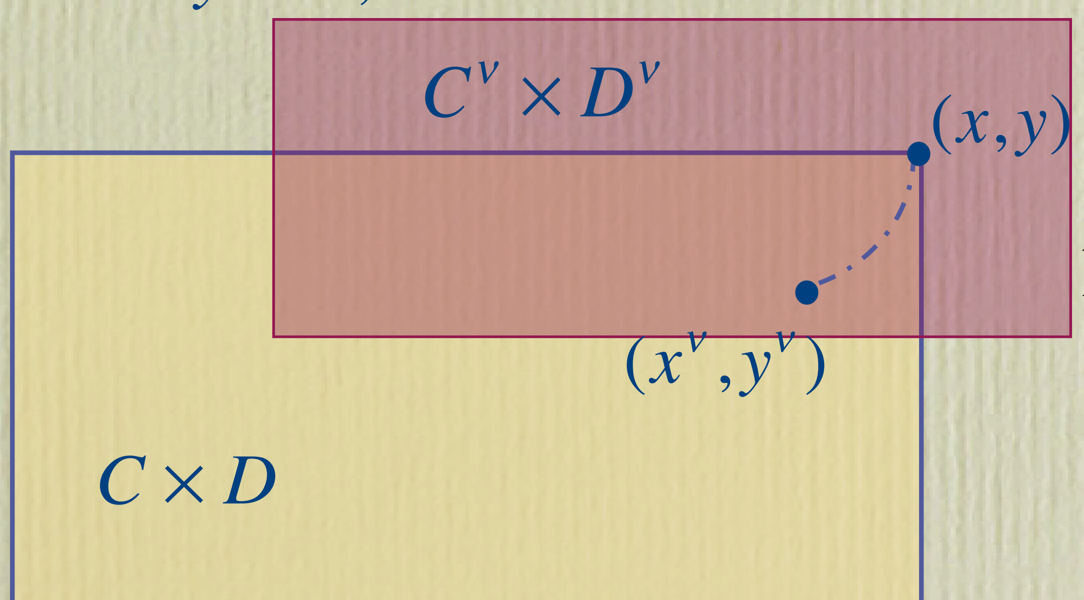
$\exists y^\nu \in D^\nu$
$\to y \in D$



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

$\forall y \in D, \forall x^\nu \in C^\nu \to x$

$\forall y^\nu \in D^\nu$
$\to y$



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

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

$\exists x^\nu \in C^\nu \to x \in C$

The maxinf "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
- solving inclusions (equivalently, generalized equations): $S(x) \ni 0$
- minimal surface problems, ... , mountain pass solutions, ...
- ... and the dynamic versions, and the stochastic (dynamic) versions.

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 $\{x^\nu \in \operatorname{maxinf} K^\nu_{C^\nu \times D^\nu}\}_{\nu\in\mathbb{N}} \Rightarrow \bar{x} \in \operatorname{maxinf} K_{C\times D}$

$K^\nu_{C^\nu \times D^\nu} \xrightarrow[\text{lop ancillary-tight}]{} K_{C\times D}$ if $K^\nu_{C^\nu \times D^\nu} \xrightarrow[lop]{} K_{C\times D}$ 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 $x^\nu \to x$):

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

certainly satisfied when $D = \Delta$ is compact

Convergence of ε-solutions

including $\varepsilon = 0$

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

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

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

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

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

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

tight-lop when $C = \Delta$ & $D = \Delta$ are compact

...back to our Walrasian

$W(p,q) = \langle q, s(p) \rangle$ on $\Delta \times \Delta$, p-usc and q-convex

Augmented Walrasian: σ augmenting function

$$\tilde{W}_r(p,q) = \inf_z \left\{ W(p, q-z) + r *_e \sigma^*(z) \right\} \qquad r\sigma(r^{-1}z)$$

$$= \sup_z \left\{ W(p,z) \,\middle|\, \|z-q\|_\square \le r \right\} \quad \sigma = |\bullet|_\square,\ \iota_{\mathbb{B}} = \sigma^*$$

as $r \to \bar{r} < \infty$, $\boxed{\tilde{W}_r \to_{\text{lop}} \tilde{W}_{\bar{r}} = W} \Rightarrow$ ε-maxinf $W_r \to$ maxinf W

choosing $|\bullet|_\square = |\bullet|_\infty, \mathbb{B} = [-1,1]^L$

or $|\bullet|_\square = |\bullet|_2, \mathbb{B} =$ euclidean unit ball

Friday, June 21, 13

augmented Walrasian strategy

$$W(p,q) = \langle q, s(p)\rangle \text{ on } \Delta \times \Delta,\ \ p\text{-usc and } q\text{-convex}$$

$$\tilde{W}_r(p,q) = \sup_z \left\{ W(p,z) \,\middle|\, \|z-q\|_\square \le r \right\}$$

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

minimizing a linear form on a ball

i.e. finding the largest element of $s(p^k)$

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

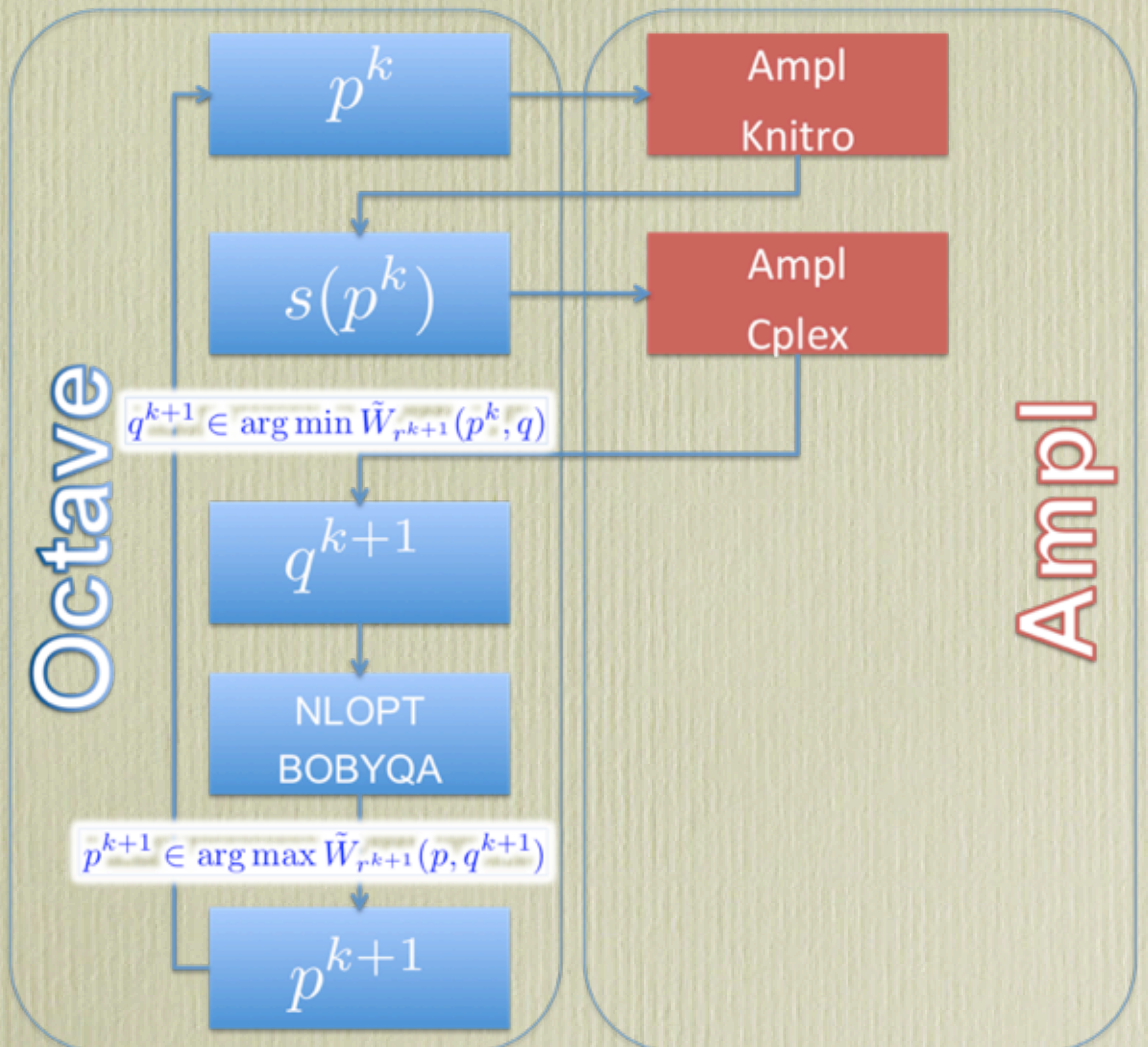
$$\text{as} \quad r_k \nearrow \infty, \quad p^k \to \bar{p}$$

Bagh, Lucero & Wets ≈ '03

first experiments: 10 agents, 150 goods (two blinks)

CMM-implementation

Center for Mathematical Modeling --- Universidad de Chile



Deride,
Jofré
& Wets
'06 -- '12

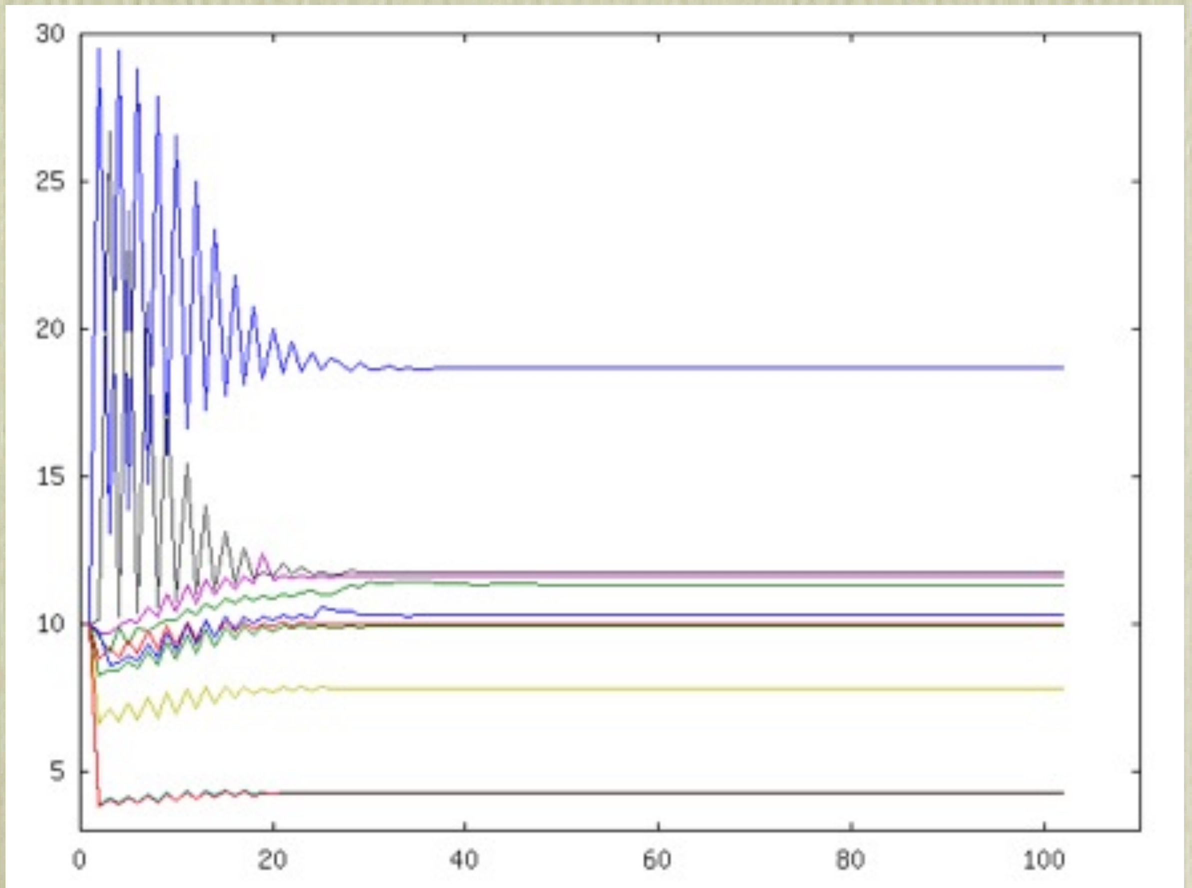
$$\sigma = \frac{1}{2}|\bullet|^2$$

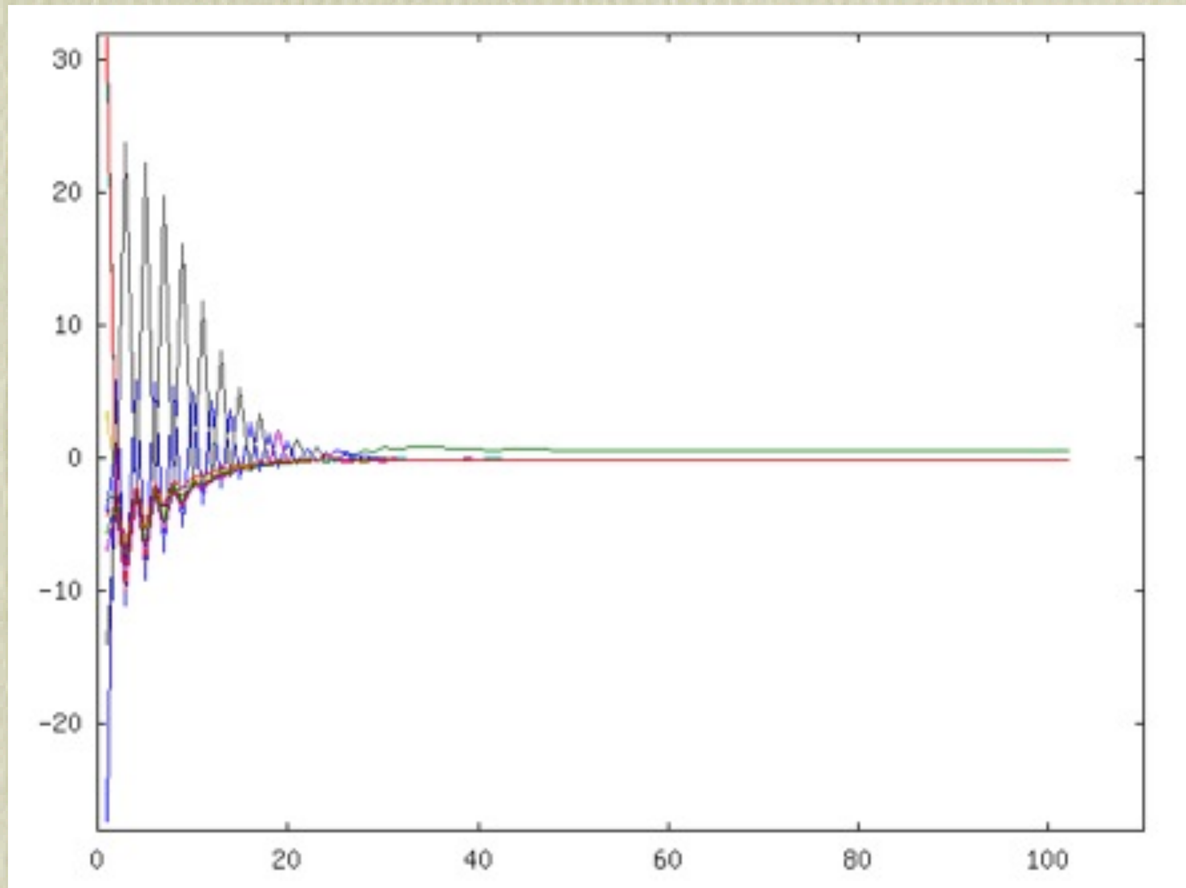
Scarf's example

$$u_i(x) = \left(\sum_{l=1}^{L} (a_{il})^{\beta_i^{-1}} (x_l)^{1-\beta_i^{-1}} \right)^{\beta_i (\beta_i - 1)^{-1}}$$ CES-utility

constant elasticity substitution

$i \in I = 5$ agents, $L = 10$ goods (2000 simplicial pivots)



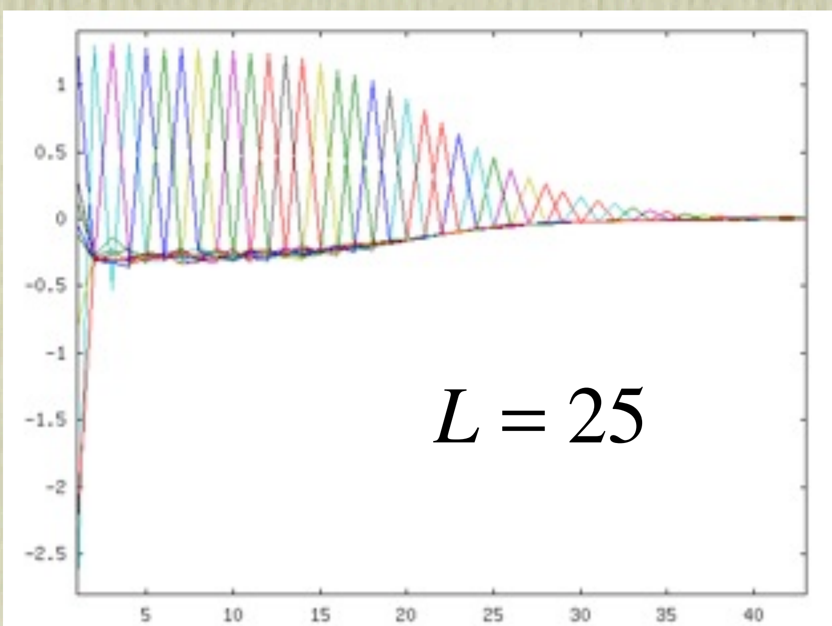


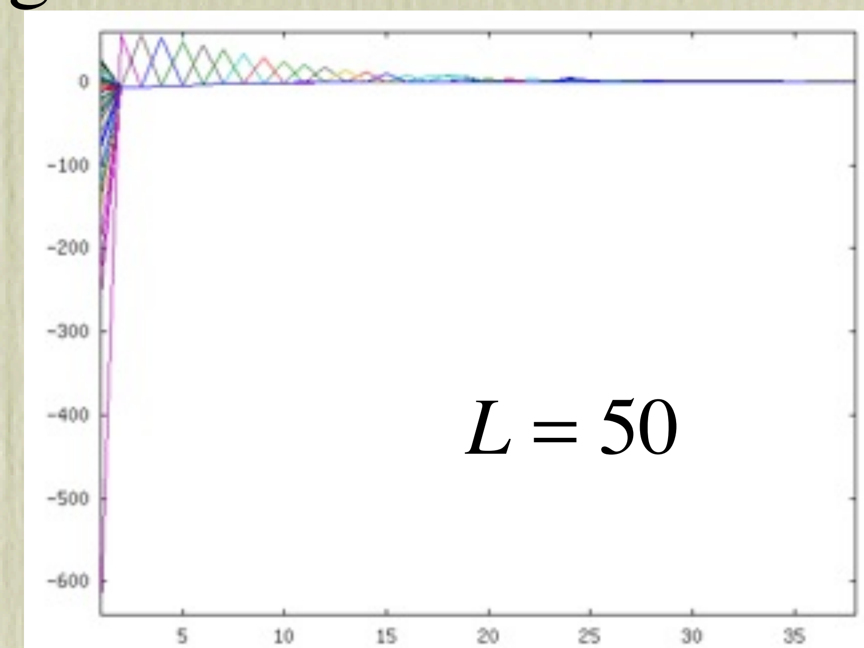
prices and excess supply convergences

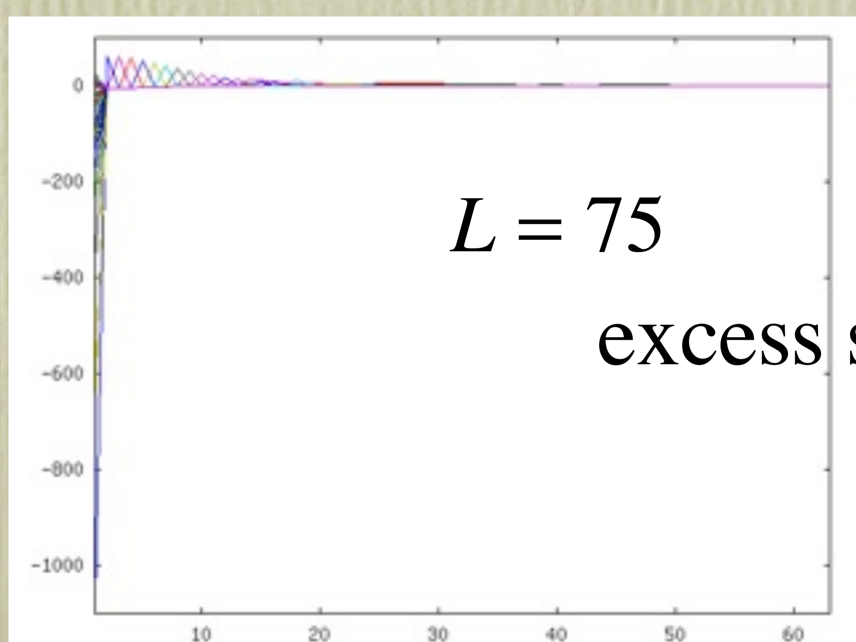
just, ... one more example

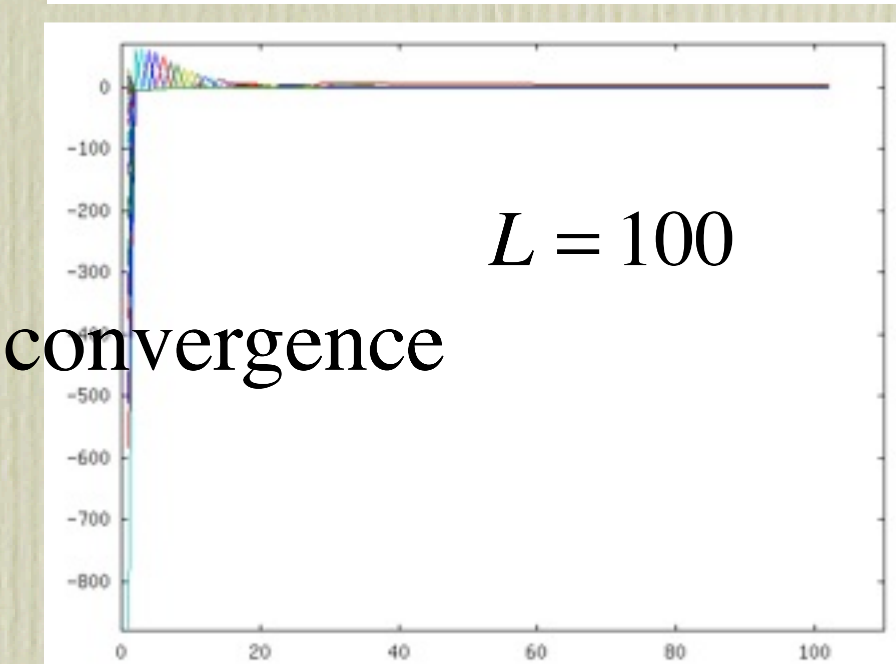
same CES-utility function ($\neq \beta_i$), $I = 10$ agents, $r_k = 1.21^k$

$L =$ # of goods









excess supply convergence

A dynamic model

$$\max_{(x^0, y, x^1)} \quad u_i^0(x^0) + u_i^1(x^1)$$

such that $\left\langle p^0, x^0 + T_i^0 y \right\rangle \le \left\langle p^0, e_i^0 \right\rangle$

$$\left\langle p^1, x^1 \right\rangle \le \left\langle p^1, e_i^1 + T_i^1 y \right\rangle$$

$$x^0 \in X_i^0,\ y \in Y_i,\ x^1 \in X_i^1$$

with solutions: $\left\{ \left(x_i^0(p), y_i(p), x_i^1(p) \right), i \in I \right\},\ p = (p^0, p^1)$

excess supply: $s^0(p) = \sum_{i \in I} \left(e_i^0 - \left(x_t(p) + T_i^0 y_i(p) \right) \right)$

$$s^1(p) = \sum_{i \in I} \left(\left(e_i^1 + T_i^1 y_i(p) \right) - x^1(p) \right)$$

equilibrium: $\bar{p} \in \Delta_L \times \Delta_L$ such that $s^0(\bar{p}) \ge 0, s^1(\bar{p}) \ge 0 \quad (=)$

Walrasian: $W(p,q) = \left\langle (q^0, q^1), \left(s^0(p), s^1(p) \right) \right\rangle$ on $\Delta_L^2 \times \Delta_L^2$

$\Rightarrow$ augmented Walrasian, ...

Exploiting separability

$$r_i(y,p) = \sup_{x^0 \in X_i^0}\left[u_i^0(x^0) \middle| \left\langle p^0, x^0 + T_i^0 y \right\rangle \le \left\langle p^0, e_i^0 \right\rangle \right]$$

$$+ \sup_{x^1 \in X_i^1}\left[u_i^1(x^1) \middle| \left\langle p^1, x^1 \right\rangle \le \left\langle p^1, e_i^1 + T_i^1 y \right\rangle \right]$$

i-agent problem: find $y_i(p) \in \mathbb{R}_+^{n_i}$ that maximizes $\left\{ r_i(y) \middle| T_i^0 y \le e_i^0 \right\}$

Example: $u_i^{0,1}$ are of Cobb-Douglas type: $u_i(x) = \prod_{l=1}^{L} x_l^{\beta_l}, \beta \in \Delta$

$$x_l^0(p^0, y) = \frac{\beta_l}{p_l^0} \sum_{k=1}^{L} p_k^0 \left(e_k^0 - \left\langle (T_i^0)_k, y \right\rangle \right),\ l = 1,\ldots,L, \quad X_i^0 = \mathbb{R}_+^L$$

$$x_l^1(p^1, y) = \frac{\beta_l}{p_l^1} \sum_{k=1}^{L} p_k^1 \left(e_k^1 + \left\langle (T_i^1)_k, y \right\rangle \right),\ l = 1,\ldots,L, \quad X_i^0 = \mathbb{R}_+^L$$

substituting $\Rightarrow r_i$ linear in y: i-agent's problem is a linear program!

substantial gain in processing time

Stochastic Environment

a 'minimal' model

Chap.7 -- Theory of Value G. Debreu, '59

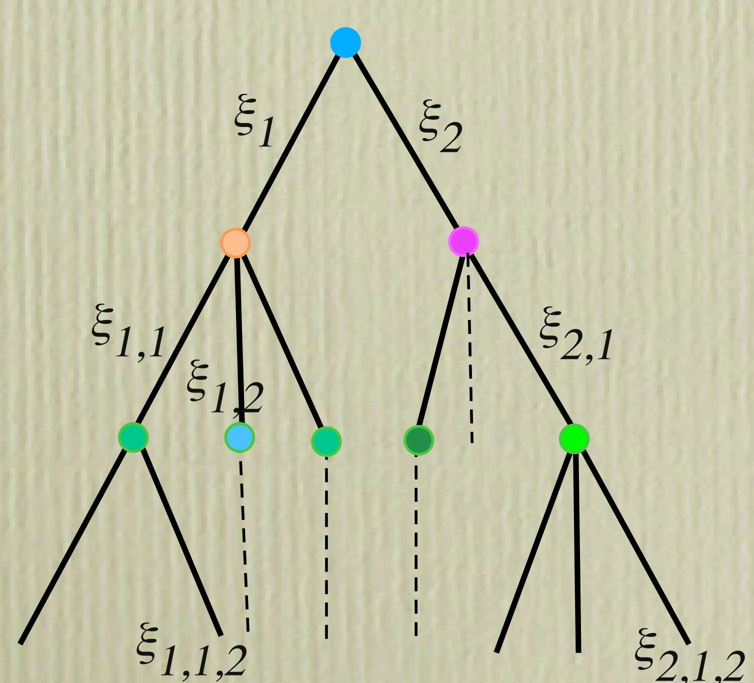
$$\max_{x^0, x^1_{\cdot} \in \mathcal{M}} E\left\{u_i(x^0, x^1_{\xi^1}, x^2_{\xi^1,\xi^2}, \ldots)\right\}$$ *i*-agent

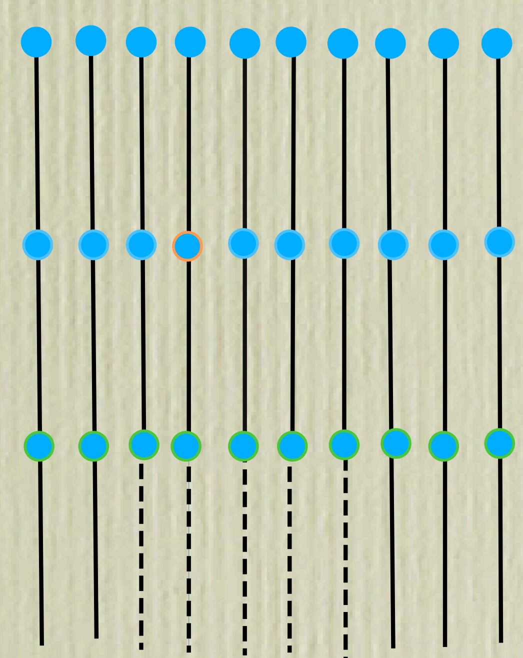
such that $\left\langle p^t_\xi, \sum_{\tau \le t}\left(e^\tau_{\xi^1,\ldots\xi^\tau} - x^\tau_{\xi^1,\ldots\xi^\tau}\right)\right\rangle \ge 0,\ \forall \xi = (\xi^1, \xi^2, \ldots),\ \ \mathrm{t} = 0,1,\ldots$

$$\sum_{i \in I}\left(\sum_{\tau \le t}\left(e^\tau_{\xi^1,\ldots\xi^\tau} - x^\tau_{\xi^1,\ldots\xi^\tau}\right)\right) \ge 0, \forall \xi, t$$ Clearing the market

Key Assumption (via K. Arrow): all contingencies available at time 0

$\Rightarrow$ complete market, i.e., all ξ's can be dealt with separately





$\exists\ \left(p_\xi = (p^0_\xi, p^1_\xi, \ldots)\right)$

$(\forall\ \xi)$ equilibrium prices.

i-Agent: stochastic case

an engineer's viewpoint?

$$\max_{x^0, y, x^1_\bullet \in \mathcal{M}} \; u_i^0(x^0) + E^i\left\{u_i^1(\xi; x_\xi^1)\right\}$$

$$\text{so that } \left\langle p^0, x^0 + T_i^0 y \right\rangle \le \left\langle p^0, e_i^0 \right\rangle$$

$$\left\langle p_\xi^1, x_\xi^1 \right\rangle \le \left\langle p_\xi^1, e_{i,\xi}^1 + T_{i,\xi}^1 y \right\rangle, \quad \forall \xi \in \Xi$$

$$x^0 \in X_i^0,\; y \in \mathbb{R}_+^{n_i},\; x_\xi^1 \in X_{i,\xi}^1, \quad \forall \xi \in \Xi$$

- $E^i\{.\}$ expectation w.r.t. i-agent beliefs

Stochastic program with recourse: 2-stage
Well-developed solution procedures
Well-developed "Approximation Theory"

Simplest-classical assumptions

Ξ finite (support)

$u_i^0 : X_i^0 \to \mathbb{R}, \quad \forall \xi \in \Xi,\ u_i^1(\xi,\bullet) : X_{i,\xi}^1 \to \mathbb{R}$ concave

continuous, numerical experiments: differentiable

$T_{\mathrm{i}}^0, T_{i,\xi}^1$: input-ouput matrices

(savings, production, investment, etc.)

$X_{\mathrm{i}}^0, X_{i,\xi}^1$: closed, convex, non-empty interior (survival sets

$e_{\mathrm{i}}^0 \in \operatorname{int} X_{\mathrm{i}}^0,\ e_{i,\xi}^1 \in \operatorname{int} X_{i,\xi}^1$ for all ξ (or as on first slide)

Market Clearing

Agents: $i \in I,\ |I|$ finite ("large"), $\quad p = \left(p^0, (p^1_\xi, \xi \in \Xi)\right)$

$$\left(\overline{x}_i^{\,0}(p), \overline{y}_i(p), \left\{\overline{x}^1_{i,\xi}(p)\right\}_{\xi\in\Xi}\right) \in \arg\max\left\{i\text{-agent problem}\right\}$$

excess supply:

$$\sum_{i\in I}\left(e_i^0 - (\overline{x}_i^{\,0}(p) + T_i^{\,0}\overline{y}_i(p))\right) = s^0(p^0, \{p^1_\xi\}_{\{\xi\in\Xi\}}) \geq\ 0$$

$\forall \xi \in \Xi:$

$$\sum_{i\in I}\left(e^1_{i,\xi} + T^1_{i,\xi}\overline{y}_i(p) - \overline{x}^1_{i,\xi}(p)\right) = s^1_\xi(p^0, \{p^1_\xi\}_{\{\xi\in\Xi\}}) \geq\ 0$$

Existence: via Ky Fan Inequality

$$W(p_\bullet, q_\bullet) = \langle q_\bullet, s(p_\bullet) \rangle$$

$$= \left\langle (q^0, \{q^1_\xi\}_{\xi\in\Xi}), \left(s^0(p^0, \{p^1_\xi\}_{\xi\in\Xi}), \left\{ s^1_\xi(p^0, \{p^1_\xi\}_{\xi\in\Xi}) \right\}_{\xi\in\Xi} \right) \right\rangle$$

$W : \prod_{1+|\Xi|} \Delta \times \prod_{1+|\Xi|} \Delta \to \mathbb{R}$ a Ky Fan function: usc, convex

linear w.r.t. $q_\bullet$, continuous w.r.t. $p_\bullet$

and also $W(p_\bullet, p_\bullet) \geq 0$.

provided $s(\bullet)$ continuous w.r.t. $p_\bullet$

another lecture series

Incomplete → '*i*-Complete' Market

$\forall \xi \in \Xi$ (separately),

i-agent's problem:

$$\left(x_i^0, y_i, x_{i,\xi}^1\right) \in \arg\max \left\{ u_i^{w_{i,\xi}}\left(\xi; x^0, y, x^1\right) \text{ on } \widehat{C}_{i,\xi}(p^0, p_\xi^1) \right\}$$

for $\{w_{i,\xi}\}_{\xi \in \Xi}$ associated with (p^0, p_ξ^1)

clearing the market:

$$s^0(p^0, p_\xi^1) \geq 0, \quad s_\xi^1(p^0, p_\xi^1) \geq 0$$

Arrow-Debreu 'stochastic' equilibrium problem

Cobb-Douglas utilities

$$u_i^t(x) = \prod_{l=1}^{L} x_l^{\beta_i^t}, \; \beta^t \in \Delta, \; t = 0,1$$

"agent's optimization" (skipping i)

$$r(\xi; y, p) = \alpha^0(p^0)\Big(\sum_{k=1}^{n} p_k^0\big(e_k^0 - \langle T_k^0, y\rangle\big)\Big)$$

taking advantage of separability

$$+\, \alpha^1(p_\xi^1)\Big(\sum_{k=1}^{n} p_{k,\xi}^1\big(e_{k,\xi}^1 - \langle (T_\xi^1)_k, y\rangle\big)\Big), \qquad \alpha(p) = \prod_{l=1}^{L}\left(\frac{\beta_l}{p_l}\right)^{\beta_l}$$

$$y_\xi^{\nu+1} \in \arg\max\left\{ r^\nu(\xi; y) - \langle w^\nu, y\rangle - \frac{\rho}{2}\left|y - \bar{y}^\nu\right|^2 \;\middle|\; T^0 y \le e^0, u \in \mathbb{R}_+^n \right\}$$

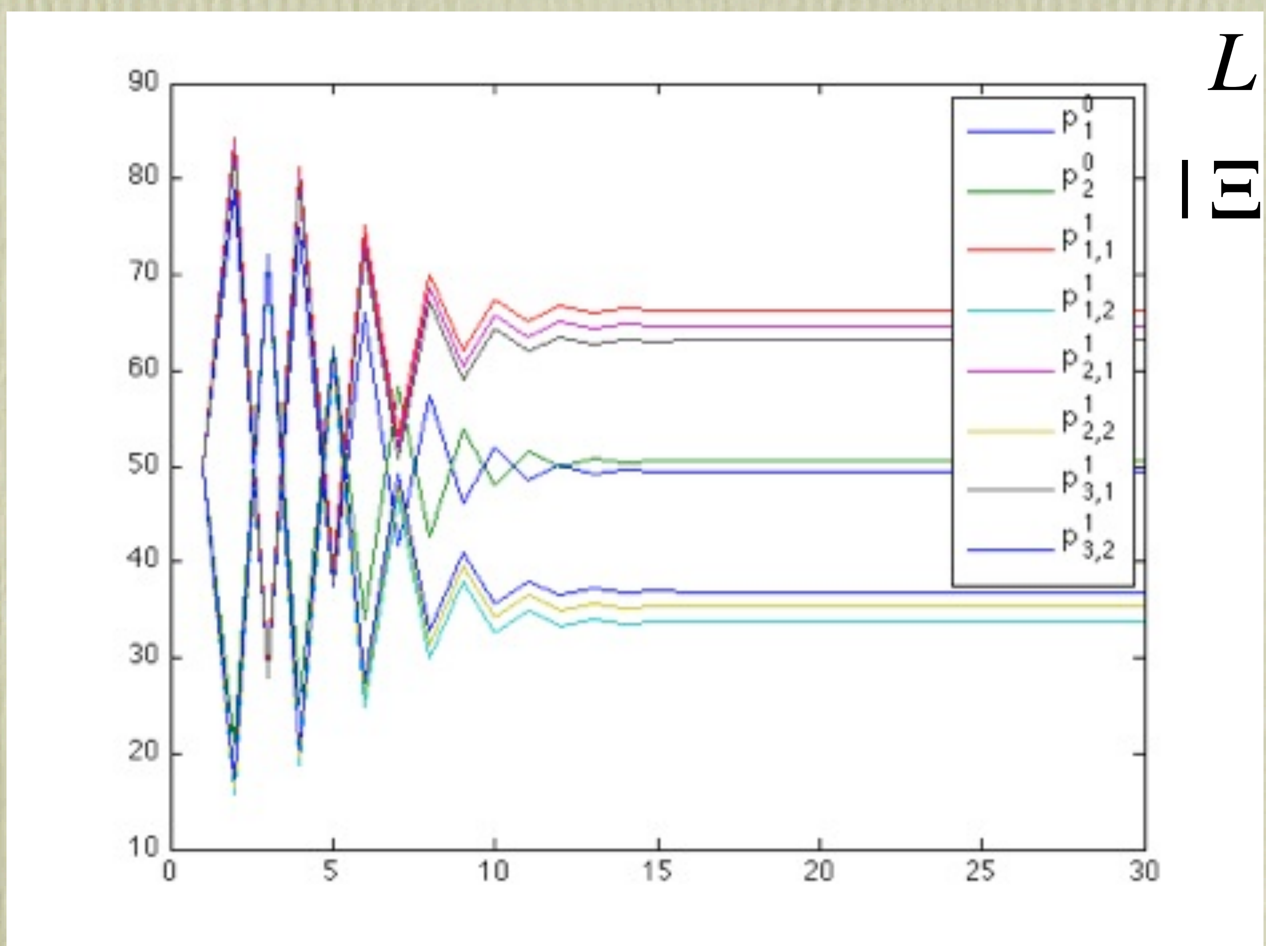
"outer loop", calculating $\big(p^0, (p_\xi^1, \xi \in \Xi)\big)$

augmented Walrasian

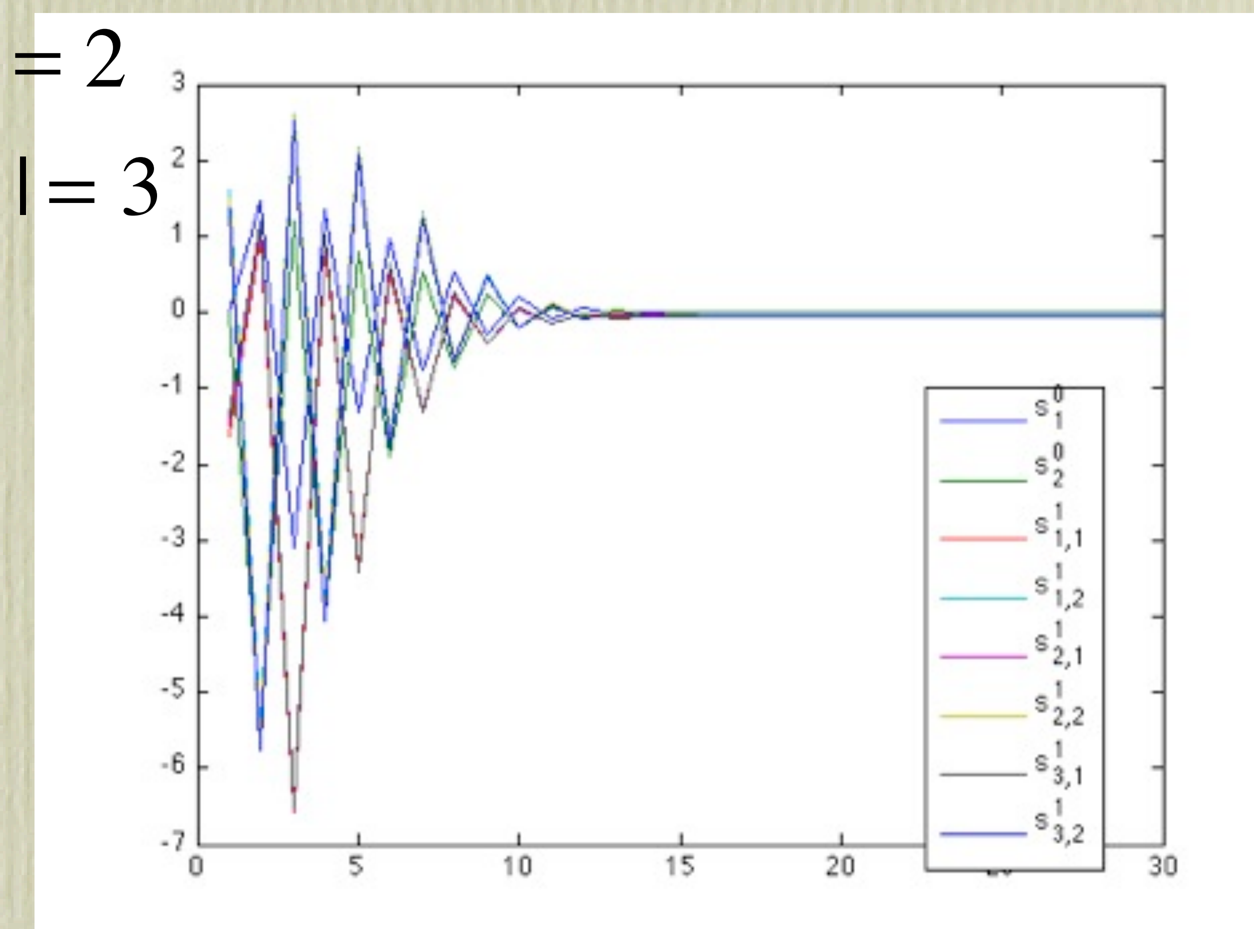
Convergence: exploiting separability

$L = 2$

$|\Xi| = 3$



prices



excess supply

4. Variational Inequality

to be dealt with <u>now</u> in glorious detail

back ... Arrow-Debreu model

i-agent demand: $x_i(p) \in \arg\max \left\{ u_i(x) \,\middle|\, \langle p,x \rangle \le \langle p,e_i \rangle \right\}$

$u_i = X_i \to \mathbb{R}^L$, usc, concave $\not\Rightarrow X_i$ closed (but convex)

excess supply: $s(p) = \sum_{i\in I} \big(e_i - x_i(p)\big)$, market clearing: $s(p) \ge 0$

(under ample survivability, indispensability, unactractiveness)

i-agent optimal $x_i(p) \iff \exists \lambda_i \ge 0$ such that

$\lambda_i \perp \langle p, e_i - x_i(p) \rangle, \quad \langle p, e_i - x_i(p) \rangle \ge 0,$

$x_i(p) \in \arg\max_{x\in X_i} u_i(x) - \lambda_i \langle p, x \rangle$

λ_i utility scaling

when $X_i = \mathbb{R}^L_+, u_i$ smooth:

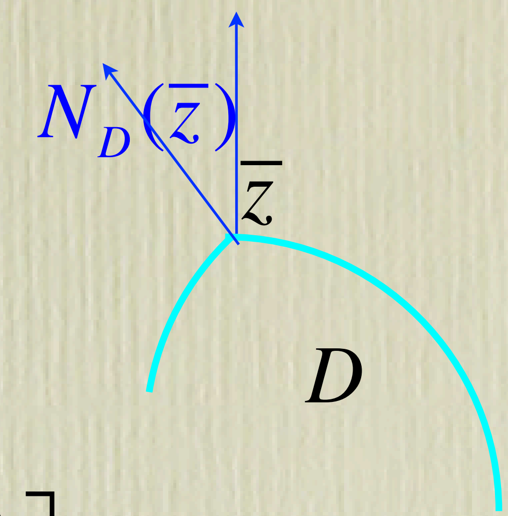
$0 \le x_i(p) \perp \lambda_i p - \nabla u_i(x_i(p) \ge 0$

via Variational Inequality

$$x_i(p) \in \arg\max\left\{u_i(x) \,\middle|\, \langle p,x\rangle \le \langle p,e_i\rangle, x \in X_i\right\}$$

$$\Updownarrow \quad \sum_i \left(e_i - x_i(p)\right) = s(p) \ge 0$$

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



$$G\left(p,(x_i),(\lambda_i)\right) = \left[\sum_i (e_i - x_i); \left(\lambda_i p - \nabla u_i(x_i)\right); \langle p, e_i - x_i\rangle\right]$$

$$D = \Delta \times \left(\prod_i X_i\right) \times \left(\prod_i \mathbb{R}_+\right)$$

$$\boxed{-G\left(\bar{p},(x_i),(\lambda_i)\right) \in N_D\left(\bar{p},(x_i),(\lambda_i)\right)}$$

suggests
LCP, NCP
J.-S. Pang,
M.Ferris, ..

D unbounded $\rightarrow \hat{D}$ bounded

geometric V.I.

via smoothing: L. Qi, X.Chen, ...

adding Firms (a parenthesis)

i-agent: $e_i \to e_i + \theta_{ij} z_j$ share θ_{ij} of production $z_j \in Z_j$ of firm $j \in J$

$$\theta_{ij} \geq 0,\ \sum_{i \in I} \theta_{ij} = 1 \quad j = 1, \ldots, J$$

$$z_j(p) \in \arg\max \left\{ \langle p, z \rangle \,\middle|\, z \in Z_j \right\}$$

excess supply: $\sum_{i \in I} \big(e_i - x_i(p)\big) + \sum_{j \in J} z_j(p) \geq 0$ (equilibrium)

functional V.I. $\boxed{-G\big(\bar{p}, (\bar{x}_i), (\bar{\lambda}_i), (\bar{z}_j)\big) \in \partial f\big(\bar{p}, (\bar{x}_i), (\bar{\lambda}_i), (\bar{z}_j)\big)}$

$$f\big(p, (x_i), (\lambda_i), (z_j)\big) = \iota_{\mathbb{R}^L_+}(p) - \sum_{i \in I} u_i(x_i) + \sum_{i \in I} \iota_{\mathbb{R}_+}(\lambda_i) + \sum_{j \in J} \iota_{Z_j}(z_j)$$

f convex $\Rightarrow \partial f$ monotone operator (yields existence of solution)

$$G\big(p, (x_i), (\lambda_i), (z_j)\big)$$

$$= \left[\sum_i (e_i - x_i) + \sum_j z_j; (\lambda_i p); \left(\left\langle p, e_i + \sum_j \theta_{ij} z_j - x_i \right\rangle \right); (-p_j) \right]$$

Jofré, Rockafellar & Wets '07

Path Solver .. (M.Ferris, D.Ralph et al)

$$-G(\bar{z}) \in N_D(\bar{z}), \quad \bar{z} = (\bar{p},(\bar{x}_i),(\bar{\lambda}_i))$$

$$D = \Delta \times \left(\prod_i X_i\right) \times \left(\prod_i \mathbb{R}_+\right) = \left\{ z \middle| Az \geq b \right\}$$

Complementarity problem:

$$-G(z) = A^T y, \ \ y \geq 0, \ \ Az - b \perp y$$

with $K = \mathbb{R}^N \times \mathbb{R}^M_+$:

$$\boxed{(z,y) \in K, \ \ H(z,y) \in -K^*, \ \ (z,y) \perp H(z,y)}$$

$$H(z,y) = \begin{bmatrix} G(z) + A^T y \\ Az \end{bmatrix} - \begin{pmatrix} 0 \\ b \end{pmatrix}$$

Equivalent nonsmooth mapping

- $0 = H(\mathrm{prj}_K(z, y)) + (z, y) - \mathrm{prj}_K(z, y)$
  - with simplified K

    (CP) $0 \le x \perp F(x) \ge 0$ Complementarity Problem

    (NS) $0 = F(x_+) + x - x_+$ Nonlinear system

- $\bar{x}$ sol'n (CP) $\Rightarrow$ $\tilde{x}$ sol'n (NS):

  $\tilde{x}_k = \bar{x}_k$ if $F_k(\bar{x}) = 0$, $\tilde{x}_k = -F_k(\bar{x}_k)$ if $F_k(\bar{x}) > 0$
- $\tilde{x}$ sol'n (NS) $\Rightarrow \tilde{x}_+$ sol'n (CP):

  $\tilde{x}_+ \ge 0, F(\tilde{x}_+) = \tilde{x}_+ - \tilde{x} \ge 0$ & $\tilde{x}_+ \perp \tilde{x}_+ - \tilde{x}$

PATH Solver:

$x = (z, y),\ x_+ = \mathrm{prj}_K(x, y)$

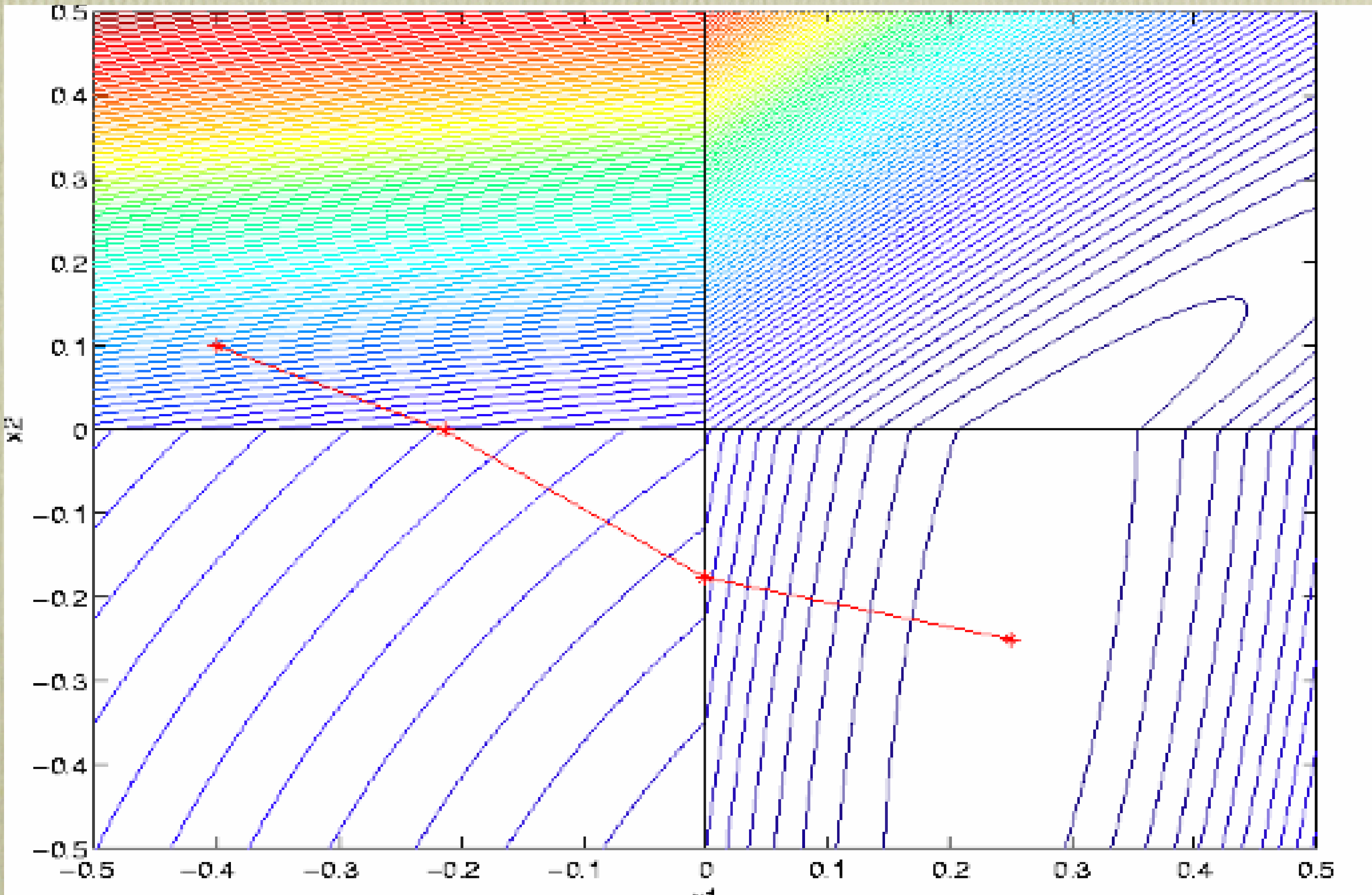
- PATH: Newton method based on nonsmooth normal mapping:

$$H(x_+) + x - x_+$$

- Newton point: solution of piecewise linearization:

$$H(x_+^k) + \left\langle \nabla H(x_+^k), x_+ - x_+^k \right\rangle + x - x_+ = 0$$

The “Newton” step



V.I.-Extensive Formulation

discrete distribution: $|\Xi|$ finite

$$G\Big(\big(p^0, p^1_{\xi\in\Xi}\big), \big(x^0_i, x^1_{i,\xi\in\Xi}\big)_{i\in I}, \big(\lambda^0_i, \lambda^1_{i,\xi\in\Xi}\big)_{i\in I}\Big) =$$

$$\Big[\Big(\sum_i (e^0_i - x^0_i), \sum_i (e^1_{i,\xi} - x^1_{i,\xi})_{\xi\in\Xi}\Big); \Big(\big(\lambda^0_i p^0 - \nabla u^0_i(x^0_i)\big)_{i\in I}, \big(\lambda^1_{i,\xi} p^1_\xi - \nabla u^1_i(\xi; x^1_{i,\xi})\big)_{i\in I,\xi\in\Xi}\Big); \Big(\langle p^0, e^0_i - x^0_i\rangle, \big(\langle p^1_\xi, e^1_{i,\xi} - x^1_{i,\xi}\rangle\big)_{i\in I,\xi\in\Xi}\Big)\Big]$$

$$D = \Big(\Delta \times \prod_{\xi\in\Xi} \Delta\Big) \times \Big(\prod_i X^0_i \times \big(\prod_i X^1_{i,\xi}\big)_{\xi\in\Xi}\Big) \times \Big(\prod_i \mathbb{R}_+ \times \big(\prod_i \mathbb{R}_+\big)_{\xi\in\Xi}\Big)$$

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

$$z = \big(p^0, p^1_{\xi\in\Xi}\big), \big(x^0_i, x^1_{i,\xi\in\Xi}\big)_{i\in I}, \big(\lambda^0_i, \lambda^1_{i,\xi\in\Xi}\big)_{i\in I}$$

"Thanks the gods (& M. Ferris) for EMP"

for a special VI handled via smoothing/sampling: later

so, let's go: PATH Solver

- Economy: (8 goods), 5 types of agents
  - Skilled & unskilled workers
  - Businesses: Basic goods & leisure
  - Banker: bonds (riskless), 2 stocks
- small # of scenarios 280,
- utilities: CES-functions (gen. Cobb-Douglas)
  - Utility in stage 2 assigned to financial instruments

unfortunately, ... PATH Solver lets us down

Back to the drawing board

Disaggregation!

initialization: $\forall\, \mathrm{i} \in I$

$w_{i,\xi}$ such that $E\left\{w_{i,\xi}\right\} = 0$

$\forall \xi \;\Rightarrow\;$ equilibrium with $\left\{u_i^{w_{i,\xi}}(\xi;\bullet)\right\}_{i\in I}$

yields: $(\bar{x}^0_{i,\xi}, \bar{y}_{i,\xi}, \bar{x}^1_{i,\xi}),\quad (\bar{p}^0_{i,\xi}, \bar{p}^1_{i,\xi})$

STOP $\forall i \in I : \left(\bar{x}^0_{i,\xi}, \bar{y}_{i,\xi}\right) \simeq E\left\{\left(\bar{x}^0_{i,\xi}, \bar{y}_{i,\xi}\right)\right\}$

otherwise adjust $\left[w_{i,\xi}, \xi \in \Xi\right]_{i\in I}$

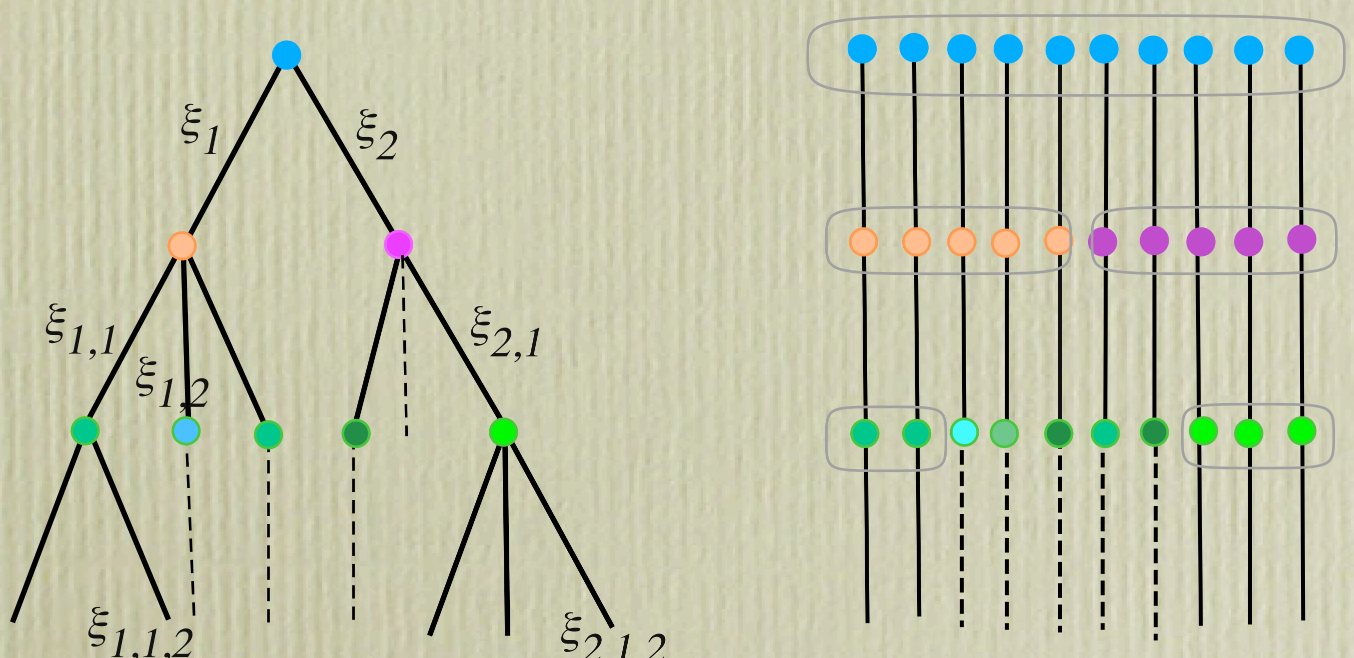
stochastic-pure exchange

for each $\xi \in \Xi$: the equilibrium problem, I agents with

i-agent's problem: $\max \left\{ u_i^{w_{i,\xi}} \left(\xi; x^0, y, x^1 \right) \text{ on } \widehat{C}_{i,\xi}(p^0, p_\xi^1) \right\}$

$$u_i^{w_{i,\xi}} = u_i^0(x^0) - \langle w_{i,\xi}, (x^0, y) \rangle - \frac{\rho}{2} \|(x^0, y) - (\bar{x}^0, \bar{y})\|^2 + u_i^1(\xi; x^1)$$

clearing the market: $s^0(p^0, p_\xi^1) \geq 0, \;\; s_\xi^1(p^0, p_\xi^1) \geq 0$



but now with $w_{i,\xi} \sim$ constraints

$x^0(\xi) \equiv x^0$ constant

$x^1(\xi | \xi^1) \equiv x_1^1, x^1(\xi | \xi^2) \equiv x_2^1$

$\ldots$

Disaggregation with PATH Solver

- Economy: (5 agents - 8 goods)
  - Skilled & unskilled workers
  - Businesses: Basic goods & leisure
  - Banker: bonds (riskless), 2 stocks

on M. Ferris
semi-slow laptop
using EMP-package
4 min + 2 min for
verification

- 2-stages, solved under # of scenarios (280)
- utilities: CES-functions (gen. Cobb-Douglas)
  - Utility in stage 2 assigned to financial instruments
  - Financial instruments only used for transfer to time 1
- used for calibration (-> stochastic model)

numerically: `blink' (5000 iterations).

Ja! scenario disaggregation, but ...

i-agent: $x_i(p) \in \arg\max \left\{ u_i(x) \,\middle|\, \langle p, x \rangle \le \langle p, e_i \rangle \right\},\ i \in I$

with excess supply $s(p)$: $\quad 0 \le p \perp s(p) \ge 0$

Multi-Optimization Problem with Equilibirum Constraint

MOPEC-class ~ maxinf family

$$x_i \in \arg\max_{x \in \mathbb{R}^{n_i}} f_i(p, x, x_{-i}), \quad i \in I, \quad x_I = (x_i, i \in I)$$

$$D(p, x_I) \in \partial g(p) \quad [\text{or } \in N_C(p)]$$

with Michael Ferris '11-'?? ... '05?

Examples: Walras, noncooperative games, ...
stochastic (dynamic): decentralized electricity markets,
joint estimation and optimization, financial equilibrium, ...

Contracts (Assets)

assets (= contract types): $k \in K$, $|K|$ finite

$z_i = z_i^+ - z_i^- = (z_i^1, \ldots, z_i^K)$ assets 'acquired' by i-agent

q_k market price of asset k

D_ξ^k bundle of goods 'delivered' by one unit of asset k

budgetary constraints:

$$\left\langle p^0, x^0 + T_i^0 y \right\rangle + \langle q, z \rangle \leq \left\langle p^0, e_i^0 \right\rangle$$

$$\left\langle p_\xi^1, x_\xi^1 \right\rangle \leq \left\langle p_\xi^1, e_{i,\xi}^1 + T_{i,\xi}^1 y + D_\xi z \right\rangle \quad \forall \xi \in \Xi$$

clearing the market:

$$s^0\left(p^0, \left(p_\xi^1\right)_{\xi \in \Xi}, q\right) \geq 0, \; s^1\left(\xi; p^0, \left(p_\xi^1\right)_{\xi \in \Xi}, q\right) \geq 0 \; \forall \xi, \; \boxed{\sum_{i \in I} z_i = 0}$$

The BDE-example

Brown-DeMarzo-Eaves (Econometrica '96)

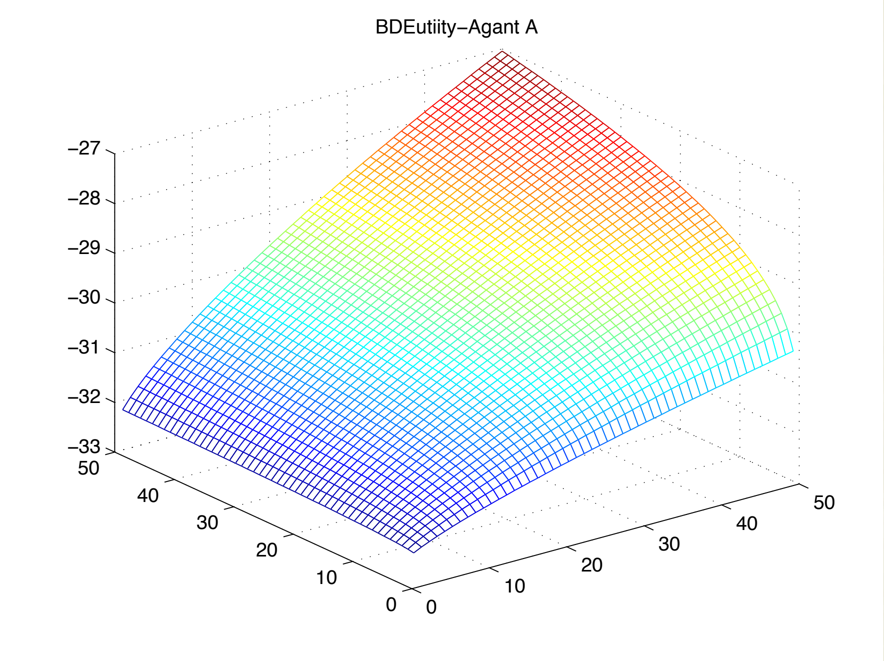
3 agents (2-agent & 3-agent of the same type)

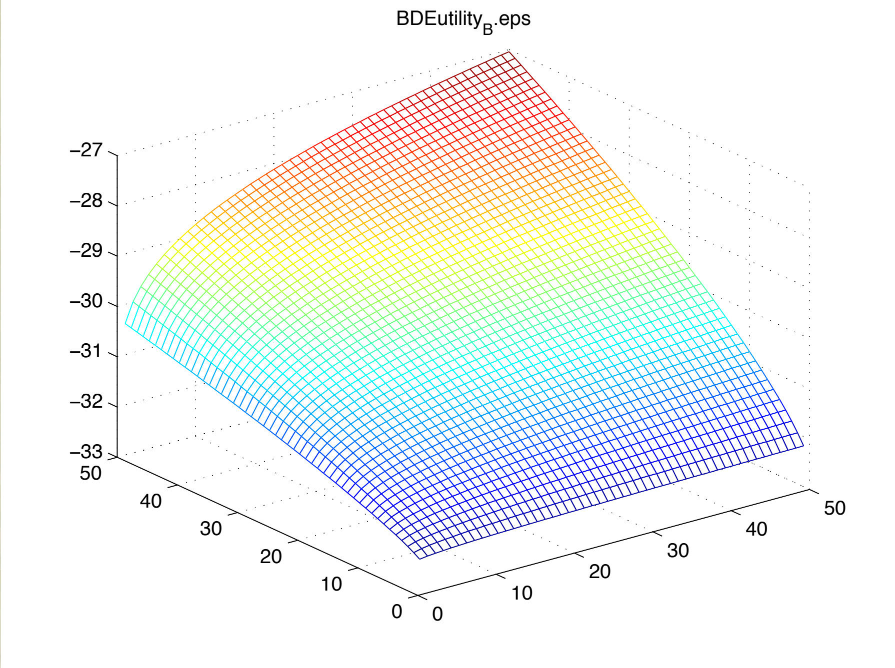
2 goods, $|\Xi| = 3$ (future states), no y-activities

$$u_i^1(\xi; x) = -\left(5.7 - \prod_{l=1}^{2} (x_l)^{\alpha_{i,l}}\right) = u_1^0(x)$$

$$\alpha_1 = (0.25, 0.75),\ \alpha_{2\&3} = (0.75, 0.25)$$

asset #1: $D_\xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, asset #2: $D_\xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ for all ξ





The BDE-example

3 agents, 2 goods, $|\Xi| = 3$, no y-activities

$$u_i^1(\xi; x) = -\left(5.7 - \prod_{l=1}^{2} (x_l)^{\alpha_{i,l}}\right) = u_1^0(x), \quad \alpha_1 = (0.25, 0.75), \ \alpha_{2\,\&\,3} = (0.75, 0.25)$$

asset #1: $D_\xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, asset #2: $D_\xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ for all ξ

Path Solver solution: $p^0 = (1,\ 0.7338)$ (with scaling & $z\_i \leq 100$)

$p_\xi^1 = (1,\ 0.7158;\ 1,\ 0.7182;\ 1,\ 0.7205)$

$q = (0.9188,\ 0.6600)$, $z = (72.9868, \text{-}100; \text{-}36.4934, 50;\ \text{-}36.4934, 50)$

sol'n time: not noticeable

value transfer for #1-agent: @$t = 0$: -1.0649,

@$t = 1$, scn-1: 1.403, scn-2: 1.168, scn-3: 0.933,

The BDE-example

3 agents, 2 goods, $|\Xi| = 3$, no y-activities

$$u_i^1(\xi;x) = -\left(5.7 - \prod_{l=1}^{2}(x_l)^{\alpha_{i,l}}\right) = u_1^0(x),\ \alpha_1 = (0.25, 0.75),\ \alpha_{2\&3} = (0.75, 0.25)$$

asset #1: $D_\xi^1 = \begin{pmatrix}1\\0\end{pmatrix}$, asset #2: $D_\xi^2 = \begin{pmatrix}0\\1\end{pmatrix}$ for all ξ

BDE- solution: $p^0 = (1,\ 0.74)$ (with scaling)

$p_\xi^1 = (1,\ 0.7375;\ 1,\ 0.7174;\ 1,\ 0.6633)$

$q = (??,\ ??),\quad z = (0.94, 0;\ 0.03, 0;\ 0.03, 0)$ ⟸ all buyers no sellers

change of variables + add unconstrained agent:

homotopy continuation method (predictor-corrector steps)

The Cass trick

or the no-arbitrage condition

$$q = \sum_{\xi\in\Xi} w_\xi \left(D_\xi^T p_\xi^1\right) \text{ for some weights } w_\xi \geq 0,\ \sum_\xi w_\xi = 1$$

$$\max u_i\left(x^0, (x_\xi^1)_{\xi\in\Xi}\right) = u_i^0(x^0) + E\left\{u_i^1(\xi; x_\xi^1)\right\}$$

$$\text{such that } \left\langle p^0, e_i^0 - x^0 \right\rangle + \langle q, z\rangle \geq 0$$

$$\left\langle p_\xi^1, e_{i,\xi}^1 + D_{i,\xi} z - x_\xi^1 \right\rangle \geq 0, \forall \xi \in \Xi$$

$$\tilde{p}^0 = p^0,\ \tilde{p}_\xi^1 = w_\xi p_\xi^1, \xi \in \Xi$$

solved by BDE

Path Solver $\Rightarrow$ BDE-sol'n

no Path Solver sol'n!

via Augmented Walrasian

for 'money' assets (Deride, Jofré & Wets '09)

cf. financial equilibrium: Hens & Pilgrim '06

$$\max u_i\left(x^0, (x_\xi^1)_{\xi\in\Xi}\right) \text{ such that}$$

$$\left\langle \tilde{p}^0, e_i^0 - x^0 \right\rangle + \sum_{\xi\in\Xi} \left\langle \tilde{p}_\xi^1, D_\xi z - x_\xi^1 \right\rangle \geq 0$$

$$\left\langle \tilde{p}_\xi^1, e_{i,\xi}^1 + D_{i,\xi} z - x_\xi^1 \right\rangle \geq 0,\ \forall \xi \in \Xi$$

$$s^0\left(\tilde{p}^0, \left(\tilde{p}_\xi^1\right)_{\xi\in\Xi}\right) \geq 0,\ s^0\left(\xi; \tilde{p}^0, \left(\tilde{p}_\xi^1\right)_{\xi\in\Xi}\right) \geq 0, \forall \xi$$

clearing the market $\boxed{\text{also: } \sum_{i\in I} z_i = 0}$

Further readings

- Jofré, A. & R. Wets, Variational convergence of bivariate functions: theoretical foundations. Mathematical Programming (2006).
- Jofré, A., R.T. Rockafellar R.T & R. Wets. Variational Inequalities and economic equilibrium. Mathematics of Operation Research (2006?)
- Jofré, A., RT. Rockafellar & R. Wets, A variational inequality scheme for determining an economic equilibrium of classical or extended type. In "Variational analysis and applications", 553--577, Nonconvex Optimisation and Applications, 79, Springer, New York, 2005.
- Jofré, A. & R. Wets, Continuity properties of Walras equilibrium points. Stochastic equilibrium problems in economics and game theory. Annals of Operations Research, 114 (2002), 229--243.
- S. P. Dirkse and M. C. Ferris. The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems. Optimization Methods and Software, 5:123-156, 1995.

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

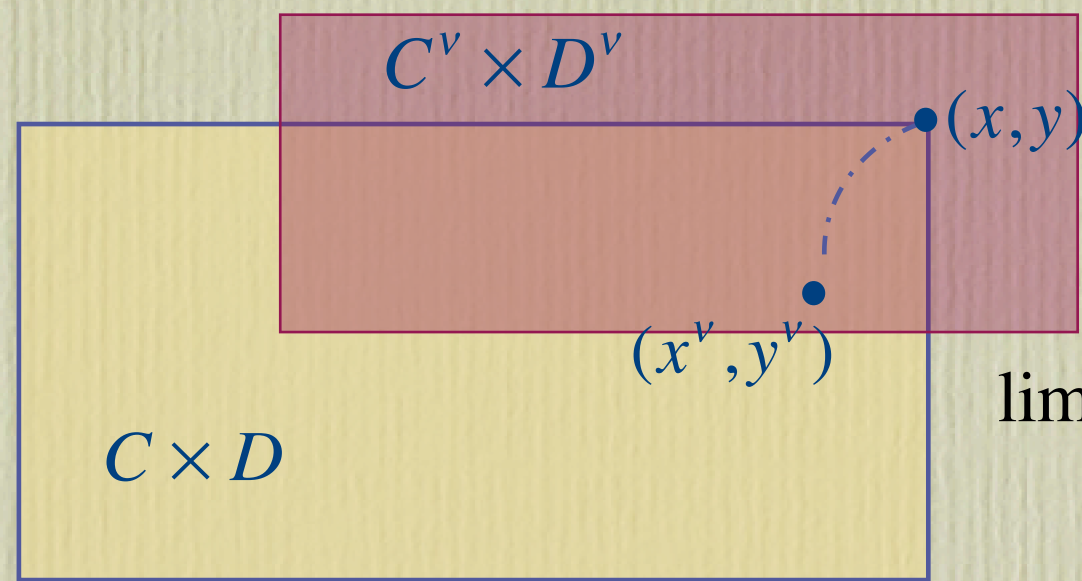
Saddle functions
Epi/hypo convergence

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

EPI/HYPO-Convergence

$\exists y^\nu \in D^\nu$

$\to y \in D$



$\forall\, y \in D, \forall\, x^\nu \in C^\nu \to x$

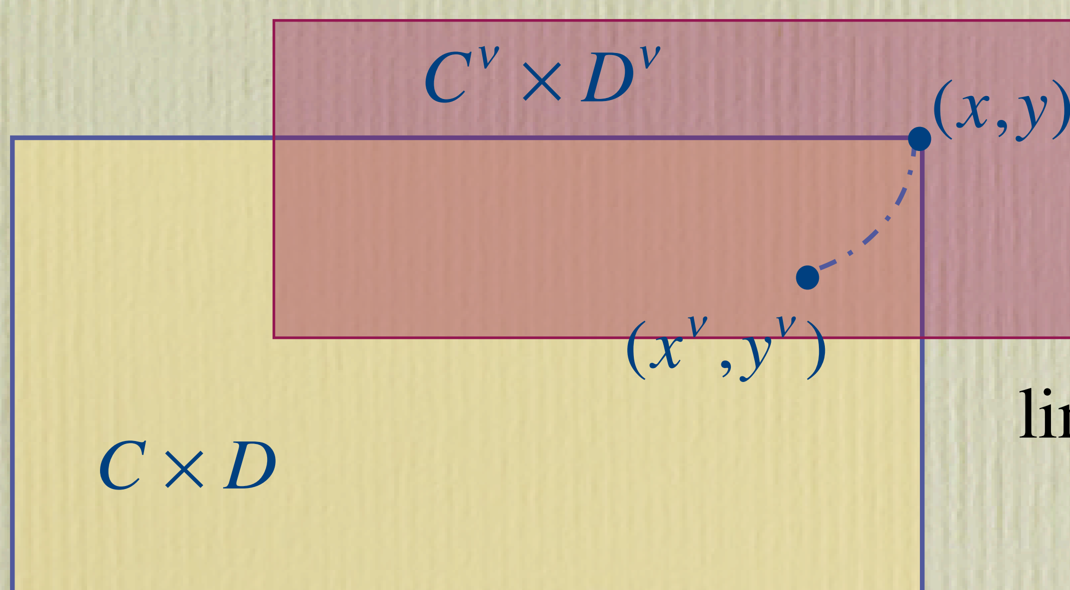
just y dependent: epi-convergence
just x dependent: hypo-convergence

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

epi- in y & hypo- in x $\Rightarrow$ epi/hypo
but not a necessary condition!

$\exists x^\nu \in C^\nu$

$\to x \in C$



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

$\forall\, x \in C, \forall\, y^\nu \in D^\nu \to x \in D$

Saddle Points: vs-Convergene

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

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

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

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

Max-Inf $\neq$ Min-Sup

Variational Inequalities

- $G : C \to \mathbb{R}^n$ $\quad C \subset \mathbb{R}^n$ 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 $C^\nu \to C$, $\quad G^\nu : C^\nu \to \mathbb{R}^n$ continuous
- S^ν solution set of approximating problems

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

V.I.: The gap function

- Let $\boxed{K(u,v) = \langle G(u), v-u \rangle}$ 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 $K^\nu(u^\nu,\bullet) \geq 0$
  - $\boxed{K^\nu \xrightarrow{?} K \quad \text{and} \quad \ldots}$
  - $\bar{u} \in$ 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

$$\Leftrightarrow \bar{x} \in \arg\operatorname{maxinf} N,\ \ N(\bar{x}, \bullet) \geq 0$$

Approximating games

- Nikaido-Isoda functions of approximating games

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

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

$$\boxed{N^\nu \xrightarrow{?} N \quad \text{and} \quad \ldots}$$

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

Ky Fan functions & 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, dom $K = C \times C$ + C compact

$$\boxed{\Rightarrow \arg\max\!-\!\inf K \neq \emptyset}$$

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

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

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, ...

MOPEC Family

Checking Lopsided Tightness

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

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

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

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

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

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

- $K^\nu \to_{lop} K$ 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

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

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

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

Continuous convergence (?):

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

Fixed Points (Set-Valued)

find $x \in C$ (convex) : $\quad x \in S(x),\ S : C \rightrightarrows \mathrm{C} \ \subset \mathbb{R}^n$, 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: $K^\nu(x,v) = \sup\left\{ \langle x - v, z - x \rangle \middle| z \in S^\nu(x) \subset C^\nu \right\}$

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

$\forall \varepsilon_\nu \searrow 0, \quad \bar{x} \in$ cluster points $\left\{ x^\nu \in \varepsilon_\nu\text{-maxinf } K^\nu \right\}$ 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 (maxinf-paradigm) - sufficient conditions
  - $f_0^\nu, f_1^\nu, \ldots, f_m^\nu$ hypo-converge to $f_0, f_1, \ldots, f_m$ on $X^\nu \to X$
  - the collection $\left\{f_i^\nu, \nu \in \mathbb{N}\right\}$ is equi-usc, $i = 0, \ldots, m$
  - Constraint Qualification: $S^\nu = \left\{x \,\middle|\, f_i^\nu \geq 0,\ i = 1, \ldots, m\right\} \to S$
- lop-limit L is unique concave-convex case (epi/hypo): int $S \neq \emptyset$

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 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 $\mathrm{x} \in S(x)$, g continuous, $S : C \rightrightarrows C$ convex

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

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

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

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

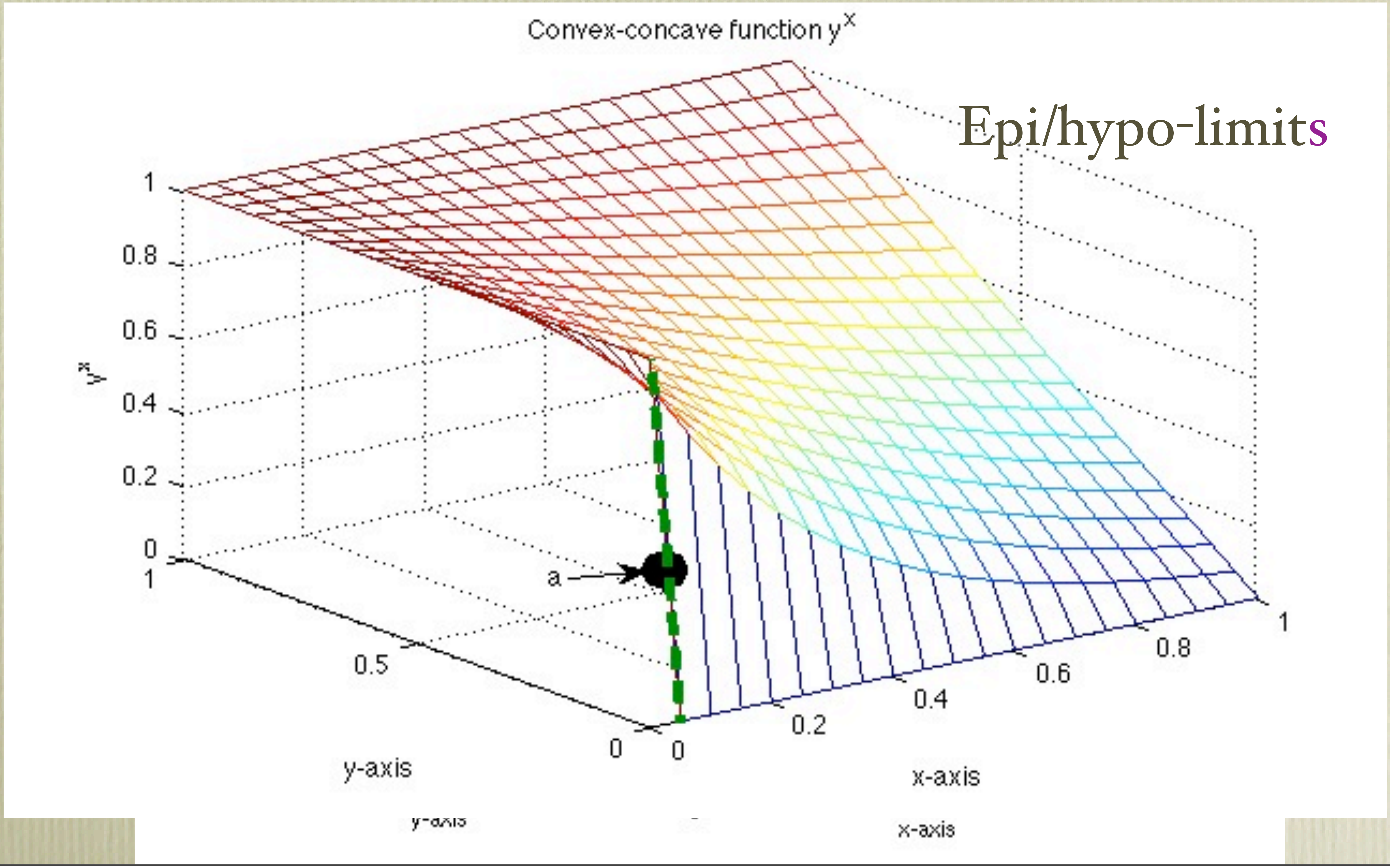
approximating bifunctions: $S^\nu : C^\nu \rightrightarrows C$

$$K^\nu(x,v) = g^\nu(x) + \sup_z \left\{ \langle x - v, z - x \rangle \middle| z \in S^\nu(x) \right\}$$

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

then $K^\nu \xrightarrow[lop]{} K$ & K unique

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



III. Random Sets & Mappings

Why?

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

$EG(x) = \mathbb{E}\{G(\boldsymbol{\xi}, x)\} = 0$ "approximated" by $G^\nu(x) = 0$

$\xi^1, \dots, \xi^\nu$ sample, $G^\nu(x) = \frac{1}{\nu} \sum_{l=1}^{\nu} G(\xi^l, x)$

$G : \Xi \times D \rightrightarrows E$, set-valued $G(\xi, x) \subset E$, inclusion $\mathbb{E}\{G(\boldsymbol{\xi}, x)\} \ni 0$

$\xi^1, \dots, \xi^\nu$ sample, approximation $\frac{1}{\nu} \sum_{l=1}^{\nu} G(\xi^l, x) \ni 0$

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

$\xi^1, \dots \xi^\nu$ sample P^ν (random) empirical measure

approx.: $\min \mathbb{E}^\nu\{f(\boldsymbol{\xi}, x)\} = \frac{1}{\nu} \sum_{l=1}^{\nu} f(\xi^l, x),\ x \in C$

Examples:

$\min f = f_0 + \iota_C$, optimality: $"0 \in \partial f(\bar{x}) = S(x)"$ $\quad \sim 0 = \nabla f(\bar{x}))$

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

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

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

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

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

When f_0, C are convex: $-\partial f_0(\bar{x}) \in N_C(\bar{x})$,

a functional variational inequality

“Variational” Approximations

(E, d) Polish, in paricular $E = \mathbb{R}^n$

$(\text{cl-sets}(E), d\!l)$ complete metric space; Polish if $E = \mathbb{R}^n$

$d\!l(C^\nu, C) \to 0 \iff C^\nu \to C$

<u>osc-mappings</u> = closed graph

(osc-maps$(S), d\!l)$ complete, metric space;

Polish if dom $\subset \mathbb{R}^n$, rge $\subset \mathbb{R}^m$

Convergence:

$S^\nu \xrightarrow{g} S$ if $d\!l(\text{gph}\, S^\nu, \text{gph}\, S) \to 0 \implies (S^\nu)^{-1}(0) \rightrightarrows_v S^-1(0)$

Why?

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

$EG(x) = \mathbb{E}\{G(\boldsymbol{\xi}, x)\} = 0$ "approximated" by $G^\nu(x) = 0$

$\xi^1, \ldots, \xi^\nu$ sample, $G^\nu(x) = \frac{1}{\nu}\sum_{l=1}^{\nu} G(\xi^l, x)$

$G : \Xi \times D \rightrightarrows E$, set-valued $G(\xi, x) \subset E$, inclusion $\mathbb{E}\{G(\boldsymbol{\xi}, x)\} \ni 0$

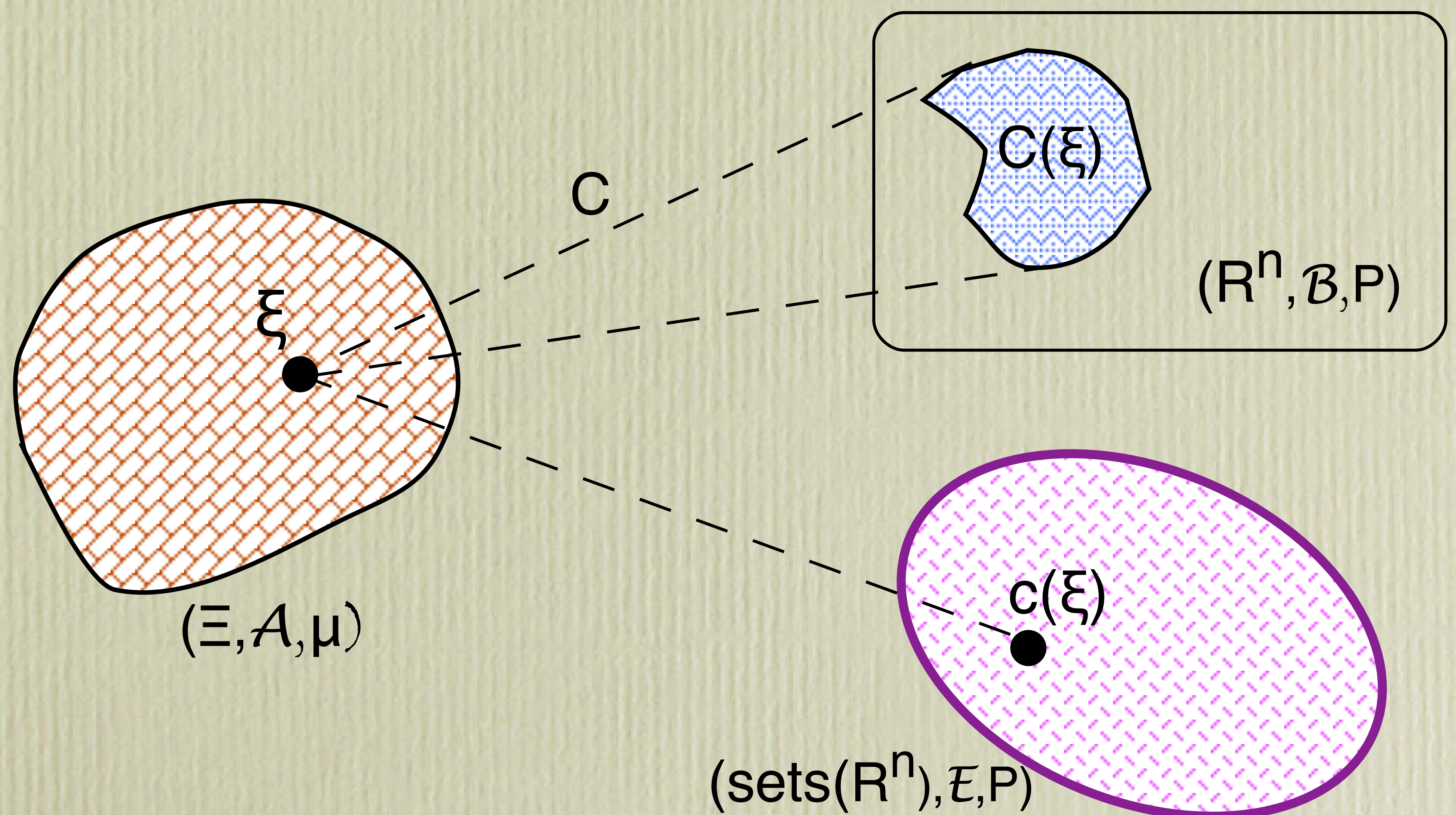
$\xi^1, \ldots, \xi^\nu$ sample, approximation $\frac{1}{\nu}\sum_{l=1}^{\nu} G(\xi^l, x) \ni 0$

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

$\xi^1, \ldots \xi^\nu$ sample P^ν (random) empirical measure

approx.: $\min \mathbb{E}^\nu\{f(\boldsymbol{\xi}, x)\} = \frac{1}{\nu}\sum_{l=1}^{\nu} f(\xi^l, x),\ x \in C$

Set- & “Single”-valued



Random Closed Sets

$(\Xi, \mathcal{A}, P)$, $\quad \Xi \subset \mathbb{R}^N$ & E Polish, for example $\mathbb{R}^n$

$C : \Xi \rightrightarrows E,\ C(\xi) \subset E$ closed set for all $\xi \in \Xi$

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

$\quad \Rightarrow$ dom $C = C^{-1}(E) \in \mathcal{A}$, measurability ~ hit open sets

==

$c : \Xi \to$ cl-sets(E), $c(\xi) \sim C(\xi)$, $\mathcal{F}_O = \{F \subset E \text{ closed} \,|\, F \cap O \neq \emptyset\}$

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

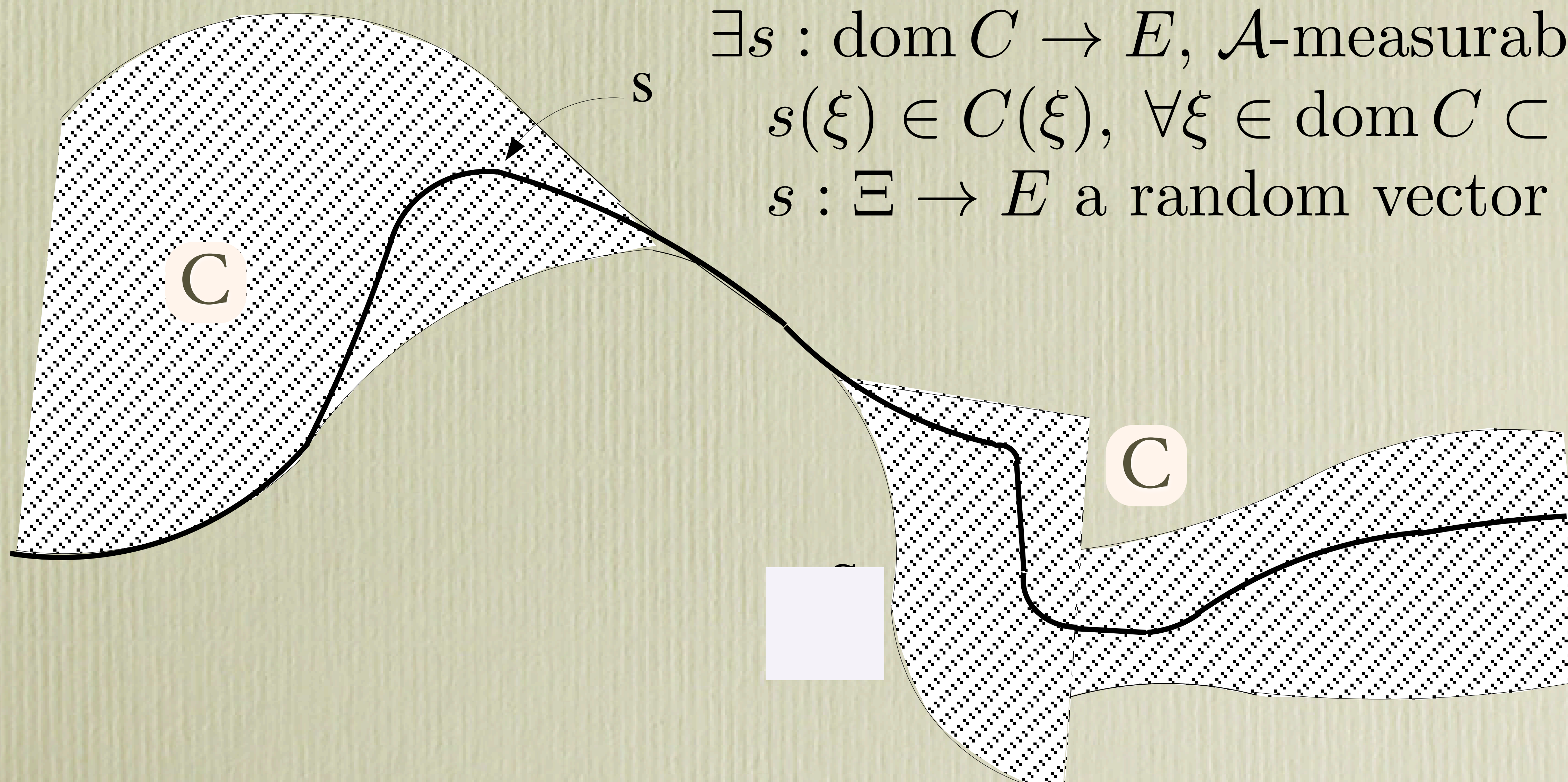
$\quad C$ measurable $\Leftrightarrow c$ measurable $[c^{-1}(\mathcal{F}_O) \in \mathcal{A}]$

$\mathcal{E} = \mathcal{B}$ Borel field when E Polish (complete separable metric space)

Measurable selection

- a random closed set C always admits a measurable selection!

$$\exists s : \operatorname{dom} C \to E,\ \mathcal{A}\text{-measurable},$$
$$s(\xi) \in C(\xi),\ \forall \xi \in \operatorname{dom} C \subset \Xi$$
$$s : \Xi \to E \text{ a random vector}$$



Castaing Representation

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

$$\left\{ s^\nu : \text{ dom } C \to E, \text{ meas.-selections} \right\}$$

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

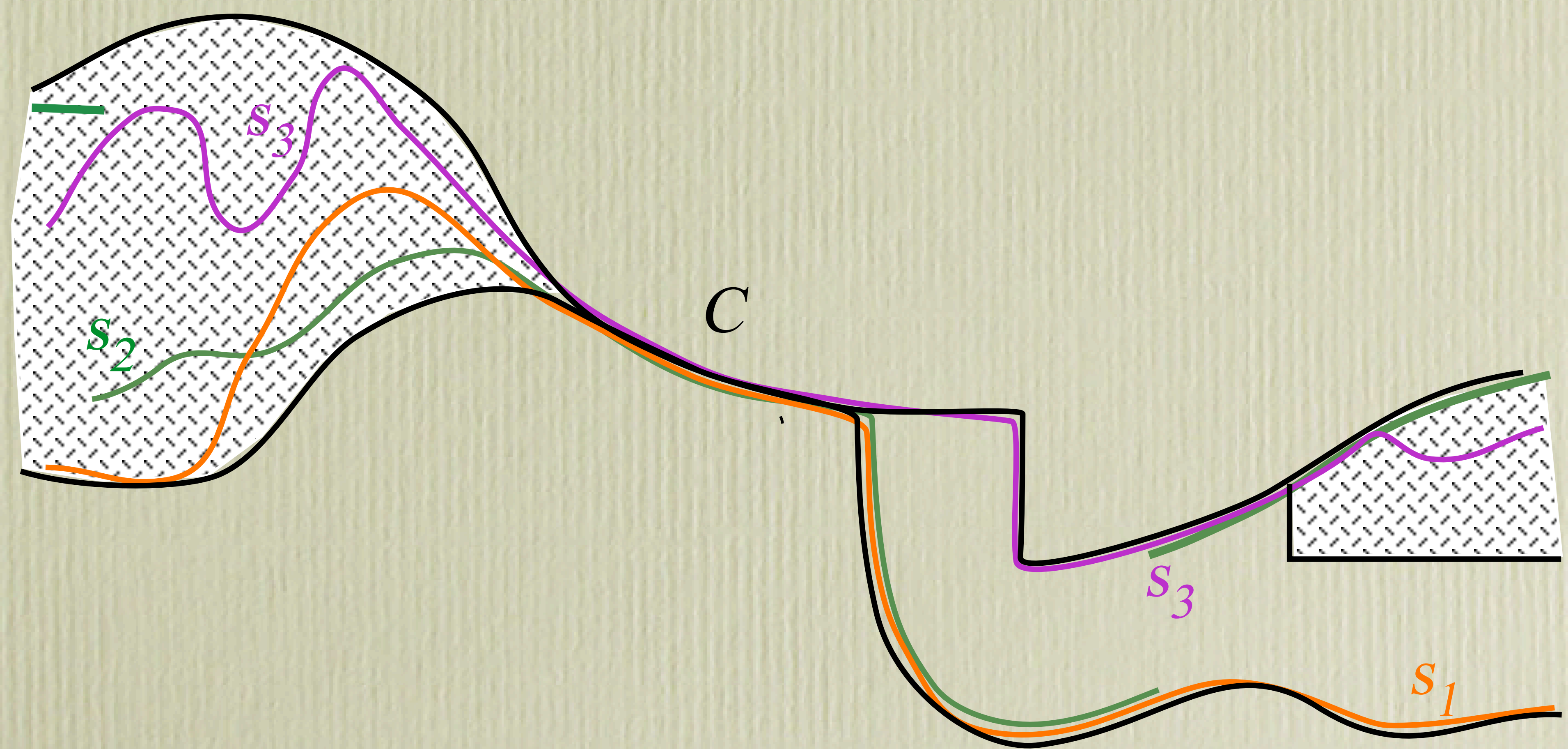
- Graph measurability

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

(negligible sets are P-measurable)

$$C \text{ random set } \Leftrightarrow \text{ gph } C \ \mathcal{A} \otimes \mathcal{B}_n\text{-measurable}$$

Castaing Representation



Convergence of Random Elements

(review)

$$\boldsymbol{\xi} : (\Omega, \mathcal{F}, \mu) \to (\Xi, \mathcal{A}, P), \quad \boldsymbol{\xi}^\nu \xrightarrow{\star} \boldsymbol{\xi}$$

- a.s. (almost sure) convergence:

$$P\big\{\xi \,\big|\, \lim_\nu \boldsymbol{\xi}^\nu(\omega) = \xi \neq \boldsymbol{\xi}(\omega),\, \omega \in \Omega\big\} = 0$$

- convergence in probability:

$$P(|\boldsymbol{\xi}^\nu - \boldsymbol{\xi}| > \varepsilon) \to 0 \text{ for all } \varepsilon > 0$$

- convergence in distribution: $P^\nu \xrightarrow{\mathcal{D}} P$

Outer/Inner Limits

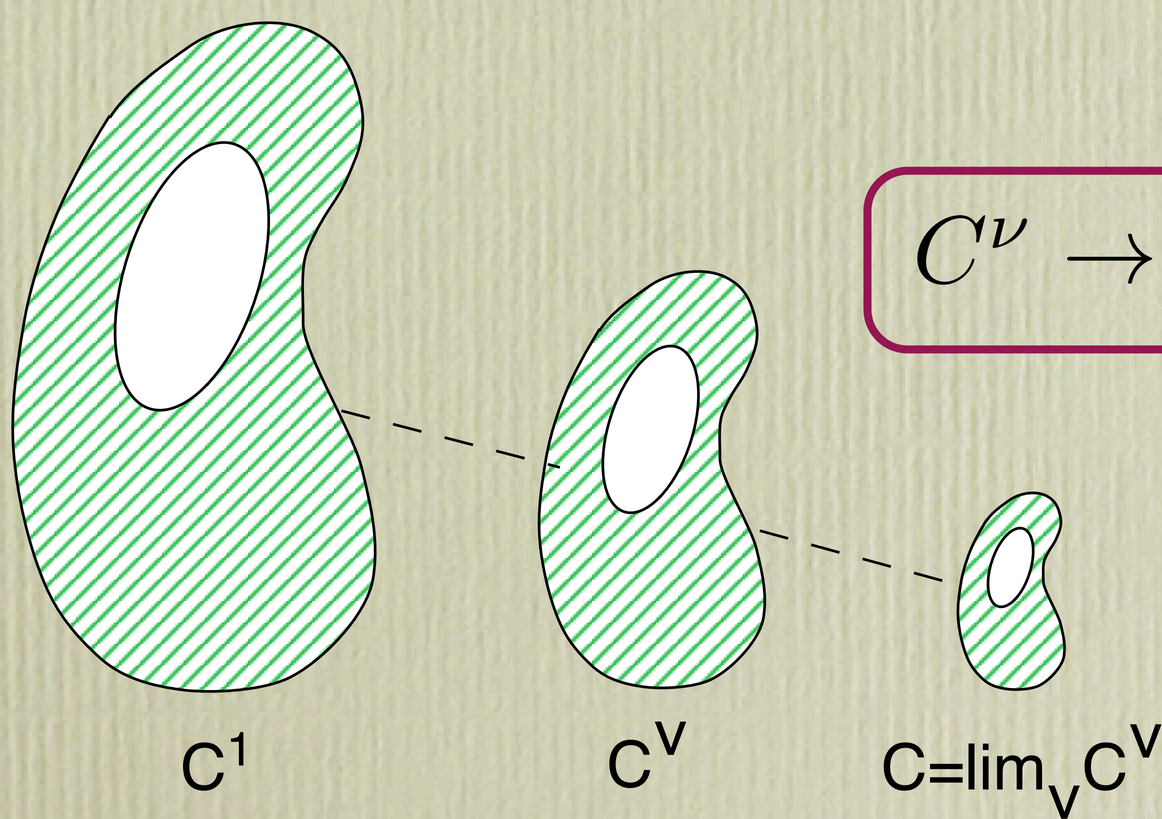
(review)

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

inner limit: $\mathrm{Li}_\nu C^\nu = \left\{ x = \lim_\nu x^\nu, x^\nu \in C^\nu \subset \mathbb{R}^n \right\} \subset \mathrm{Lo}_\nu C^\nu$

limit: $C^\nu \to C$ if $C = \mathrm{Li}_\nu C^\nu = \mathrm{Lo}_\nu C^\nu$ (Painlevé - Kuratowski)

All limit sets are closed



$$\boxed{C^\nu \to C \iff d\!l(C^\nu, C) \to 0}$$

Characterizing a.s. convergence

$\Big\{C; C^\nu : \Xi \rightrightarrows \mathbb{R}^n,\ \nu \in \mathbb{N}\Big\}$ random closed sets. Then,

1. $C^\nu \to C$ *a.s.*, $d\!\!l(C^\nu, C) \to 0$ *a.s.*, $\mathrm{Lo}_\nu(C^\nu) \subset C \subset \mathrm{Li}_\nu(C^\nu)$ *a.s.*,

2. $\forall x \in \mathbb{R}^n$ and $\xi \in \Xi_1$ with $P(\Xi_1) = 1$, $d(x, C^\nu(\xi)) \to d(x, C(\xi))$,

3. $\forall x \in \mathbb{R}^n$ and $\xi \in \Xi_1$ with $P(\Xi_1) = 1$,

$$\lim_{\rho \nearrow \infty} \mathrm{Lo}_\nu \big(C^\nu(\xi) \cap \mathbb{B}(x, \rho)\big) \subset C(\xi) \subset \lim_{\rho \nearrow \infty} \mathrm{Li}_\nu \big(C^\nu(\xi) \cap \mathbb{B}(x, \rho)\big).$$

"Proof 1. $\Leftrightarrow$ 2."

$C^\nu \to C \iff \forall x \in \mathbb{R}^n,\ d(x, C^\nu) \to d(x, C)$ provided $E = \mathbb{R}^n$.

$C^\nu \to C$ if and only if the hit-miss criterion is satisfied

C hits $\mathbb{B}^o(x,\rho)$ then C^ν hits $\mathbb{B}^o(x,\rho)$ for $\nu \geq \nu_{x,\rho}$
so, $C \subset \mathrm{Li}_\nu C^\nu \iff d(x,C) \geq \mathrm{limsup}_\nu\, d(x, C^\nu), \forall x$

C misses $\mathbb{B}(x,\rho)$ then C^ν misses $\mathbb{B}(x,\rho)$ for $\nu \geq \nu_{x,\rho}$
so, $C \supset \mathrm{Lo}_\nu C^\nu \iff d(x,C) \geq \mathrm{liminf}_\nu\, d(x, C^\nu), \forall x$

a.s.-Convergence via Castaing Representations

- $\left\{C^\nu : \Xi \rightrightarrows \mathbb{R}^d, \nu \in \mathbb{N}\right\}$ random closed sets

- $C^\nu \to C$ P-a.s. and dom $C^\nu =$ dom C. Then,

  $\exists$ Castaing representations of $C^\nu \to$ a Castaing representation of C

  If $s : \Xi \to E$ is a measurable selection of C, then

  $\exists\, s^\nu : \Xi \to E$ selections of C^ν converging P-a.s. to s

- ('Egorov's Theorem': $C^\nu \to C$ μ-a.s. $\Leftrightarrow C^\nu \to C$ almost uniformly)

Building Castaing representations

$C : \Xi \rightrightarrows \mathbb{R}^n$, a random closed set. Let

$$A = \left\{ a_k = (a_k^1, \ldots, a_k^n, a_k^{n+1}) \,\middle|\, a_k^i \in \mathbb{Q}^n \text{ \& aff. independent } \right\}$$

for $\emptyset \neq D = D^0$ closed, define $\operatorname{prj}_D a_k = \operatorname{prj}_{D^n} a_k^{n+1}$
where $D^l = \operatorname{prj}_{D^{l-1}} a_k^l$ for $l = 1, \ldots, n$

$\operatorname{prj}_D a_k$ is a singleton: intersection of n+1 "aff. independent" spheres.
Moreover, $\left\{ \operatorname{prj}_D a_k, a_k \in A \right\}$ also dense in D

$s_k : \Xi \to \mathbb{R}^n$ with $s_k(\xi) = \operatorname{prj}_{C(\xi)} a_k$ is a measurable selection of C

$\square$ When D is a random closed set, so is $\xi \mapsto \operatorname{prj}_{D(\xi)} a$, $a \in \mathbb{R}^n$
repeat the argument $n+1$ times to obtain s_k measurable. $\square$

Converging Castaing representations

$C^\nu : \Xi \rightrightarrows \mathbb{R}^n$ random closed sets converging P-$a.s.$ to C, $\operatorname{dom} C^\nu = \operatorname{dom} C$.
Then, $\exists \{ s_k^\nu, k \in \mathbb{N} \}$ Castaing representations of C^ν converging for each k to a Castaing representation $\{ s_k, k \in \mathbb{N} \}$ of C.

$\square$ All Castaing representations are built via our earlier "projections".
Then, $\forall \xi \in \Xi_1, s_k^\nu(\xi) \to s_k(\xi)$, $P(\Xi_1) = 1$ the set of $a.s.$-convergence.
Since, P-$a.s.$ convergence of $C^\nu \to C \Longrightarrow$ (rely on 2. earlier)

$$d(a_k^1, s_k^\nu(\xi)) = d(a_k^1, C^\nu(\xi)) \to d(a_k^1, C(\xi)) = d(a_k^1, s_k(\xi)), \forall \xi \in \Xi_1. \quad \square$$

(a) Convergence of Castaing representations $\not\Rightarrow$ convergence of random sets!
(b) v meas-selection of $C \Rightarrow \exists v^\nu$ meas-selection of C^ν converging $a.s.$ to v.

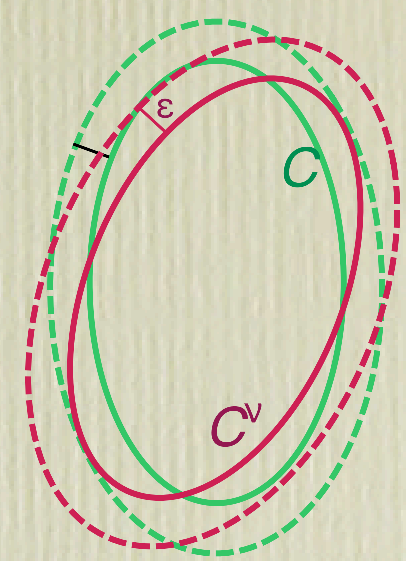
Convergence in probability

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

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

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

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



C^ν converges to C in probability

$\Leftrightarrow P(dl(C^\nu, C) > \varepsilon) \to 0$ for all $\varepsilon > 0$

$\Leftrightarrow$ 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 $\left[\int h(\xi)\, dl(C^\nu(\xi), C(\xi))\, P(d\xi) \to 0 \right]$

$P^\nu \xrightarrow{\mathcal{D}} P \;\sim$ distribution fcns converge

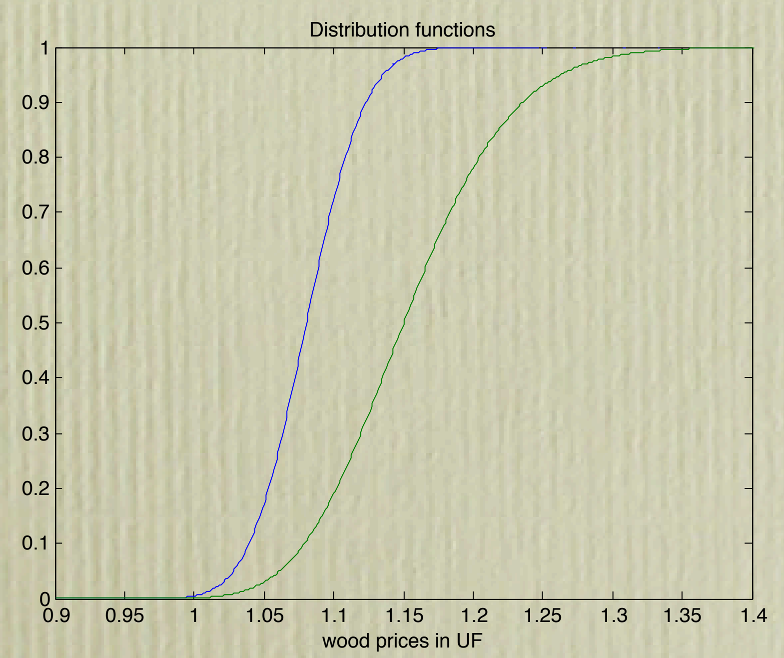
(review)

P^ν, P defined on $(\mathbb{R}, \mathcal{B})$

$P^\nu \xrightarrow{\mathcal{D}} P \iff \int h(\xi)\, P^\nu(d\xi) \to \int h(\xi)\, P(d\xi)\ \forall h$ continuous

$F^\nu(z) = P^\nu\big((-\infty, z)\big), \quad F(z) = P\big((-\infty, z)\big)$, cumulative distributions

$P^\nu \xrightarrow{\mathcal{D}} P \iff F^\nu \xrightarrow{p} F$ on cont $F = \{$ all continuity points of $F\}$



$$\boxed{P^\nu \xrightarrow{\mathcal{D}} P \iff F^\nu \xrightarrow{h} F}$$

Friday, June 21, 13

$P^\nu \xrightarrow{\mathcal{D}} P \;\sim\;$ distribution fcns converge

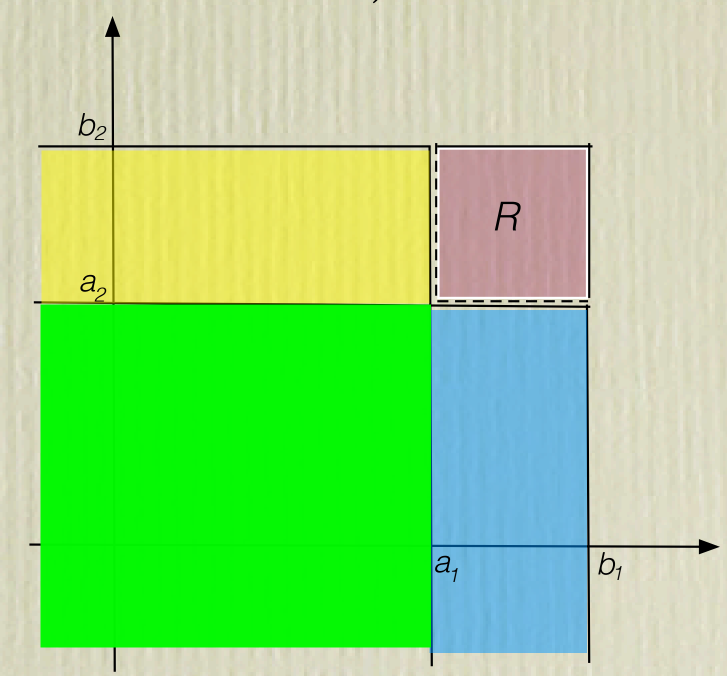
(review)

P^ν, P defined on $(\mathbb{R}^n, \mathcal{B}_n)$ random vectors $\boldsymbol{\xi}^\nu, \boldsymbol{\xi}$

$P^\nu \xrightarrow{\mathcal{D}} P \iff \int h(\xi)\,P^\nu(d\xi) \to \int h(\xi)\,P(d\xi)\;\; \forall h$ continuous

$F^\nu(z) = P^\nu(\xi_i \le z_i, i = 1, \dots, n), \quad F(z) = P(\xi_i \le z_i, i = 1, \dots, n)$

1. $z \le \tilde{z} \implies F(z) \le F(\tilde{z})$ "increasing"
2. $\lim_{z\to\infty} F(z) = 1,\;\; \lim_{z_j \to -\infty} F(z) \to 0,$
3. F is usc (upper sc) $\limsup_{z' \to z} F(z') \le F(z),$



4. $R = (a_1, b_1] \times \cdots \times (a_n, b_n], \quad V = \{a_1, b_1\} \times \cdots \times \{a_n, b_n\}$ vertices of R
   $\forall R \subset \mathbb{R}^n,\; P(\boldsymbol{\xi} \in R) = \sum_{v \in V} \mathrm{sgn}(v) F(v), \quad \mathrm{sgn}(v \in V) = (-1)^{\#a \text{ in } v}$

$$\boxed{P^\nu \xrightarrow{\mathcal{D}} P \iff -F^\nu \xrightarrow{e} -F}$$

Distribution of a random set

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

Distribution $(P, \mathcal{B})$ regular, $\quad \mathcal{K}$ compact subsets E

determined by values on $\{\mathcal{F}^K \mid K \in \mathcal{K}\}$ or $\{\mathcal{F}_K \mid K \in \mathcal{K}\}$

Distribution function (Choquet capacity):

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

a) $T(K^\nu) \searrow T(K)$ when $K^\nu \searrow K \qquad$ (~ usc on $\mathbb{R}^n$) *(1,3)*

(4) b) $\{D_\nu : \mathcal{K} \to [0,1]\}_{\nu \in \mathbb{N}}$ 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 $\nu = 2, \ldots$

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

(~ rectangle condition on $\mathbb{R}^n$)

Existence-Uniqueness of 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)*

(refined) via probabilistic arguments *(Salinetti-Wets)*

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

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

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

Convergence in Distribution

random sets C^ν converge in distribution to C when

induced P^ν narrow-converge to $P: P^\nu \to_n P \; = P^\nu \overset{\mathcal{D}}{\to} \; P$

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

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

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

b) $K^\nu \nearrow K =$ cl $\bigcup_\nu K^\nu$ regularly if 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 discontinuities

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) $\lambda(\emptyset) = 0, \lambda(E) = 1$

(v) 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)

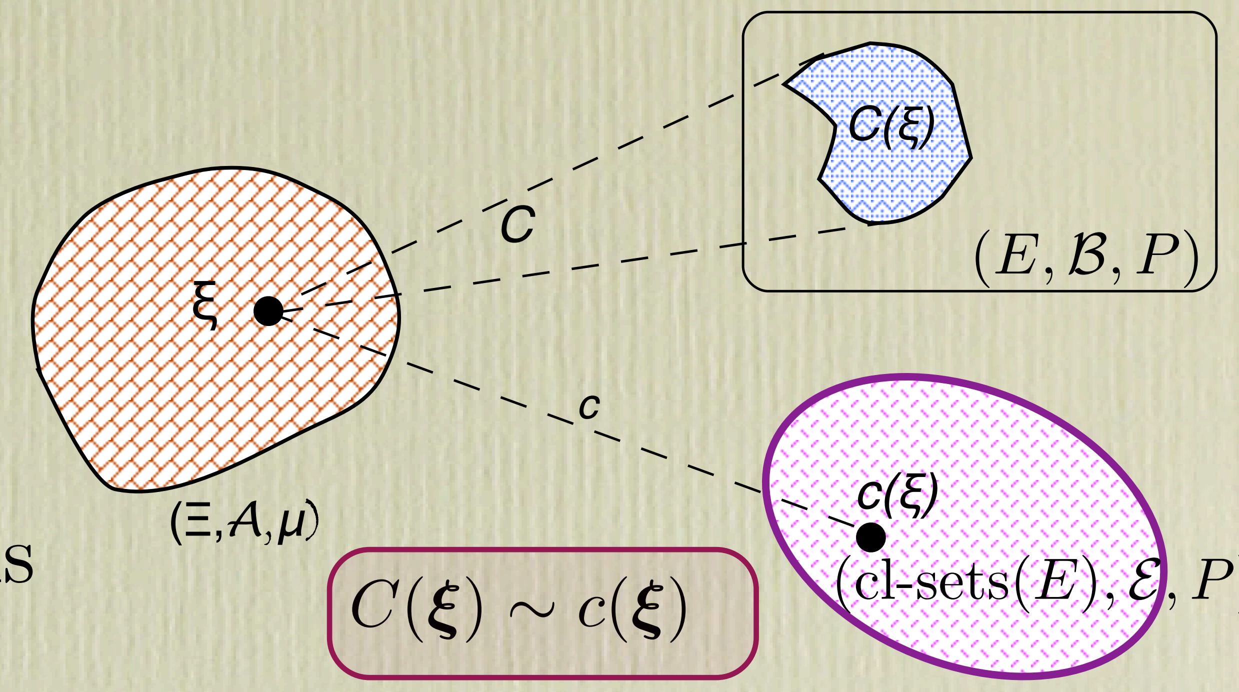
$\{P^\nu, \nu \in \mathbb{N}\}$ tight: $P^\nu \to_n P \Leftrightarrow \lambda^\nu \to_h \lambda$ ($\sim -\lambda^\nu \to_e -\lambda$) on cl-sets($E$)

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., "~" Skorohod distance)

Random sets

(review)



C 'covered' by countable selections
Castaing represnetation

a.s convergence: $P\{\xi \,|\, d\!l(C^\nu(\xi), C(\xi)) \to 0\} = 0$

$\Rightarrow$ in probability: $\forall \varepsilon > 0,\ P\{\xi \,|\, d\!l(C^\nu(\xi), C(\xi)) > \varepsilon\} \to 0$

$\Rightarrow$ in distribution $T : \text{cpct-sets}(E) \to [0,1],\ T(\emptyset) = 0,$
(a) $T(K^\nu) \searrow T(K)$ for $K^\nu) \searrow K$, (b) 'rectangle cond'n'
$P^\nu \overset{\mathcal{D}}{\to} P \iff T^\nu \to T$ on cpct-sets$(\mathbb{R}^n)$
or, even, on finite union of closed rational balls.

Expectation

Artstein, Vitale, Hart, Wets,
Cressis, Hiai, Weyl, ...

“Simple” random sets

$C : \Xi \rightrightarrows \mathbb{R}^n$ is a *simple* random set if $\operatorname{rge} C$ is finite.
C is a closed random set $\iff$ $C = P$-$a.s.$ limit of simple random sets.

$\square$ $\Leftarrow$: the limit of a sequence of random sets is a random set
$\Rightarrow$: let $C^\nu = C \cap \nu\mathbb{B}$, unif. bounded closed random set, $C = \operatorname{Lm}_\nu C^\nu$
build (via "prj") Castaing representations $\{r_k^\nu\}_{k\in\mathbb{N}}$ of the C^ν
let $\{s_k^\nu\}_{k\in\mathbb{N}'} = \bigcup_{\upsilon\le\nu} \{r_k^\upsilon\}_{\upsilon\in\mathbb{N}}$, also Castaing for C^ν
$D_k^\nu = \bigcup_{j\le k} s_j^\nu$ $d\!l$-converge uniformly to C^ν as $k\to\infty$
since each $s_k^\nu = \lim_{l\to\infty} s_{kl}^\nu$ uniformly, s_{kl}^ν simple random variables
$\Delta_{kl}^\nu = \bigcup_{j\le k} s_{jl}^\nu$ is a simple random set, $C(\xi) = \operatorname{Lm}_\nu \operatorname{Lm}_k \operatorname{Lm}_l \Delta_{kl}^\nu(\xi)$

$$\Delta_{kl}^\nu \xrightarrow{u} D_k^\nu \xrightarrow{u} C^\nu \text{ allows diagonalization to find } \Delta_{k^\nu l^\nu}^\nu \to C. \quad \square$$

Sierpiński-Lyapunov Theorems

$(\Xi, \mathcal{A})$ a measure space

Sierpiński (1922). Suppose P is an atomless probability measure.
Given $A_0, A_1 \in \mathcal{A}$ with $0 \leq P(A_0) \leq P(A_1) \leq 1$, then
$\forall \lambda \in [0,1]$, $\exists$ $A_\lambda \in \mathcal{A}$ such that $P(A_\lambda) = (1-\lambda)P(A_0) + \lambda P(A_1)$.
In particular, it implies $\forall\, \lambda \in [0,1]$, $\exists$ $A \in \mathcal{A}$ such that $P(A) = \lambda$;
choose $A_0 = \emptyset$ and $A_1 = \Xi$.

Lyapunov (1940) $\mu : \mathcal{A} \to \mathbb{R}^n$ atomless, σ-additive measure.
For $A \in \mathcal{A}$, define $\operatorname{rge} \mu(A) = \left\{ \mu(B) \,\middle|\, B \subset A \cap \mathcal{A} \right\}$. Then,
$\operatorname{rge} \mu(\Xi) \subset \mathbb{R}^n$ is convex and if μ is also bounded, it's compact.

Expectation: simple random set

$C : \Xi \rightrightarrows \mathbb{R}^n$ a simple random set, i.e., $\operatorname{rge} C = \left\{ z^k \in \mathbb{R}^n \,\middle|\, k \in K, |K| \text{ finite } \right\}$
Given $\bar{r}, \bar{s} \in EC = \mathbb{E}\{C(\boldsymbol{\xi})\} \quad \Longrightarrow$
$\exists$ simple selections $r, s : \Xi \to \mathbb{R}^n$ with $\mathbb{E}\{r(\boldsymbol{\xi})\} = \bar{r}, \mathbb{E}\{s(\boldsymbol{\xi})\} = \bar{s}$.
Let $\lambda \in [0,1]$. Define $v : \Xi \to \mathbb{R}^n$ as follows:

1. partition Ξ into subsets $A_=$ and $\mathcal{A}_{\neq}$
2. $A_= = \left\{ \xi \in \Xi \,\middle|\, r(\xi) = s(\xi) \right\} \in \mathcal{A}$
3. $A = \left\{ \xi \in \Xi \,\middle|\, r(\xi) = z_k, s(\xi) = z_l, k \neq l \right\} \in \mathcal{A}_{\neq}$, a finite collection
4. split each $A \in \mathcal{A}_{\neq}$, $P(A_r) = \lambda P(A)$ & $A_s = A \setminus A_r$ (Sierpiński)

$$\text{set } v(\xi) = \begin{cases} r(\xi) & \text{on } \bigcup_{A \in \mathcal{A}_{\neq}} A_r \,\cup\, A_= \\ s(\xi) & \text{on } \bigcup_{A \in \mathcal{A}_{\neq}} A_s \end{cases}$$

then $\bar{v} = \mathbb{E}\{v(\xi)\} = \lambda \bar{r} + (1-\lambda)\bar{s} \quad \Longrightarrow \quad EC$ convex.
Clearly EC is bounded and it's easy to show it's also closed $\Longrightarrow$ compact.

Expectation of random set

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

$\Longrightarrow C = P$-$a.s.$ limit of simple random sets,

say $C^\nu \underset{a.s.}{\to} C$ with $C^\nu \nearrow$ without loss of generality

$EC^\nu = \mathbb{E}\{C^\nu(\boldsymbol{\xi})\} \nearrow$ are convex, compact $\Rightarrow$

$EC = \mathbb{E}\{C(\boldsymbol{\xi})\} = \bigcup_\nu EC^\nu$

$\Longrightarrow EC$ convex

$\Longrightarrow EC$ closed if C is integrably bounded

$\Longrightarrow$ compact if $\operatorname{rge} C$ is bounded

Random set: Expectation

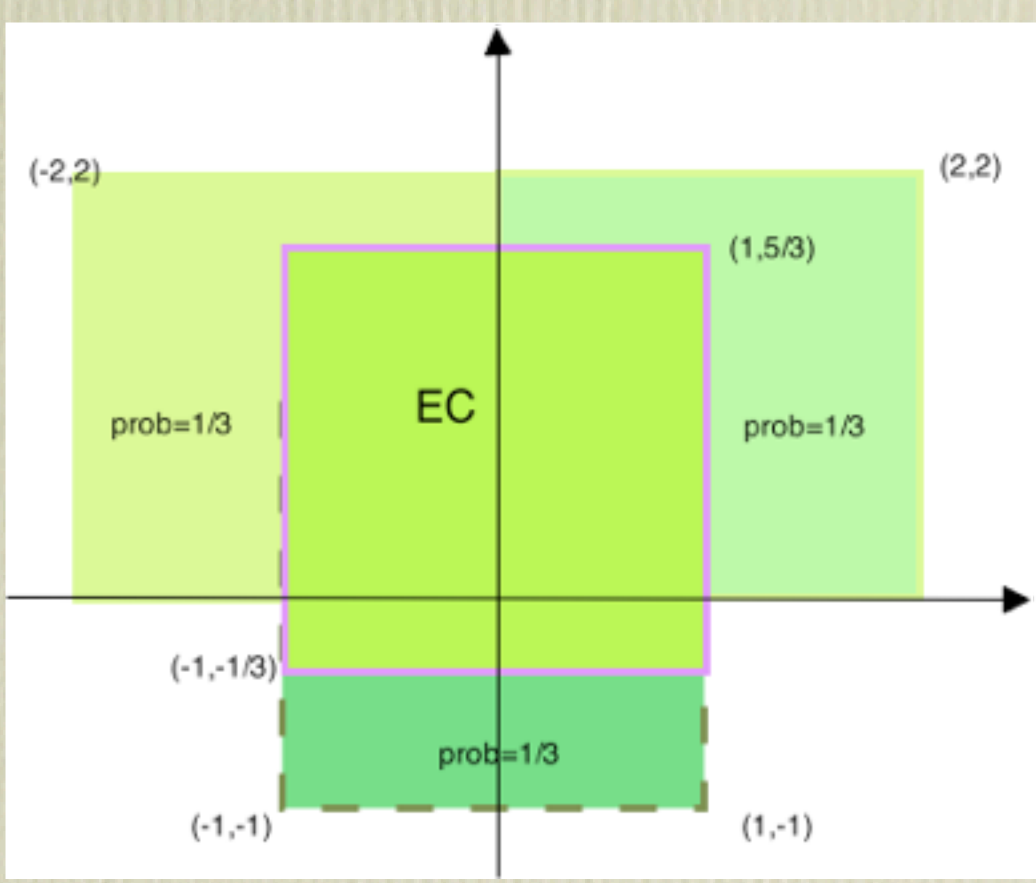
$$EC = \mathbb{E}\{C(\xi)\} = \left\{ \int_{\Xi} s(\xi)\,P(d\xi) \,\middle\|\, s(\bullet)\ P\text{-summable selection} \right\}$$

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

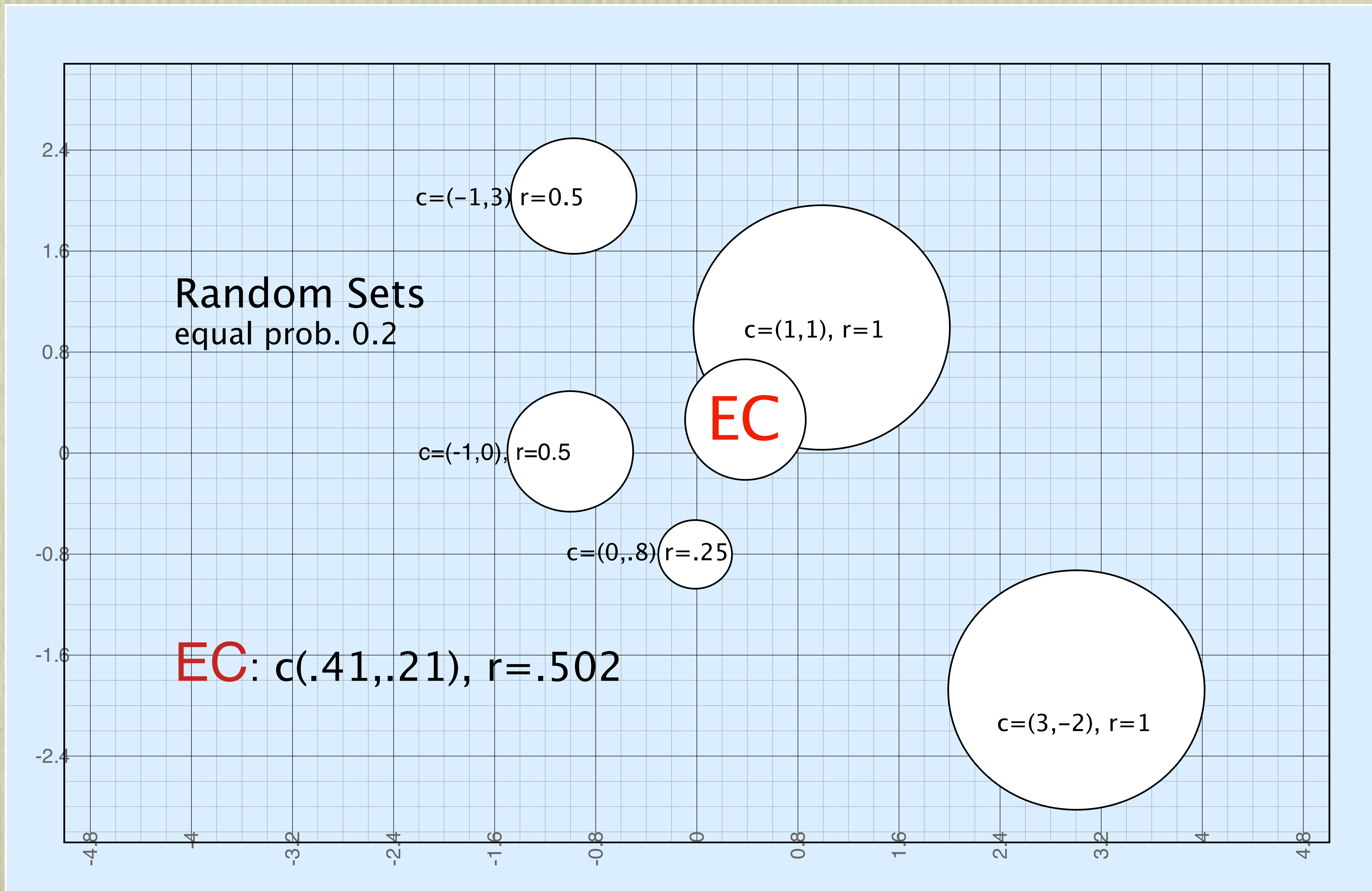
Convexity:

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

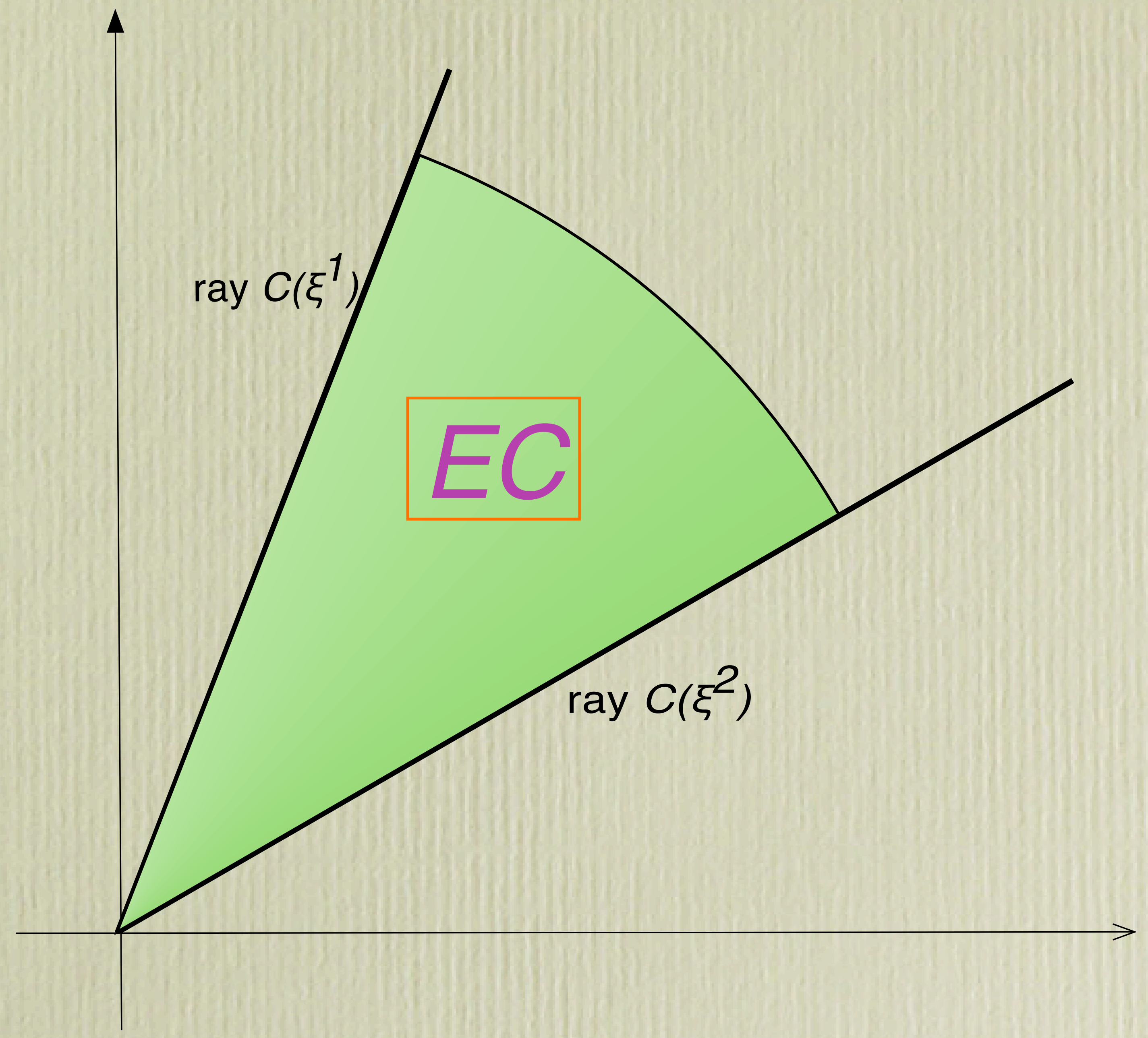
(certainly when P is atomless).



Expectation: Bounded Random Sets



Expectation: Unbounded Random Sets



Some properties: $\mathbb{E}\{C(\boldsymbol{\xi})\}$

- measure P atomless, then $EC = \mathbb{E}\{C(\boldsymbol{\xi})\}$ is convex *(Richter, Lyapounov,...)*
- P is P-atom convex $\Longrightarrow$ EC is convex; [an atom contains no (measurable) subset of positive probability]
- C a random set, $\emptyset \neq EC = \mathbb{E}\{C(\boldsymbol{\xi})\}$ contains no line, then

  $$\operatorname{con} EC = \mathbb{E}\{\operatorname{con} X(\boldsymbol{\xi})\}$$

  this essentially requires that $C(\xi) \subset$ a pointed cone
- in general, the expectation of a (closed-valued) random set is *not* closed
- if $|C| = \mathbb{E}\big\{\sup\left[\,|s(\xi)|\,\big|\, s(\xi) \in C(\xi)\,\right]\big\} < \infty$ then EC is closed; C is then *integrably bounded.*

LLN: Random Sets (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

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

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

$\text{Lo}_\nu C^\nu(\xi^\infty) \subset \bar{C} \iff \limsup_\nu \sigma_{C^\nu} \leq \sigma_{\bar{C}}$ support functions

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

Random mappings

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

$\mathcal{A} \otimes \mathcal{B}^n$-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 $\xi \mapsto S(\xi, \cdot)$ P-atom convex

Law of Large Numbers for random sets

applies pointwise

Sample Average Approximation (SAA)

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') $\quad \bar{x} \in \bar{S}^{-1}(0)$

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

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

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

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

A few examples ...
(from the MOPEC fanily)

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

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

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

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

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

Stochastic V.I. (variational inequality)

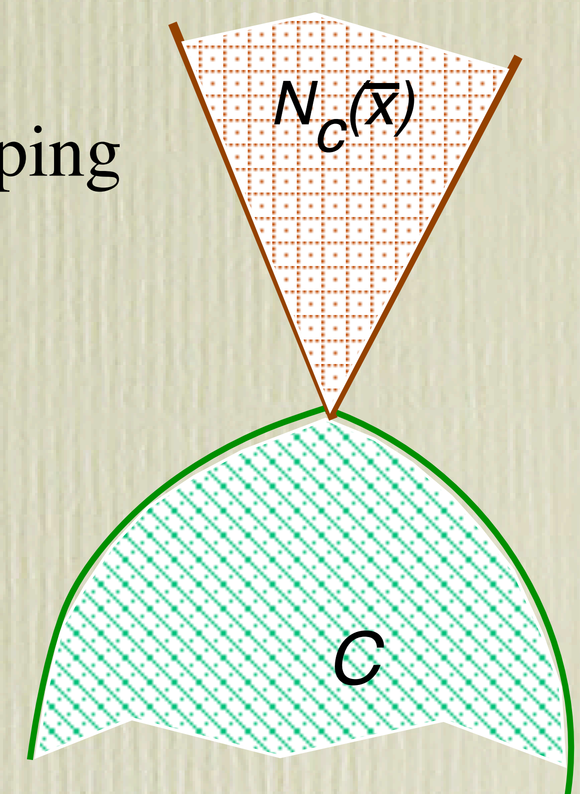
Network flow equilibrium with stochastic demand and link capacities
Economic equilibrium in a stochastic environment

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

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

$N_{C(x)} + G^\nu(\xi, x) = S^\nu(x) \ni 0$, S^ν closed set-valued mapping

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



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

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

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

Walras Equilibrium: stochastic environment

$$(c_a^1, y_a, c_{a,\xi}^2) = \arg\max_{x^1, y \in \mathbb{R}^L, x_\cdot^2 \in \mathcal{M}} u_a^1(x^1) + \mathbb{E}^a \left\{ u_a^2(\boldsymbol{\xi}, x^2(\boldsymbol{\xi})) \right\}$$

$$\text{such that } \left\langle p^1, x_a^1 + T_a^1 y \right\rangle \leq \left\langle p^1, e_a^1 \right\rangle$$

$$\left\langle p_\xi^2, x_{a,\xi}^2 \right\rangle \leq \left\langle p_\xi^2, e_{a,\xi}^2 + T_{a,\xi}^2 y \right\rangle, \ \forall \xi \in \Xi$$

$$x_a^1 \in X_a^1, \ x_{a,\xi}^2 \in X_{a,\xi}^2, \quad \forall \xi \in \Xi$$

$\mathbb{E}^a \left\{ \bullet \right\}$ 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)

$e_a^1 \in \operatorname{int} X_a^1, \ e_{a,\xi}^2 \in \operatorname{int} X_{a,\xi}^2$ for all ξ

Market Clearing -Equilibrium

excess supply: agent-a: $\left(c_a^1, y_a^1, \left\{ c_{a,\xi}^2 \right\}_{\xi\in\Xi} \right)$

$$\sum_{a\in\mathcal{A}} \left(e_a^1 - (c_a^1 + T_a^1 y_a) \right) = s^1\left(p^1, \{p_\xi^2\}_{\xi\in\Xi} \right) \geq 0$$

$$\forall \xi,\ \sum_{a\in\mathcal{A}} \left((e_{a,\xi}^2 + T_{a,\xi}^2) - c_{a,\xi}^2 \right) = s_\xi^2\left(p^1, \{p_\xi^2\}_{\xi\in\Xi} \right) \geq 0$$

Variational inequality: $-G(p,(x_a),(\lambda_a)) \in N_D(p,(x_a),(\lambda_a))$,

$$p = \left(p^1, \{p_\xi^2\}_{\xi\in\Xi} \right),\ x = \left(x^1, \{x_\xi^2\}_{\xi\in\Xi} \right),\ \lambda = \left(\lambda^1, \{\lambda_\xi^2\}_{\xi\in\Xi} \right)$$

$$S(\xi,(p,x,\lambda)) = G(\xi,(x,p,\lambda)) + N_{D(\xi)}(p,x,\lambda)),$$

$$\mathbb{E}\left\{ S(\xi,(p,x,\lambda)) \right\} \ni 0$$

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

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

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

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. □

So far ...

➪ *generalized equations*

$S : \Xi \times D \rightrightarrows E$, set-valued $S(\xi, x) \subset E$, inclusion $\mathbb{E}\{S(\boldsymbol{\xi}, x)\} \ni 0$

iid-sample $\vec{\xi}^{\,\nu} = \xi^1, \dots, \xi^\nu$ and $x \mapsto S(\xi, x)$ osc

SAA-mapping $S^\nu : \Xi^\infty \times D \rightrightarrows E$, random osc mappings

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

$\forall x \in D$, $S(\cdot, x)$, closed random set,

let $\bar{S} = \operatorname{cl} \operatorname{con} ES$, $\quad ES(x) = \mathbb{E}\{S(x, \boldsymbol{\xi})\}$

Artstein-Hart LLN applies: $S^\nu \xrightarrow{p} \bar{S}$ $a.s.$ when $E = \mathbb{R}^m$

but $\xrightarrow{p}$ $\not\Rightarrow$ $(S^\nu)^{-1}(0) \rightrightarrows \bar{S}^{-1}(0)$. Needed $S^\nu \xrightarrow{g} \bar{S}$

recall: $\bar{S}(x) = \operatorname{cl} ES(x)$ when P-atom convex, $ES(x)$ closed if $\xi \mapsto S(\xi, x)$ is integrably bounded and compact if $\operatorname{rge} S(\cdot, x)$ is bounded.

Consistent Approximations ?

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

sometimes!

graphical rather than pointwise convergence is required

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

relationship between graphical and pointwise convergence?

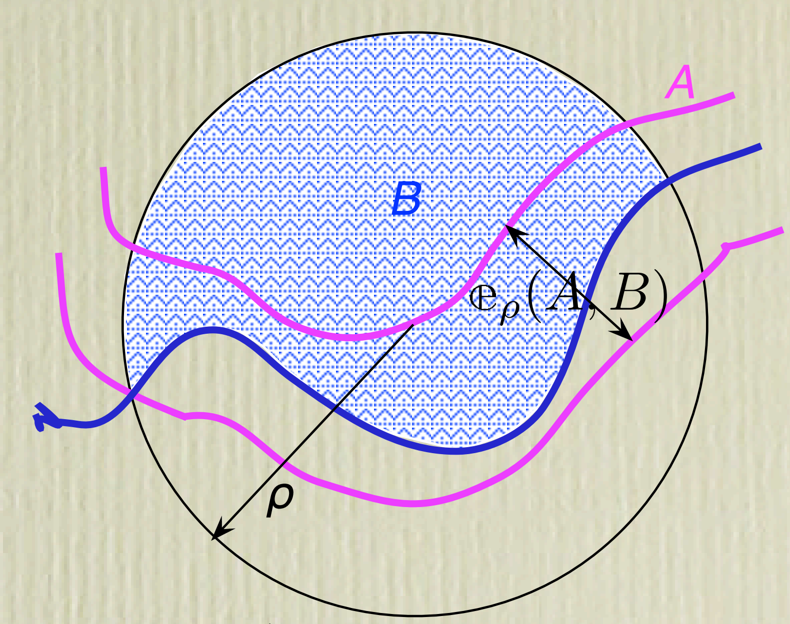
Semicontinuity: osc/isc

(review)

$S : D \rightrightarrows \mathbb{R}^m$ continuous at $\bar{x}$ if $\lim_{x^\nu \to \bar{x}} \mathbb{d}\big(S(x^\nu), S(\bar{x})\big) \to 0$

$$\mathbb{d}(S(x^\nu), S(\bar{x})) \to 0 \iff \mathbb{d}_\rho(S(x^\nu), S(\bar{x})) \to 0$$

$$\iff \hat{\mathbb{d}}_\rho(S(x^\nu), S(\bar{x})) \to 0. \forall\, \rho > \bar{\rho} \geq 0$$



$$\hat{\mathbb{d}}_\rho\big(S(x^\nu), S(\bar{x})\big) = \max\big[\,\mathbb{e}_\rho\big(S(x^\nu), S(\bar{x})\big), \mathbb{e}_\rho\big(S(\bar{x}), S(x^\nu)\big)\,\big]$$

S is osc (outer semicontinuous) at $\bar{x}$ if $\mathbb{e}_\rho\big(S(x^\nu), S(\bar{x})\big) \to 0$ as $x^\nu \to \bar{x}$

S is isc (inner semicontinuous) at $\bar{x}$ if $\mathbb{e}_\rho\big(S(\bar{x}), S(x^\nu)\big) \to 0$ as $x^\nu \to \bar{x}$

Equi-osc mappings

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

osc at $\bar{x}$: given any $\rho > 0, \epsilon > 0$

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

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

given any $\rho > 0, \epsilon > 0$

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

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

Graphical versus Pointwise Convergence

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

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

~ Arzela-Ascoli Theorem for set-valued mappings

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

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

gph-convergence of SAA-mappings

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

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

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

Sufficient condition: P^∞-a.s.

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

Law of large Numbers for Random Mappings

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

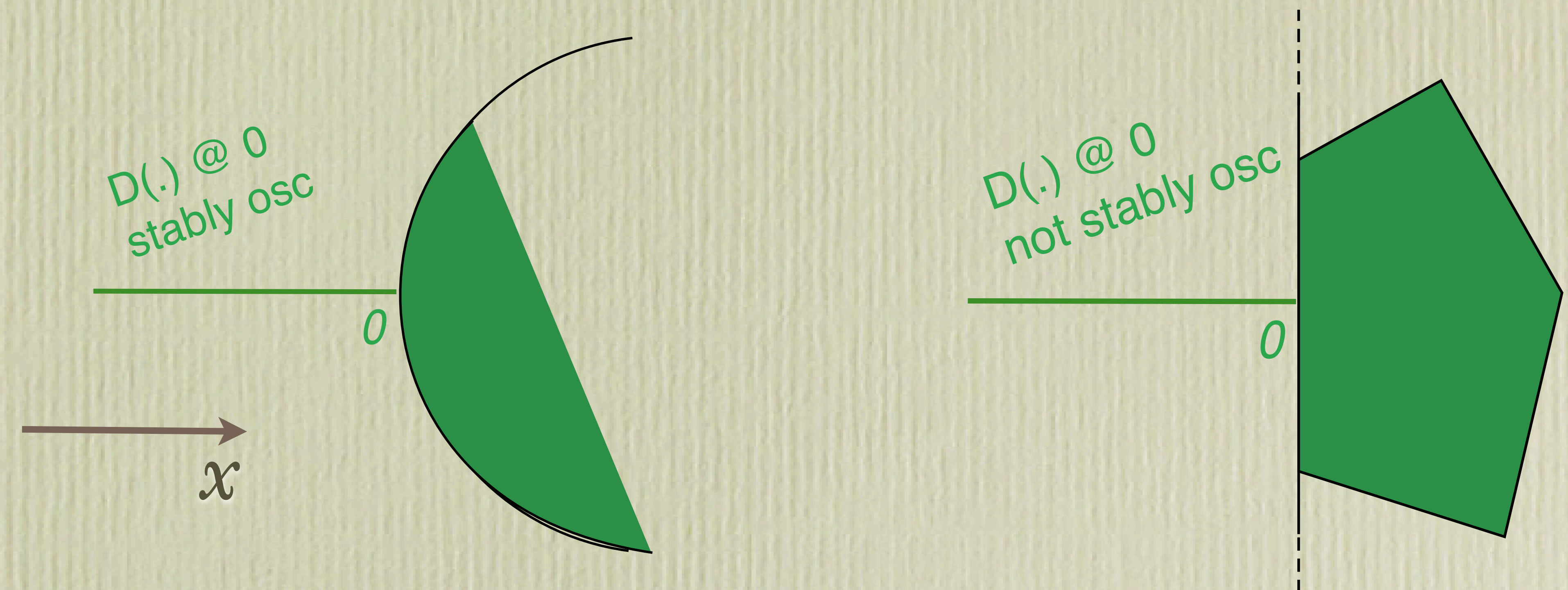
stably osc & steady under averaging

$\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \ldots,$ iid random variables (values in Ξ), distribution P

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

Friday, June 21, 13

Stably osc

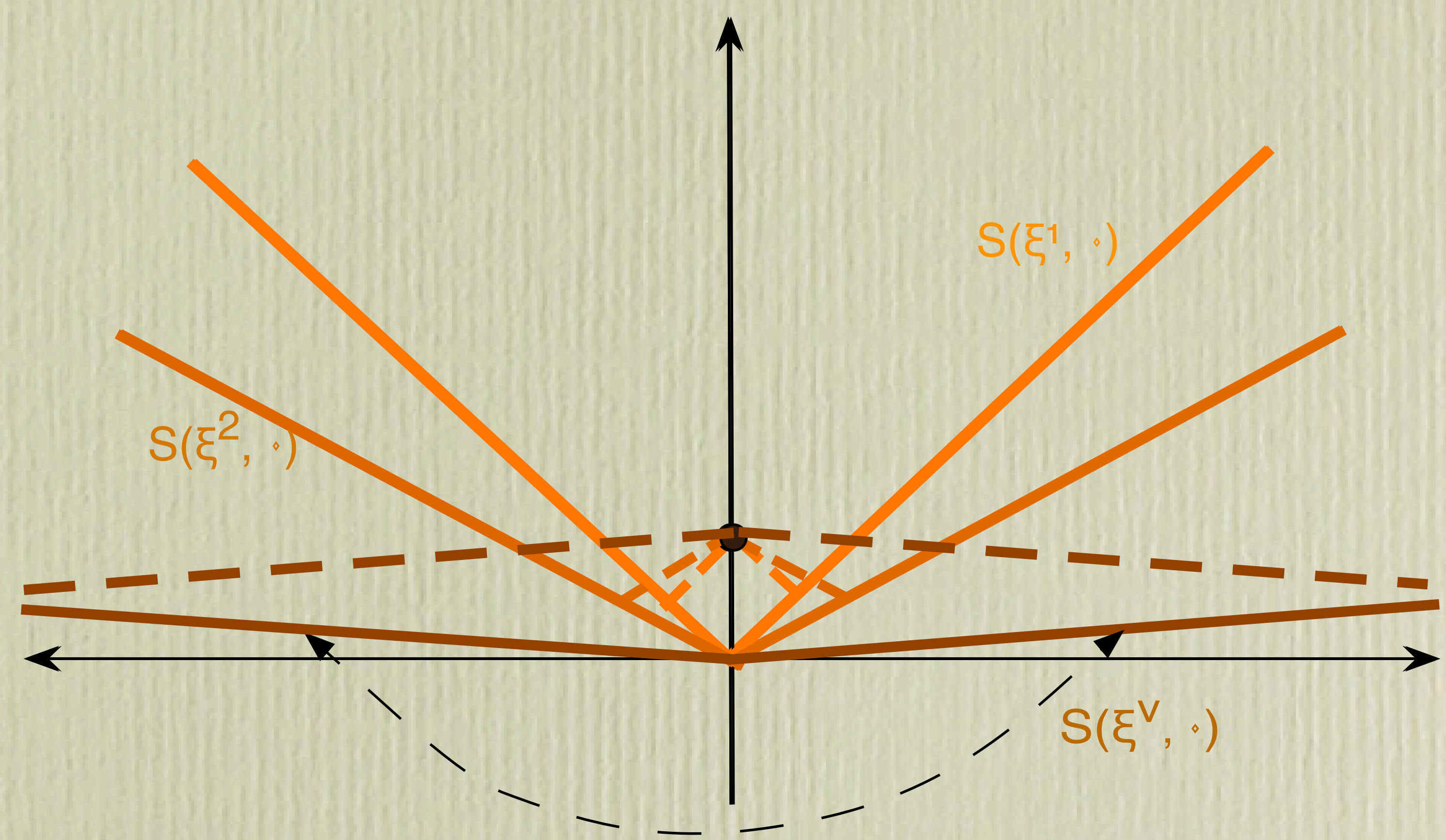


S stably osc near $\bar{x}$ if μ-a.s.,

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

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

Steady under averaging



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

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

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 $\quad x \mapsto N_{\mathbb{B}}(x)$ osc

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

Implementing SAA ** locally

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

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

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

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

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

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

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

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

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

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

Implementing SAA ** example

stochastic program with recourse (simple): ξ uniform on [1,2]

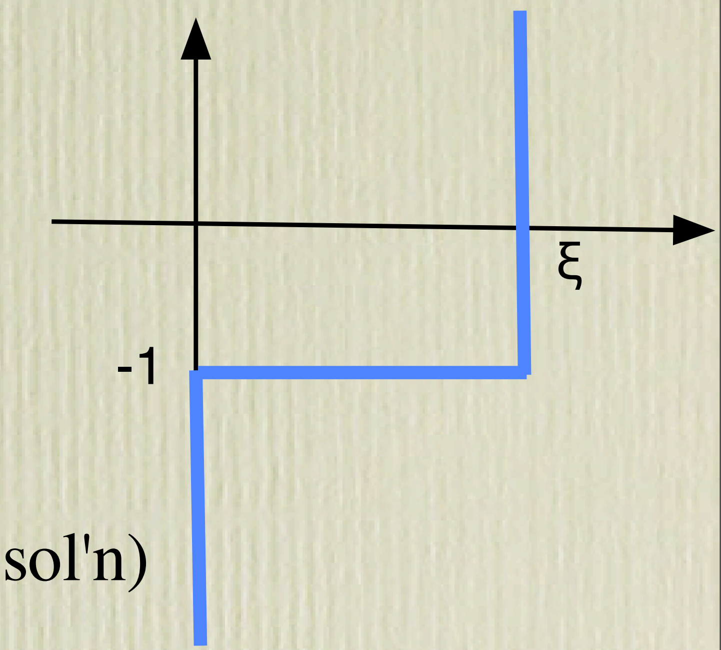
$$\min_{x,y_\xi} \mathbb{E}\left\{-x \,\middle|\, x+y_\xi \le \xi, x\in[0,2], y_\xi \ge 0\right\} = \min\left(Ef(x) = \mathbb{E}\{f(\boldsymbol{\xi},x)\}\right)$$

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

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

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

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



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