Minimax Stochastic Programs and beyond

Jitka Dupačová

Dept. of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic
www.karlin.mff.cuni.cz/ ∼dupacova

International Colloquium on Stochastic Modelling and Optimization
dedicated to the 80th birthday of András Prékopa
Rutcor, Rutgers, December 1, 2009

Content

- The origins of minimax approach
- Class $\mathcal{P}$ defined by generalized moment conditions
- Stability with respect to input information
- Extensions
- Conclusions
- Main quoted references

Basic SP model

CHOOSE the best $\mathbf{x} \in \mathcal{X}$, its outcome depends on realization of random parameter ω and is quantified as $f(\mathbf{x}, \omega)$.

Reformulation $\longrightarrow$ basic SP model

$$\min_{\mathbf{x} \in \mathcal{X}(P)} E_P f(\mathbf{x}, \omega) \tag{1}$$

is identified by

- known probability distribution P of random parameter ω whose support belongs to Ω – a closed subset of $\mathbb{R}^s$;
- a given, nonempty, closed set $\mathcal{X}(P) \subset \mathbb{R}^n$ of decisions $\mathbf{x}$;
  mostly, $\mathcal{X}$ does not depend on P,
  probability (chance) constraints considered separately;
- preselected random objective $f : \mathcal{X}(P) \times \Omega \to \mathbb{R}$ — loss or cost caused by decision $\mathbf{x}$ when scenario ω occurs. Structure of f may be quite complicated (e.g. for multistage problems).

Need to study properties of the model (existence of expectation, convexity, etc.), to get $f, P, \mathcal{X}(P)$, to solve, interprete.

Origins of Minimax SP

Assumption of known P is not realistic, its origin is not clear: wish (test of software, comparisons), generally accepted model (finance, water resources), data, experts opinion, etc.

Suggestion of Jaroslav Hájek $\sim$ 1964 – "what if you try minimax."

ASSUME: P belongs to a specified class $\mathcal{P}$ of probability distributions & APPLY game theoretical approach, or worst-case analysis of SP (1).

FORMULATION (differs from Iosifescu&Theodorescu (1963)):

Incomplete knowledge of P included into the SP model & hedging, e.g.

$$\min_{\mathbf{x}\in\mathcal{X}} \max_{P\in\mathcal{P}} E_P f(\mathbf{x}, \omega). \tag{2}$$

Study specification of $\mathcal{P}$ and sensitivity of results on its choice.

1966 Congress of Econometric Society in Warsaw $\longrightarrow$
discussed with Peter Kall and Roger Wets.

1966 Paper on minimax in English.

$\sim$ 1975 András Prékopa and Mátrafüred.

The minimax paper

Časopis pro pěstování matematiky, roč. 91 (1966), Praha

ON MINIMAX SOLUTIONS OF STOCHASTIC LINEAR PROGRAMMING PROBLEMS

JITKA ŽÁČKOVÁ, Praha

(Received October 12, 1965)

1. Our starting point is the formulation of a stochastic linear program as a strategic game. This formulation differs only slightly from that given by IOSIFESCU and THEODORESCU [3]. Secondly, we state a minimax theorem for that game and study the methods of solution. In some special but important cases it is shown that the minimax solution of a stochastic linear program is equivalent to the solution of an ordinary linear program (of greater dimension, in general). The existence of a finite solution is also discussed.

2. Let E_n^+ denote the non-negative orthant of the n-dimensional Euclidean space.

Three levels of uncertainties

1. Unknown values of coefficients in optimization problems modeled as random, with a **known** probability distribution P – basic SP model
2. Incomplete knowledge of $P \longrightarrow$
   • output analysis wrt. P (e.g. Römisch in Handbook)
   • **minimax approach**, $P \in \mathcal{P}$
   e.g. J.D. 1966–1987, Ben-Tal&Hochman 1972, Jagannathan 1977, Birge & Wets 1987, Bühler $\sim$ 1980, Ermoliev & Gaivoronski $\sim$ 1985, Gaivoronski 1991, Klein Haneveld 1986,
   Shapiro, Takriti, Ahmed, Kleiwegt 2002–2004, Riis 2002–2003, Čerbáková 2003–2008, Popescu 2005–2008, Thiele 2008, Pflug, Wozabal 2007-2008,
   and host of papers related to moment bounds applied in stochastic programming algorithms and output analysis, e.g.
   Edirisinghe&Ziemba, Frauendorfer, Kall, Prékopa and many others.
3. **Vague specification** of $\mathcal{P}$.

Minimax bounds

ASSUME:

1. $\mathcal{X}$ is independent of P, $\mathcal{P}$ independent of decisions $\mathbf{x}$,
2. optimal value $\varphi(P)$ in (1) exists for all $P \in \mathcal{P}$.

Given class $\mathcal{P}$ we want to construct bounds

$$L(\mathbf{x}) = \inf_{P \in \mathcal{P}} E_P f(\mathbf{x}, \omega) \text{ and } U(\mathbf{x}) = \sup_{P \in \mathcal{P}} E_P f(\mathbf{x}, \omega)$$

for values of objective functions (exploited in numerical procedures) and/or minimax and maximax bounds for the optimal value $\varphi(P)$

$$L = \inf_{\mathbf{x} \in \mathcal{X}} \inf_{P \in \mathcal{P}} E_P f(\mathbf{x}, \omega) \leq \varphi(P) \leq \inf_{\mathbf{x} \in \mathcal{X}} \sup_{P \in \mathcal{P}} E_P f(\mathbf{x}, \omega) = U$$

valid for all probability distributions $P \in \mathcal{P}$.
Applicability depends on f and $\mathcal{P}$;

EXPLOIT EXISTING (AVAILABLE) INFORMATION

KEEP THE MINIMAX PROBLEM NUMERICALLY TRACTABLE

Choice of $\mathcal{P}$

Various possibilities have been suggested and elaborated.

Convenient situations:

1. $\mathcal{P}$ is a finite set
2. $\mathcal{P}$ is convex compact;
   then the (linear in P) objective functions $E_P f(\mathbf{x}, \omega)$ attain their infimum and supremum on $\mathcal{P}$,
   the best case and the worst case probability distributions are extremal points of $\mathcal{P}$.

Borel measurability of all functions and sets, as well as existence of expectations will be assumed and we shall focus mostly on $\mathcal{P}$ identified by (generalized) moment conditions and a given "carrier" set Ω; see e.g. Prékopa 1995 for collection and discussion of relevant results.

Frequent choices of $\mathcal{P}$ I.

The listed classes are not strictly separated!

- $\mathcal{P}$ consists of probability distributions on Borel set $\Omega \subseteq R^m$ which fulfill certain moment conditions, e.g.,

$$\mathcal{P}_y = \{P : E_P g_j(\omega) = y_j, j = 1, \dots, J\}$$

  with prescribed values $y_j \; \forall j$, mostly 1st and 2nd order moments.
- Inequalities in (3).
  Interesting idea (Delage&Ye 2008): identify $\mathcal{P}$ by bounds on expectations (μ) and bounds on the covariance matrix, e.g.

$$E_P[(\omega - \mu)(\omega - \mu)^\top] \preceq \gamma\Sigma_0 \text{ for all } P \in \mathcal{P}$$

  and apply approaches of semi-definite programming.
- $\mathcal{P}$ contains probability distributions on Ω with prescribed marginals (Klein Haneveld);
- Additional qualitative information, e.g. unimodality, symmetry or bounded density of P, taken into account, e.g. J.D., Popescu, Shapiro;

Frequent choices of $\mathcal{P}$ II.

- $\mathcal{P}$ consists of probability distributions carried by specified finite set Ω. To get P means to fix the worst case probabilities of considered atoms (scenarios) taking into account a prior knowledge about their partial ordering (Bühler, Čerbáková) or their pertinence to an uncertainty set (Thiele), etc.;
- $\mathcal{P}$ is a neighborhood of a hypothetical, nominal or sample probability distribution P_0 such as the empirical distribution. This means that

$$\mathcal{P} := \{P : d(P, P_0) \leq \varepsilon\}$$

  with $\varepsilon > 0$ and d a suitable distance between P_0 and P. Naturally, results are influenced by choice of d and ε.
  See Calafiore for the Kullback-Leibler divergence between discrete P, P_0 :

$$d_{KL}(P, P_0) := \sum_i p_i \log(p_i / p_i^0)$$

  or Pflug&Wozabal, Wozabal for the Kantorovich distance.
- $\mathcal{P}$ consists of finite number of specified probability distributions $P_1, \dots, P_k$, e.g. Shapiro&Kleywegt; the problem is

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{i=1,\dots,k} F(\mathbf{x}, P_i).$$

$\mathcal{P}_y$ defined by moment conditions

$\mathcal{P} = \mathcal{P}_y$ consists of probability distributions $\Omega \subseteq R^m$ which fulfill certain moment conditions, e.g.

$$\mathcal{P}_y = \{P : E_P g_j(\omega) = y_j, j = 1, \dots, J\} \tag{3}$$

with prescribed values $y_j \; \forall j$, mostly 1st and 2nd order moments.

Also with inequalities in (3), with additional qualitative information, e.g. unimodality, symmetry or bounded density of P, taken into account; cf. J.D., Čerbáková, Popescu, Shapiro.

Allows to exploit classical results on moment problems.

SUGGESTED SOLUTION TECHNIQUES: generalized simplex method, stochastic quasigradients, L-shaped method ...

Basic assumptions

For simplicity assume:
Ω is compact, $g_j\ \forall j$ continuous, $f(\mathbf{x},\bullet)$ upper semicontinuous on Ω, $\mathbf{y} \in \mathcal{Y} := \operatorname{conv}\{\mathbf{g}(\omega), \omega \in \Omega\} \Longrightarrow \mathcal{P}_{\mathbf{y}}$ is convex compact (in weak topology), $\exists$ extremal probability distributions, have finite supports and solution of the "inner" problem

$$U(\mathbf{x},\mathbf{y}) := \max_{P \in \mathcal{P}_y} E_P f(\mathbf{x},\omega) := \int_\Omega f(\mathbf{x},\omega)dP(\omega) \quad \text{subject to}$$

$$\int_\Omega dP(\omega) = 1, \quad \int_\Omega g_j(\omega)dP(\omega) = y_j, \quad j = 1,\dots,J$$

with prescribed $\mathbf{y} \in \mathcal{Y}$ reduces to solution of generalized linear program $\longrightarrow$ atoms of the worst-case probability distribution & their probabilities.

Dual program:

$$\min_{\mathbf{d}} \sum_{j=1}^J d_j y_j + d_0 \quad \text{subject to}$$

$$d_0 + \sum_{j=1}^J d_j g_j(\mathbf{z}) \geq f(\mathbf{x},\mathbf{z}) \quad \forall \mathbf{z} \in \Omega.$$

$U(\mathbf{x},\mathbf{y})$ is a concave function of $\mathbf{y}$ on $\mathcal{Y}$, to be minimized over $\mathbf{x} \in \mathcal{X}$.

Example: Special convex case

For $f(\mathbf{x}, \bullet)$ *convex* function on bounded convex polyhedron $\Omega = \text{conv}\{\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(H)}\} \subset \mathbb{R}^m$, g_j linear,

$$\mathcal{P}_y = \{P : P(\Omega) = 1, E_P\omega_j = y_j, j = 1, \dots, m\}$$

$\mathbf{y}$ given *interior* point of Ω. Constraints of dual problem

$$d_0 + \sum_{j=1}^{m} d_j z_j \geq f(\mathbf{x}, \mathbf{z})$$

hold true $\forall\, \mathbf{z} \in \Omega \iff$ they are fulfilled for all extreme points $\mathbf{z}^{(h)}$ of Ω. By LP duality, the moment problem reduces to linear program

$$\max_{\mathbf{p}} \sum_{h=1}^{H} p_h f(\mathbf{x}, \mathbf{z}^{(h)}) \text{ subject to} \tag{4}$$

$$\sum_{h=1}^{H} p_h z_j^{(h)} = y_j,\ j = 1, \dots, m, \quad \sum_{h=1}^{H} p_h = 1,\ p_h \geq 0\ \forall h.$$

Set of feasible solutions is bounded convex polyhedron $\Longrightarrow \exists$ finite number of worst-case (decision-dependent) probability distributions P^* satisfying moment conditions & carried by extreme points of Ω (cf. Edmunson-Madansky).

EXAMPLE 1. $f(\mathbf{x},\mathbf{z}) = \sum_{j=1}^m f_j(\mathbf{x}, z_j)$, f_j convex functions of $z_j\ \forall j$, $\mathbf{x}$, $\mathcal{P}_y$ is defined by conditions on marginal distributions of ω_j :
carried by given intervals $[a_j, b_j]$ (Ω is their Cartesian product),
$E_P\omega_j = y_j$, with prescribed values $y_j \in (a_j, b_j)\ \forall j \Longrightarrow$

$$\max_{P\in\mathcal{P}_y} f(\mathbf{x},\omega) = \sum_{j=1}^m \lambda_j f_j(\mathbf{x}, a_j) + \sum_{j=1}^m (1-\lambda_j) f_j(\mathbf{x}, b_j) \qquad (5)$$

with $\lambda_j = (b_j - y_j)/(b_j - a_j)$.

$\exists$ extensions to inequality constrained moment problems, non-compact Ω
$\exists$ relaxations to moment problems with given range and expectation etc.

ORIGIN OF PRESCRIBED MOMENTS VALUES AND/OR OF Ω?
estimated, given by regulations, fixed ad hoc ..., vague definition of $\mathcal{P}$
$\longrightarrow$ interest in robustness wrt. changes of $\mathcal{P}$

3rd LEVEL OF UNCERTAINTY

IDEA: Exploit parametric optimization and asymptotic statistics in output analysis wrt. $\mathcal{P}_y$ developed for ordinary stochastic programs. Different techniques needed for nonparametric classes.

ASSUMPTIONS

- $\mathcal{X} \subset \mathbb{R}^n$ is a nonempty convex compact set,
- $\Omega \subset \mathbb{R}^m$ is a nonempty compact set,
- $g_1, \ldots, g_K$ are given continuous functions on Ω,
- $f : \mathcal{X} \times \Omega \to \mathbb{R}^1$ is continuous on Ω for an arbitrary fixed $\mathbf{x} \in \mathcal{X}$ and for every $\omega \in \Omega$ it is a closed convex function of $\mathbf{x}$,
- the interior of the moment space $\mathcal{Y} := \mathrm{conv}\,\{\mathbf{g}(\Omega)\}$ is nonempty.

Basic assumptions $\Longrightarrow$ for fixed $\mathbf{y} \in \mathcal{Y}$ class $\mathcal{P}_{\mathbf{y}}$ is convex and compact (in weak topology) and
for fixed $\mathbf{x}$, $U(\mathbf{x}, \mathbf{y})$ is concave function of $\mathbf{y}$ on $\mathcal{Y}$.

Additional convexity assumptions $\Longrightarrow$ convex - concave function $U(\mathbf{x}, \mathbf{y})$.

Deterministic stability wrt. prescribed moments values

Stability of minimax bound $U(\mathbf{y}) := \min_{\mathbf{x}\in\mathcal{X}} U(\mathbf{x},\mathbf{y})$ follows from results for *nonlinear parametric programming*.
Denote $\mathcal{X}(\mathbf{y})$ set of minimax solutions.

PROPOSITION
If $\mathcal{X} \subset \mathbb{R}^n$ is nonempty, compact, convex $\Rightarrow$

- $U(\mathbf{y})$ is concave on $\mathcal{Y}$,
- mapping $\mathbf{y} \to \mathcal{X}(\mathbf{y})$ is upper semicontinuous on $\mathcal{Y}$.

Directional derivatives exist on int$\mathcal{Y}$ in all directions and gradients of $U(\mathbf{y})$ exist almost everywhere there.
Explicit formulas are available under additional assumptions.

Stability wrt. choice of Ω

Direct analysis of explicit formulas in Example 1 shows that due to changes of Ω the upper bound function $U(\mathbf{x}, \mathbf{y})$ may change substantially.

For probability distributions carried by given finite set of scenarios and in the special convex case, worst case probabilities are obtained as solutions of LP of the type (4) with compact set of feasible solutions. Changes of scenarios or vertices $\mathbf{z}^{(h)}$ influence matrix of coefficients and coefficients in the objective function. Classical stability analysis for LP applies:
For $\mathbf{y}$ interior point of Ω, set of optimal solutions of LP dual to (4)

$$\inf_{\mathbf{d} \in \mathcal{D}} d_0 + \sum_{k=1}^{K} d_k y_k \tag{6}$$

$$\mathcal{D} = \{\mathbf{d} \in \mathbb{R}^{K+1} \, : \, d_0 + \sum_{k=1}^{K} d_k z_k^{(h)} \geq f(\mathbf{x}, \mathbf{z}^{(h)}), h = 1, \ldots, H\} \tag{7}$$

is nonempty and bounded (cf. Kemperman)
$\Longrightarrow$ the LP (4) is stable $\Longrightarrow$
local continuity of its optimal value U wrt. all input coefficients.
(cf. Robinson)
Covers the case of unique, nondegenerated optimal solution of (4).

Stability wrt. choice of Ω – cont.

Another possibility: allow some uncertainty in selection of $\mathbf{z}^{(h)}$:
vertices $\mathbf{z}^h$ belong to ellipsoid around $\mathbf{z}^{(h)}$,

$$\mathbf{z}^{(h)} \to \mathbf{z}^h = \mathbf{z}^{(h)} + \mathbf{E}_h \delta^h, \quad \|\delta^h\|_2 \leq \varrho, \tag{8}$$

best solution of dual LP (6)–(7) which is feasible for all choices of $\mathbf{z}^h$ obtained by perturbations (8). In the simplest case $\mathbf{E}_h = \mathbf{I}$ h-th constraint of (7) is fulfilled if

$$d_0 + \mathbf{d}^\top \mathbf{z}^{(h)} + \mathbf{d}^\top \delta^h - f(\mathbf{x}, \mathbf{z}^{(h)} + \delta^h) \geq 0 \quad \forall \|\delta^h\|_2 \leq \varrho. \tag{9}$$

Lipschitz property of $f(\mathbf{x}, \bullet)$ on neighborhood (8) $\longrightarrow$ $\exists$ constant l s. t.

$$|f(\mathbf{x}, \mathbf{z}^h) - f(\mathbf{x}, \mathbf{z}^{(h)})| \leq l \|\delta^h\|_2 \leq l\varrho.$$

$\Longrightarrow$ To satisfy constraint (9), it is sufficient that

$$d_0 + \mathbf{d}^\top \mathbf{z}^{(h)} - f(\mathbf{x}, \mathbf{z}^{(h)}) - \varrho \sqrt{\|\mathbf{d}\|_2^2 + l^2} \geq 0. \tag{10}$$

When the optimal solution of unperturbed LP (6)–(7) is unique and nondegenerated, then $\exists\, \varrho_{\max} > 0$ such that for all problems with perturbed constraints (10) with $0 < \varrho < \varrho_{\max}$ optimal solutions are unique and nondegenerated.

Comments

• Similar analysis applies to the case of probability distributions $\mathcal{P}_y$ carried by a fixed finite support (under suitable assumptions about the mapping $\mathbf{g}$).

• Even for classes $\mathcal{P}$ which do not assume a known support various assumptions about Ω are exploited in output analysis, e.g.

> *there is a ball of radius R that contains the entire support of the unknown probability distribution; the magnitude of R may follow from "an educated and conservative guess".*

Approximated support – Example

Convergence properties can be given for finite supports which are consecutively improved to approximate the unknown support, cf. Riis&Anderson.

EXAMPLE 2. $\mathcal{P}_y$ is class of probability distributions (3) carried by compact set $\Omega \subset \mathbb{R}^m$, $f : \mathbb{R}^n \times \Omega$ is convex in $\mathbf{x}$ and bounded, continuous in ω.

$\{\Omega^\nu\}_{\nu \geq 1}$ – sequence of finite sets in $\mathbb{R}^m$ such that $\Omega^\nu \subseteq \Omega^{\nu+1} \subseteq \Omega$. (Use additional sample scenarios.) Choose ν_0 such that $\mathbf{y} \in$ int conv$\{\mathbf{g}(\Omega^{\nu_0})\}$. For $\nu \geq \nu_0$ consider classes $\mathcal{P}_y^\nu$ of probability distributions carried by Ω^ν for which moment conditions (3) are fulfilled.
Application of Proposition 2.1 of Riis&Anderson $\Longrightarrow$

If for every $P \in \mathcal{P}_y$ $\exists$ subsequence of $\{P^\nu\}_{\nu \geq \nu_0}$, $P^\nu \in \mathcal{P}_y^\nu$ which converges weakly to P, then for $\nu \to \infty$,

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{P \in \mathcal{P}_y^\nu} E_P f(\mathbf{x}, \omega) \to \min_{\mathbf{x} \in \mathcal{X}} \max_{P \in \mathcal{P}_y} E_P f(\mathbf{x}, \omega)$$

and upper semicontinuity of sets of minimax solutions with respect to the considered convergence of classes $\mathcal{P}_y^\nu$ to $\mathcal{P}_y$ holds true.

Additional input information

Qualitative information such as unimodality – removed by transformation of probability distributions and functions $\longrightarrow$ basic moment problem. Approach for unimodal probability distributions on $\mathbb{R}^1$; and all expectations finite; general case cf. Popescu, Shapiro.

$\mathcal{P}_y^M$ – class of unimodal probability distributions on $\mathbb{R}^1$ with given mode M & moment conditions (3) $\Longrightarrow$ extremal points of $\mathcal{P}_y^M$ are mixtures of uniform distributions over (u, M) or (M, u'), $-\infty < u < M < u' < +\infty$ and of degenerated distribution concentrated at M; support Ω is kept.

h – real function on $\mathbb{R}^1$, integrable over any finite interval of (u, M) and (M, u), h^* – transform of h defined as follows:

$$h^*(z) = \frac{1}{z - M} \int_M^z h(u) du \text{ for } z \neq M \text{ and } h^*(z) = h(z) \text{ for } z = M. \tag{11}$$

Then

$$\tilde{U}(\mathbf{x}, y, M) := \max_{P \in \mathcal{P}_y^M} E_P f(\mathbf{x}, \omega) = \max_P \{E_P f^*(\mathbf{x}, \omega) \; : \; E_P g_j^*(\omega) = y_j, \; \forall j\}. \tag{12}$$

Transform h^* of a convex function h is convex. See J.D.

Additional input information – Example

EXAMPLE 3 – Example 1 with $m = 1$ for unimodal probability distributions with given mode M : Define $\mu = 2y - M$. For $g(u) = u$, $g^*(z) = 1/2(z + M)$, $E_P g^*(\omega) = 1/2(y + M)$. Then $E_P\omega = \mu$ and the transformed moment problem on rhs. of (12) reads

$$U(\mathbf{x}, \mu) = \max_P \{E_P f^*(\mathbf{x}, \omega) \ : \ E_P\omega = \mu, \ P(\omega \in [a, b] = 1)\} := \tilde{U}(\mathbf{x}, y, M)$$

i.e. the usual moment problem over class $\mathcal{P}_\mu$. Transformed objective $f^*(\mathbf{x}, z)$ is convex in $z \Longrightarrow$ maximal expectation $E_P f^*(\mathbf{x}, \omega)$ over $\mathcal{P}_\mu$ of probability distributions on $[a, b]$ with fixed expectation $E_P\omega = 2y - M$ is

$$U(\mathbf{x}, \mu) = \lambda f^*(\mathbf{x}, a) + (1 - \lambda) f^*(\mathbf{x}, b) = \tilde{U}(\mathbf{x}, y, M)$$

with $\lambda = \frac{b-\mu}{b-a} = \frac{b-2y+M}{b-a}$. Substitution for $f^*(\mathbf{x}, z)$ according to (11) $\Longrightarrow$

$$\tilde{U}(\mathbf{x}, y, M) = \frac{b - 2y + M}{(b-a)(M-a)} \int_a^M f(\mathbf{x}, u)du + \frac{2y - M - a}{(b-a)(b-M)} \int_M^b f(\mathbf{x}, u)du.$$

Two densities of uniform distributions weighted by λ resp. $(1 - \lambda)$. Unknown mode – additional maximization wrt. $M \in [a, b]$: worst-case probability distribution is uniform on $[a, b]$ if $y = 1/2(a + b)$ or mixture of uniform distribution over $[a, b]$ and degenerated one concentrated at a or b for $y > 1/2(a + b)$ resp. $y < 1/2(a + b)$.

Stability wrt. estimated moments values

For compact Ω, for $g_j \forall j$ continuous in ω and for $\mathbf{y} \in \mathcal{Y}$ class $\mathcal{P}_y$ is nonempty, convex, compact (in weak topology).
Expectations $E_P f(\mathbf{x}, \omega)$ attain their maxima and minima at extremal points of $\mathcal{P}_y$, i.e. for discrete probability distributions from $\mathcal{P}_y$ carried by no more than $J+1$ points of Ω and for $\mathbf{y} \in \mathcal{Y}$,
$U(\mathbf{x}, \mathbf{y})$ is a finite concave function of $\mathbf{y}$ on $\mathcal{Y}$.
Assume that sample information was used to estimate moments values, or other parameters which identify class $\mathcal{P}_y$. Assume that these parameters $\mathbf{y}$ were consistently estimated e.g. using sequence of iid observations of ω. Let $\mathbf{y}^\nu$ be based on the first ν observations. Using continuity of function $U(\mathbf{x}, \bullet)$ and theorems about transformed random variables, we get for consistent estimates $\mathbf{y}^\nu$ of true parameter $\mathbf{y}$ pointwise convergence

$$u^\nu(\mathbf{x}) := U(\mathbf{x}, \mathbf{y}^\nu) \to U(\mathbf{x}, \mathbf{y}) \text{ a.s.} \tag{13}$$

valid at an arbitrary element $\mathbf{x} \in \mathcal{X}$.
In general, pointwise convergence does neither imply consistency of optimal values $U(\mathbf{y}^\nu) := \min_{\mathbf{x} \in \mathcal{X}} U(\mathbf{x}, \mathbf{y}^\nu)$ nor consistency of minimax solutions.

Use epi-convergence.

Epi-convergence

DEFINITION – Epi-convergence.

A sequence of functions $\{u^\nu : \mathbb{R}^n \to \bar{\mathbb{R}}, \nu = 1, \dots\}$ is said to epi-converge to $u : \mathbb{R}^n \to \bar{\mathbb{R}}$ if for all $\mathbf{x} \in \mathbb{R}^n$ the two following properties hold true:

$$\liminf_{\nu \to \infty} u^\nu(\mathbf{x}^\nu) \geq u(\mathbf{x}) \text{ for all } \{\mathbf{x}^\nu\} \to \mathbf{x} \tag{14}$$

and for some $\{\mathbf{x}^\nu\}$ converging to $\mathbf{x}$

$$\limsup_{\nu \to \infty} u^\nu(\mathbf{x}^\nu) \leq u(\mathbf{x}). \tag{15}$$

Pointwise convergence implies condition (15), additional assumptions are needed to get validity of condition (14).

Fortunately, pointwise convergence of closed, convex functions u, u^ν with $\text{int dom}(u) \neq \emptyset$ implies epi-convergence.
In such case, we also have $\limsup\{\arg\min u^\nu\} \subset \arg\min u$.

Convexity is important!

Stability wrt. estimated moments values – Consistency

Apply this approach to class $\mathcal{P}_y$ defined by generalized moment conditions. $U(\mathbf{x}, \bullet)$ is concave and finite on $\mathcal{Y} \Longrightarrow$ continuous on int$\mathcal{Y}$ $\Longrightarrow$ almost sure pointwise convergence of $u^\nu(\mathbf{x}) \to U(\mathbf{x}, \mathbf{y})$.
Boundedness, continuity and convexity of $f(\mathbf{x}, \omega)$ wrt. $\mathbf{x} \Longrightarrow$ expectations $E_P f(\mathbf{x}, \omega)$ are convex functions of $\mathbf{x}$ for all $P \in \mathcal{P}$.

THEOREM:
Under consistency of estimates $\mathbf{y}^\nu$, continuity properties of $U(\mathbf{x}, \mathbf{y})$ with respect to $\mathbf{y}$, convexity with respect to $\mathbf{x}$, convexity and compactness of $\mathcal{X}$ (see ASSUMPTIONS) approximate objectives $u^\nu(\mathbf{x})$ epi-converge almost surely to $U(\mathbf{x}, \mathbf{y})$ as $\nu \to \infty \Longrightarrow$ with probability 1 all cluster points of sequence of minimizers $\mathbf{x}^\nu$ of $u^\nu(\mathbf{x})$ on $\mathcal{X}$ are minimizers of $U(\mathbf{x}, \mathbf{y})$ on $\mathcal{X}$ and $\min_{x \in \mathcal{X}} u^\nu(\mathbf{x}) \to \min_{x \in \mathcal{X}} U(\mathbf{x}, \mathbf{y})$.

EXAMPLES:
•$\mathcal{P}_y$ defined by moment conditions (3), fixed compact Ω & f convex in $\mathbf{x}$;
• Special convex case for f convex in $\mathbf{x}$ with perturbed $\mathbf{y} \in$ int$\mathcal{Y}$ & Ω;
• Similar result holds true also for the "data-driven" version of Example 3 (unimodal probability distributions with estimated mean).

Consistency – Example

Stability wrt. choice of Ω in one-dimensional case; exploit epi-consistency ideas under special circumstances:

EXAMPLE 4 – related to Example 1.
Parameters a, b, μ identifying the class of one-dimensional probability distributions on interval $[a, b]$ with mean value μ are known to belong to the interior of a compact set and their values can be obtained by estimation procedure based on a sample path of iid observations of ω from the true probability distribution P. Their consistent estimates based on a sample size ν are the minimal/maximal sample values and the arithmetic mean, i.e. $\omega_{\nu:1}, \omega_{\nu:\nu}$ and $\mu^\nu = 1/\nu \sum_{i=1}^{\nu} \omega_i$. We know explicit form of all approximate objective functions

$$u^\nu(\mathbf{x}) := \lambda^\nu f(\mathbf{x}, \omega_{\nu:1}) + (1 - \lambda^\nu) f(\mathbf{x}, \omega_{\nu:\nu})$$

with $\lambda^\nu = (\omega_{\nu:\nu} - \mu^\nu)/(\omega_{\nu:\nu} - \omega_{\nu:1})$; see Example 1 for $m = 1$. This is continuous function of parameters provided that $\omega_{\nu:1} < \omega_{\nu:\nu}$. For convex $f(\bullet, \omega)$, $u^\nu(\mathbf{x})$ are convex in $\mathbf{x}$ and epi-converge to $u(\mathbf{x})$. For compact set $\mathcal{X}$, existence of the true minimax solution $\mathbf{x}$ follows from continuity of $f(\bullet, a)$ and $f(\bullet, b)$.

Extensions

Stochastic programs whose set of feasible solutions depends on P :

$$\text{minimize } F(\mathbf{x}, P) := E_P f(\mathbf{x}, \omega) \text{ on the set } \mathcal{X}(P) \tag{16}$$

where $\mathcal{X}(P) = \{\mathbf{x} \in \mathcal{X} \,:\, G(\mathbf{x}, P) \leq 0\}$, e.g. probabilistic programs, risk or stochastic dominance constraints, VaR.
Incomplete knowledge of P – solve "robustified" version of (16), cf. Pflug& Wozabal, Dentcheva& Ruszczyński:

$$\min_{\mathbf{x} \in \mathcal{X}} \max\{F(\mathbf{x}, P) \,:\, P \in \mathcal{P}\} \tag{17}$$

subject to $G(\mathbf{x}, P) \leq 0 \, \forall P \in \mathcal{P}$ or equivalently, subject to

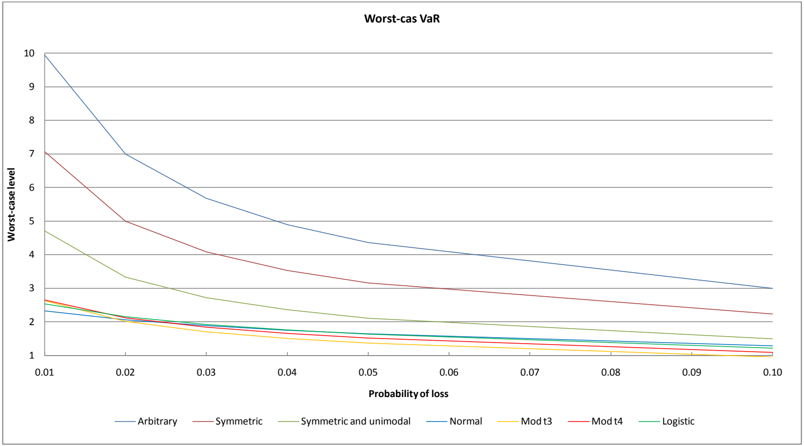
$$\max_{P \in \mathcal{P}} G(\mathbf{x}, P) \leq 0. \tag{18}$$

Results of moment problem apply again when $G(\mathbf{x}, P)$ is convex in $\mathbf{x}$ and linear in P, e.g. portfolio optimization with CVaR constraints, calculation of worst-case VaR

$$\text{VaR}_\alpha^{wc}(\omega, \mathcal{P}) = \min k \text{ subject to } \sup_{P \in \mathcal{P}} P(\omega \geq k) \leq \alpha.$$

Depends on choice of class $\mathcal{P}$, additional information (unimodality, symmetry) – Example:

Worst-case VaR; mean 0, variance 1



Functionals nonlinear in P

For convex, compact class $\mathcal{P}$ and for fixed $\mathbf{x}$, the maxima in (17), (18) are attained at extremal points of $\mathcal{P}$; hence for the class $\mathcal{P}_y$ identified by moment conditions (3) (and under mild assumptions), it is possible to work with discrete distributions $P \in \mathcal{P}$. This property carries over also to $G(\mathbf{x}, P)$ in (18) and/or $F(\mathbf{x}, P)$ in (17) convex in P.

WARNING:
Whereas expected utility functions or $\mathrm{CVaR}_\alpha(\mathbf{x}, P)$ are linear in P, other popular portfolio characteristics are even not convex in P: the variance is concave in P, mean absolute deviation is neither convex nor concave in P. This means that extensions to risk functionals nonlinear in P carry over only under special circumstances.

EXAMPLE 5.
ω – random vector of unit returns of assets included to portfolio, $f(\mathbf{x}, \omega) = -\omega^\top \mathbf{x}$ quantifies random loss of investment $\mathbf{x}$.
Probability distribution P of ω is known to belong to a class $\mathcal{P}$ of distributions for which i.a. expectation $E_P\omega = \mu$ is fixed (independent of P). Then for fixed $\mathbf{x}$, $\mathrm{var}_P f(\mathbf{x}, \omega) = E_P(\omega^\top \mathbf{x})^2 - (\mu^\top \mathbf{x})^2$ and mean absolute deviation $\mathrm{MAD}_P f(\mathbf{x}, \omega) = E_P|\omega^\top \mathbf{x} - \mu^\top \mathbf{x}|$ are linear in P.

Conclusions

The presented approach to stability analysis of minimax stochastic programs with respect to input information was elaborated for class $\mathcal{P}$ defined by generalized moment conditions (3) and a given carrier set Ω. It is suitable also for other "parametric" classes $\mathcal{P}$. Stability for "nonparametric" classes, e.g. Pflug&Wozabal, would require different techniques.
We did not aim at the most general statements and results on stability and sensitivity of minimax bounds and minimax decisions with respect to the model input. Various convexity assumptions were exploited:

- convexity and compactness of class $\mathcal{P}_{\mathbf{y}}$,
- convexity of random objective function $f(\mathbf{x}, \omega)$ with respect to $\mathbf{x}$ on a compact convex set of feasible decisions,
- convexity of functionals $F(\mathbf{x}, P)$, $G(\mathbf{x}, P)$ with respect to probability distribution P.

Convexity of random objective with respect to $\mathbf{x}$ can be replaced by saddle property and under suitable conditions, also unbounded sets $\mathcal{X}$ can be treated. Open question: under what assumptions the presented approach can be applied to minimax problems with functionals nonconvex in P.

Selected references

1. J. Dupačová (1987), The minimax approach to stochastic programming and an illustrative application, *Stochastics* **20** 73–88.
2. J. Dupačová (1997), Moment bounds for stochastic programs in particular for recourse problems. In: V. Beneš, J. Štěpán (eds.) *Distributions with given Marginals and Moment Problems*, Kluwer, pp. 199–204.
3. J. Dupačová (2001), Stochastic programming: minimax approach. In: Ch. A. Floudas, P. M. Pardalos (eds.) *Encyclopedia of Optimization*, Vol. V., pp.327–330 and references therein.
4. J. Dupačová (2009), Uncertainties in minimax stochastic programs, SPEPS 9.
5. G. Ch. Pflug, D. Wozabal (2007), Ambiguity in portfolio selection, *Quant. Fin.* **7**, 435–442.
6. I. Popescu (2007), A semidefinite programming approach to optimal moment bounds for convex classes of distributions, *Math. Oper. Res.* **30**.
7. M. Riis, K. A. Andersen (2005), Applying the minimax criterion in stochastic recourse programs, *European J. of Oper. Res.* **165**, 569–584.
8. A. Shapiro, S. Ahmed (2004), On a class of minimax stochastic programs, *SIAM J. Optim.* **14**, 1237–1249.
9. A. Shapiro, A. Kleywegt (2002), Minimax analysis of stochastic programs, *GOMS* **17**, 532–542.