On scenario reduction with respect to polyhedral discrepancies

René Henrion   Christian Küchler   Werner Römisch

FS Numerik stochastischer Modelle

March 14, 2007

Stability in Stochastic Programming

- Optimization problems including uncertainty $\boldsymbol{\xi} \sim \mathbb{P}$ - often not numerically solvable until $\operatorname{supp} \mathbb{P}$ is finite and sufficiently small.
- Discretization / Scenario reduction - substitute the initial measure $\mathbb{P}$ by $\mathbb{Q}$, such that the problem becomes tractable.
- How to choose representative scenarios?
- Stability-based: If there is a distance α on the space of probability measures such that the optimal value $\theta(\cdot)$ behaves continuously w.r.t. α, i.e.

$$|\vartheta(\mathbb{P}) - \vartheta(\mathbb{Q})| \leq C \cdot \alpha(\mathbb{P}, \mathbb{Q}),$$

find a measure $\mathbb{Q}$ such that $\alpha(\mathbb{P}, \mathbb{Q})$ is small.

Stability of two-stage problems

Linear problems:

$$\begin{aligned} \vartheta(\mathbb{P}) &\triangleq \min \langle c, x\rangle + \int_{\Xi\subset\mathbb{R}^s} \langle q(\xi), y(\xi)\rangle \; \mathbb{P}(d\xi) \\ \text{s.t.} & \\ Wy(\xi) &= h(\xi) - T(\xi)x, \\ y(\xi) &\geq 0, \\ x &\in X. \end{aligned}$$

Stability[1] of the optimal value w.r.t. perturbations of $\mathbb{P}$:

$$|\vartheta(\mathbb{P}) - \vartheta(\mathbb{Q})| \leq L \; \hat{\mu}_2(\mathbb{P}, \mathbb{Q}),$$

with the Kantorovich functional

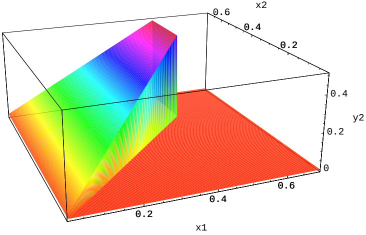
$$\hat{\mu}_2(\mathbb{P}, \mathbb{Q}) \triangleq \inf_{\eta:\pi_1\eta=\mathbb{P},\pi_2\eta=\mathbb{Q}} \int_{\Xi\times\Xi} \max\left\{1, \|\xi\|, \|\tilde{\xi}\|\right\} \|\xi-\tilde{\xi}\| \; \eta(d(\xi, \tilde{\xi})).$$

[1] e.g. Römisch(2003), Thm. 23

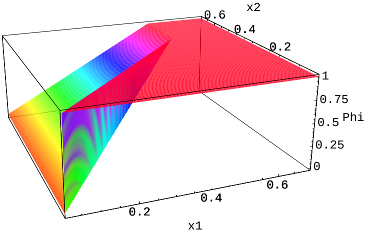
Stability of two-stage problems

Mixed integer problems - Example:

$$\begin{array}{rrcl} & \vartheta(\mathbb{P}) & \triangleq & \displaystyle\int_{\Xi\subset\mathbb{R}^2} \min y_1(\xi) + 2y_2(\xi)\ \mathbb{P}(d\xi) \\ \text{s.t.} & y_1(\xi) + y_2(\xi) & \geq & \xi_1, \\ & y_2(\xi) & \leq & \xi_2, \\ & y_1 & \in & \mathbb{Z}_+, y_2 \in \mathbb{R}_+. \end{array}$$

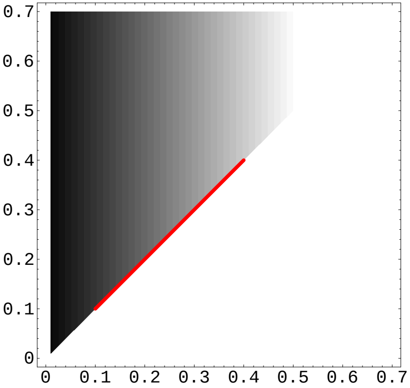


Optimal control $y_2^*(\xi)$.



Optimal value $\Phi(\xi)$.

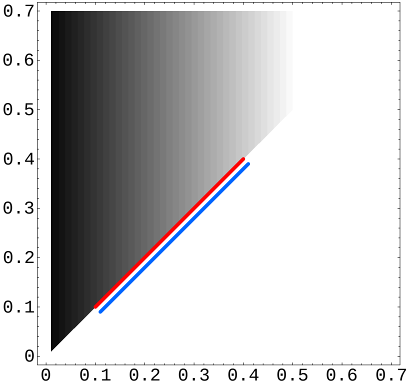
Stability of two-stage problems



Contour plot of $\Phi(\xi)$.

Support of $\mathbb{P}$.

Stability of two-stage problems



Contour plot of $\Phi(\xi)$.

Support of $\mathbb{P}$. Support of $\mathbb{Q}_\varepsilon$ with $\mathbb{Q}_\varepsilon[A] \triangleq \mathbb{P}[A + \binom{\varepsilon}{-\varepsilon}]$.

Stability of two-stage problems

Mixed integer problems - Example (with slack variables)

$$\vartheta(\mathbb{P}) \triangleq \int_{\Xi \subset \mathbb{R}^2} \min y_1(\xi) + 2y_2(\xi)\ \mathbb{P}(d\xi)$$

s.t.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} y_1(\xi) + \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_2(\xi) \\ y_3(\xi) \\ y_4(\xi) \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$$

$$y_1 \in \mathbb{Z}_+,\ \ y_2, y_3, y_4 \in \mathbb{R}_+.$$

Observation[2]: the continuity regions of the integrand can be described via $\mathrm{pos}\, W = W(\mathbb{R}^3_+)$.

[2] Blair and Jeroslow (1977), Bank et al. (1982)

Stability of two-stage problems

Mixed integer problems :

$$\begin{aligned}
\vartheta(\mathbb{P}) &\triangleq \min \langle c, x\rangle + \int_{\Xi\subset\mathbb{R}^s} \langle q, y(\xi)\rangle + \langle \tilde{q}, \tilde{y}(\xi)\rangle \;\; \mathbb{P}(d\xi) \\
\text{s.t.} & \\
Wy(\xi) + \tilde{W}\tilde{y}(\xi) &= \xi - Tx, \\
y(\xi), \tilde{y}(\xi) &\geq 0, \;\; y \in \mathbb{R}^{s'}, \tilde{y} \in \mathbb{Z}^{\tilde{s}} \\
x &\in X.
\end{aligned}$$

Stability[3] of the optimal value w.r.t. perturbations of $\mathbb{P}$:

$$|\vartheta(\mathbb{P}) - \vartheta(\mathbb{Q})| \leq L \; \alpha_{\mathcal{B}_{\text{poly(W)}}}(\mathbb{P}, \mathbb{Q})^{\frac{1}{s+1}},$$

with the polyhedral discrepancy

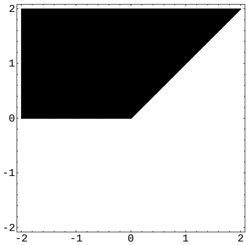
$$\alpha_{\mathcal{B}_{\text{poly(W)}}}(\mathbb{P}, \mathbb{Q}) \triangleq \sup_{B\in\mathcal{B}_{\text{poly(W)}}} |\mathbb{P}(B) - \mathbb{Q}(B)|.$$

[3] Schultz (1996), Römisch and Vigerske (2007)

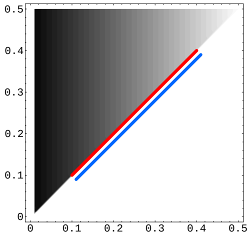
The polyhedral discrepancy

$$\alpha_{\mathcal{B}_{\text{poly(W)}}}(\mathbb{P},\mathbb{Q}) \triangleq \sup_{B\in\mathcal{B}_{\text{poly(W)}}} |\mathbb{P}(B)-\mathbb{Q}(B)|,$$

where $\mathcal{B}_{\text{poly(W)}}$ denotes the class of all closed bounded polyhedra in $\mathbb{R}^s$ each of whose facets (i.e. $(s-1)$-dimensional faces) parallels a facet of $\mathrm{pos}\, W$ or a facet of the unit cube $\times_{i=1}^{s}[0,1]$.



$\mathrm{pos}\, W,\ W = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$



Support of $\mathbb{P}$. Support of $\mathbb{Q}_\varepsilon$.

Objective

- Purely integer recourse ($W = 0$): $\alpha_{\mathcal{B}_{\text{poly(W)}}} = \alpha_{\mathcal{B}_{\text{rect}}}$.
- Chance constrained models: Kolmogorov-Smirnov distance $\alpha_{\mathcal{B}_{\text{cell}}}$.

Objective: Development of suitable **techniques for stability-based scenario reduction w.r.t. these discrepancies** for chance constrained and mixed-integer two-stage models.

Scenario reduction

Let $\mathbb{P}$ be a discrete probability measure on $\mathbb{R}^s$ with support $\left\{\xi^1, \ldots, \xi^N\right\}$ and $p_i \triangleq \mathbb{P}(\xi^i) > 0, i = 1, \ldots, N$.

Problem

Find another probability measure $\mathbb{Q}$ on $\mathbb{R}^s$ with

$$\begin{aligned} \operatorname{supp} \mathbb{Q} &\subset \operatorname{supp} \mathbb{P}, \\ \# \operatorname{supp} \mathbb{Q} &= n < N, \end{aligned}$$

which deviates from $\mathbb{P}$ as little as possible w.r.t. the discrepancy $\alpha_{\mathcal{B}}$, i.e.

$$\begin{aligned} &\text{minimize } \alpha_{\mathcal{B}}(\mathbb{P}, \mathbb{Q}) = \alpha_{\mathcal{B}}\left(\sum_{i=1}^{N} p_i \cdot \delta_{\xi^i}, \sum_{j=1}^{n} q_j \cdot \delta_{\eta^j}\right) \\ &\text{s.t.} \{\eta^1, \ldots, \eta^n\} \subset \left\{\xi^1, \ldots, \xi^N\right\}, \\ &\qquad q_j \geq 0, \ \sum_{j=1}^{n} q_j = 1 \end{aligned}$$

Scenario reduction

This optimization problem can be decomposed into two subproblems:

- determine the scenario set $\operatorname{supp} \mathbb{Q} = \eta \triangleq \{\eta^1, \ldots, \eta^n\}$,
- fix the weights $q = (q_1, \ldots, q_n)$:

$$\Delta_{\mathcal{B}} \triangleq \min_{\eta} \{ \inf_{q \in S_n} \alpha_{\mathcal{B}} (\mathbb{P}, (\eta, q)) \mid \eta \subset \{\xi^1, \ldots, \xi^N\}, \#\eta = n\},$$

with the standard simplex

$$S_n \triangleq \{q \in \mathbb{R}^n | \; q_j \geq 0, \, j = 1, \ldots, n, \, \sum_{j=1}^n q_j = 1\}.$$

Solution techniques

$$\Delta_{\mathcal{B}} = \min_{\eta} \{ \inf_{q \in S_n} \alpha_{\mathcal{B}} (\mathbb{P}, (\eta, q)) \mid \eta \subset \{\xi^1, \ldots, \xi^N\}, \#\eta = n\},$$

Bilevel approach:

- outer iteration - choose support η, NP-hard combinatorial problem.
- Heuristics or branch-and-bound
- inner iteration - determine optimal probabilities q, given the fixed support η.
- This can be formulated as a linear optimization problem.
- We assume that the support is given by $\{\eta^1, \ldots, \eta^n\} = \{\xi^1, \ldots, \xi^n\}$.

Critical index sets

For $B \in \mathcal{B}$, we define a critical index set $I(B)$ by the relation

$$I(B) = \left\{ i \in \{1, \ldots, N\} : \xi^i \in B \right\}.$$

We obtain

$$|\mathbb{P}(B) - \mathbb{Q}(B)| = \left| \sum_{i \in I(B)} p_i - \sum_{j \in I(B) \cap \{1, \ldots, n\}} q_j \right|.$$

Thus, we can define the system of critical index sets

$$\mathcal{I}_{\mathcal{B}} := \{ I \subseteq \{1, \ldots, N\} | \exists B \in \mathcal{B} : I = I(B) \},$$

and arrive at

$$\alpha_{\mathcal{B}}(\mathbb{P}, \mathbb{Q}) = \max_{I \in \mathcal{I}_{\mathcal{B}}} \left| \sum_{i \in I} p_i - \sum_{j \in I \cap \{1, \ldots, n\}} q_j \right|.$$

Critical index sets

$$\alpha_{\mathcal{B}}(\mathbb{P},\mathbb{Q}) = \max_{I\in\mathcal{I}_{\mathcal{B}}} \left| \sum_{i\in I} p_i - \sum_{j\in I\cap\{1,\dots,n\}} q_j \right|$$

Minimizing this w.r.t. $q = (q_1, \dots, q_n)$ is equivalent to

$$\text{minimize} \quad t \quad \text{subject to} \quad q \in S_n,$$

$$\left.\begin{array}{rcl} -\sum_{j\in I\cap\{1,\dots,n\}} q_j & \leq & t - \sum_{i\in I} p_i \\ \sum_{j\in I\cap\{1,\dots,n\}} q_j & \leq & t + \sum_{i\in I} p_i \end{array}\right\} I \in \mathcal{I}_{\mathcal{B}}.$$

Problem: $\mathcal{I}_{\mathcal{B}}$ is very large, in general. $(\leq 2^N)$

Idea: Many different index sets $I \in \mathcal{I}_{\mathcal{B}}$ may lead to the same intersection $I \cap \{1, \dots, n\}$. Then only the r.h.s. of the corresponding inequalities differ. $\mathcal{I}^*_{\mathcal{B}} \triangleq \{I \cap \{1, \dots, n\} | I \in \mathcal{I}_{\mathcal{B}}\}$.

Critical index sets

For $J \in \mathcal{I}_{\mathcal{B}}^*$ we set

$$\gamma^J \triangleq \max_{\substack{I \in \mathcal{I}_{\mathcal{B}} \\ J = I \cap \{1,\ldots,n\}}} \sum_{i \in I} p_i \quad \text{and} \quad \gamma_J \triangleq \min_{\substack{I \in \mathcal{I}_{\mathcal{B}} \\ J = I \cap \{1,\ldots,n\}}} \sum_{i \in I} p_i,$$

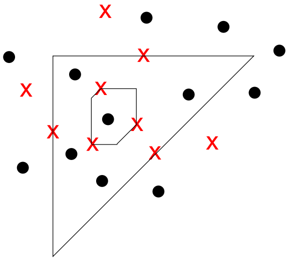
and obtain the problem

$$\begin{array}{lcl} \text{minimize} & t & \text{subject to} \quad q \in S_n, \\ -\sum_{j \in J} q_j & \leq & t - \gamma^J \\ \sum_{j \in J} q_j & \leq & t + \gamma_J \end{array} \left.\vphantom{\begin{array}{l} a \\ b \end{array}}\right\} J \in \mathcal{I}_{\mathcal{B}}^*.$$

How to determine $\mathcal{I}_{\mathcal{B}}^* = \{I \cap \{1, \ldots, n\} | I \in \mathcal{I}_{\mathcal{B}}\}, \gamma_J, \gamma^J$?

Supporting Polyhedra

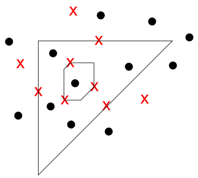
How to determine $\mathcal{I}^*_{\mathcal{B}} = \{I \cap \{1, \ldots, n\} | I \in \mathcal{I}_{\mathcal{B}}\}, \gamma_J, \gamma^J$?



Observation: $\mathcal{I}^*_{\mathcal{B}}, \gamma_J, \gamma^J$ are determined by those polyhedra $\mathcal{P}$, each of whose facets contains an element of $\{\xi^1, \ldots, \xi^n\}$, such that $\mathcal{P}$ can not be enlarged without changing its interior's intersection with $\{\xi^1, \ldots, \xi^n\}$.

These polyhedra $\mathcal{P}$ are called supporting.

Supporting Polyhedra



Proposition

$$\mathcal{I}^*_{\mathcal{B}_{\text{poly(W)}}} = \{J \subseteq \{1,..,n\} : \exists\ \mathcal{P} \text{ with } \cup_{j\in J}\{\xi^j\} = \{\xi^1,..,\xi^n\} \cap \operatorname{int}\mathcal{P}\}$$

$$\gamma^J = \max\{\ \mathbb{P}(\operatorname{int}\mathcal{P}) : \cup_{j\in J}\{\xi^j\} = \{\xi^1,..,\xi^n\} \cap \operatorname{int}\mathcal{P}\}$$

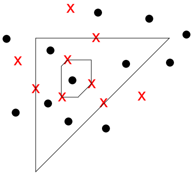
$$\gamma_J = \sum_{i\in I} p_i \quad \text{with } I := \{i \in \{1,\dots,N\} : \xi^i \in [\![\{\xi^j : j \in J\}]\!]\}.$$

Optimal redistribution algorithm

Algorithm

1. Set $\mathcal{I}_{\mathcal{B}}^* = \emptyset$.
2. For every supporting polyhedron $\mathcal{P}$:
   1. Define J via $\cup_{j \in J}\{\xi^j\} = \{\xi^1, .., \xi^n\} \cap \operatorname{int} \mathcal{P}$
   2. If $J \notin \mathcal{I}_{\mathcal{B}}^*$ then update $\mathcal{I}_{\mathcal{B}}^* \triangleq \mathcal{I}_{\mathcal{B}}^* \cup \{J\}$ and γ_J.
   3. Update γ^J.
3. With the additional data $\mathcal{I}_{\mathcal{B}}^*$ and γ_J, γ^J for $J \in \mathcal{I}_{\mathcal{B}}^*$:
   Solve the linear optimization problem.

Supporting Polyhedra



Supporting polyhedra $\mathcal{P}$ = polyhedra each of whose facets is parallel to $\operatorname{pos} W$ and $\times_{i=1}^{s}[0,1]$'s k facets and contains an element of $\xi^1, \ldots, \xi^n$ in its interior .

Using a k-tupel $(m_1, \ldots, m_k)$ of associated normal vectors, each $\mathcal{P}$ can be written as a $(k \times 2)$ matrix $[\underline{a}, \bar{a}]$, where the entries of the j−th row are contained in $\{\langle m_j, \xi^i \rangle, i = 1, \ldots, n\} \cup \pm\infty$, or, equivalently, in $\{1, \ldots, n\} \cup \pm\infty$.

$\binom{n+2}{2}^k$ potential supporting polyhedra $\Rightarrow$ recursive construction, verifying the supporting-condition at each step.

The polyhedral discrepancy $\alpha_{\mathcal{B}_{poly(W)}}$ - optimal redistribution

	k	n=5	n=10	n=15	n=20
	cell	0.01	0.01	0.01	0.05
$\mathbb{R}^3$	3	0.01	0.04	0.56	6.02
N=100	6	0.03	1.03	14.18	157.51
	9	0.15	7.36	94.49	948.17
	cell	0.01	0.01	0.05	0.30
$\mathbb{R}^4$	4	0.01	0.19	1.83	17.22
N=100	8	0.11	5.66	59.28	521.31
	12	0.67	39.86	374.15	3509.34
	cell	0.01	0.01	0.01	0.07
$\mathbb{R}^3$	3	0.01	0.05	0.53	4.28
N=200	6	0.03	0.76	11.80	132.21
	9	0.12	4.22	78.49	815.79
	cell	0.01	0.01	0.06	0.29
$\mathbb{R}^4$	4	0.01	0.20	2.56	41.73
N=200	8	0.11	4.44	73.70	1042.78
	12	0.74	28.29	473.72	6337.68

Running times [sec] of the optimal redistribution algorithm.

The polyhedral discrepancy $\alpha_{\mathcal{B}_{poly(W)}}$ - optimal redistribution

	k	n=5	n=10	n=15	n=20
	cell	0.42	0.28	0.20	0.17
$\mathbb{R}^3$	3	0.66	0.48	0.41	0.36
N=100	6	0.71	0.48	0.42	0.36
	9	0.71	0.48	0.42	0.39
	cell	0.65	0.28	0.22	0.22
$\mathbb{R}^4$	4	0.85	0.53	0.38	0.31
N=100	8	0.86	0.53	0.38	0.31
	12	0.86	0.53	0.38	0.31
	cell	0.35	0.27	0.21	0.20
$\mathbb{R}^3$	3	0.54	0.47	0.35	0.32
N=200	6	0.56	0.47	0.37	0.34
	9	0.56	0.48	0.37	0.34
	cell	0.54	0.40	0.28	0.20
$\mathbb{R}^4$	4	0.80	0.55	0.46	0.40
N=200	8	0.80	0.56	0.50	0.46
	12	0.80	0.56	0.50	0.46

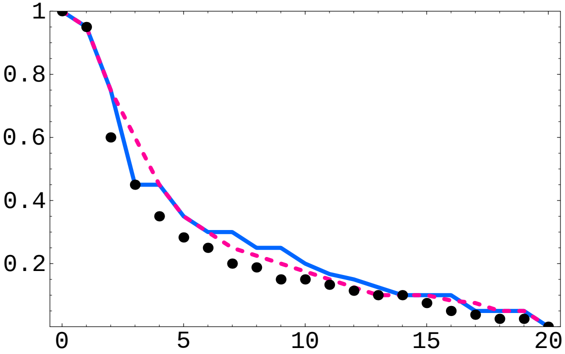
Discrepancies resulting from optimal redistribution.

The polyhedral discrepancy $\alpha_{\mathcal{B}_{poly(W)}}$ - selection heuristics

Rectangular discrepancies resulting from

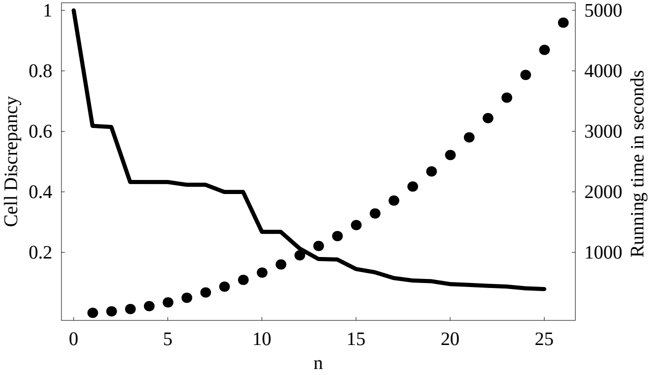
- forward selection (solid line, 3*s*),
- backward reduction (dashed line, 10*s*), and
- complete enumeration (dots, 34*min*!),

depending on the number of remaining scenarios *n*. The initial measure consists of 20 equally weighted scenarios on $\mathbb{R}^2$.



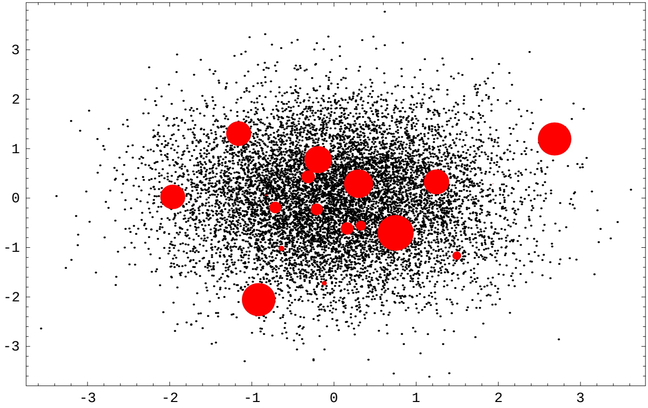
The polyhedral discrepancy $\alpha_{\mathcal{B}_{poly(W)}}$ - selection heuristics

Cell discrepancy and running time in the course of *forward selection*. The initial measure consists of 10 000 equally weighted points in $\mathbb{R}^2$, sampled from a standard normal distribution.



The polyhedral discrepancy $\alpha_{\mathcal{B}_{poly(W)}}$ - selection heuristics

Results of forward selection w.r.t. the cell discrepancy of 20 out of 10 000 points in $\mathbb{R}^2$, sampled from a standard normal distribution. The resulting cell discrepancy is 0.0951.



- *More appropriate heuristics* for the outer problem.
- Stability of *multistage* mixed-integer stochastic programs?
- Comparison with other scenario generation methods.

Thank you very much.