Success Stories in Stochastic Programming

Andrzej Ruszczyński

Rutgers University

Optimization models with uncertain data

$$\begin{aligned} \text{minimize } & f(x, Y) \\ \text{subject to } & g(x, Y) \leq 0 \\ & x \in \mathcal{D} \end{aligned}$$

$f : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$ objective function

$\mathcal{D} \subseteq \mathbb{R}^n$ deterministic constraints (nonempty closed set)

Y - an uncertain vector in $\mathbb{R}^s$

$g : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^m$ constraint function

Questions

- What does "uncertain" mean?
- What is a "better" value of the objective?
- What is a "feasible" solution?

Stochastic programming uses probabilistic models of uncertainty:

Y is a random vector on a probability space $(\Omega, \mathcal{F}, P)$

Probabilistic Models of Uncertainty

Benefits

- The wealth of the theory of probability
- Connection to real data via statistics
- Common language with other areas of science and technology, where probabilistic models are common

Disadvantages

- Theoretical difficulty
- Computational complexity

Non-Stochastic Approaches

Worst-case models (robust optimization)

$$
\begin{aligned}
\text{minimize} \quad & \max_{y \in \mathcal{Y}} f(x, y) \\
\text{subject to} \quad & \max_{y \in \mathcal{Y}} g(x, y) \le 0 \\
& x \in \mathcal{D} \\
& \mathcal{Y} - \text{ ``uncertainty set'' to be specified in advance}
\end{aligned}
$$

Fuzzy set theory, etc...

Expected value model

$$\begin{aligned} \text{min } & \mathbb{E}[f(x,Y)] && \text{(optimization on average)} \\ \text{subject to } & \mathbb{E}[g_i(x,Y)] \le 0, \quad i = 1,\dots,m && \text{(feasibility on average)} \\ & x \in \mathcal{D} \end{aligned}$$

Probabilistic (chance) constraints

$$\begin{aligned} & \mathbb{P}\{g_i(x,Y) \le 0\} \ge p_i, \quad i = 1,\dots,m && \text{(individual constraints)} \\ & \mathbb{P}\{g_i(x,Y) \le 0 \quad i = 1,\dots,m\} \ge p && \text{(joint constraints)} \end{aligned}$$

Penalty for violating the constraints

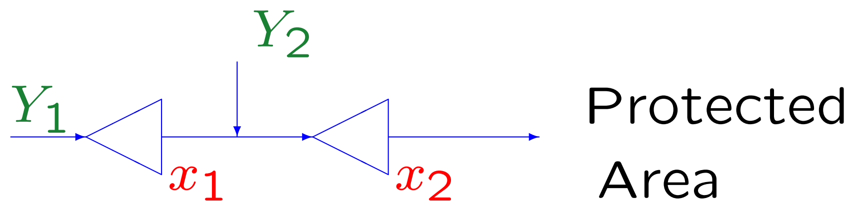
Penalty function $Q : \mathbb{R}^m \to \mathbb{R}_+$ non-decreasing and $Q(t) = 0$ for $t \le 0$

$$\min_{x \in \mathcal{D}} \; \mathbb{E}[f(x,Y) + Q(g_1(x,Y), g_2(x,Y), \dots, g_m(x,Y))].$$

Example: $Q(z) = \sum_{j=1}^m q_j \max(0, z_j)$

Water reservoir system

A system of two reservoirs has to retain the flood in the protected area.
Two random inflows, Y_1 and Y_2 cause flood danger simultaneously.
Objective: Determine suitable reservoir capacities x_1 and x_2.



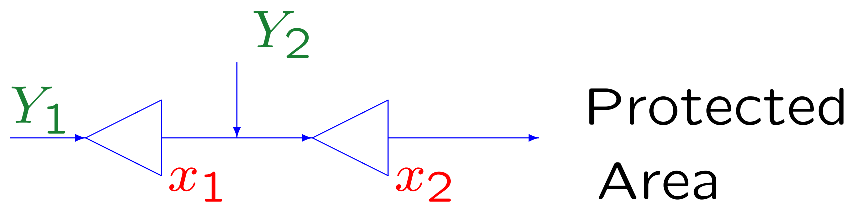
Model with expected value constraints

Annual cost of the reservoirs: $f(x_1, x_2)$
The size of the flood: $z = \max\{0, Y_1 + Y_2 - x_1 - x_2, Y_2 - x_2\}$.
The damage from flood of size $z \geq 0$ is modeled by a *convex nondecreasing* function $L(z)$ with $L(0) = 0$.

$$
\begin{array}{ll}
\min & f(x_1, x_2) \\
\text{s.t.} & \mathbb{E}\Big[L\left(\max\{0, Y_1 + Y_2 - x_1 - x_2, Y_2 - x_2\}\right)\Big] \leq b \\
 & x_1 \geq 0,\ x_2 \geq 0
\end{array}
$$

Water reservoir system

A system of two reservoirs has to retain the flood in the protected area.
Two random inflows, Y_1 and Y_2 cause flood danger simultaneously.
Objective: Determine suitable reservoir capacities x_1 and x_2.

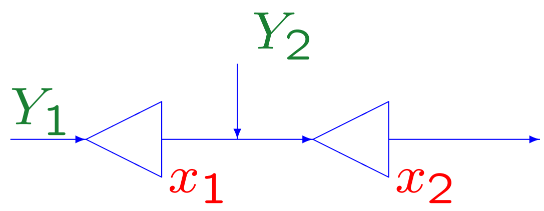


Model with probabilistic constraints

Cost for building the reservoirs $f(x_1, x_2)$.
The size of the flood $Z = \max\{0, Y_1 + Y_2 - x_1 - x_2, Y_2 - x_2\}$.

$$\begin{aligned} \min \quad & f(x_1, x_2) \\ \text{s.t.} \quad & \mathbb{P}\Big\{ \max\{0, Y_1 + Y_2 - x_1 - x_2, Y_2 - x_2\} \le \hat{z} \Big\} \ge p \\ & x_1 \ge 0, \; x_2 \ge 0 \end{aligned}$$

Water reservoir system



Protected
Area

Robust optimization model

Cost for building the reservoirs $f(x_1, x_2)$.
The size of the flood $Z = \max\{0, Y_1 + Y_2 - x_1 - x_2, Y_2 - x_2\}$.

$$\begin{aligned} \min \ & f(x_1, x_2) \\ \text{s.t.} \ & \max_{(Y_1,Y_2)\in\mathcal{Y}} \max\{0, Y_1 + Y_2 - x_1 - x_2, Y_2 - x_2\} \le \hat{z} \\ & x_1 \ge 0, \ x_2 \ge 0 \end{aligned}$$

Solution:

$$\hat{x}_2 = \Big[\max_{(Y_1,Y_2)\in\mathcal{Y}} Y_2 - \hat{z}\Big]_+$$

$$\hat{x}_1 = \Big[\max_{(Y_1,Y_2)\in\mathcal{Y}} (Y_1 + Y_2) - \hat{x}_2 - \hat{z}\Big]_+$$

Probabilistic optimization

(Charnes, Cooper, Symonds (1958), Prékopa (1970...))

$$\begin{array}{lrl} & \text{min} & f(x) \\ \text{(P)} & \text{subject to} & G(x) := \mathbb{P}\{g(x,Y) \geq 0\} \geq p \\ & & x \in \mathcal{D} \end{array}$$

Special case - $g(x, Y)$ is separable: $\mathbb{P}\{g(x) \geq Y\} \geq p$

$$\text{(P)} \qquad \ldots \text{subject to } F_Y(G(x)) \geq p$$

$F_Y(\cdot)$ - the distribution function of Y

Major challenges

- Structure of the feasible set:
  convexity, connectedness
  properties of G or F_Y ("concavity", differentiability)

- Numerical evaluation:
  how to calculate $G(x)$ or $F_Y(g(x))$?
  how to determine the level sets of G or F_Y?

Convexity Analysis: Generalized concave probability measures

Definition: A probability measure μ on $\mathbb{R}^s$ is called r-concave if

$$\mu(\lambda A + (1-\lambda)B) \geq \Big[\lambda[\mu(A)]^r + (1-\lambda)[\mu(B)]^r\Big]^{1/r} \qquad \text{(for } r \neq 0)$$

$$\mu(\lambda A + (1-\lambda)B) \geq [\mu(A)]^\lambda [\mu(B)]^{(1-\lambda)} \qquad \text{(for } r = 0)$$

for all $A, B \subset \mathbb{R}^s$ Borel measurable convex sets and all $\lambda \in [0,1]$

Theorem (Prékopa (1973), Rinott, Borell): A measure μ is r-concave ($r \leq 1/s$) iff it has $r/(1-sr)$-concave density

Examples: uniform, normal, beta, Dirichlet, gamma, Pareto, Cauchy

Consider the feasible set:

$$\mathcal{Z}_p = \Big\{x \in \mathbb{R}^n : \mathbb{P}\{\ g(x, Y) \geq 0\} \geq p\Big\}$$

Theorem (Prékopa): If the measure μ_Y is r-concave and $g(\cdot,\cdot)$ is quasi-concave, then the set $\mathcal{Z}_p$ is convex

Separable constraints: p-efficient points

Consider the feasible set:

$$\mathcal{Z}_p = \left\{ x \in \mathbb{R}^n : \mathbb{P}\{\ g(x) \geq Y\} \geq p \right\}$$

Definition: A point $v \in \mathbb{R}^s$ is called p-efficient if $\mathbb{P}\{Y \leq v\} \geq p$, and there is no $w \leq v$, $w \neq v$, such that $\mathbb{P}\{Y \leq w\} \geq p$.

Disjunctive representation of the feasible set:

$$\mathcal{Z}_p = \left\{ x \in \mathbb{R}^n : \exists\ p\text{-efficient } v \text{ such that } g(x) \geq v \right\}$$

Solution approach (Dentcheva–R. 2002...):
- generate a subset of "reasonable" p-efficient points $v^1, \ldots, v^N$
- solve the disjunctive problem

$$\begin{aligned} \min\ & f(x) \\ \text{subject to } & g(x) \geq v^k \quad \text{for at least one } k \qquad \text{(or convex combination)} \\ & x \in \mathcal{D} \end{aligned}$$

-find a new p-efficient point v^{N+1} (column generation), etc....

$$\begin{aligned} &\min\ f(x) \\ &\text{subject to } \mathbb{P}\{g(x,Y) \le 0\} \ge p \\ &x \in \mathcal{D} \end{aligned}$$

Monte Carlo Approximation: Sample $Y^1, \ldots, Y^N$ and solve

$$\begin{aligned} &\min\ f(x) \\ &\text{subject to } g(x,Y^k) \le 0 \quad \text{for **at least** } \lceil pN \rceil \text{ sample points } Y^k \\ &x \in \mathcal{D} \end{aligned}$$

- Mixed-integer reformulation is obvious
- Precedence-constrained knapsack structure (valid inequalities)
- Huge symmetry and combinatorial difficulty

Almost-Guaranteed Monte Carlo (Calafiore-Campi, 2004):

$$\text{subject to } g(x,Y^k) \le 0 \quad \text{for **all** sample points } Y^k$$

If N is **very** large, the probability that the probabilistic constraint will be satisfied is high

Linear Constraints With Random Coefficients

$$\mathbb{P}\left\{ \sum_{i=1}^{n} x_i Y_i \le b \right\} \ge p \qquad \text{(P)}$$

In general, non-convex feasible set. Focus on **scalar** Y_i's and b.

Floating Body Theorem (Ball 1988) If the joint distribution of Y is log-concave and symmetric (about some point) then the feasible set defined by (P) is convex for every $\frac{1}{2} \le p < 1$

(Rather) Conservative Convex Approximations of (P):

$$F_{\langle x,Y\rangle}^{(-1)}(p) \le b \qquad \text{which is (P)}$$

- Integrated chance constraints (Klein Haneveld, 1986)
- CVaR (Uryasev, Rockafellar, 2001)

$$\frac{1}{1-p}\int_p^1 F_{\langle x,Y\rangle}^{(-1)}(\alpha)\, d\alpha \le b \qquad \text{which implies (P)}$$

- Generating function approximations for *independent* Y_i's (Pinter, 1989, Nemirovski-Shapiro 2006).

Why not Central Limit Theorem?

Two-stage problems (Beale, Dantzig, 1955)

First stage (here-and-now decision)

$$\begin{array}{ll} \min_x & c^T x + \mathbb{E}[Q(x, D)] \\ \text{s.t.} & Ax = b,\ x \geq 0, \end{array}$$

$Q(x, D)$ is the optimal value of the second stage problem with random data $D := (q, h, T, W)$

Second stage (recourse decision)

$$\begin{array}{ll} \min_y & q^T y \\ \text{s.t.} & Tx + Wy = h,\ y \geq 0. \end{array}$$

New issues

- Time
- Information
- Nested Optimization

Network Capacity Expansion

Directed graph with node set $\mathcal{N}$ and arc set $\mathcal{A}$.
Random demand D^{mn} for each pair of nodes $(m, n) \in \mathcal{N} \times \mathcal{N}$
Objective: Decide on capacity x_a to be installed on each arc $a \in \mathcal{A}$ now to carry the random demand later, minimizing the installation and shipping cost. Denoting y_a^{mn} amount of shipment from m to n.

First stage

$$\min_{x \geq 0} \sum_{a \in \mathcal{A}} c_a x_a + \mathbb{E}[Q(x, D)].$$

Second stage: multi-commodity network flow problem

$$\begin{aligned}
\min \quad & \sum_{m,n \in \mathcal{N}} \sum_{a \in \mathcal{A}} q_a y_a^{mn} \\
\text{s.t.} \quad & \sum_{a \in \mathcal{A}_+(i)} y_a^{mn} - \sum_{a \in \mathcal{A}_-(i)} y_a^{mn} = \begin{cases} -D^{mn}, & \text{if } i = m, \\ D^{mn}, & \text{if } i = n, \\ 0, & \text{otherwise,} \end{cases} \\
& \sum_{m,n \in \mathcal{N}} y_a^{mn} \leq x_a, \quad a \in \mathcal{A}, \\
& y_a^{mn} \geq 0, \quad a \in \mathcal{A},\ i, m, n \in \mathcal{N}.
\end{aligned}$$

$\mathcal{A}_+(i)$ and $\mathcal{A}_-(i)$ are the set of arcs entering and leaving node i

Discrete Distributions

The random data D have a discrete distribution with possible realizations (*scenarios*) $D_s = (q_s, h_s, T_s, W_s)$ with probabilities p_s, $s = 1..S$.

$$\mathbb{E}[Q(x,D)] = \sum_{s=1}^{S} p_s Q(x, D_s),$$

where

$$Q(x, D_s) = \inf \left\{ q_s^T y_s : T_s x + W_s y_s = h_s,\ y_s \geq 0 \right\} \qquad \text{(nonsmooth!)}$$

Equivalent Linear Programming Problem

$$\begin{aligned}
\min_{x, y_1, \ldots, y_S} \quad & c^T x + \sum_{s=1}^{S} p_s q_s^T y_s \\
\text{s.t.} \quad & Ax = b \\
& T_s x + W_s y_s = h_s, \qquad s = 1, \ldots, S \\
& x \geq 0,\ y_s \geq 0 \qquad s = 1, \ldots, S
\end{aligned}$$

Nonanticipativity constraints (Rockafellar-Wets, 1975 ...)

Create copies of the first stage decision x for all scenarios: x_s

$$
\begin{array}{ll}
\min\limits_{\substack{x_1,\ldots,x_S \\ y_1,\ldots,y_S}} & \sum\limits_{s=1}^{S} p_s(c^T x_s + q_s^T y_s) \\
\text{s.t.} & A x_s = b \\
& T_s x_s + W_s y_s = h_s \\
& x_s = x_j \quad \text{for all } 1 \leq s < j \leq S \quad \text{(NA)} \\
& x_s \geq 0,\ y_s \geq 0 \quad s = 1, \ldots, S
\end{array}
$$

(NA) are called nonanticipativity constraints.

Numerical solution: direct approach

Solve as large scale linear problem (Beware! Huge dimension!)

Capacity expansion example with number of nodes ν
Assume independent components of the demand vector
with r realizations each, i.e., $r^{\nu(\nu-1)}$ scenarios.

The two-stage problem has
$|\mathcal{A}| + \nu(\nu-1)|\mathcal{A}|r^{\nu(\nu-1)}$ variables
$(\nu^2(\nu-1) + |\mathcal{A}|)r^{\nu(\nu-1)}$ constraints

Numerical solution: Decomposition

Primal methods – solve the first stage problem dealing with the second stage cost by non-smooth optimization methods

- Different versions of cutting plane methods (VanSlyke–Wets, Birge ,...)
- Regularized decomposition (bundle) method (R.)
- Trust-region method (Linderoth–Wright)

Dual methods – Lagrange relaxation associated with the nonanticipativity constraints causes the dual functional to decompose into scenario terms

- Dual cutting plane methods
- Progressive hedging–operator splitting (Rockafellar–Wets)
- Bundle methods (Robinson)
- Augmented Lagrangian methods (R.)

Currently, problems with millions of scenarios and thousands of variables and constraints per scenario can be solved on networks of supercomputers
But if integer variables are present, the problem becomes very difficult

Multistage Models

Sequential decision process x_t in stages $t = 1, \dots, T$ with learning (observations) of the random data D_t:

$$\begin{array}{rr} & \text{decision } (x_1) \\ \text{observation } D_2 := (c_2, A_{21}, A_{22}, b_2) & \text{decision } (x_2) \\ \vdots & \\ \text{observation } D_T := (c_T, A_{T,T-1}, A_{TT}, b_T) & \text{decision } (x_T). \end{array}$$

Objective: design the decision process so that the expected value of the total cost is minimized.

$$\begin{array}{llllllllll} \min & \mathbb{E}\big[c_1 x_1 & + & c_2 x_2 & + & c_3 x_3 & + & \dots & + & c_T x_T\big] \\ \text{s.t.} & A_{11} x_1 & & & & & & & & & = & b_1, \\ & A_{21} x_1 & + & A_{22} x_2 & & & & & & & = & b_2, \\ & & & A_{32} x_2 & + & A_{33} x_3 & & & & & = & b_3, \\ & \multicolumn{11}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ & & & & & & & A_{T,T-1} x_{T-1} & + & A_{TT} x_T & = & b_T, \\ & x_1 \geq 0, & & x_2 \geq 0, & & x_3 \geq 0, & & \dots & & x_T \geq 0. \end{array}$$

assume c_1, A_{11} and b_1 are known.

Nonanticipativity (Rockafellar–Wets, 1975 ...)

$$
\begin{array}{llllllllll}
\min \sum_{s=1}^{S} & p_s\Big[c_1 x_1^s & + & c_2^s x_2^s & + & c_3^s x_3^s & + \quad \dots & + & c_T^s x_T^s\Big] & \\
\text{s.t.} & A_{11} x_1^s & & & & & & & & = b_1 \\
 & A_{21}^s x_1^s & + & A_{22}^s x_2^s & & & & & & = b_2^s \\
 & & & A_{32}^s x_2^s & + & A_{33}^s x_3^s & & & & = b_3^s \\
 & \multicolumn{9}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\
 & & & & & & A_{T,T-1}^s x_{T-1}^s & + & A_{TT}^s x_T^s & = b_T^s \\
 & x_1^s \geq 0, & & x_2^s \geq 0, & & x_3^s \geq 0, & \dots & & x_T^s \geq 0 & \\
 & & & & & & s = 1, \dots, S & & &
\end{array}
$$

Feasibility constraints

$$(x_1^s, x_2^s, \dots, x_T^s) \in F^s \quad \text{(feasible set for scenario } s\text{)}$$

Nonanticipativity constraints (NA)

$$x_t^s = x_t^\sigma \text{ for all } s, \sigma \text{ for which } D_{[1,t]}^s = D_{[1,t]}^\sigma, \quad t = 1, \dots, T$$

They define a subspace $\mathcal{W}$ of implementable policies

Main Ideas

- Associate with (NA) multipliers λ
- Optimality and duality theory
- Scenario decomposition methods

Risk Modeling

Random outcome (e.g., cost):

$$Z_x(\omega) = f(x,\omega), \quad \omega \in \Omega$$

$(\Omega, \mathcal{F}, P)$ - probability space, x - decision vector

Minimization of the expected value $\mathbb{E}[Z_x]$
optimizes the outcome on average

Expected Utility Models (von Neumann and Morgenstern, 1944)

$$\min_{x\in X} \mathbb{E}\left[u(Z_x)\right] \qquad \left(= \int_\Omega u\Big(f(x,\omega)\Big)\, dP(\omega) \right)$$

$u(\cdot)$ is a (dis)utility function: nondecreasing, convex, ...

Rank Dependent Utility (Quiggin, 1984, Yaari)

$$\min_{x\in X} \int_0^1 F_{Z_x}^{(-1)}(\alpha)\, dw(\alpha)$$

$w(\cdot)$ is a rank dependent (dis)utility function: nondecreasing, convex, ...

Utility functions are very difficult to elicit

Elementary Mean–Risk Models

Objective: Consider random outcome $Z_x(\omega) = f(x, \omega)$
and choose decision x to
- minimize the expected outcome, the mean $\mathbb{E}[Z_x]$,
- minimize a scalar measure of uncertainty of Z_x, the risk $r[Z_x]$

Examples of risk measures:

$r_1[Z] := \text{Var}[Z]$ (Markowitz' model)

$r_2[Z] := \mathbb{E}[(Z - \alpha)_+]$ (expected excess over the target level α)

$r_3[Z] := \left(\mathbb{E}[(Z - \mathbb{E}Z)_+^q]\right)^{1/q}$ (semideviation, Ogryczak-R.)

Optimization models:

$$\min_{x \in X} \mathbb{E}[Z_x] + \kappa\, r[Z_x], \qquad 0 \le \kappa \le \kappa_{\text{max}}$$

Interesting application of parametric optimization (R.-Vanderbei)

Note: Risk measures depend on the entire distribution, $r(Z)$ is *nonlinear in probability* and possibly *nonconvex* function of the decisions x.

Consistency with Stochastic Ordering (Ogryczak–R., 1997 ...)

$(\Omega, \mathcal{F}, P)$ - probability space, $Z : \Omega \to \mathbb{R}$ - random outcome
Stochastic order with generator $\mathcal{U}$ (set of real functions):

$$Z_1 \preceq Z_2 \quad \text{iff} \quad \mathbb{E}[u(Z_1)] \le \mathbb{E}[u(Z_2)] \quad \forall\ u \in \mathcal{U}$$

If $\mathcal{U}$ contains all nondecreasing functions - first order stochastic dominance
If $\mathcal{U}$ contains all convex nondecreasing functions - increasing convex order

Definition. A mean-risk model is consistent with stochastic order $\preceq$ if there exists $\kappa_{\max} > 0$ such that

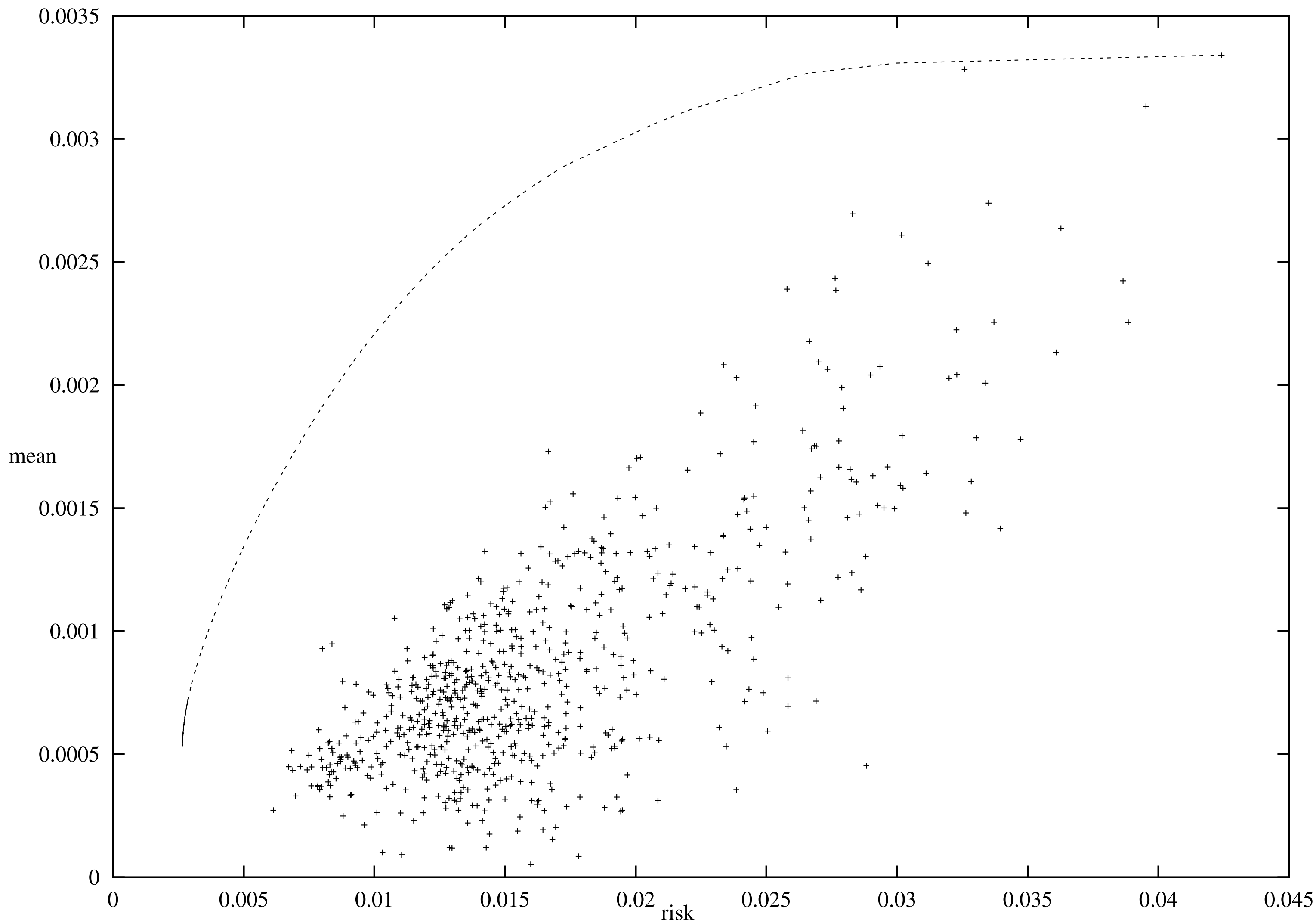
$$Z_1 \preceq Z_2 \;\Rightarrow\; \mathbb{E}[Z_1] + \kappa\, r[Z_1] \le \mathbb{E}[Z_2] + \kappa\, r[Z_2]$$

for all $0 \le \kappa \le \kappa_{\max}$.

Known risk measures consistent with the increasing convex order (second order stochastic dominance):
- expected excess over target ($\kappa_{\max} = +\infty$)
- semideviation ($\kappa_{\max} = 1$, Ogryczak–R.)
- average (conditional) value at risk ($\kappa_{\max} = +\infty$, Pflug, Ogryczak–R.))

but not variance.

Optimal solutions of mean–risk optimization models cannot be dominated



The efficient frontier for the mean-semideviation model ($0 \leq \kappa \leq 1$) in a portfolio problem with 719 securities and 3080 realizations.

General Theory of Risk Functionals

$(\Omega, \mathcal{F})$ - measurable space
$\mathcal{Z}$ - linear space of $\mathcal{F}$-measurable functions $Z : \Omega \to \mathbb{R}$
Risk functional is a $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ satisfying the conditions:

(A1) **Convexity**:

$$\rho(\alpha Z_1 + (1-\alpha) Z_2) \leq \alpha \rho(Z_1) + (1-\alpha)\rho(Z_2) \text{ for all } Z_1, Z_2 \in \mathcal{Z} \text{ and } \alpha \in [0,1]$$

(A2) **Monotonicity**: If $Z_1, Z_2 \in \mathcal{Z}$ and $Z_1(\cdot) \leq Z_2(\cdot)$, then $\rho(Z_1) \leq \rho(Z_2)$

(A3) **Translation Equivariance**: If $a \in \mathbb{R}$ and $Z \in \mathcal{Z}$, then
$\rho(Z + a) = \rho(Z) + a$

(A4) **Positive homogeneity**: If $t > 0$ and $Z \in \mathcal{Z}$, then $\rho(tZ) = t\rho(Z)$

Artzner-Delbaen-Eber-Heath (1999), Föllmer-Schied (2002), *et al.*

Conjugate Duality of Risk Functionals

Pairing of a linear topological space $\mathcal{Z}$ with a linear topological space $\mathcal{Y}$ of measures (set functions) on Ω :

$$\langle \mu, Z \rangle := \int_\Omega Z(\omega)\, d\mu(\omega).$$

Theorem (R.–Shapiro): Suppose that $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ is proper, lower semicontinuous and convex. Then

$$\rho(Z) = \sup_{\mu \in \mathcal{A}} \Big\{ \langle \mu, Z \rangle - \rho^*(\mu) \Big\}, \ \ \forall Z \in \mathcal{Z},$$

with $\mathcal{A} := \text{dom}(\rho^*) \subset \mathcal{Y}$. Moreover, we have that:
(i) Monotonicity (A2) holds iff every measure $\mu \in \mathcal{A}$ is nonnegative;
(ii) Translation equivariance (A3) holds iff $\mu(\Omega) = 1$ for every $\mu \in \mathcal{A}$;
(iii) Homogeneity (A4) holds iff $\rho^*(\mu) = 0$ for all $\mu \in \mathcal{A}$

(extends Delbaen, Föllmer–Schied, Rockafellar–Uryasev–Zabarankin, Cheridito–Delbaen–Kupper)

Optimization of Risk Functionals

Consider an uncertain outcome $Z_x(\omega) = f(x,\omega)$, $\omega \in \Omega$.
The optimization problem:

$$\min_{x\in X} \rho(Z_x) \qquad\qquad \text{(P)}$$

Theorem (R.–Shapiro): Let $f(\cdot,\omega)$ be convex for all $\omega \in \Omega$ and let $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ satisfy conditions (A1)–(A3). Suppose that a point $\hat{x} \in X$ is an optimal solution of the optimization problem (P) and that $\rho(\cdot)$ is continuous at $\hat{Z} := Z_{\hat{x}}$. Then there exists a probability measure $\hat{\mu} \in \partial\rho(\hat{Z})$ such that $\hat{x}$ is an optimal solution of the problem

$$\min_{x\in X} \left\{ \mathbb{E}_{\hat{\mu}}[Z_x] - \rho^*(\hat{\mu}) \right\} = \min_{x\in X} \max_{\mu\in\mathcal{A}} \left\{ \mathbb{E}_{\mu}[Z_x] - \rho^*(\hat{\mu}) \right\}$$

Further refinements:

- duality theory
- nonanticipativity analysis

Main Research Directions

- Optimality conditions and duality theory for optimization problems with risk functionals (R.–Shapiro)

- Analysis of subclasses of risk functionals (deviation measures, etc..) (Rockafellar–Uryasev–Zabarankin)

- Construction of dynamic risk measures (Cheridito-Delbaen-Kupper, Pflug–R., Eichhorn–Römisch)

- Axiomatic approach to dynamic risk measures, representation theorems (Riedel, Artzner–Delbaen–Eber–Heath–Ku, R.–Shapiro)

- Optimality conditions for multistage stochastic optimization problems involving risk measures (R.–Shapiro)

- Numerical methods for two-stage problems involving risk measures (Ahmed)

Risk Aversion by Stochastic Ordering Constraints

(Dentcheva–R., 2003 ...)

$Z_x(\omega) = f(x,\omega)$ - random outcome

Y - benchmark random outcome, e.g. $Y(\omega) = f(\bar{x},\omega)$ for some $\bar{x} \in X$

New Model:

$$\min \ \mathbb{E}[Z_x] \qquad \text{(or some other objective)}$$
$$\text{subject to } Z_x \preceq Y \qquad \text{(stochastic ordering constraint)}$$
$$x \in X$$

The ordering constraint reflects risk aversion:

$$\mathbb{E}[u(Z_x)] \le \mathbb{E}[u(Y)] \quad \text{for all} \quad u \in \mathcal{G} \quad \text{(generator)}$$

Lagrangian:

$$L(x,u) = \mathbb{E}[Z_x + u(Z_x) - u(Y)]$$

with $u \in \mathcal{G}$. The function $t \to t + u(t)$ is the implied (dis)utility.

- optimality, duality, first- vs second-order constraints, inverse forms ...

Plenary talk by Darinka Dentcheva, Monday 14:30

Reference

Stochastic Programming

Handbook in Operations Research and Management Science

Andrzej Ruszczyński and Alexander Shapiro (Editors)

Elsevier Science, Amsterdam 2003