Stability of Stochastic Programming Problems

W. Römisch

Humboldt-University Berlin
Institute of Mathematics
10099 Berlin, Germany

http://www.math.hu-berlin.de/~romisch

Spring School Stochastic Programming, Bergamo, April 13, 2007

DFG Research Center MATHEON
Mathematics for key technologies

Contents:

1. Introduction

Consider the stochastic programming model

$$\min \left\{ \int_\Xi f_0(\xi, x) P(d\xi) : x \in M(P) \right\}$$

$$M(P) := \left\{ x \in X : \int_\Xi f_j(\xi, x) P(d\xi) \leq 0, j = 1, ..., d \right\}$$

where f_j from $\Xi \times \mathbb{R}^m$ to the extended reals $\overline{\mathbb{R}}$ are normal integrands, X is a nonempty closed subset of $\mathbb{R}^m$, Ξ is a closed subset of $\mathbb{R}^s$ and P is a Borel probability measure on Ξ.

(f is a normal integrand if it is Borel measurable and $f(\xi, .)$ is lower semicontinuous $\forall \xi \in \Xi$.)

Let $\mathcal{P}(\Xi)$ the set of all Borel probability measures on Ξ and by

$$v(P) = \inf_{x \in M(P)} \int_\Xi f_0(\xi, x) P(d\xi) \quad \text{(optimal value)}$$

$$S_\varepsilon(P) = \left\{ x \in M(P) : \int_\Xi f_0(\xi, x) P(d\xi) \leq v(P) + \varepsilon \right\}$$

$$S(P) = S_0(P) = \arg \min_{x \in M(P)} \int_\Xi f_0(\xi, x) P(d\xi) \quad \text{(solution set).}$$

The underlying probability distribution P is often incompletely known in applied models and/or has to be approximated (estimated, discretized).

$\longrightarrow$ stability behaviour of stochastic programs becomes important when changing (perturbing, estimating, approximating) $P \in \mathcal{P}(\Xi)$.

Here, stability refers to (quantitative) continuity properties of the optimal value function $v(.)$ and of the set-valued mapping $S_\varepsilon(.)$ at P, where both are regarded as mappings given on certain subset of $\mathcal{P}(\Xi)$ equipped with some convergence of probability measures and some probability metric, respectively.

(The corresponding subset of probability measures is determined such that certain moment conditions are satisfied that are related to growth properties of the integrands f_j with respect to ξ.)

Examples: Two-stage stochastic programs, chance constrained stochastic programs.

Survey:

W. Römisch: Stability of stochastic programming problems, in: Stochastic Programming (A. Ruszczynski, A. Shapiro eds.), Handbook, Elsevier, 2003.

Weak convergence in $\mathcal{P}(\Xi)$

$$P_n \to_w P \text{ iff } \int_\Xi f(\xi)P_n(d\xi) \to \int_\Xi f(\xi)P(d\xi) \quad (\forall f \in C_b(\Xi)),$$
$$\text{iff } P_n(\{\xi \le z\}) \to P(\{\xi \le z\}) \text{ at continuity points } z \text{ of } P(\{\xi \le \cdot\}).$$

Probability metrics on $\mathcal{P}(\Xi)$ (Monographs: Rachev 91, Rachev/Rüschendorf 98)

Metrics with ζ-structure:

$$d_{\mathcal{F}}(P,Q) = \sup\left\{\left|\int_\Xi f(\xi)P(d\xi) - \int_\Xi f(\xi)Q(d\xi)\right| : f \in \mathcal{F}\right\}$$

where $\mathcal{F}$ is a suitable set of measurable functions from Ξ to $\overline{\mathbb{R}}$ and P, Q are probability measures in some set $\mathcal{P}_{\mathcal{F}}$ on which $d_{\mathcal{F}}$ is finite.

Examples (of $\mathcal{F}$): Sets of locally Lipschitzian functions on Ξ or of piecewise (locally) Lipschitzian functions.

There exist canonical sets $\mathcal{F}$ and metrics $d_{\mathcal{F}}$ for each specific class of stochastic programs!

2. General quantitative stability results

To simplify matters, let X be compact (otherwise, consider localizations).

$$
\begin{aligned}
\mathcal{F} &:= \{f_j(.,x) \,:\, x \in X, j = 0,\ldots,d\}, \\
\mathcal{P}_{\mathcal{F}} &:= \{Q \in \mathcal{P}(\Xi) : \int_{\Xi} \inf_{x \in X} f_j(\xi,x)Q(d\xi) > -\infty, \\
&\qquad \sup_{x \in X} \int_{\Xi} f_j(\xi,x)Q(d\xi) < \infty,\ j = 0,\ldots,d\},
\end{aligned}
$$

and the probability (semi-) metric on $\mathcal{P}_{\mathcal{F}}$:

$$
d_{\mathcal{F}}(P,Q) = \sup_{x \in X} \max_{j=0,\ldots,d} \left| \int_{\Xi} f_j(\xi,x)(P-Q)(d\xi) \right|.
$$

Lemma:

The functions $(x,Q) \mapsto \int_{\Xi} f_j(\xi,x)Q(d\xi)$ are lower semicontinuous on $X \times \mathcal{P}_{\mathcal{F}}$.

Theorem: (Rachev-Römisch 02)

If $d \geq 1$, let the function $x \mapsto \int_\Xi f_0(\xi,x)P(d\xi)$ be Lipschitz continuous on X, and, let the function

$$(x,y) \mapsto d\left(x, \left\{\tilde{x} \in X : \int_\Xi f_j(\xi,\tilde{x})P(d\xi) \leq y_j, j = 1,...,d\right\}\right)$$

be locally Lipschitz continuous around $(\bar{x},0)$ for every $\bar{x} \in S(P)$ (regularity condition).

Then there exist constants L, $\delta > 0$ such that

$$\begin{aligned} |v(P) - v(Q)| &\leq Ld_{\mathcal{F}}(P,Q) \\ S(Q) &\subseteq S(P) + \Psi_P(Ld_{\mathcal{F}}(P,Q))\mathbb{B} \end{aligned}$$

holds for all $Q \in \mathcal{P}_{\mathcal{F}}$ with $d_{\mathcal{F}}(P,Q) < \delta$.

Here $\Psi_P(\eta) := \eta + \psi^{-1}(\eta)$ and $\psi : \mathbb{R}_+ \to \mathbb{R}_+$

$$\psi(\tau) := \min\left\{\int_\Xi f_0(\xi,x)P(d\xi) - v(P) : d(x,S(P)) \geq \tau, x \in M(P)\right\}.$$

(Proof by appealing to general perturbation results of Klatte 94 and Rockafellar/Wets 98.)

Convex case and $d := 0$:

Assume that $f_0(\xi, \cdot)$ is convex on $\mathbb{R}^m$ $\forall \xi \in \Xi$.

Theorem: (Römisch-Wets 06)

Then there exist constants L, $\bar{\varepsilon} > 0$ such that

$$d\!l_\infty(S_\varepsilon(P), S_\varepsilon(Q)) \leq \frac{L}{\varepsilon} d_{\mathcal{F}}(P, Q)$$

for every $\varepsilon \in (0, \bar{\varepsilon})$ and $Q \in \mathcal{P}_{\mathcal{F}}$ such that $d_{\mathcal{F}}(P, Q) < \varepsilon$.

Here, $d\!l_\infty$ is the Pompeiu-Hausdorff distance of nonempty closed subsets of $\mathbb{R}^m$, i.e.,

$$d\!l_\infty(C, D) = \inf\{\eta \geq 0 : C \subseteq D + \eta \mathbb{B}, D \subseteq C + \eta \mathbb{B}\}.$$

Proof using a perturbation result by Rockafellar/Wets 98.

The (semi-) distance $d_{\mathcal{F}}$ plays the role of a minimal probability metric implying quantitative stability.

Furthermore, the result remains valid when bounding $d_{\mathcal{F}}$ from above by another distance and when reducing the set $\mathcal{P}_{\mathcal{F}}$ to a subset on which this distance is defined and finite.

Idea: Enlarge $\mathcal{F}$, but maintain the analytical (e.g., (dis)continuity) properties of $f_j(\cdot, x)$, $j = 0, \ldots, d$!

This idea may lead to well-known probability metrics, for which a well developed theory is available !

Example: (Fortet-Mourier-type metrics)
We consider the following classes of locally Lipschitz continuous functions (on Ξ)

$$\mathcal{F}_H := \{f : \Xi \to \mathbb{R} \ : \ f(\xi) - f(\tilde{\xi}) \le \max\{1, H(\|\xi\|), H(\|\tilde{\xi}\|)\} \cdot \|\xi - \tilde{\xi}\|, \forall \xi, \tilde{\xi} \in \Xi\}$$

are of particular interest, where $H : \mathbb{R}_+ \to \mathbb{R}_+$ is nondecreasing, $H(0) = 0$. The corresponding distances are

$$d_{\mathcal{F}_H}(P, Q) = \sup_{f \in \mathcal{F}_H} \left| \int_\Xi f(\xi) P(d\xi) - \int_\Xi f(\xi) Q(d\xi) \right| =: \zeta_H(P, Q)$$

are so-called Fortet-Mourier-type metrics defined on

$$\mathcal{P}_H(\Xi) := \{Q \in \mathcal{P}(\Xi) : \int_\Xi \max\{1, H(\|\xi\|)\} \|\xi\| Q(d\xi) < \infty\}$$

Important special case: $H(t) := t^{p-1}$ for $p \ge 1$.
The corresponding classes of functions and measures, and the distances are denoted by $\mathcal{F}_p$, $\mathcal{P}_p(\Xi)$ and ζ_p, respectively.
(Convergence with respect to ζ_p means weak conergence of the probability measures and convergence of the p-th order moments (Rachev 91))

Application: Convergence of empirical estimates

Let $P \in \mathcal{P}(\Xi)$ and let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be independent, identically distributed Ξ-valued random variables on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ having the common distribution P.
Let $P_n(\omega) := \frac{1}{n}\sum_{i=1}^{n} \delta_{\xi_i(\omega)}$ be the empirical measures $\forall n \in \mathbb{N}$.
We consider the empirical estimates or sample average approximation (replacing P by $P_n(\cdot)$):

$$\min \left\{ \frac{1}{n}\sum_{i=1}^{n} f_0(\xi_i, x) : x \in X, \frac{1}{n}\sum_{i=1}^{n} f_j(\xi_i, x) \le 0, j = 1, \ldots, d \right\}$$

Then results on the convergence in probability of

$$d_{\mathcal{F}}(P, P_n(\cdot))$$

and, hence, of

$$|v(P) - v(P_n(\cdot))|$$

may be obtained using the general stability results, empirical process theory and covering numbers for $\mathcal{F}$ as subsets of $L_p(\Xi, P)$.

3. Two-stage stochastic programming models

We consider the two-stage stochastic program

$$\min \left\{ \langle c, x\rangle + \int_{\Xi} \hat{\Phi}(\xi, x) P(d\xi) : x \in X \right\},$$

where

$$\hat{\Phi}(\xi, x) := \inf\{\langle q(\xi), y\rangle : y \in Y, W(\xi)y = h(\xi) - T(\xi)x\}$$

$P := \mathbb{P}\xi^{-1} \in \mathcal{P}(\Xi)$ is the probability distribution of the random vector ξ, $c \in \mathbb{R}^m$, $X \subseteq \mathbb{R}^m$ is a bounded polyhedron, $q(\xi) \in \mathbb{R}^{\overline{m}}$, $Y \in \mathbb{R}^{\overline{m}}$ is a polyhedral cone, $W(\xi)$ a $r \times \overline{m}$-matrix, $h(\xi) \in \mathbb{R}^r$ and $T(\xi)$ a $r \times m$-matrix.

We assume that $q(\xi)$, $h(\xi)$, $W(\xi)$ and $T(\xi)$ are affine functions of ξ (e.g., some of their components or elements are random).

Example:(two-stage model with simple recourse)

$m = s = 1,\ d = 0,\ f_0(\xi, x) := \max\{0, \xi - x\},$
$\Xi := \mathbb{R},\ X := [-1, 1],$

$P := \delta_0$ (unit mass at 0),
$P_n := (1 - \frac{1}{n})\delta_0 + \frac{1}{n}\delta_{n^2}\ ,\ n \in \mathbb{N}.$

$$\int_\Xi f_0(\xi, x)P(d\xi) = \begin{cases} -x & , x \in [-1, 0) \\ 0 & , x \in [0, 1] \end{cases}$$
$v(P) = 0,\ S(P) = [0, 1],$

$$\int_\Xi f_0(\xi, x)P_n(d\xi) = (1 - \tfrac{1}{n})\max\{0, -x\} + \tfrac{1}{n}\max\{0, n^2 - x\}$$
$v(P_n) = n - \frac{1}{n},\ S(P_n) = \{1\}$ $(n \in \mathbb{N})$.

Note: $P_n \to^w P$, but first order moments do not converge !

$$\int_\Xi f(\xi)P_n(d\xi) = (1 - \tfrac{1}{n})f(0) + \tfrac{1}{n}f(n^2) \longrightarrow f(0),\ \forall f \in C_b(\Xi).$$
But, $\int_\Xi |\xi|P_n(d\xi) = \frac{1}{n}n^2 = n$ $(n \in \mathbb{N})$ and $\int_\Xi |\xi|P(d\xi) = 0$

Structural properties of two-stage models

We consider the infimum function of the parametrized linear (second-stage) program and the dual feasible set of the second-stage program, namely,

$$\Phi(\xi,u,t):=\inf\{\langle u,y\rangle : W(\xi)y=t, y\in Y\}\ ((\xi,u,t)\in\Xi\times\mathbb{R}^{\overline{m}}\times\mathbb{R}^r)$$
$$D(\xi):=\{z\in\mathbb{R}^r : W(\xi)^\top z-q(\xi)\in Y^*\}\ (\xi\in\Xi),$$

where $W(\xi)^\top$ is the transposed of $W(\xi)$ and Y^* the polar cone of Y (i.e., $Y^*=\{y^*:\langle y^*,y\rangle\leq 0,\forall y\in Y\}$). Then we have

$$\hat{\Phi}(\xi,x)=\Phi(\xi,q(\xi),h(\xi)-T(\xi)x)=\sup\{\langle h(\xi)-T(\xi)x,z\rangle : z\in D(\xi)\}.$$

Theorem: (Walkup/Wets 69)
For any $\xi\in\Xi$, the function $\Phi(\xi,\cdot,\cdot)$ is finite and continuous on the polyhedral set $D(\xi)\times W(\xi)Y$. Furthermore, the function $\Phi(\xi,u,\cdot)$ is piecewise linear convex on the polyhedral set $W(\xi)Y$ for fixed $u\in D(\xi)$, and $\Phi(\xi,\cdot,t)$ is piecewise linear concave on $D(\xi)$ for fixed $t\in W(\xi)Y$.

Assumptions:

(A1) *relatively complete recourse*: for any $(\xi, x) \in \Xi \times X$, $h(\xi) - T(\xi)x \in W(\xi)Y$;

(A2) *dual feasibility*: $D(\xi) \neq \emptyset$ holds for all $\xi \in \Xi$.

Note that (A1) is satisfied if $W(\xi)Y = \mathbb{R}^r$ (complete recourse). In general, (A1) and (A2) impose a condition on the support of P.

Proposition:

Then the deterministic equivalent of the two-stage model represents a finite convex program (with polyhedral constraints) if the integrals $\int_\Xi \Phi(\xi, q(\xi), h(\xi) - T(\xi)x) P(d\xi)$ are finite for all $x \in X$.

For fixed recourse ($W(\xi) \equiv W$), it suffices to assume

$$\int_\Xi \|\xi\|^2 P(d\xi) < \infty.$$

Convex subdifferentials, optimality conditions, conditions for differentiability, duality results are well known.

(Ruszczyński/Shapiro, Handbook, 2003)

Towards stability

We define the integrand $f_0 : \Xi \times \mathbb{R}^m \to \overline{\mathbb{R}}$ by

$$f_0(\xi, x) = \begin{cases} \langle c, x\rangle + \Phi(\xi, q(\xi), h(\xi) - T(\xi)x) & \text{if } h(\xi) - T(\xi)x \in W(\xi)Y,\ D(\xi) \neq \emptyset, \\ +\infty & \text{otherwise,} \end{cases}$$

and note that f_0 is a convex random lsc function with $\Xi \times X \subseteq \operatorname{dom} f_0$ if (A1) and (A2) are satisfied.

The two-stage stochastic program can thus be expressed as

$$\min \left\{ \int_\Xi f_0(\xi, x) P(d\xi) : x \in X \right\}.$$

Then the general stability theory applies !

Simple examples of two-stage stochastic programs show that, in general, the set-valued mapping $S(.)$ is not inner semicontinuous at P. Furthermore, explicit descriptions of conditioning functions ψ_P of stochastic programs (like linear or quadratic growth at solution sets) are only known in some specific cases.

Stability w.r.t. the Fortet-Mourier metric ζ_H

Proposition:
Suppose the stochastic program satisfies (A1) and (A2). Assume that the mapping $\xi \mapsto D(\xi)$ is bounded-valued and there exists a constant $L > 0$, and a nondecreasing function $h : \mathbb{R}_+ \to \mathbb{R}_+$ with $h(0) = 0$ such that

$$d\!\!l_\infty(D(\xi), D(\tilde{\xi})) \leq L \max\{1, h(\|\xi\|), h(\|\tilde{\xi}\|)\}\|\xi - \tilde{\xi}\|$$

holds for all $\xi, \tilde{\xi} \in \Xi$.
Then there exist $\hat{L} > 0$ such that

$$\begin{aligned} |f_0(\xi, x) - f_0(\tilde{\xi}, x)| &\leq \hat{L} \max\{1, H(\|\xi\|), H(\|\tilde{\xi}\|)\}\|\xi - \tilde{\xi}\| \\ |f_0(\xi, x) - f_0(\xi, \tilde{x})| &\leq \hat{L} \max\{1, H(\|\xi\|)\|\xi\|\}\|x - \tilde{x}\| \end{aligned}$$

for all $\xi, \tilde{\xi} \in \Xi$, $x, \tilde{x} \in X$, where H is defined by

$$H(t) := h(t)t, \; \forall t \in \mathbb{R}_+.$$

Note that $h(t) = \begin{cases} 1 & \text{, fixed recourse} \\ t^k & \text{, lower diagonal randomness with } k \text{ blocks.} \end{cases}$

Discrete approximations of two-stage stochastic programs

Replace the (original) probability measure P by measures P_n having (finite) discrete support $\{\xi_1, \dots, \xi_n\}$ ($n \in \mathbb{N}$), i.e.,

$$P_n = \sum_{i=1}^{n} p_i \delta_{\xi_i},$$

and insert it into the infinite-dimensional stochastic program:

$$\min\{\langle c, x\rangle + \sum_{i=1}^{n} p_i \langle q(\xi_i), y_i\rangle : x \in X, y_i \in Y, \; i = 1, \dots, n,$$

$$\begin{array}{cccccc} W(\xi_1)y_1 & & & +T(\xi_1)x & = & h(\xi_1) \\ & W(\xi_2)y_2 & & +T(\xi_2)x & = & h(\xi_2) \\ & & \ddots & \vdots & = & \vdots \\ & & W(\xi_n)y_n & +T(\xi_n)x & = & h(\xi_n)\} \end{array}$$

Hence, we arrive at a (finite-dimensional) large scale block-structured linear program which allows for specific decomposition methods.

(Ruszczyński/Shapiro, Handbook, 2003)

How to choose the discrete approximation ?

The quantitative stability results suggest to determine P_n such that it forms the best approximation of P with respect to the semi-distance $d_{\mathcal{F}}$ or the probability metric ζ_p, i.e., given $n \in \mathbb{N}$ solve

$$\min\left\{\zeta_p\left(P, \frac{1}{n}\sum_{i=1}^{n}\delta_{\xi_i}\right) : \xi_i \in \Xi,\, i = 1, \ldots, n\right\}$$

Such best approximations $P_n^* = \frac{1}{n}\sum_{i=1}^{n}\delta_{\xi_i^*}$ are known as optimal quantizations of the probability distribution P (Graf/Luschgy, Lecture Notes Math. 1730 2000).

Convergence properties of optimal quantizations in case of the ℓ_p-minimal metrics (or Wasserstein metrics)

$$\ell_p(P,Q) := \left(\inf\left\{\int_{\Xi\times\Xi}\|\xi - \tilde{\xi}\|^p \eta(d\xi, d\tilde{\xi}) \mid \pi_1\eta = P,\ \pi_2\eta = Q\right\}\right)^{\frac{1}{p}},$$

are already known. Here, π_i is the projection onto the i-th component. Note that $\zeta_p(P,Q) \le (1 + \int_{\Xi}\|\xi\|^p(P+Q)(d\xi))^{\frac{p-1}{p}}\ell_p(P,Q)$.

Scenario reduction

We consider discrete distributions P with scenarios ξ_i and probabilities p_i, $i = 1, \ldots, N$, and Q being supported by a given subset of scenarios ξ_j, $j \notin J \subset \{1, \ldots, N\}$, of P.

Optimal reduction of a given scenario set J:
The best approximation of P with respect to ζ_r by such a distribution Q exists and is denoted by Q^*. It has the distance

$$\begin{aligned} D_J \;:=\; & \zeta_r(P, Q^*) = \min_Q \zeta_r(P, Q) = \sum_{i \in J} p_i \min_{j \notin J} \hat{c}_r(\xi_i, \xi_j) \\ = \;& \sum_{i \in J} p_i \min\{\sum_{k=1}^{n-1} c_r(\xi_{l_k}, \xi_{l_{k+1}}) : n \in \mathbb{N}, l_k \in \{1, \ldots, N\}, \\ & \qquad\qquad\qquad\qquad l_1 = i, l_n = j \notin J\} \end{aligned}$$

and the probabilities $q_j^* = p_j + \sum_{i \in J_j} p_i$, $\forall j \notin J$, where

$J_j := \{i \in J : j = j(i)\}$ and $j(i) \in \arg\min_{j \notin J} \hat{c}_r(\xi_i, \xi_j)$, $\forall i \in J$.

(Dupačová-Gröwe-Kuska-Römisch 03, Heitsch-Römisch 07)

We needed the following notation:

$$c_r(\xi,\tilde{\xi}) := \max\{1, \|\xi\|^{r-1}, \|\tilde{\xi}\|^{r-1}\}\|\xi-\tilde{\xi}\| \quad (\xi,\tilde{\xi}\in\Xi).$$

Proposition: (Rachev/Rüschendorf 98)

$$\zeta_r(P,Q) = \inf\left\{\int_{\Xi\times\Xi}\hat{c}_r(\xi,\tilde{\xi})\eta(d\xi,d\tilde{\xi}) : \pi_1\eta = P, \pi_2\eta = Q\right\}$$

where $\hat{c}_r \le c_r$ and $\hat{c}_r$ is the metric (reduced cost)

$$\hat{c}_r(\xi,\tilde{\xi}) := \inf\left\{\sum_{i=1}^{k-1} c_r(\xi_{l_i},\xi_{l_{i+1}}) : k\in\mathbb{N}, \xi_{l_i}\in\Xi, \xi_{l_1}=\xi, \xi_{l_k}=\tilde{\xi}\right\}.$$

Determining the optimal scenario index set with prescribed cardinality n is, however, a combinatorial optimization problem of set covering type:

$$\min\{D_J = \sum_{i\in J} p_i \min_{j\notin J}\hat{c}_r(\xi_i,\xi_j) : J\subset\{1,\ldots,N\}, \#J = N-n\}$$

Hence, the problem of finding the optimal set J to delete is $\mathcal{NP}$-hard and polynomial time solution algorithms do not exist.

Fast reduction heuristic

Starting point ($n = 1$): $\min\limits_{u \in \{1,\ldots,N\}} \sum\limits_{k=1}^{N} p_k \hat{c}_r(\xi_k, \xi_u)$

Algorithm: (Forward selection)

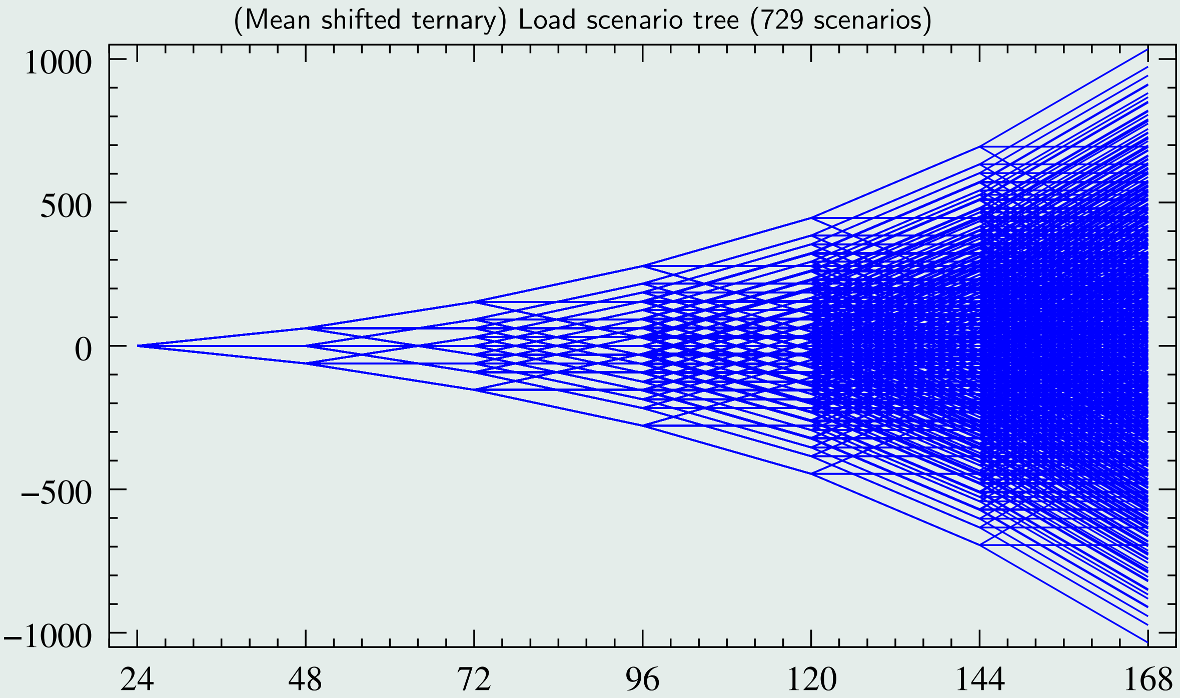
Step [0]: $J^{[0]} := \{1, \ldots, N\}$.

Step [i]: $u_i \in \arg \min\limits_{u \in J^{[i-1]}} \sum\limits_{k \in J^{[i-1]} \setminus \{u\}} p_k \min\limits_{j \notin J^{[i-1]} \setminus \{u\}} \hat{c}_r(\xi_k, \xi_j),$

$J^{[i]} := J^{[i-1]} \setminus \{u_i\}$.

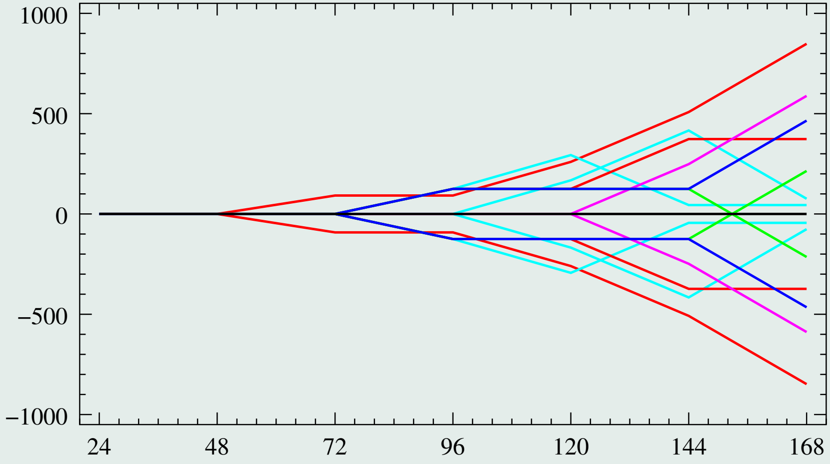
Step [n+1]: Optimal redistribution.

Example: (Electrical load scenario tree)



<Start Animation>

Reduced load scenario tree obtained by the forward selection method (15 scenarios)



4. Chance constrained stochastic programs

We consider the chance constrained model

$$\min\{\langle c, x\rangle : x \in X, P(\{\xi \in \Xi : T(\xi)x \geq h(\xi)\}) \geq p\},$$

where $c \in \mathbb{R}^m$, X and Ξ are polyhedra in $\mathbb{R}^m$ and $\mathbb{R}^s$, respectively, $p \in (0,1)$, $P \in \mathcal{P}(\Xi)$, and the right-hand side $h(\xi) \in \mathbb{R}^d$ and the (d,m)-matrix $T(\xi)$ are affine functions of ξ.

By specifying the general (semi-) distance we obtain

$$\begin{aligned} d_{\mathcal{F}}(P,Q) &:= \sup_{x \in X} \max_{j=0,1} \left| \int_{\Xi} f_j(x,\xi)(P-Q)(d\xi) \right| \\ &= \sup_{x \in X} |P(H(x)) - Q(H(x))|, \end{aligned}$$

where $f_0(\xi, x) = \langle c, x\rangle$, $f_1(\xi, x) = p - \chi_{H(x)}(\xi)$ and $H(x) = \{\xi \in \Xi : T(\xi)x \geq h(\xi)\}$ (polyhedral subsets of Ξ). The relevant probability metrics are polyhedral discrepancies:

$$\alpha_{\mathrm{ph}}(P,Q) = \sup_{B \in \mathcal{B}_{\mathrm{ph}}(\Xi)} |P(B) - Q(B)|$$

5. Two-stage mixed-integer stochastic programs

$$\min\left\{\langle c,x\rangle+\int_{\Xi}\Phi(q(\xi),h(\xi)-T(\xi)x)P(d\xi):x\in X\right\},$$

where Φ is given by

$$\Phi(u,t):=\inf\left\{\langle u_1,y\rangle+\langle u_2,\bar{y}\rangle : Wy+\bar{W}\bar{y}\le t, y\in\mathbb{Z}^{\hat{m}},\bar{y}\in\mathbb{R}^{\bar{m}}\right\}$$

for all pairs $(u,t)\in\mathbb{R}^{\hat{m}+\bar{m}}\times\mathbb{R}^r$, and $c\in\mathbb{R}^m$, X is a closed subset of $\mathbb{R}^m$, Ξ a polyhedron in $\mathbb{R}^s$, W and $\bar{W}$ are $(r,\hat{m})$- and $(r,\bar{m})$-matrices, respectively, $q(\xi)\in\mathbb{R}^{\hat{m}+\bar{m}}$, $h(\xi)\in\mathbb{R}^r$, and the (r,m)-matrix $T(\xi)$ are affine functions of ξ, and $P\in\mathcal{P}_2(\Xi)$.

Probability metric on $\mathcal{P}_2(\Xi)$:

$$\begin{aligned}\zeta_{2,\mathrm{ph}}(P,Q) &:= \sup\left\{\left|\int_B f(\xi)(P-Q)(d\xi)\right| \;\middle|\; \begin{array}{l} f\in\mathcal{F}_2(\Xi) \\ B\in\mathcal{B}_{\mathrm{ph}}(\Xi)\end{array}\right\}\\ &\le C\alpha_{\mathrm{ph}}(P,Q)^{\frac{1}{s+1}} \quad (\text{if } \Xi \text{ is bounded})\end{aligned}$$

Here, the set $\mathcal{F}_2(\Xi)$ contains all functions $f:\Xi\to\mathbb{R}$ such that

$$|f(\xi)|\le\max\{1,\|\xi\|\}\|\xi\|\,,\ |f(\xi)-f(\tilde{\xi})|\le\max\{1,\|\xi\|,\|\tilde{\xi}\|\}\|\xi-\tilde{\xi}\|.$$

6. Multistage stochastic programs

Let $\{\xi_t\}_{t=1}^T$ be a discrete-time stochastic data process defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and with ξ_1 deterministic. The stochastic decision x_t at period t is assumed to be measurable with respect to $\mathcal{A}_t(\xi) := \sigma(\xi_1, \ldots, \xi_t)$ (nonanticipativity).

Multistage stochastic optimization model:

$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T} \langle b_t(\xi_t), x_t \rangle\right] \middle| \begin{array}{l} x_t \in X_t, t = 1, \ldots, T, \\ x_t \text{ is } \mathcal{A}_t(\xi)\text{-measurable}, t = 1, \ldots, T, \\ A_{t,0}x_t + A_{t,1}(\xi_t)x_{t-1} = h_t(\xi_t), t = 2, ., T \end{array} \right\}$$

where X_1 is bounded polyhedral and X_t, $t = 2, \ldots, T$, are polyhedral cones, the vectors $b_t(\cdot)$, $h_t(\cdot)$ and $A_{t,1}(\cdot)$ are affine functions of ξ_t, where ξ varies in a polyhedral set Ξ.

If the process $\{\xi_t\}_{t=1}^T$ has a finite number of scenarios, they exhibit a scenario tree structure.

To have the model well defined, we assume
$x \in L_{r'}(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^m)$ and $\xi \in L_r(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^s)$,
where $r \geq 1$ and

$$r' := \begin{cases} \frac{r}{r-1} & , \text{ if only costs are random} \\ r & , \text{ if only right-hand sides are random} \\ 2 & , \text{ if costs and right-hand sides are random} \\ \infty & , \text{ if all technology matrices are random and } r = T. \end{cases}$$

Then nonanticipativity may be expressed as

$$x \in \mathcal{N}_{r'}(\xi)$$

$$\mathcal{N}_{r'}(\xi) = \{x \in \times_{t=1}^{T} L_{r'}(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^{m_t}) : x_t = \mathbb{E}[x_t | \mathcal{A}_t(\xi)]\,,\ \forall t\},$$

i.e., as a subspace constraint, by using the conditional expectation $\mathbb{E}[\cdot | \mathcal{A}_t(\xi)]$ with respect to the σ-algebra $\mathcal{A}_t(\xi)$.

For $T = 2$ we have $\mathcal{N}_{r'}(\xi) = \mathbb{R}^{m_1} \times L_{r'}(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^{m_2})$.

$\longrightarrow$ infinite-dimensional optimization problem

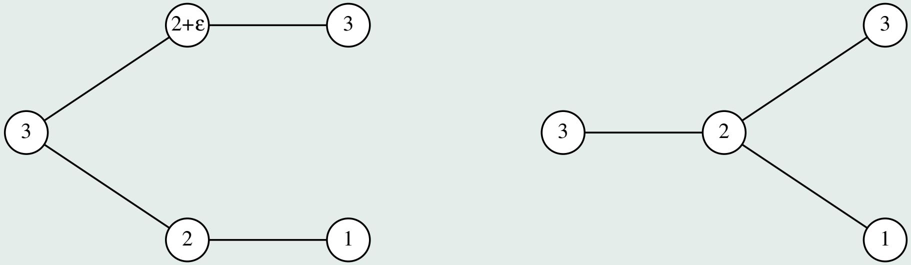
Example: (Optimal purchase under uncertainty)

The decisions x_t correspond to the amounts to be purchased at each time period with uncertain prices are ξ_t, $t = 1, \dots, T$, and such that a prescribed amount a is achieved at the end of a given time horizon. The problem is of the form

$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T} \xi_t x_t\right] \;\middle|\; \begin{array}{l} (x_t, s_t) \in X_t = \mathbb{R}^2_+, \\ (x_t, s_t) \text{ is } (\xi_1, \dots, \xi_t)\text{-measurable}, \\ s_t - s_{t-1} = x_t,\ t = 2, \dots, T, \\ s_1 = 0,\ s_T = a. \end{array} \right\},$$

where the state variable s_t corresponds to the amount at time t.

Let $T := 3$ and ξ_ε denote the stochastic price process having the two scenarios $\xi_\varepsilon^1 = (3, 2 + \varepsilon, 3)$ $(\varepsilon \in (0, 1))$ and $\xi_\varepsilon^2 = (3, 2, 1)$ each endowed with probability $\frac{1}{2}$. Let $\tilde{\xi}$ denote the approximation of ξ_ε given by the two scenarios $\tilde{\xi}^1 = (3, 2, 3)$ and $\tilde{\xi}^2 = (3, 2, 1)$ with the same probabilities $\frac{1}{2}$.



Scenario trees for ξ_ε (left) and $\tilde{\xi}$

We obtain

$$\begin{aligned} v(\xi_\varepsilon) &= \frac{1}{2}((2+\varepsilon)a + a) = \frac{3+\varepsilon}{2}a \\ v(\tilde{\xi}) &= 2a\,, \qquad \text{but} \\ \|\xi_\varepsilon - \tilde{\xi}\|_1 &\le \frac{1}{2}(0+\varepsilon+0) + \frac{1}{2}(0+0+0) = \frac{\varepsilon}{2}. \end{aligned}$$

Hence, the multistage stochastic purchasing model is not stable with respect to $\|\cdot\|_1$.

Quantitative Stability

Let F denote the objective function defined on $L_r(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^s) \times L_{r'}(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^m) \to \mathbb{R}$ by $F(\xi, x) := \mathbb{E}[\sum_{t=1}^{T} \langle b_t(\xi_t), x_t \rangle]$, let

$$\mathcal{X}_t(x_{t-1}; \xi_t) := \{x_t \in X_t : A_{t,0}x_t + A_{t,1}(\xi_t)x_{t-1} = h_t(\xi_t)\}$$

denote the t-th feasibility set for every $t = 2, \ldots, T$ and

$$\mathcal{X}(\xi) := \{x \in L_{r'}(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m) : x_1 \in X_1, x_t \in \mathcal{X}_t(x_{t-1}; \xi_t)\}$$

the set of feasible elements with input ξ.
Then the multistage stochastic program may be rewritten as

$$\min\{F(\xi, x) \, : \, x \in \mathcal{X}(\xi) \cap \mathcal{N}_{r'}(\xi)\}.$$

Let $v(\xi)$ denote its optimal value and, for any $\alpha \geq 0$,

$$\begin{aligned} S_\alpha(\xi) &:= \{x \in \mathcal{X}(\xi) \cap \mathcal{N}_{r'}(\xi) : F(\xi, x) \leq v(\xi) + \alpha\} \\ S(\xi) &:= S_0(\xi) \end{aligned}$$

denote the α-approximate solution set and the solution set of the stochastic program with input ξ.

Assumptions:

(A1) $\xi \in L_r(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^s)$ for some $r \geq 1$.

(A2) There exists a $\delta > 0$ such that for any $\tilde{\xi} \in L_r(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^s)$ with $\|\tilde{\xi} - \xi\|_r \leq \delta$, any $t = 2, \ldots, T$ and any $x_1 \in X_1$, $x_\tau \in \mathcal{X}_\tau(x_{\tau-1}; \tilde{\xi}_\tau)$, $\tau = 2, \ldots, t-1$, the set $\mathcal{X}_t(x_{t-1}; \tilde{\xi}_t)$ is nonempty (relatively complete recourse locally around ξ).

(A3) Assume that the optimal values $v(\tilde{\xi})$ are finite if $\|\xi - \tilde{\xi}\|_r \leq \delta$ and that the objective function F is level-bounded locally uniformly at ξ, i.e., for some $\alpha > 0$ there exists a bounded subset B of $L_{r'}(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^m)$ such that $S_\alpha(\tilde{\xi})$ is contained in B if $\|\tilde{\xi} - \xi\|_r \leq \delta$.

Theorem: (Heitsch-Römisch-Strugarek 06)
Let (A1) – (A3) be satisfied and X_1 be bounded.
Then there exist positive constants L and δ such that

$$|v(\xi) - v(\tilde{\xi})| \leq L(\|\xi - \tilde{\xi}\|_r + d_{\mathrm{f},T-1}(\xi, \tilde{\xi}))$$

holds for all $\tilde{\xi} \in L_r(\Omega, \mathcal{A}, \mathbb{P}; \mathbb{R}^s)$ with $\|\tilde{\xi} - \xi\|_r \leq \delta$.
If $1 < r' < \infty$ and $(\xi^{(n)})$ converges to ξ in L_r and with respect to $d_{\mathrm{f},T}$, then any sequence $x_n \in S(\xi^{(n)})$, $n \in \mathbb{N}$, contains a subsequence converging weakly in $L_{r'}$ to some element of $S(\xi)$.

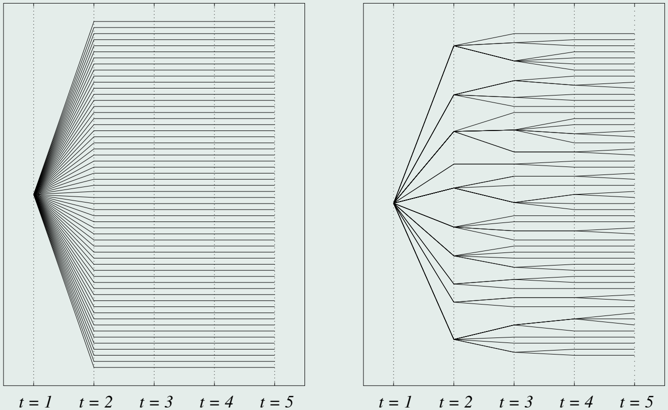
Here, $d_{\mathrm{f},\tau}(\xi, \tilde{\xi})$ denotes the filtration distance of ξ and $\tilde{\xi}$ defined by

$$d_{\mathrm{f},\tau}(\xi, \tilde{\xi}) := \sup_{\|x\|_{r'} \leq 1} \sum_{t=2}^{\tau} \|\mathbb{E}[x_t|\mathcal{A}_t(\xi)] - \mathbb{E}[x_t|\mathcal{A}_t(\tilde{\xi})]\|_{r'}.$$

Remark:
For $T = 2$ we obtain the same result for the optimal values as in the two-stage case ! However, we obtain weak convergence of subsequences of (random) second-stage solutions, too !

Consequences for designing scenario trees

- If ξ_{tr} is a scenario tree process approximating ξ, one has to take care that $\|\xi - \xi_{\mathrm{tr}}\|_r$ and $d_{\mathrm{f},T}(\xi, \xi_{\mathrm{tr}})$ are small. This is achieved for the generation of scenario trees by recursive scenario reduction (Heitsch-Römisch 05).



<Start Animation>

- Are there specific approximations $\tilde{\xi}$ of ξ such that an estimate of the form $|v(\xi) - v(\tilde{\xi})| \leq L\|\xi - \tilde{\xi}\|_r$ is valid ? Recently, such approximations $\tilde{\xi}$ were characterized by Küchler 07 ! The conditions on ξ and approximation schemes developed by Kuhn 05, Pennanen 05, Mirkov-Pflug 07 also avoid filtration distances.

Home Page
Title Page
Contents
Page 34 of 35
Go Back
Full Screen
Close
Quit

References:

Dupačová, J., Gröwe-Kuska, N., Römisch, W.: Scenario reduction in stochastic programming: An approach using probability metrics, *Mathematical Programming* 95 (2003), 493–511.

Heitsch, H., Römisch, W., Strugarek, C.: Stability of multistage stochastic programs, *SIAM Journal Optimization* 17 (2006), 511–525.

Heitsch, H., Römisch, W.: Scenario tree modelling for multistage stochastic programs, Preprint 296, DFG Research Center MATHEON, Berlin 2005, and submitted.

Heitsch, H., Römisch, W.: Stability and scenario trees for multistage stochastic programs, Preprint 324, DFG Research Center MATHEON, Berlin 2006, and submitted to Stochastic Programming - The State of the Art (G. Dantzig, G. Infanger eds.).

Heitsch, H., Römisch, W.: A note on scenario reduction for two-stage stochastic programs, Operations Research Letters (2007) (to appear).

Henrion, R., Römisch, W.: Hölder and Lipschitz stability of solution sets in programs with probabilistic constraints, *Mathematical Programming* 100 (2004), 589–611.

Rachev, S. T., Römisch, W.: Quantitative stability in stochastic programming: The method of probability metrics, *Mathematics of Operations Research* 27 (2002), 792–818.

Römisch, W., Wets, R. J-B: Stability of ε-approximate solutions to convex stochastic programs, *SIAM Journal Optimization* (to appear)

Römisch, W., Vigerske, S.: Quantitative stability of fully random mixed-integer two-stage stochastic programs, Preprint 367, DFG Research Center MATHEON, Berlin 2007, and submitted.