Scenario Reduction:
Old and New Results

W. Römisch

Humboldt-University Berlin
Institute of Mathematics
10099 Berlin, Germany

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

(H. Heitsch, R. Henrion, C. Küchler)

Workshop in honor of Jitka Dupačová

Charles University Prague, November 20, 2014

Introduction

Most approaches for solving stochastic programs of the form

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

with a probability measure P on $\Xi \subset \mathbb{R}^d$ and a (normal) integrand f_0, require numerical integration techniques, i.e., replacing the integral by some quadrature formula

$$\int_{\Xi} f_0(x,\xi)P(d\xi) \approx \sum_{i=1}^{n} p_i f_0(x,\xi_i),$$

where $p_i = P(\{\xi_i\})$, $\sum_{i=1}^{n} p_i = 1$ and $\xi_i \in \Xi$, $i = 1, \ldots, n$.

Since f_0 is often expensive to compute, the number n should be as small as possible.

With $v(P)$ and $S(P)$ denoting the optimal value and solution set of the stochastic program, respectively, the following estimates are known

$$
\begin{aligned}
|v(P)-v(Q)| &\le \sup_{x\in X}\left|\int_\Xi f_0(x,\xi)(P-Q)(d\xi)\right| \\
\emptyset \neq S(Q) &\subseteq S(P)+\Psi_P\left(\sup_{x\in X}\left|\int_\Xi f_0(x,\xi)(P-Q)(d\xi)\right|\right),
\end{aligned}
$$

where X is assumed to be compact, Q is a probability distribution approximating P and the function Ψ_P is the inverse of the growth function of the objective near the solution set, i.e.,

$$
\Psi_P^{-1}(t) := \inf\left\{\int_\Xi f_0(x,\xi)P(d\xi)-v(P) : x\in X, d(x,S(P))\ge t\right\}.
$$

Hence, the distance $d_{\mathcal{F}}$ with $\mathcal{F} := \{f_0(x,\cdot) : x\in X\}$

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

becomes important when approximating P.

For given $n \in \mathbb{N}$ and for the special case $p_i = \frac{1}{n}$, $i = 1, \ldots, n$, the best possible choice of elements $\xi_i \in \Xi$, $i = 1, \ldots, n$ (scenarios), is obtained by minimizing

$$\sup_{x \in X} \left| \int_{\Xi} f_0(x, \xi) P(d\xi) - n^{-1} \sum_{i=1}^{n} f_0(x, \xi_i) \right|,$$

i.e., by solving the **best approximation problem**

$$\min_{Q \in \mathcal{P}_n(\Xi)} d_{\mathcal{F}}(P, Q)$$

where

$$\mathcal{P}_n(\Xi) := \{Q : Q \text{ is a uniform probability measure with } n \text{ scenarios}\}.$$

It may be reformulated as a semi-infinite program. and is known as optimal quantization of P with respect to the function class $\mathcal{F}$. Such optimal quantization problems of probability measures are often extremely difficult to solve.

Idea: Enlarging the class $\mathcal{F}$ to the class of all Lipschitz functions with a uniform constant. But, then

$$\min_{Q \in \mathcal{P}_n(\Xi)} d_{\mathcal{F}}(P, Q) = O(n^{-\frac{1}{d}}).$$

If the functions $f_0(x, \cdot)$ belong to mixed Sobolev spaces, then the convergence rate

$$\min_{Q \in \mathcal{P}_n(\Xi)} d_{\mathcal{F}}(P, Q) = O(n^{-1+\delta}) \quad (\delta \in (0, 0.5])$$

can be achieved by certain randomized Quasi-Monte Carlo methods.

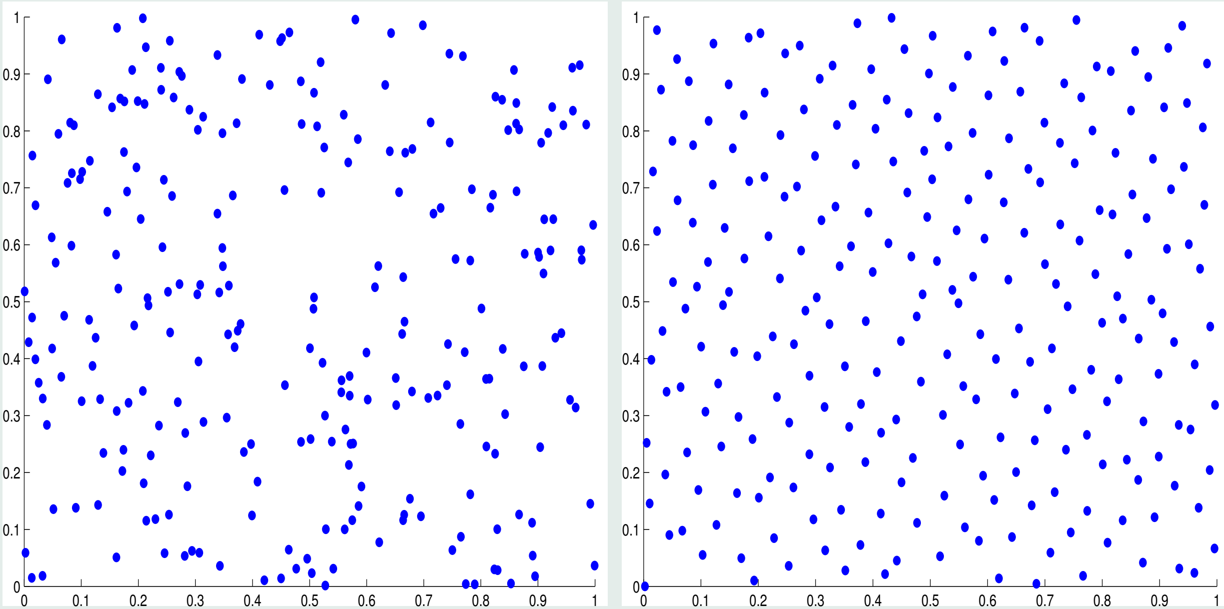
The convergence rate can be improved if the functions $f \in \mathcal{F}$ satisfy a higher degree of smoothness.

Aim of the talk:

Solving the best approximation problem for discrete probability measures P having many scenarios and for function classes $\mathcal{F}$, which are relevant for two-stage stochastic programs (scenario reduction).

Additional motivation:

Scenario reduction methods may be important for generating scenario trees form multistage stochastic programs.



Comparison of $N = 256$ Monte Carlo Mersenne Twister points and randomly binary shifted Sobol' points in dimension $d = 500$, projection (8,9)

Linear two-stage 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 $c\in\mathbb{R}^m$, Ξ and X are polyhedral subsets of $\mathbb{R}^d$ and $\mathbb{R}^m$, respectively, P is a probability measure on Ξ and the $s\times m$-matrix $T(\cdot)$, the vectors $q(\cdot)\in\mathbb{R}^{\overline{m}}$ and $h(\cdot)\in\mathbb{R}^s$ are affine functions of ξ.

Furthermore, Φ and D denote the infimum function of the linear second-stage program and its dual feasibility set, i.e.,

$$\begin{aligned}\Phi(u,t) &:= \inf\{\langle u,y\rangle : Wy=t, y\in Y\}\;((u,t)\in\mathbb{R}^{\overline{m}}\times\mathbb{R}^s)\\ D &:= \{u\in\mathbb{R}^{\overline{m}}:\{z\in\mathbb{R}^s: W^\top z-u\in Y^*\}\neq\emptyset\},\end{aligned}$$

where $q(\xi)\in\mathbb{R}^{\overline{m}}$ are the recourse costs, W is the $s\times\overline{m}$ recourse matrix, $W^\top$ the transposed of W and Y^* the polar cone to the polyhedral cone Y.

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

Assumptions:

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

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

(A3) *existence of second moments:* $\int_\Xi \|\xi\|^2 P(d\xi) < +\infty$.

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

Extensions to certain random recourse models, i.e., to $W(\xi)$, exist.

Idea: Extend the class $\mathcal{F}$ such that it covers all two-stage models.

Fortet-Mourier metrics: (as canonical distances for two-stage models)

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

where $r \geq 1$ ($r \in \{1,2\}$ if $W(\xi) \equiv W$)

$$\mathcal{F}_r(\Xi) := \{ f : \Xi \mapsto \mathbb{R} : f(\xi) - f(\tilde{\xi}) \leq c_r(\xi,\tilde{\xi}),\ \forall \xi, \tilde{\xi} \in \Xi \},$$

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

Proposition: (Rachev-Rüschendorf 98)

If Ξ is bounded, ζ_r may be reformulated as transportation problem

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

where $\hat{c}_r$ is a metric (reduced cost) with $\hat{c}_r \leq c_r$ and given by

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

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

$$D_J := \zeta_r(P, Q^*) = \min_Q \zeta_r(P, Q) = \sum_{i \in J} p_i \min_{j \notin J} \hat{c}_r(\xi_i, \xi_j)$$

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

(optimal redistribution).

Determining the optimal index set J with prescribed cardinality $N-n$ is, however, a combinatorial optimization problem: (n-median problem)

$$\min\{D_J : J \subset \{1,...,N\}, |J| = N-n\}$$

Hence, the problem of finding the optimal set J for deleting scenarios is $\mathcal{NP}$-hard and polynomial time algorithms are not available.

First idea: Reformulation as linear mixed-integer program

$$\min \quad n^{-1}\sum_{i,j=1}^{N} p_j x_{ij}\hat{c}_r(\xi_i,\xi_j) \quad \text{s.t.}$$

$$\sum_{j=1, j\neq i}^{N} x_{ij} + y_i = 1 \quad (i=1,\dots,N), \quad \sum_{i=1}^{N} y_i = n\,,$$

$$x_{ij} \le y_i \quad 0 \le x_{ij} \le 1 \quad (i,j=1,\dots,N)\,,$$

$$y_i \in \{0,1\} \quad (1,\dots,N).$$

and application of standard software or of specialized algorithms.

Solution: $x_{ij} = \begin{cases} \frac{\min_{i\in J}\hat{c}_r(\xi_i,\xi_j)}{n\hat{c}_r(\xi_i,\xi_j)} & , i \notin J, j \in J \\ 0 & , \text{else.} \end{cases} \qquad y_i = \begin{cases} 1 & , i \notin J \\ 0 & , i \in J. \end{cases}$

Fast reduction heuristics

Second idea: Application of (randomized) greedy heuristics.

Starting point ($n = N-1$): $\min_{l\in\{1,\ldots,N\}} p_l \min_{j\neq l} \hat{c}_r(\xi_l, \xi_j)$

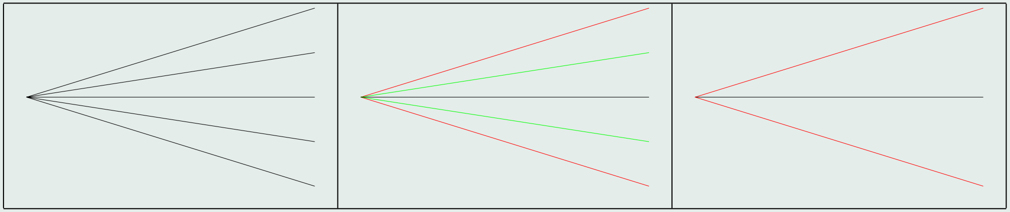
Algorithm 1: (Backward reduction)

Step [0]: $J^{[0]} := \emptyset\,.$

Step [i]: $l_i \in \arg\min_{l\notin J^{[i-1]}} \sum_{k\in J^{[i-1]}\cup\{l\}} p_k \min_{j\notin J^{[i-1]}\cup\{l\}} \hat{c}_r(\xi_k, \xi_j).$

$J^{[i]} := J^{[i-1]} \cup \{l_i\}\,.$

Step [N-n+1]: Optimal redistribution.



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

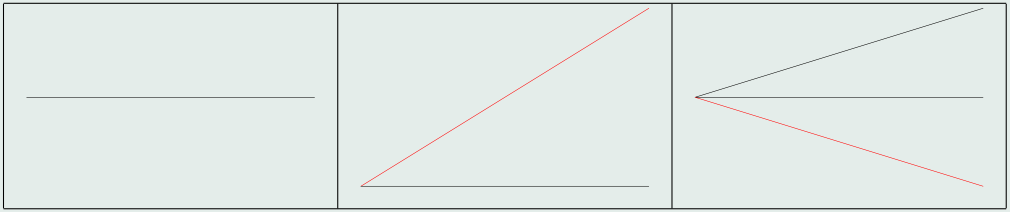
Algorithm 2: (Forward selection)

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

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

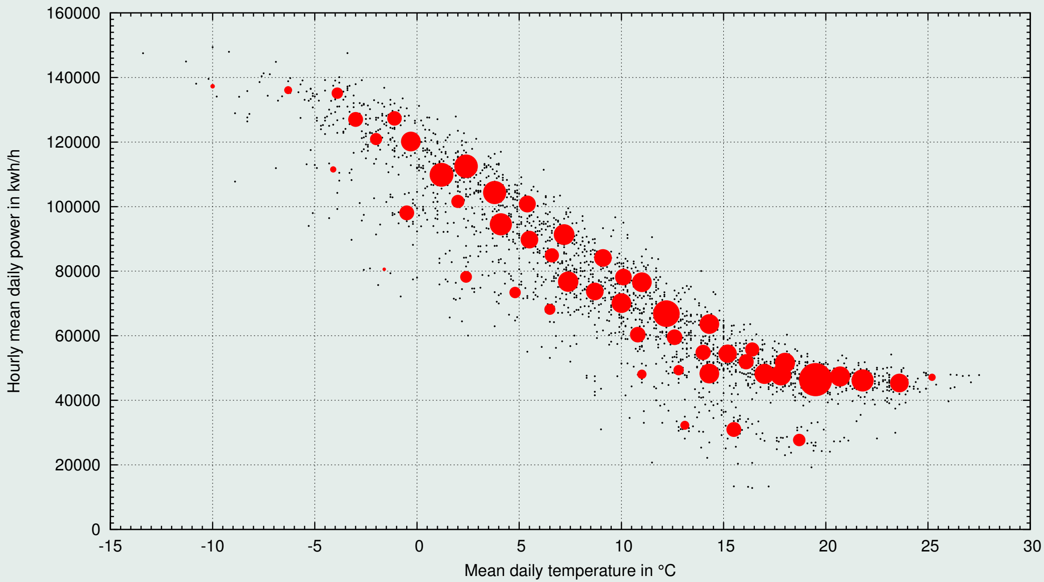
$J^{[i]} := J^{[i-1]} \setminus \{u_i\}$.

Step [n+1]: Optimal redistribution.



Application:

Optimization of gas transport in a huge transportation network including hundreds of gas delivery nodes. A stationary situation is considered; more than 8 years of hourly data available at all delivery nodes; multivariate probability distribution for the gas output in certain temperature classes is estimated; 10^3 samples based on randomized Quasi-Monte Carlo methods are generated and later reduced by scenario reduction to 50 scenarios.



Mixed-integer two-stage stochastic programs

We consider

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

where Φ is given by

$$\Phi(u,t) := \inf \left\{ \langle u_1, y_1 \rangle + \langle u_2, y_2 \rangle \;\middle|\; \begin{array}{l} W_1 y_1 + W_2 y_2 \le t \\ y_1 \in \mathbb{R}_+^{m_1}, y_2 \in \mathbb{Z}_+^{m_2} \end{array} \right\}$$

for all pairs $(u,t) \in \mathbb{R}^{m_1+m_2} \times \mathbb{R}^r$, and $c \in \mathbb{R}^m$, X is a closed subset of $\mathbb{R}^m$, Ξ a polyhedron in $\mathbb{R}^s$, $W_1 \in \mathbb{Q}^{r\times m_1}$, $W_2 \in \mathbb{Q}^{r\times m_2}$, and $T(\xi) \in \mathbb{R}^{r\times m}$, $q(\xi) \in \mathbb{R}^{m_1+m_2}$ and $h(\xi) \in \mathbb{R}^r$ are affine functions of ξ, and P is a probability measure.

We again assume (A1) for $W = (W_1, W_2)$ (relatively complete recourse), (A2) (dual feasibility) and (A3).

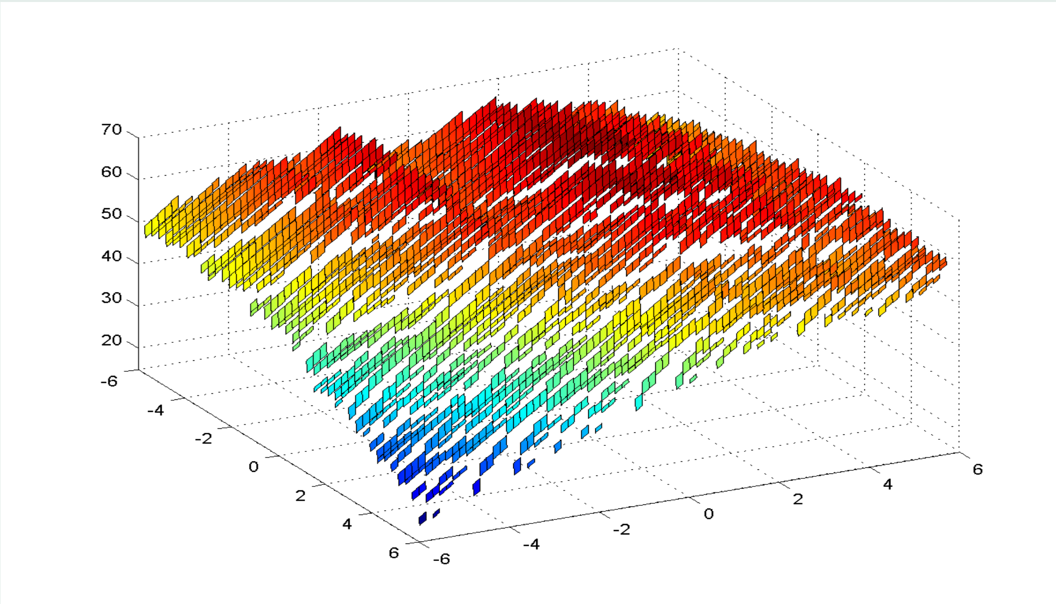
Example: (Schultz-Stougie-van der Vlerk 98)

Stochastic multi-knapsack problem:

$\min = \max$, $m = 2$, $m_1 = 0$, $m_2 = 4$, $c = (1.5, 4)$, $X = [-5, 5]^2$, $h(\xi) = \xi$, $q(\xi) \equiv q = (16, 19, 23, 28)$, $y_i \in \{0, 1\}$, $i = 1, 2, 3, 4$, $P \sim \mathcal{U}(5, 5.5, \ldots, 14.5, 15\}$ (discrete)

Second stage problem: MILP with 1764 0-1 variables and 882 constraints.

$$T = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \qquad W = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 6 & 1 & 3 & 2 \end{pmatrix}$$



The function Φ is well understood and the function class

$$\mathcal{F}_{r,\mathcal{B}}(\Xi) := \{f \mathbb{1}_B : f \in \mathcal{F}_r(\Xi), B \in \mathcal{B}\},$$

is relevant, where $r \in \{1, 2\}$, $\mathcal{B}$ is a class of (convex) polyhedra in Ξ and $\mathbb{1}_B$ denotes the characteristic function of the set B.

The class $\mathcal{B}$ contains all polyhedra of the form

$$B = \{\xi \in \Xi : h(\xi) - T(\xi)x \in D\},$$

where $x \in X$ and D is a polyhedron in $\mathbb{R}^s$ each of whose facets, i.e., $(s-1)$-dimensional faces, is parallel to a facet of the cone $W_1(\mathbb{R}^{m_1}_+)$ or of the unit cube $[0, 1]^s$. Hence, $\mathcal{B}$ is very problem-specific.

Therefore, we consider alternatively the class of rectangular sets

$$\mathcal{B}_{\text{rect}} = \{I_1 \times I_2 \times \cdots \times I_d : \emptyset \neq I_j \text{ is a closed interval in } \mathbb{R}\}$$

covering the situation of pure integer programs and serving as heuristic for the general case.

Proposition:

In case $\mathcal{F} = \mathcal{F}_{r,\mathcal{B}_{\text{rect}}}(\Xi)$, the metric $d_{\mathcal{F}}$ allows the estimates

$$\begin{aligned} d_{\mathcal{F}}(P,Q) &\geq \max\{\alpha_{\mathcal{B}_{\text{rect}}}(P,Q), \zeta_r(P,Q)\} \\ d_{\mathcal{F}}(P,Q) &\leq C\left(\zeta_r(P,Q) + \alpha_{\mathcal{B}_{\text{rect}}}(P,Q)^{\frac{1}{s+1}}\right) \end{aligned}$$

where C is some constant only depending on Ξ and $\alpha_{\mathcal{B}_{\text{rect}}}$ is the rectangular discrepancy given by

$$\alpha_{\mathcal{B}_{\text{rect}}}(P,Q) := \sup_{B \in \mathcal{B}_{\text{rect}}} |P(B) - Q(B)|$$

If the set Ξ is bounded, even the estimate holds

$$\alpha_{\mathcal{B}_{\text{rect}}}(P,Q) \leq d_{\mathcal{F}}(P,Q) \leq C\alpha_{\mathcal{B}_{\text{rect}}}(P,Q)^{\frac{1}{s+1}}.$$

Since $\alpha_{\mathcal{B}_{\text{rect}}}$ has a stronger influence on $d_{\mathcal{F}}$ than ζ_r, we consider the composite distance

$$d_{\lambda}(P,Q) = \lambda\,\alpha_{\mathcal{B}_{\text{rect}}}(P,Q) + (1-\lambda)\,\zeta_r(P,Q)$$

with $\lambda \in [0,1]$ close to 1.

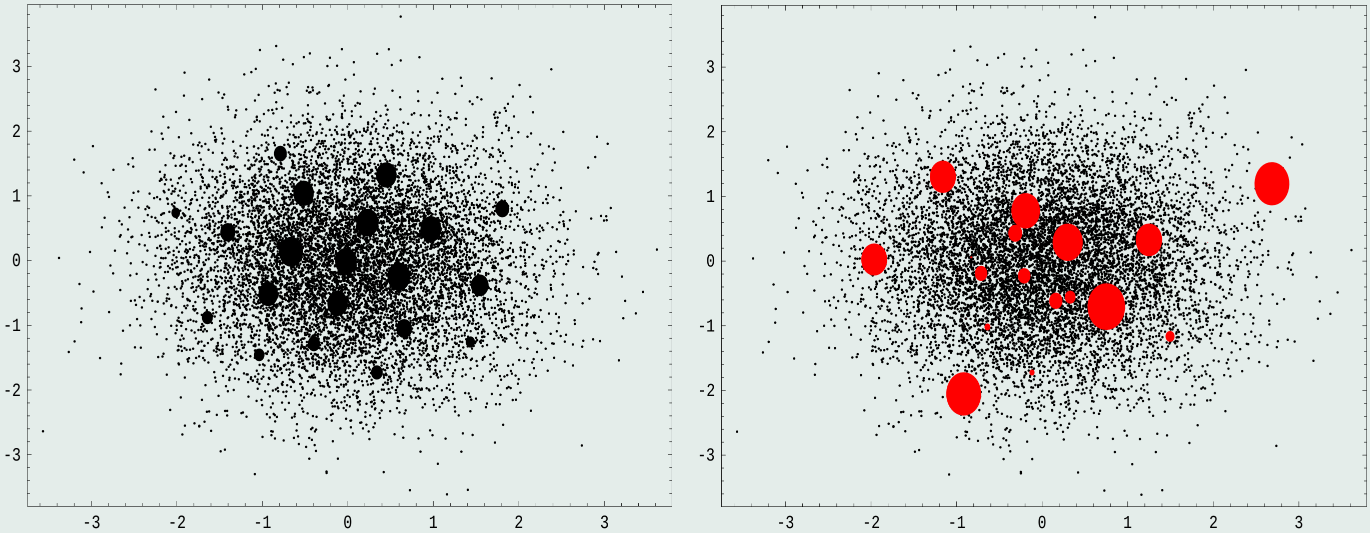
Scenario reduction

We consider again discrete distributions P with scenarios ξ_i and probabilities p_i, $i = 1, \ldots, N$, and Q being supported by a subset of scenarios ξ_j, $j \notin J \subset \{1, \ldots, N\}$, of P with weights q_j, $j \notin J$, where J has cardinality $N - n$.

The problem of optimal scenario reduction consists in determining such a probability measure Q deviating from P as little as possible with respect to d_λ. It can be written as

$$\min \left\{ d_\lambda \left(P, \sum_{j \notin J} q_j \delta_{\xi_j} \right) \middle| \begin{array}{l} J \subset \{1, \ldots, N\}, |J| = N - n \\ q_j \geq 0 \, j \notin J, \sum_{j \notin J} q_j = 1 \end{array} \right\}.$$

This optimization problem may be decomposed into an outer problem for determining the index set J and an inner problem for choosing the probabilities q_j, $j \notin J$.



From 10^4 Monte Carlo samples of a two-dimensional standard normal distribution 20 scenarios are selected that represent best approximations with respect to the first order Fortet-Mourier metric (left) and the Kolmogorov distance (right), i.e., the uniform distance of probability distribution functions.

To this end, we denote

$$d(P,(J,q)) := d_\lambda\left(P, \sum\nolimits_{j\notin J} q_j \delta_{\xi_j}\right)$$
$$S_n := \{q \in \mathbb{R}^n : q_j \geq 0, j \notin J, \sum\nolimits_{j\notin J} q_j = 1\}.$$

Then the optimal scenario reduction problem may be rewritten as

$$\min_J \{\min_{q\in S_n} d(P,(J,q)) : J \subset \{1,\ldots,N\}, |J| = N-n\}$$

with the inner problem (optimal redistribution)

$$\min\{d(P,(J,q)) : q \in S_n\}$$

for fixed index set J. The outer problem is a $\mathcal{NP}$ hard combinatorial optimization problem while the inner problem may be reformulated as a linear program.

Again a reformulation as linear mixed-integer program is possible.

An explicit formula for $D_J := \min_{q\in S_n} d(P,(J,q))$ is no longer available, but the inner problem may be rewritten.

For $B \in \mathcal{B}_{\mathrm{rect}}$ we define the system of critical index sets $I(B)$ by

$$\mathcal{I}_{\mathrm{rect}} := \{I(B) = \{i \in \{1,\ldots,N\} : \xi_i \in B\} : B \in \mathcal{B}_{\mathrm{rect}}\}$$

and write

$$|P(B) - Q(B)| = \left| \sum_{i \in I(B)} p_i - \sum_{j \in I(B) \setminus J} q_j \right|.$$

Then, the rectangular discrepancy between P and Q is

$$\alpha_{\mathcal{B}_{\mathrm{rect}}}(P,Q) = \max_{I \in \mathcal{I}_{\mathrm{rect}}} \left| \sum_{i \in I} p_i - \sum_{j \in I \setminus J} q_j \right|.$$

Using the reduced system of critical index sets

$$\mathcal{I}^*_{\mathrm{rect}}(J) := \{I \setminus J : I \in \mathcal{I}_{\mathrm{rect}}\},$$

every $I^* \in \mathcal{I}^*_{\mathrm{rect}}(J)$ is associated with a family $\varphi(I^*) \subset \mathcal{I}_{\mathrm{rect}}$:

$$\varphi(I^*) := \{I \in \mathcal{I}_{\mathrm{rect}} : I^* = I \setminus J\} \quad (I^* \in \mathcal{I}^*_{\mathrm{rect}}(J)).$$

With the quantities

$$\gamma^{I^*} := \max_{I \in \varphi(I^*)} \sum_{i \in I} p_i \quad \text{and} \quad \gamma_{I^*} := \min_{I \in \varphi(I^*)} \sum_{i \in J} p_i \quad (I^* \in \mathcal{I}^*_{\text{rect}}(J)),$$

we obtain D_J as infimum of the linear program

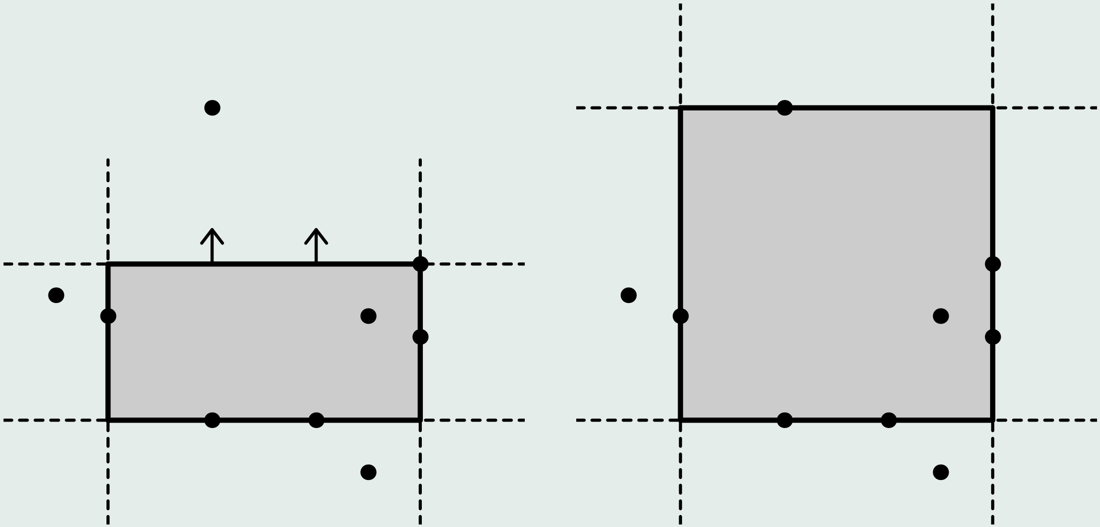
$$\min \left\{ \lambda t_\alpha + (1-\lambda) t_\zeta \;\middle|\; \begin{array}{l} t_\alpha, t_\zeta \geq 0,\; q_j \geq 0,\; \sum_{j \notin J} q_j = 1, \\ \eta_{i,j} \geq 0, i = 1, \ldots, N,\; j \notin J, \\ t_\zeta \geq \;\; \sum_{i=1,\ldots,N, j \notin J} \hat{c}_r(\xi_i, \xi_j) \eta_{i,j}, \\ \sum_{j \notin J} \eta_{i,j} = \;\; p_i,\; i = 1, \ldots, N, \\ \sum_{i=1}^N \eta_{i,j} = \;\; q_j,\; j \notin J, \\ -\sum_{j \in I^*} q_j \leq \;\; t_\alpha - \gamma^{I^*},\; I^* \in \mathcal{I}^*_{\text{rect}}(J) \\ \quad \sum_{j \in I^*} q_j \leq \;\; t_\alpha + \gamma_{I^*},\; I^* \in \mathcal{I}^*_{\text{rect}}(J) \end{array} \right\}$$

We have $|\mathcal{I}^*_{\text{rect}}(J)| \leq 2^n$ and, hence, the LP should be solvable at least for moderate values of n.

How to determine $\mathcal{I}^*_{\text{rect}}(J)$, γ_{I^*} and γ^{I^*}?

Observation:

$\mathcal{I}^*_{\text{rect}}(J)$, γ_{I^*} and γ^{I^*} are determined by those rectangles $B \in \mathcal{R}$, each of whose facets contains an element of $\{\xi_j : j \notin J\}$, such that it can not be enlarged without changing its interior's intersection with $\{\xi_j : j \notin J\}$. The rectangles in $\mathcal{R}$ are called supporting.



Non supporting rectangle (left) and supporting rectangle (right). The dots represent the remaining scenarios ξ_j, $j \notin J$.

Proposition:

It holds that

$$\mathcal{I}^*_{\text{rect}}(J) = \bigcup_{B \in \mathcal{R}} \{I^* \subseteq \{1, \ldots, N\} \setminus J : \cup_{j \in I^*}\{\xi_j\} = \{\xi_j : j \notin J\} \cap \operatorname{int} B\}$$

and, for every $I^* \in \mathcal{I}^*_{\text{rect}}(J)$,

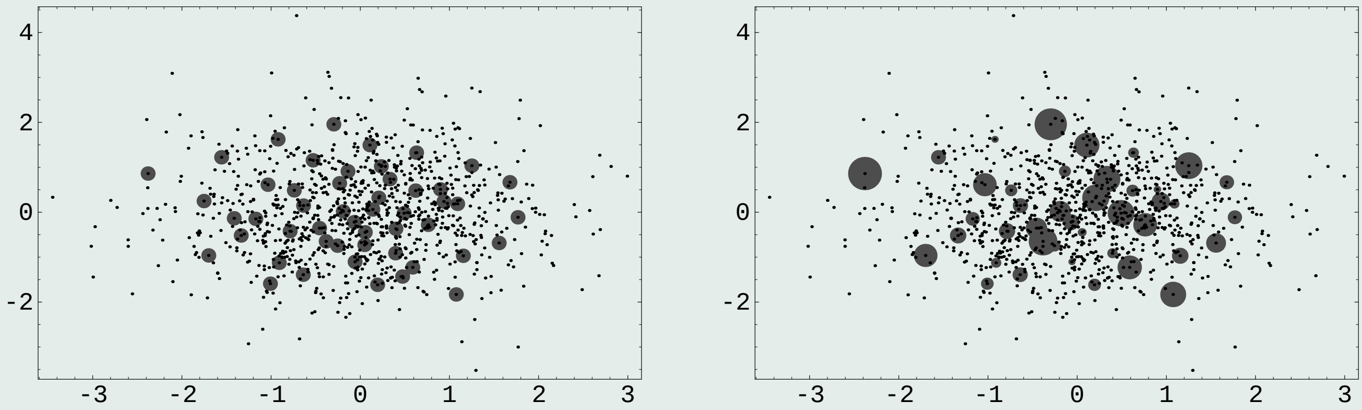
$$\gamma^{I^*} = \max \{P(\operatorname{int} B) : B \in \mathcal{R}, \cup_{j \in I^*}\{\xi_j\} = \{\xi_j : j \notin J\} \cap \operatorname{int} B\}$$
$$\gamma_{I^*} = \sum_{i \in \underline{I}} p_i,$$

where

$$\underline{I} := \{i \in \{1, \ldots, N\} : \min_{j \in I^*} \xi_{j,l} \leq \xi_{i,l} \leq \max_{j \in I^*} \xi_{j,l}, l = 1, \ldots, d\}.$$

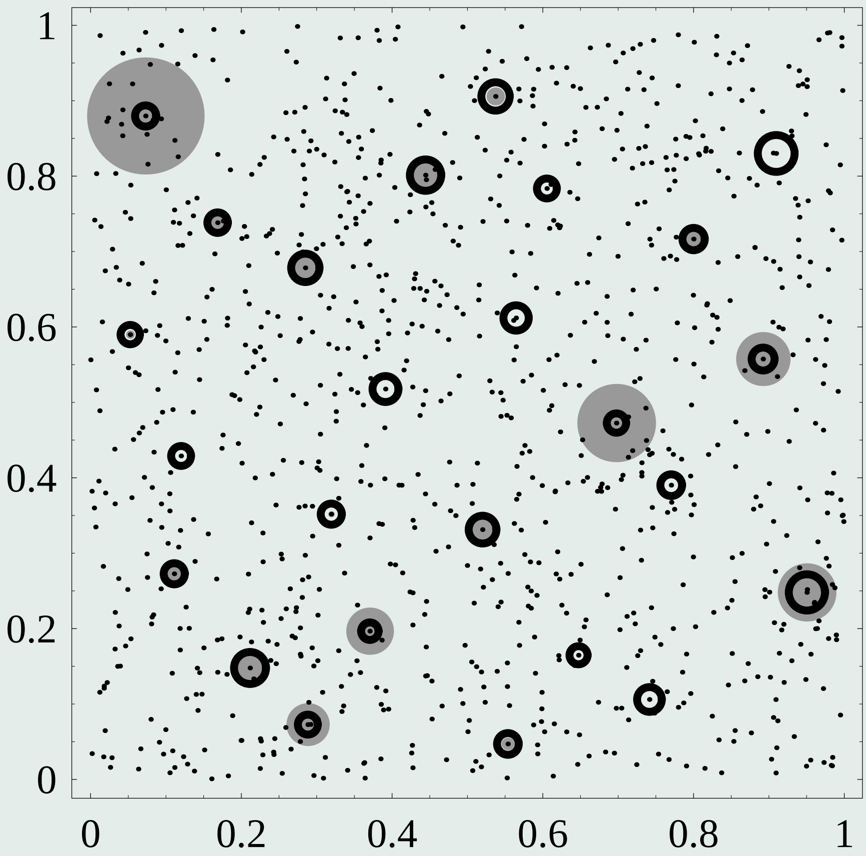
Note that $|\mathcal{R}| \leq \binom{n+2}{2}^d$!

Numerical results



Comparison of scenario reduction for a Fortet-Mourier metric and a composite distance including α_{rect}.

Optimal redistribution: $\alpha_{\mathcal{B}_{\text{rect}}}$ versus ζ_2



25 scenarios chosen by Quasi Monte Carlo out of 1000 samples from the uniform distribution on $[0,1]^2$ and optimal probabilities adjusted w.r.t. $\lambda\alpha_{\mathcal{B}_{\text{rect}}} + (1-\lambda)\zeta_2$ for $\lambda = 1$ (gray balls) and $\lambda = 0.9$ (black circles)

Optimal redistribution w.r.t. the rectangular discrepancy $\alpha_{\mathcal{B}_{\mathrm{rect}}}$:

	d	n=5	n=10	n=15	n=20
N=100	3	0.01	0.04	0.56	6.02
	4	0.01	0.19	1.83	17.22
N=200	3	0.01	0.05	0.53	4.28
	4	0.01	0.20	2.56	41.73

Running times [sec] of the optimal redistribution algorithm

The majority of the running time is spent for determining the supporting rectangles, while the time needed to solve the linear program is insignificant.

Optimal scenario reduction

Forward selection:

Step [0]: $J^{[0]} := \emptyset\,.$

Step [i]: $l_i \in \mathrm{argmin}_{l \notin J^{[i-1]}} \inf_{q \in S_i} d_\lambda \left(P, \sum_{j \in J^{[i-1]} \cup \{l\}} q_j \delta_{\xi_j} \right),$

$J^{[i]} := J^{[i-1]} \cup \{l_i\}.$

Step [n+1]: Minimize $d_\lambda \left(P, \sum_{j \in J^{[n]}} q_j \delta_{\xi_j} \right)$ s.t. $q \in S_n$.

N=100	n=5	n=10	n=15
$d = 2$	0.21	2.07	17.46
$d = 3$	0.33	8.40	230.40
$d = 4$	0.61	33.69	1944.94

Growth of running times (in seconds) of forward selection for $\lambda = 1$

$\longrightarrow$ Search for more efficient heuristics

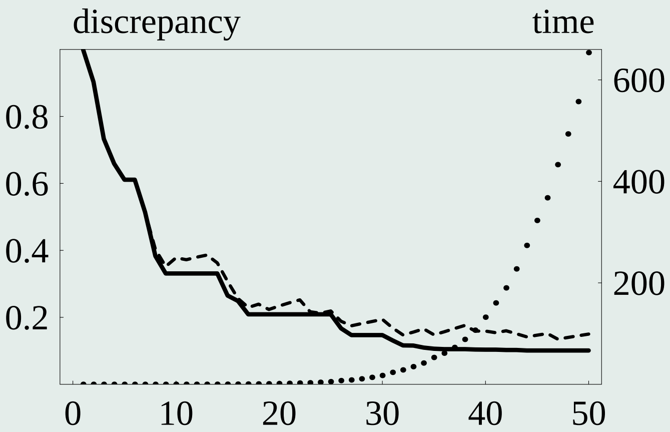
Alternative heuristics (for P with independent marginals):

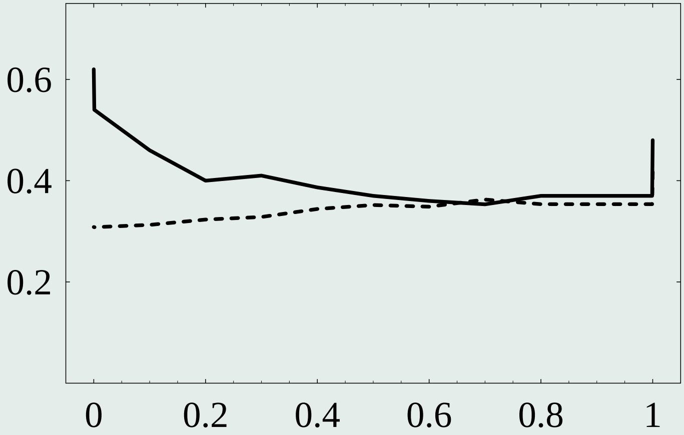
- **(next neighbor) Quasi Monte Carlo**: The first n numbers of randomized QMC sequences provide n equally weighted points. The closest scenarios are determined and the resulting discrepancy to the initial measure is computed for *fixed* probability weights.

- **(next neighbor) adjusted Quasi Monte Carlo**: The probabilities of the closest scenarios are adjusted by the optimal redistribution algorithm to obtain a minimal rectangular discrepancy to P.

For general distributions P with densities transformation formulas are needed (e.g. Hlawka-Mück 71).

Conclusion: (Next neighbor) readjusted QMC decreases significantly the approximation error. Forward selection provides good results, but is very slow due to the optimal redistribution in each step.





Left: The distance d_λ ($\lambda = 1$) between P and uniform (next neighbor) QMC points (dashed line) and (next neighbor) readjusted QMC points (solid line), and running time in seconds of optimal redistribution.
Right: Distances α_{rect} (solid) and ζ_2 (dashed) of 10 out of 100 scenarios, resulting from forward selection for several $\lambda \in [0, 1]$.

Conclusions and outlook

- There exist reasonably fast heuristics for scenario reduction in linear two-stage stochastic programs,
- It may be worth to study and compare exact solution methods with heuristics,
- Recursive application of the heuristics apply to generating scenario trees for multistage stochastic programs,
- For scenario tree reduction the heuristics have to be modified.
- For mixed-integer two-stage stochastic programs heuristics exist, but have to be based on different arguments. They are more expensive and so far restricted to moderate dimensions. This motivates to study exact approaches.
- There is hope for generating scenario trees for mixed-integer multistage models, but it is not yet supported by stability results.

References

Avella, P., Sassano, A, Vasil'ev, I.: Computational study of large-scale p-median problems, *Mathematical Programming* 109 (2007), 89–114.

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.: Scenario reduction algorithms in stochastic programming, *Computational Optimization and Applications* 24 (2003), 187–206.

Heitsch, H., Römisch, W.: A note on scenario reduction for two-stage stochastic programs, *Operations Research Letters* 35 (2007), 731–736.

Heitsch, H., Römisch, W.: Scenario tree modeling for multistage stochastic programs, *Mathematical Programming* 118 (2009), 371–406.

Heitsch, H., Römisch, W.: Scenario tree reduction for multistage stochastic programs, *Computational Management Science* 6 (2009), 117–133.

Henrion, R., Küchler, C., Römisch, W.: Scenario reduction in stochastic programming with respect to discrepancy distances, *Computational Optimization and Applications* 43 (2009), 67–93.

Henrion, R., Küchler, C., Römisch, W.: Discrepancy distances and scenario reduction in two-stage stochastic mixed-integer programming, *Journal of Industrial and Management Optimization* 4 (2008), 363–384.

Li, Z., Floudas, C. A.: Optimal scenario reduction framework based on distance of uncertainty distribution and output performance:I. Single reduction via mixed integer linear optimization, *Computers and Chemical Engineering* 70 (2014), 50–66.

Römisch, W.: Stability of Stochastic Programming Problems, in: *Stochastic Programming* (A. Ruszczyński and A. Shapiro eds.), Handbooks in Operations Research and Management Science, Volume 10, Elsevier, Amsterdam 2003, 483–554.

Römisch, W., Vigerske, S.: Quantitative stability of fully random mixed-integer two-stage stochastic programs, *Optimization Letters* 2 (2008), 377–388.

Römisch, W.; Wets, R. J-B: Stability of ε-approximate solutions to convex stochastic programs, *SIAM Journal on Optimization* 18 (2007), 961–979.

Koch, T., Hiller, B., Pfetsch, M. E., Schewe, L. (Eds.): Evaluating Gas Network Capacities, SIAM-MOS (in preparation).