# Problem-based scenario generation for two-stage stochastic programs using semi-infinite optimization

W. Römisch

Humboldt-University Berlin
Institute of Mathematics

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

R. Henrion (WIAS Berlin)

Applications of SIP, Online Workshop, 20.5./21.5. 2021

# Introduction

A number of stochastic programming models may be traced back to minimizing an expectation functional on some closed subset of a Euclidean space. A general form is

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

where $X$ is a closed subset of $\mathbb{R}^m$, $\Xi$ a closed subset of $\mathbb{R}^s$, $P$ is a Borel probability measure on $\Xi$ abbreviated by $P \in \mathcal{P}(\Xi)$. The function $f_0$ from $\mathbb{R}^m \times \Xi$ to the extended reals $\overline{\mathbb{R}} = [-\infty, \infty]$ is a normal integrand.

For example, typical integrands in linear two-stage stochastic programming models are

$$f_0(x,\xi) = \begin{cases} g(x) + \Phi(q(\xi), h(x,\xi)) & , q(\xi) \in D \\ +\infty & , \text{else} \end{cases},$$

where $X$ and $\Xi$ are convex polyhedral, $g(\cdot)$ is a linear function, $q(\cdot)$ is affine, $D = \{q \in \mathbb{R}^{\bar{m}} : \{z \in \mathbb{R}^r : W^\top z - q \in Y^\star\} \neq \emptyset\}$ denotes the convex polyhedral dual feasibility set, $h(\cdot,\xi)$ is affine for fixed $\xi$ and $h(x,\cdot)$ is affine for fixed $x$, and $\Phi$ denotes the infimal function of the linear (second-stage) optimization problem

$$\Phi(q,t) := \inf\{\langle q, y\rangle : Wy = t, y \in Y\}$$

with $(r, \bar{m})$ matrix $W$, convex polyhedral cone $Y \subset \mathbb{R}^{\bar{m}}$ and its polar cone $Y^\star \subset \mathbb{R}^{\bar{m}}$.

For general continuous multivariate probability distributions $P$ such stochastic optimization models are not solvable in general. The computation of the objective of linear two-stage stochastic programs is #P-hard.

Many approaches for solving such optimization models computationally are based on discrete approximations of the probability measure $P$, i.e., on finding a discrete probability measure $P_n$ in

$$\mathcal{P}_n(\Xi) := \Big\{ \sum_{i=1}^{n} p_i \delta_{\xi^i} : \xi^i \in \Xi,\, p_i \geq 0,\, i = 1, \ldots, n,\, \sum_{i=1}^{n} p_i = 1 \Big\}$$

for some $n \in \mathbb{N}$, which approximates $P$ in a *suitable* way. Here, $\delta_\xi$ denotes the Dirac measure placing unit mass to $\xi$ and zero elsewhere.
The atoms $\xi^i$, $i = 1, \ldots, n$, of $P_n$ are often called scenarios in this context. Of course, the notion *suitable* should at least include that the distance of infima and solution sets

$$|v(P) - v(P_n)| \quad \text{and} \quad \sup_{x \in S(P_n)} d(x, S(P))$$

become reasonably small, where $v(P)$ and $S(P)$ denote the infimum and solution set of (SP).

## Stability-based scenario generation

We are interested in the continuous dependence of infima and solution sets on the underlying probability distribution $P$ in terms of a suitable metric.

To state a corresponding result we introduce the following sets of functions and of probability distributions (both defined on $\Xi$)

$$\mathcal{F} = \{f_0(x,\cdot\,) : x \in X\},$$
$$\mathcal{P}_{\mathcal{F}} = \left\{Q \in \mathcal{P}(\Xi) : -\infty < \int_{\Xi} \inf_{x\in X} f_0(x,\xi)Q(d\xi), \sup_{x\in X} \int_{\Xi} f_0(x,\xi)Q(d\xi) < +\infty\right\}$$

and the (pseudo-) metric on $\mathcal{P}_{\mathcal{F}}$

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

For typical applications, like for linear two-stage models, the sets $\mathcal{P}_{\mathcal{F}}$ allow a simple characterization, for example, as subsets of $\mathcal{P}(\Xi)$ satisfying certain moment conditions. The (pseudo) metric $d_{\mathcal{F}}$ is called problem-based or minimal information distance.

**Proposition:**

Consider (SP) for $P \in \mathcal{P}_{\mathcal{F}}$, assume that $X$ is compact. Then the estimates

$$\begin{aligned} |v(P) - v(Q)| &\leq d_{\mathcal{F}}(P, Q) \\ \sup_{x \in S(Q)} d(x, S(P)) &\leq \psi_P^{-1}(d_{\mathcal{F}}(P, Q)) \end{aligned}$$

hold whenever $Q \in \mathcal{P}_{\mathcal{F}}$, where $\psi_P : \mathbb{R}_+ \to \mathbb{R}_+$ is the growth function

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

For given $n \in \mathbb{N}$ the above result suggests to choose discrete approximations from $\mathcal{P}_n(\Xi)$ for solving (SP) such that they solve the best approximation problem

$$\text{(OSG)} \qquad \min_{P_n \in \mathcal{P}_n(\Xi)} d_{\mathcal{F}}(P, P_n)\,.$$

Determining the scenarios of a solution to (OSG) is called optimal scenario generation.

# Monte Carlo and (randomized) Quasi-Monte Carlo

**Monte Carlo:** Let $\xi^i(\cdot)$, $i \in \mathbb{N}$, denote independent and identically distributed random vectors in $\Xi$ with common distribution $P$. Empirical measure:

$$P_n(\cdot) = \frac{1}{n}\sum_{i=1}^{n} \delta_{\xi^i(\cdot)} \quad (n \in \mathbb{N})$$

defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Best possible mean convergence rate:

$$\mathbb{E}[d_{\mathcal{F}}(P_n(\cdot), P)] = O(n^{-\frac{1}{2}}).$$

under strong assumptions on $\mathcal{F}$.

**Quasi-Monte Carlo:** The basic idea of Quasi-Monte Carlo (QMC) methods is to use deterministic points $\xi^i$, $i = 1, \ldots, n$, that are (in some way) uniformly distributed in $[0,1]^s$ and to consider the approximate computation of

$$I_s(f) = \int_{[0,1]^s} f(\xi) d\xi \quad \text{by} \quad Q_{n,s}(f) = \frac{1}{n}\sum_{i=1}^{n} f(\xi^i).$$

There exist randomized points $\xi^i(\cdot) \in [0,1]^s$, $i = 1, \ldots, n$, such that

$$(\mathbb{E}[|Q_{n,s}(\cdot)(f) - I_s(f)|^2])^{\frac{1}{2}} \leq C(\delta)\, n^{-1+\delta} \quad (\delta \in (0, 0.5])$$

if the integrand $f$ is sufficiently smooth.

## Problem-based scenario generation for linear two-stage models

We consider linear two-stage stochastic programs as introduced earlier and impose the following conditions:

**(A0)** $X$ is a bounded polyhedron and $\Xi$ is convex polyhedral.

**(A1)** $h(x,\xi) \in W(Y)$ and $q(\xi) \in D$ are satisfied for every pair $(x,\xi) \in X \times \Xi$.

**(A2)** $P$ has a second order absolute moment.

Then the infima $v(P)$ and $v(P_n)$ are attained and the estimate

$$\begin{aligned} |v(P) - v(P_n)| &\le \sup_{x\in X} \left| \int_\Xi f_0(x,\xi)P(d\xi) - \int_\Xi f_0(x,\xi)P_n(d\xi) \right| = d_{\mathcal{F}}(P,P_n) \\ &= \sup_{x\in X} \left| \int_\Xi \Phi(q(\xi),h(x,\xi))P(d\xi) - \int_\Xi \Phi(q(\xi),h(x,\xi))P_n(d\xi) \right| \end{aligned}$$

holds due to the Proposition for every $P_n \in \mathcal{P}_n(\Xi)$.

Optimal scenario generation problem (OSG):

Determine $P_n^* \in \mathcal{P}_n(\Xi)$ such that it solves the best approximation problem

$$\min_{\substack{(\xi^1,\ldots,\xi^n)\in\Xi^n \\ p_i\ge 0, \sum_{i=1}^n p_i=1}} \sup_{x\in X} \left| \int_\Xi \Phi(q(\xi),h(x,\xi))P(d\xi) - \sum_{i=1}^n p_i \Phi(q(\xi^i),h(x,\xi^i)) \right|.$$

The class of functions $\{\Phi(q(\cdot), h(x,\cdot)) : x \in X\}$ from $\Xi$ to $\overline{\mathbb{R}}$ enjoys specific properties. All functions are finite, continuous and piecewise linear-quadratic on $\Xi$. They are linear-quadratic on each convex polyhedral set

$$\Xi_j(x) = \{\xi \in \Xi : (q(\xi), h(x,\xi)) \in \mathcal{K}_j\} \quad (j = 1, \ldots, \ell),$$

where the convex polyhedral cones $\mathcal{K}_j$, $j = 1, \ldots, \ell$, represent a decomposition of the domain $\operatorname{dom}\Phi$ of $\Phi$, which is itself a convex polyhedral cone in $\mathbb{R}^{\bar{m}+r}$.

**Theorem 1:** Assume (A0)–(A2).
Then (OSG) is equivalent to the generalized semi-infinite program (GSIP)

$$\min_{\substack{t\geq 0, (\xi^1,\ldots,\xi^n)\in\Xi^n \\ p_i\geq 0, \sum_{i=1}^n p_i = 1}} \left\{ t \;\middle|\; \begin{array}{c} \sum_{i=1}^n p_i \langle h(x,\xi^i), z_i\rangle \leq t + F_P(x) \\ F_P(x) \leq t + \sum_{i=1}^n p_i \langle q(\xi^i), y_i\rangle \\ \forall (x,y,z) \in \mathcal{M}(\xi^1,\ldots,\xi^n) \end{array} \right\},$$

where the set $\mathcal{M} = \mathcal{M}(\xi^1,\ldots,\xi^n)$ and the function $F_P : X \to \mathbb{R}$ are given by

$$\mathcal{M} = \{(x,y,z) \in X \times Y^n \times \mathbb{R}^{rn} : Wy_i = h(x,\xi^i), W^\top z_i - q(\xi^i) \in Y^*, \forall i\},$$

$$F_P(x) = \int_\Xi \Phi(q(\xi), h(x,\xi)) P(d\xi).$$

The latter is the convex expected recourse function of the two-stage model.
If the function $h$ is affine, the feasible set of (GSIP) is closed.

## Generalized semi-infinite programming

Generalized semi-infinite optimization problems are of the form

$$\min\{f(x) : x \in M\} \quad \text{with} \quad M = \{x \in \mathbb{R}^n : g_i(x,y) \leq 0,\, y \in Y(x),\, i \in I\},$$

where

$$Y(x) = \{y \in \mathbb{R}^m : h_j(x,y) \leq 0,\, j \in J\}$$

and all functions $f$, $g_i$, $i \in I$, $h_j$, $j \in J$, are real-valued and continuous and $I$ and $J$ are finite index sets. Moreover, the set-valued mapping $Y : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is assumed to be (Berge) upper semicontinuous.

**Proposition:** (Stein 03)
$M$ is closed if, in addition, the set-valued mapping $Y$ is lower semicontinuous.

**Proposition:** (Still 01)
Assume that $g_i$, $i \in I$, are convex in $(x,y)$ on $\mathbb{R}^{n+m}$ and that for all $x$, $\tilde{x}$ in $\mathbb{R}^n$ and $0 < \alpha < 1$ holds that

$$Y(\alpha x + (1-\alpha)\tilde{x}) \subseteq \alpha Y(x) + (1-\alpha) Y(\tilde{x}).$$

Then the feasible set $M$ is convex.

## Convexity of problem-based scenario generation for two-stage models

**Theorem 2:**
Assume (A0)–(A2), let the function $h$ be affine, the weights $p_i$, $i = 1, \ldots, n$, be fixed and either $h$ or $q$ be random.
Then the set-valued mapping $\mathcal{M} : \Xi^n \rightrightarrows \mathbb{R}^m \times Y^n \times \mathbb{R}^{rn}$ has convex polyhedral graph and is Hausdorff Lipschitz continuous. In particular, the feasible set of the (GSIP) is closed and convex. Furthermore, if the infimum of the (GSIP) is positive, then the optimal value behaves Lipschitz continuous with respect to changes of the function $F_P$ in terms of the supremum-norm on $X$.

We note that $F_P(x)$ can only be calculated approximately even if the probability measure $P$ is completely known. For example, this could be done by Monte Carlo or Quasi-Monte Carlo methods with a large sample size $N > n$, i.e.

$$F_P(x) \approx \frac{1}{N} \sum_{j=1}^{N} \Phi(q(\hat{\xi}^j), h(x, \hat{\xi}^j)),$$

where $\hat{\xi}^j \in \Xi$, $j = 1, \ldots, N$.

## Problem-based scenario generation via semi-infinite optimization

In some cases the problem (GSIP) can be transformed into a semi-infinite program inspired by the recent paper (Schwientek-Seidel-Küfer 21).

Let only costs $q(\cdot)$ be random, $Y = \mathbb{R}^{\bar{m}}_+$ and the transformation

$$t : \Xi \times \mathcal{Z} \to \mathbb{R}^r \qquad t(\xi, z) = z + (W^+)^\top (q(\xi) - \bar{q})$$

be given, where $\mathcal{Z} = \{z \in \mathbb{R}^r : W^\top z \leq \bar{q}\}$ and the $(\bar{m}, r)$-matrix $W^+$ denotes the Moore-Penrose inverse of $W$.

**Theorem 3:** Assume (A0) and (A2).

Let $h(x) \in W(\mathbb{R}^{\bar{m}}_+)$ for all $x \in X$ and $\bar{q}, q(\xi) \in W^\top(\mathbb{R}^r)$ for all $\xi \in \Xi$.

Then (GSIP) is equivalent to the semi-infinite program

$$\min_{\substack{t \geq 0 \\ (\xi^1,\ldots,\xi^n) \in \Xi^n \\ p_i \geq 0, \sum_{i=1}^n p_i = 1}} \left\{ t \,\middle|\, \begin{array}{c} \sum_{i=1}^n p_i \langle h(x), z_i + (W^+)^\top (q(\xi^i) - \bar{q}) \rangle \leq t + F_P(x) \\ F_P(x) \leq t + \sum_{i=1}^n p_i \langle q(\xi^i), y_i \rangle \\ \forall (x, y, z) \in X \times \mathcal{Y}(x)^n \times \mathcal{Z}^n \end{array} \right\},$$

where $\mathcal{Y}(x) = \{y \in \mathbb{R}^{\bar{m}}_+ : Wy = h(x)\}$ for each $x \in X$. If the weights $p_i$, $i = 1, \ldots, n$, are fixed, the semi-infinite program is linear.

## Problem-based scenario reduction for two-stage model

Let $\xi^i$, $i = 1, \ldots, N$, be a large set of scenarios with probabilities $p_i$, $i = 1, \ldots, N$, that define a discrete probability measure

$$P = \sum_{i=1}^{N} p_i \delta_{\xi^i}.$$

For prescribed $n \in \mathbb{N}$, $n < N$, we intend to determine an index set $J \subset \{1, \ldots, N\}$ of cardinality $|J| = n$ and new weights $\bar{\pi}_j$, $j \in J$, such that

$$P_J^* = \sum_{j \in J} \bar{\pi}_j \delta_{\xi^j}$$

is a probability measure and solves the optimal scenario reduction problem (OSR)

$$\min \left\{ \sup_{x \in X} \left| \sum_{j \in J} \pi_j \varphi_j(x) - \sum_{i=1}^{N} p_i \varphi_i(x) \right| : J \subset \{1, \ldots, N\}, |J| = n, \pi \in \mathcal{S}_n(J) \right\},$$

where the functions $\varphi_i(x) = \Phi(q(\xi^i), h(x, \xi^i))$, $i = 1, \ldots, N$, are convex polyhedral on $X$ and $\mathcal{S}_n(J) = \{\pi : \pi_j \geq 0, \sum_{j \in J} \pi_j = 1\}$. Problem (OSR) represents a mixed-integer semi-infinite program.

Problem (OSR) decomposes into finding the optimal index set $J$ of remaining scenarios and into determining the optimal weights $\pi_j$, $j \in J$, given $J$. The outer combinatorial optimization problem

$$\min\left\{D(J,P) : J \subset \{1,\dots,N\}, |J| = n\right\},$$

determines the index set $J$ and can be reformulated as $0$-$1$ program. The objective function $D(J,P)$ denotes the infimum of the inner program

$$\min_{\pi \in \mathcal{S}_n(J)} \sup_{x \in X} \left| \sum_{j \in J} \pi_j \varphi_j(x) - \sum_{i=1}^{N} p_i \varphi_i(x) \right|.$$

Any evaluation of the objective in the $0$-$1$ program requires the solution of the inner program, which represents a best approximation problem and is of the form

$$\min_{t \geq 0, \pi \in \mathcal{S}_n} \left\{ t \,\middle|\, \begin{array}{c} \sum_{j \in J} \pi_j \langle h(x,\xi^j), z_j\rangle \leq t + \sum_{i=1}^{N} p_i \langle q(\xi^i), y_i\rangle \\ \sum_{i=1}^{N} p_i \langle h(x,\xi^i), z_i\rangle \leq t + \sum_{j \in J} \pi_j \langle q(\xi^j), y_j\rangle \\ \forall (x,y,z) \in \mathcal{M}(\xi^1,\dots,\xi^N) \end{array} \right\},$$

where the set $\mathcal{M}(\xi^1,\dots,\xi^N)$ is defined as before but with $n$ replaced by $N$. It represents a linear semi-infinite program with only $n+1$ variables, but with a polyhedral index set of dimension $m + (\bar{m} + r)N$.

## Conclusions

- Quantitative stability results motivate the best approximation of the underlying probability distribution by discrete measures from $\mathcal{P}_n(\Xi)$ in terms of the minimal information metric $d_{\mathcal{F}}$.
- Problem-based scenario generation for two-stage models is reformulated as a (convex) generalized semi-infinite optimization problem.
- In important specific cases problem-based scenario generation allows a transformation into a (linear) semi-infinite optimization model with $n(s+1)$ variables and a $(m+(\bar{m}+r)n)$-dimensional polyhedral index set.
- Problem-based optimal scenario reduction requires solving a combinatorial program, where in each step a linear semi-infinite program with $n+1$ variables and a polyhedral index set of dimension $m+(\bar{m}+r)N$ has to be solved. The combinatorial program represents an $n$-median problem which is known to be NP-hard but for which good heuristics exist.

# References

V. Arya, N. Garg, R. Khandekar, A. Meyerson, K. Munagala, and V. Pandit: Local search heuristics for $k$-median and facility location problems, *SIAM Journal on Computing* 33 (2004), 544–562.

F. Guerra Vázquez, J.-J. Rückmann, O. Stein and G. Still: Generalized semi-infinite programming: A tutorial, *Journal of Computational and Applied Mathematics* 217 (2008), 394–419.

G. A. Hanasusanto, D. Kuhn and W. Wiesemann: A comment on "computational complexity of stochastic programming problems", *Mathematical Programming* 159 (2016), 557–569.

R. Henrion and W. Römisch: Problem-based optimal scenario generation and reduction in stochastic programming, *Mathematical Programming*, Series B (appeared online 2019).

O. Kariv, S. L. Hakimi: An algorithmic approach to network location problems, II: The $p$-medians, *SIAM Journal Applied Mathematics* 37 (1979), 539–560.

H. Leövey and W. Römisch: Quasi-Monte Carlo methods for linear two-stage stochastic programming problems, *Mathematical Programming* 151 (2015), 315–345.

W. Römisch: Stability of stochastic programming problems, in: *Stochastic Programming* (A. Ruszczyński, A. Shapiro eds.), Handbooks in Oper. Res. and Managm. Sci., Volume 10, Elsevier, Amsterdam 2003, 483–554.

J. Schwientek, T. Seidel and K.-H. Küfer: A transformation-based discretization method for solving general semi-infinite optimization problems, *Mathematical Methods of Operations Research* 93 (2021), 83–114.

O. Stein: *Bi-level Strategies in Semi-infinite Programming*, Kluwer, Boston, 2003.

G. Still: Generalized semi-infinite programming: Numerical aspects, *Optimization* 49 (2001), 223–242.

## Example: The newsboy problem

A newsboy must place a daily order for a number $x$ of copies of a newspaper. He has to pay $r$ dollars for each copy and sells a copy at $c$ dollars, where $0 < r < c$. The daily demand $\xi$ is a real random variable with (discrete) probability distribution $P \in \mathcal{P}(\mathbb{N})$, $\Xi = \mathbb{R}$, and the remaining copies $y(\xi) = \max\{0, x - \xi\}$ have to be removed. The newsboy might wish that decision $x$ maximizes his expected profit or, equivalently, minimizes his expected costs, i.e.,

$$f_0(x, \xi) = (r - c)x + c \max\{0, x - \xi\} \quad ((x, \xi) \in \mathbb{R} \times \mathbb{R}).$$

The model may be reformulated as a linear two-stage stochastic program with the optimal value function $\Phi(t) = \max\{0, -t\}$, $h(x, \xi) = \xi - x$, dual feasible set $[0, c]$ and

$$\int_{\mathbb{R}} f_0(x, \xi) dP(\xi) = rx - cx \sum_{\substack{k \in \mathbb{N} \\ k \geq x}} \pi_k - \sum_{\substack{k \in \mathbb{N} \\ k < x}} \pi_k k \,,$$

where $\pi_k$ is the probability of demand $k \in \mathbb{N}$. The unique (integer) solution is the minimal $k \in \mathbb{N}$ such that $\sum_{i=k}^{\infty} \pi_i \geq \frac{r}{c}$.

The corresponding optimal scenario generation problem (OSG) is of the form

$$\min_{t\geq 0,(\xi^1,\dots,\xi^n)\in\mathbb{R}^n}\left\{ t \;\middle|\; \begin{array}{c} \frac{1}{n}\sum_{i=1}^n(\xi^i - x)z_i \leq t + F_P(x) \\ F_P(x) \leq t + \frac{c}{n}\sum_{i=1}^n y_{2i} \\ \forall (x,y,z)\in\mathbb{R}_+\times\mathbb{R}_+^{2n}\times\mathbb{R}^n : \\ y_{2i} - y_{1i} = \xi^i - x,\ 0\leq z_i \leq c,\ i=1,\dots,n \end{array} \right\},$$

where

$$F_P(x) = \sum_{k=1}^{\infty}\pi_k\, c\, \max\{0, x-k\}\,.$$

If $\xi^i - x \geq 0$ one has $y_{2i} = \xi^i - x$, $y_{1i} = 0$, else in case $\xi^i - x \leq 0$, one has $y_{2i} = 0$, $y_{1i} = -(\xi^i - x)$. Hence, (OSG) is equivalent with

$$\min_{t\geq 0,(\xi^1,\dots,\xi^n)\in\mathbb{R}^n}\left\{ t \;\middle|\; \begin{array}{c} \frac{c}{n}\sum_{i=1}^n \max\{0, x-\xi^i\} \leq t + F_P(x) \\ F_P(x) \leq t + \frac{c}{n}\sum_{i=1}^n \max\{0, x-\xi^i\} \\ \forall x\in\mathbb{R}_+ \end{array} \right\}.$$

and

$$\min_{(\xi^1,\dots,\xi^n)\in\mathbb{R}^n}\sup_{x\in\mathbb{R}_+}\left| F_P(x) - \frac{c}{n}\sum_{i=1}^n \max\{0, x-\xi^i\}\right|.$$