Stochastic Orders in Risk-averse Optimization

Darinka Dentcheva

Stevens Institute of Technology Hoboken, New Jersey, USA

Research supported by NSF award DMS-1311978

June 1, 2016

Paris, France

Motivation

Risk-Averse Optimization Models

Choose a decision $z \in Z$, which results in a random outcome $G(z) \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ with “good" characteristics; special attention to low probability-high impact events.

- Utility models apply a nonlinear transformation to the realizations of $G(z)$ (expected utility) or to the probability of events (rank dependent utility/distortion). Expected utility models optimize $\mathbb{E}[u(G(z))]$
- Probabilistic / chance constraints impose prescribed probability on some events: $P[G(z) \geq \eta]$
- Mean–risk models optimize a composite objective of the expected performance and a scalar measure of undesirable realizations $\mathbb{E}[G(z)] - \varrho[G(z)]$ (risk/ deviation measures)
- Stochastic-order constraints compare random outcomes using stochastic orders and random benchmarks

Outline

References

1. D. Dentcheva, A. Ruszczyński: Stochastic optimization with dominance constraints, *SIAM J. Optimization*, 14 (2003) 548–566.
2. — " — Inverse stochastic dominance constraints and rank dependent utility theory, *Mathematical Programming* 108 (2006), 297–311.
3. — " — Optimization with multivariate stochastic dominance constraints, *Mathematical Programming*, 117 (2009) 111–127.
4. — " — Duality between coherent risk measures and stochastic dominance constraints in risk-averse optimization, *Pacific Journal on Optimization*, 4 (2008), No. 3, 433–446.
5. — " — Inverse cutting plane methods for optimization problems with second-order stochastic-dominance constraints, *Optimization*, 59 (2010) 323–338.
6. D. Dentcheva, G. Martinez, Two-stage stochastic optimization problems with stochastic-order constraints on the recourse, *European Journal of Operations Research* 1 (2012), 1, 1–8.
7. D. Dentcheva, E. Wolfhagen: Optimization with multivariate stochastic dominance constraints, *SIAM J. Optimization*, 25 (2015) 1, 564–588.
8. D. Dentcheva, W. Römisch: Stability and sensitivity of stochastic dominance constrained optimization models, *SIAM J. Optimization*, 23 (2014) 3, 1672–1688.

For $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$

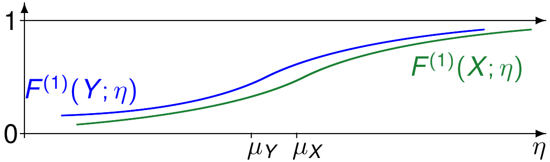
$$X \succeq_{\mathcal{U}} Y \Leftrightarrow \int_{\Omega} u(X(\omega))\, P(d\omega) \geq \int_{\Omega} u(Y(\omega))\, P(d\omega) \quad \forall\, u(\cdot) \in \mathcal{U}$$

Collection of functions $\mathcal{U}$ is the generator of the order.

Generators:

- $\mathcal{U}_1 = \{$nondecreasing functions $u : \mathbb{R} \to \mathbb{R}\}$ generates the usual stochastic order or first order stochastic dominance ($X \succeq_{(1)} Y$)
  Mann and Whitney (1947), Blackwell (1953), Lehmann (1955)
- $\mathcal{U}_2 = \{$nondecreasing concave $u : \mathbb{R} \to \mathbb{R}\}$ generates the second order stochastic dominance relation ($X \succeq_{(2)} Y$)
  Quirk and Saposnik (1962), Fishburn (1964), Hadar and Russell (1969)
- $\bar{\mathcal{U}}_2 = \{$nondecreasing convex $u : \mathbb{R} \to \mathbb{R}\}$ generates the increasing convex order ($X \preceq_{\text{ic}} Y$)

First Order Stochastic Dominance



Distribution function $F(X;\eta) = \int_{-\infty}^{\eta} P_X(dt) = P\{X \le \eta\},\ \eta \in \mathbb{R}$

Quantile function $F^{(-1)}(X;p) = \inf\{\eta : F(X;\eta) \ge p\},\ p \in (0,1)$

Survival function $\bar{F}(X;\eta) = 1 - F(X;\eta) = P\{X > \eta\},\ \eta \in \mathbb{R}$

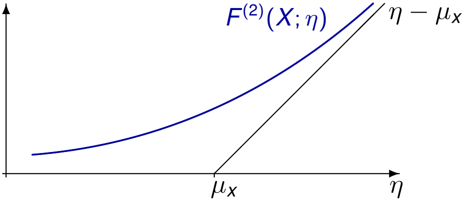
The usual stochastic order

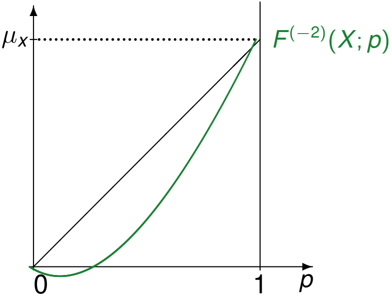
$$
\begin{aligned}
X \succeq_{(1)} Y \quad &\Leftrightarrow \quad F(X;\eta) \le F(Y;\eta) \quad \text{for all } \eta \in \mathbb{R} \\
&\Leftrightarrow \quad F^{(-1)}(X;p) \ge F^{(-1)}(Y;p) \quad \text{for all} \quad 0 < p < 1. \\
&\Leftrightarrow \quad \bar{F}(X;\eta) \ge \bar{F}(Y;\eta) \quad \text{for all} \quad \eta \in \mathbb{R}.
\end{aligned}
$$

Second-Order Stochastic Dominance

Shortfall function $F^{(2)}(X,\eta) = \int_{-\infty}^{\eta} F(X,t)\,dt = \mathbb{E}[(\eta - X)_+] \quad \eta \in \mathbb{R}.$

Lorenz function: $F^{(-2)}(X;p) = \int_0^p F^{(-1)}(X;t)\,dt \quad p \in (0,1].$





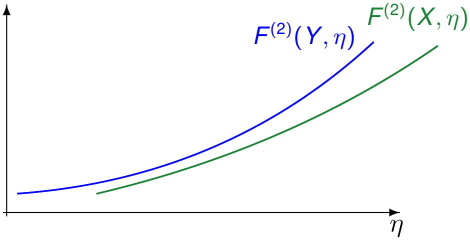
Fenchel conjugate function $\quad F^*(p) = \sup_u \{pu - F(u)\}.$

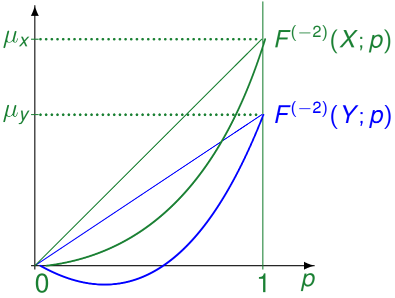
Theorem (Ogryczak and Ruszczyński, 2002)

$$F^{(-2)}(X;\cdot) = [F^{(2)}(X;\cdot)]^* \quad \text{and} \quad F^{(2)}(X;\cdot) = [F^{(-2)}(X;\cdot)]^*$$

Second-Order Stochastic Dominance

$$X \succeq_{(2)} Y \Leftrightarrow \mathbb{E}[(\eta - X)_+] \leq \mathbb{E}[(\eta - Y)_+]$$
$$\Leftrightarrow F^{(-2)}(X;p) \geq F^{(-2)}(Y;p) \quad \forall p \in [0,1].$$





Increasing convex order

Characterization by the integrated survival function

For $X, Y \in \mathcal{L}_1$, the relation $X \preceq_{\text{ic}} Y$ holds if and only if

$$\int_\eta^\infty P(X > t)\,dt \le \int_\eta^\infty P(Y > t)\,dt \quad \text{for all } \eta \in \mathbb{R}.$$

The excess function and its Fenchel conjugate

$$H(Z,\eta) = \int_\eta^\infty \overline{F}(Z,t)\,dt = \mathbb{E}(Z-\eta)_+$$

$$L(Z,q) = -\int_{1+q}^1 F^{(-1)}(Z,t)\,dt \quad \text{for } -1 \le q < 0,$$

$$L(Z,0) = 0,\ L(Z,q) = \infty \text{ for } q \notin [-1,0]$$

Increasing convex order vs. Second order dominance

$$X \preceq_{\text{ic}} Y \quad \Leftrightarrow \quad -X \succeq_{(2)} -Y.$$

Higher order relation

For any $X \in \mathcal{L}_k(\Omega, \mathcal{F}, \mathbb{P})(\Omega, \mathcal{F}, P)$, $\|X\|_k = \left(\mathbb{E}(|X|^k)\right)^{\frac{1}{k}}$ and

$$F^{(k+1)}(X;\eta) = \int_{-\infty}^{\eta} F^{(k)}(X;t)dt = \frac{1}{k!}\int_{-\infty}^{\eta}(\eta - t)^k P_X(dt) = \frac{1}{k!}\|\max\left(0, \eta - X\right)\|_k^k \quad \forall \eta \in \mathbb{R},$$

kth degree Stochastic Dominance (kSD), $k \geq 2$

$$X \succeq_{(k)} Y \quad \Leftrightarrow \quad F^{(k)}(X,\eta) \leq F^{(k)}(Y,\eta) \quad \text{for all } \eta \in \mathbb{R},$$
$$\|\max\left(0, \eta - X\right)\|_{k-1}^{k-1} \leq \|\max\left(0, \eta - Y\right)\|_{k-1}^{k-1}$$

The generator

$\mathcal{U}_k$ contains all functions $u : \mathbb{R} \to \mathbb{R}$ such that a non-increasing, left-continuous, bounded function $\varphi : \mathbb{R} \to \mathbb{R}_+$ exists such that $u^{(k-1)}(\eta) = (-1)^k \varphi(\eta)$ for a.a. $\eta \in \mathbb{R}$.

Multivariate and Dynamic Orders

Consider $X = (X_1, \dots, X_m)$ and $Y = (Y_1, \dots, Y_m)$ in $\mathcal{L}_1^m(\Omega, \mathcal{F}, P)$.

Coordinate Order

$$X \succeq_{(2)}^{\text{sep}} Y \;\Leftrightarrow\; X_t \succeq_{(2)} Y_t,\; t = 1, \dots, m$$

The generator consists of all functions $u(X) = \sum_{t=1}^{m} u_i(X_i)$ with concave nondecreasing $u_i : \mathbb{R} \to \mathbb{R}$.

Ignores temporal structure and dependency

Increasing Convex Order

$$X \succeq Y \;\Leftrightarrow\; \mathbb{E}\big[u(X)\big] \geq \mathbb{E}\big[u(Y)\big] \quad \forall\, u \in \mathfrak{F}$$

The generator consists of all concave nondecreasing functions $u : \mathbb{R}^m \to \mathbb{R}$

Hard to treat analytically, the generator is too rich

Adopted approach

Define the mutlivariate order via a family of univariate orders

Multivariate Orders: Definition

Definition

Given a closed convex set $C \in \mathbb{R}^m_+$ and a mapping $\mathfrak{M} : c \mapsto \mathcal{U}_c$, $c \in C$, where $\mathcal{U}_c$ are univariate generators, a random vector $X \in \mathcal{L}_1^m$ is stochastically larger than a random vector $Y \in \mathcal{L}_1^m$ with respect to $\mathfrak{M}$ and C if

$$c^\top X \succeq_{\mathcal{U}_c} c^\top Y \text{ for all } c \in C.$$

Example

- Set $\mathfrak{M}(c) \subset \mathbb{R}$ and $S = \{c \in \mathbb{R}^m_+ : \|c\|_1 = 1\}$. For $X, Y \in \mathcal{L}_1^m$,

$$X \succeq^{\mathfrak{M}}_{(2)} Y \Leftrightarrow \mathbb{E}[(\eta - c^\top X)_+] \leq \mathbb{E}[(\eta - c^\top Y)_+] \;\; \forall (c, \eta) \in \text{graph}\, \mathfrak{M}.$$

- If $\mathfrak{M}(c) = [a, b]$, the order is known as linear second order dominance.

Other definitions: A. Müller, D. Stoyan, Homem-de-Mello and Mehrotra: linear dominance with S a polyhedron, or a compact convex set.

Multivariate Stochastic Dominance: Generator of the order

The set $\Psi(\mathfrak{M})$ contains all mappings $\phi : c \in S \mapsto \mathcal{U}_2(\mathfrak{M}(c))$ such that $(c, x) \to [\phi(c)](c^\top x)$ is Lebesgue measurable on $S \times \mathbb{R}^m$.

$\mathcal{M}(S)$ is the space of regular countably additive measures on S with finite variation; $\mathcal{M}_+(S)$ is its subset of nonnegative measures.

With every mapping $\phi \in \Psi(\mathfrak{M})$ and every finite measure $\mu \in \mathcal{M}_+(S)$ we associate a function $\varphi_{\phi,\mu} : \mathbb{R}^m \to \mathbb{R}$ as follows:

$$\varphi_{\phi,\mu}(x) = \int_S [\phi(c)](c^\top x)\, \mu(dc).$$

Define $\mathcal{U}^m_{\mathfrak{M}} = \{\varphi_{\phi,\mu} : \phi \in \Psi(\mathfrak{M}), \mu \in \mathcal{M}_+(S)\}$.

Theorem

For each $X, Y \in \mathcal{L}^m_1$ the relation $X \succeq^{\mathfrak{M}}_{(2)} Y$ is equivalent to

$$\mathbb{E}[\varphi(X)] \geq \mathbb{E}[\varphi(Y)] \quad \text{for all } \varphi \in U^m_{\mathfrak{M}}.$$

Moreover, $\mathcal{U}^m_{\mathfrak{M}}$ is the maximal generator of the order $\succeq^{\mathfrak{M}}_{(2)}$.

Discounted Dynamic Stochastic Order

Definition

A sequence $X \in \mathcal{L}_1^T$ is stochastically larger than a sequence $Y \in \mathcal{L}_1^T$ $X \succeq_{(k)}^{\text{dis}} Y$ if

$$\langle \lambda, X \rangle \succeq_{(k)} \langle \lambda, Y \rangle \quad \forall \lambda \in \mathcal{D} \subseteq \left\{ \lambda \in \mathbb{R}^T : 1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_T \geq 0 \right\}.$$

Example Finite set $\mathcal{D}$ consisting of T elements λ^ℓ, $\ell = 1, \ldots, T$, with

$$\lambda_t^\ell = \begin{cases} 1 & \text{if } \; t \leq \ell, \\ 0 & \text{if } \; t > \ell. \end{cases}$$

Then $X \succeq_{(k)}^{\text{dis}} Y$ is equivalent to the dominance of partial sums

$$\sum_{t=1}^{\ell} X_t \succeq_{(k)} \sum_{t=1}^{\ell} Y_t, \quad \ell = 1, \ldots, T.$$

Rank Dependent Utility Functions/Distortions

$\mathcal{W}_1 = \{w : [0,1] \to \mathbb{R} : \ w \text{ is continuous nondecreasing}\}$.
$\mathcal{W}_2 = \{w \in \mathcal{W}_1 : w \text{ concave subdifferentiable at } 0\}$.

Theorem

For all random variables $X, Y \in \mathcal{L}_1$, the relation $X \succeq_{(i)} Y$, $i = 1, 2$ holds if and only if for all $w \in \mathcal{W}_i$

$$\int_0^1 F^{(-1)}(X;p)\, dw(p) \geq \int_0^1 F^{(-1)}(Y;p)\, dw(p). \tag{1}$$

Corollary: $X \preceq_{\text{ic}} Y \Leftrightarrow$ (1) for all convex functions w.

Let $\mathfrak{M}^- : S \rightrightarrows (0,1)$ have closed convex images.

Multivariate distortions

For $X, Y \in \mathcal{L}_1^m$, the relation $X \succeq_{(i)} Y$, $i = 1, 2$ holds if and only if for all measurable $\vartheta : c \in S \to \mathcal{W}_i(\mathfrak{M}^-(c))$ and all $\mu \in \mathcal{M}_+(S)$

$$\int_S \int_0^1 F^{(-1)}(c^\top X;p)\, d\vartheta_c(p)\, d\mu(c) \geq \int_S \int_0^1 F^{(-1)}(c^\top Y;p)\, d\vartheta_c(p)\, d\mu(c) \tag{2}$$

For all $k \geq 1$, Y - benchmark outcome in $\mathcal{L}_{k-1}(\Omega, \mathcal{F}, P)$.

Acceptance sets $A_k(Y) = \{X \in \mathcal{L}_{k-1} : X \succeq_{(k)} Y\}$

Theorem

The set $A_k(Y)$ is convex and closed for all Y, and $k \geq 2$. Its recession cone has the form

$$A_k^\infty(Y) = \left\{H \in \mathcal{L}_{k-1}(\Omega, \mathcal{F}, P) : H \geq 0 \text{ a.s. } \right\}$$

$A_1(Y)$ is *closed* and $A_k(Y) \subseteq A_{k+1}(Y) \quad \forall k \geq 1$. $A_k(Y)$ is a cone pointed at Y if and only if Y is a constant.

Theorem

If $(\Omega, \mathcal{F}, P)$ is atomless, then $A_2(Y) = \overline{\text{conv}}\, A_1(Y)$ If $\Omega = \{1..N\}$, and $P[k] = 1/N$, then $A_2(Y) = \text{conv } A_1(Y)$

The result is not true for general probability spaces

Dominance Relation as Constraints in Optimization

$$\begin{aligned} &\min f(z) \\ &\text{subject to } G_i(z) \succeq_{\mathcal{U}_i} Y_i, \quad i = 1..m \\ &z \in Z \end{aligned}$$

Y_i - benchmark random outcome

The stochastic order constraints reflect risk aversion

- $G_i(z)$ is preferred over Y_i by all risk-averse decision makers with utility functions in the generator $\mathcal{U}_i$;
- Easier consensus on a benchmark rather than a utility function;
- Data of a benchmark is readily available.

$G(z) \succeq_{(1)} Y \equiv$ continuum of chance constraints;
$G(z) \succeq_{(2)} Y \equiv$ continuum of CV@R constraints.

Introduced by Dentcheva and Ruszczyński (2003)

Second Order Multivariate Dominance Constraints

Given $Y \in \mathcal{L}_1^m$ - benchmark random vector

Direct Stochastic Order Constraints

$$(\mathscr{P}) \qquad \min f(z)$$
$$\text{s. t. } \mathbb{E}[(\eta - c^\top G(z))_+] \le \mathbb{E}[(\eta - c^\top Y)_+]$$
$$\forall\, (c,\eta) \in \operatorname{graph} \mathfrak{M},$$
$$z \in Z.$$

Let $\mathfrak{M}^- : S \rightrightarrows (0,1)$ have closed convex images.

Inverse Stochastic Order Constraints

$$(\mathscr{Q}) \qquad \min f(z)$$
$$\text{s. t. } F^{(-2)}(c^\top G(z); p) \ge F^{(-2)}(c^\top Y; p)$$
$$\forall\, (c,p) \in \operatorname{graph} \mathfrak{M}^-,$$
$$z \in Z.$$

Z is a closed subset of a Banach space $\mathcal{X}$; $\mathfrak{M}(c)$, resp. $\mathfrak{M}^-(c)$ are compact, $G : \mathcal{X} \to \mathcal{L}_1^m$ is continuous and for P-almost all $\omega \in \Omega$ the functions $[G_i(\cdot)](\omega)$ are concave and continuous. $f : \mathcal{X} \to \mathbb{R}$ is convex and continuous.

Optimality Conditions Using von Neumann Utility Functions

The Lagrangian-like functional $L : \mathcal{X} \times \mathcal{U}^m_{\mathfrak{M}} \to \mathbb{R}$

$$L(z, u) = f(z) - \mathbb{E}\big[u(G(z)) - u(Y)\big]$$

Uniform Dominance Condition (UDC) for problem $(\mathscr{P})$

$$\exists \tilde{z} \in Z : \inf_{(\eta, c) \in \operatorname{graph} \mathfrak{M}} \big\{ F^{(2)}(c^\top Y; \eta) - F^{(2)}(c^\top G(\tilde{z}); \eta) \big\} > 0.$$

Theorem

Assume UDC. If $\hat{z}$ is an optimal solution of $(\mathscr{P})$ then $\hat{u} \in \mathcal{U}^m_{\mathfrak{M}}$ exists:

$$L(\hat{z}, \hat{u}) = \min_{z \in Z} L(z, \hat{u}) \tag{3}$$

$$\mathbb{E}\big[\hat{u}(G(\hat{z})\big] = \mathbb{E}\big[\hat{u}(Y)\big] \tag{4}$$

If for some $\hat{u} \in \mathcal{U}^m_{\mathfrak{M}}$ an optimal solution $\hat{z}$ of (3) satisfies the dominance constraints and (4), then $\hat{z}$ solves $(\mathscr{P})$.

Optimality Conditions Using Rank Dependent Utility Function

Lagrangian-like functional

$$\Lambda(z,\vartheta,\mu) = f(z) - \int_S \int_0^1 F^{(-1)}(c^\top G(z);p) - F^{(-1)}(c^\top Y;p)\; d\vartheta_c(p)\, d\mu(c)$$

Uniform inverse dominance condition (UIDC) for $(\mathscr{Q})$

$$\exists \tilde{z} \in Z \quad \inf_{(p,c)\in \operatorname{graph} \mathfrak{M}^-} \left\{ F^{(-2)}(c^\top G(\tilde{z});p) - F^{(-2)}(c^\top Y;p) \right\} > 0.$$

Theorem

Assume UIDC. If $\hat{z}$ is a solution of $(\mathscr{Q})$, then $\hat{\vartheta} : S \to \mathcal{W}_2(\mathfrak{M}^-)$ and $\hat{\mu} \in \mathcal{M}_+(S)$ exist:

$$\Lambda(\hat{z},\hat{\vartheta},\hat{\mu}) = \min_{z\in Z} \Lambda(z,\hat{\vartheta},\hat{\mu}) \tag{5}$$

$$\int_S \int_0^1 F^{(-1)}(G(\hat{z});p)\, d\hat{\vartheta}_c(p)\, d\hat{\mu}(c) = \int_S \int_0^1 F^{(-1)}(Y;p)\, d\hat{\vartheta}_c(p)\, d\hat{\mu}(c) \tag{6}$$

If for some $\hat{\vartheta} : S \to \mathcal{W}(\mathfrak{M}^-)$, $\hat{\mu} \in \mathcal{M}_+(S)$ and a solution $\hat{z}$ of (5) the order constraint and (6) are satisfied, then $\hat{z}$ is a solution of $(\mathscr{Q})$.

Duality Relations to Utility Theories

The Dual Functionals

$D(u) = \inf_{z \in Z} L(z, u) \qquad \Delta(\vartheta, \mu) = \inf_{z \in Z} \Lambda(z, \vartheta, \mu)$

The Dual Problems

$(\mathfrak{D}_2) \quad \max_{u \in \mathcal{U}^m(\mathfrak{M})} D(u) \qquad (\mathfrak{D}_{-2}) \quad \max_{\vartheta, \mu} \Phi(\vartheta, \mu).$

Theorem

Under UDC/UIDC, if problem $(\mathscr{P})$ resp. $(\mathscr{Q})$ has an optimal solution, then the corresponding dual problem has an optimal solution and the same optimal value. The optimal solutions of the dual problem are the utility functions $\hat{u} \in \mathcal{U}^m(\mathfrak{M})$ satisfying (3)–(4) for an optimal solution $\hat{z}$ of problem $(\mathscr{P})$. The optimal solutions of $(\mathfrak{D}_{-2})$ provide rank dependent utility functions $\hat{\vartheta}_c \in \mathcal{W}(\mathfrak{M}^-)$ and non-negative measures on S satisfying (5)–(6) for an optimal solution $\hat{z}$ of problem $(\mathscr{Q})$.

A coherent measure of risk for random variables expressing cost or losses is a functional $\varrho : \mathcal{L}_1(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ satisfying the axioms:

- Convexity: $\varrho(\alpha X + (1-\alpha)Y) \le \alpha\varrho(X) + (1-\alpha)\varrho(Y)$ for all $X, Y \in \mathcal{L}_1$, $\forall \alpha \in [0,1]$.
- Monotonicity: If $X(\omega) \le Y(\omega)$ $\forall \omega \in \Omega$, then $\varrho(X) \le \varrho(Y)$.
- Translation Equivariance : $\varrho(X + a) = \varrho(X) + a \quad \forall a \in \mathbb{R}$.
- Positive homogeneity: $\varrho(tX) = t\varrho(X) \quad \forall t > 0$.

A coherent measure of risk for random variables expressing gains is a functional $\varrho : \mathcal{L}_1(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ satisfying the convexity, positive homogeneity, and the following modified axioms:

- Monotonicity: If $X(\omega) \le Y(\omega)$ $\forall \omega \in \Omega$, then $\varrho(X) \ge \varrho(Y)$.
- Translation Equivariance : $\varrho(X + a) = \varrho(X) - a \quad \forall a \in \mathbb{R}$.

Definition

A risk measure $\varrho : \mathcal{Z} \to \overline{\mathbb{R}}$ is law invariant, if it assigns the same value to random variables which have the same distribution.

Relations to Measures of Risk

Consistency of law-invariant risk measure with orders

Suppose the probability space is nonatomic. A risk measure is consistent with the usual stochastic order iff it satisfies the monotonicity axiom. A proper lower semicontinuous convex risk measure $\varrho : \mathcal{Z} \to \overline{\mathbb{R}}$ for variables expressing losses (gains) is consistent with the increasing convex order (second-order stochastic dominance), i.e.

$$X \preceq_{\text{ic}} Y \;\; (X \succeq_{(2)} Y) \quad \Rightarrow \quad \varrho(X) \le \varrho(Y).$$

$$\max_{z,\sigma} \{ f(z) - \lambda\sigma \; : \; z \in Z, \; G(z) + \sigma \succeq_{(2)} Y \}$$

$\lambda > 0$ is a tradeoff between $f(\cdot)$ and the error in dominating.

Proposition

The optimal value of σ is a coherent measure of risk expressing losses.

Mean-risk models as Lagrangian Relaxation

Kusuoka representation

If Ω is atomless, then for every law invariant measure of risk on $\varrho : \mathcal{L}_\infty(\Omega, \mathcal{F}, P) \to \mathbb{R}$ a convex set $\mathcal{M}_\varrho$ of probability measures on $(0, 1]$ exists such that

$$\varrho(X) = \sup_{\mu \in \mathcal{M}_\varrho} -\int_0^1 \frac{1}{p} F^{(-2)}(X; p)\, \mu(dp) \quad \forall X \in \mathcal{L}_\infty.$$

$$(\mathcal{Q}') \quad \max \Big\{ f(X) \ : \ F^{(-2)}(X; p) \geq F^{(-2)}(Y; p),\ p \in [\alpha, \beta], X \in C \Big\}.$$

Theorem

Under the UIDC, if $\hat{X}$ is an optimal solution of ($\mathscr{Q}$'), then a law-invariant coherent risk measure $\hat{\varrho}$ (with $\mathcal{M}_{\hat{\varrho}}$ being a singleton) and $\kappa \geq 0$ exist such that $\hat{X}$ is a solution of the mean-risk problem

$$\max_{X \in C} \{ f(X) - \kappa \hat{\varrho}(X) \} \quad \text{and} \quad \kappa \hat{\varrho}(X) = \kappa \hat{\varrho}(Y).$$

If the dominance constraint is active, then $\hat{\varrho}(X) = \hat{\varrho}(Y)$.

The Implied Dominance Constraint

Given the problem

$$(\mathcal{R}) \qquad \min_{X \in C} \left\{ f(X) - \kappa \varrho(X) \right\}$$

$\varrho(\cdot)$ a coherent law invariant measure of risk and $\kappa > 0$

Theorem If $\mathcal{M}_\varrho$ is compact* in $\mathcal{M}([0,1])$ and $\hat{X}$ is a solution of problem $(\mathcal{R})$, then $\exists\, \hat{\mu} \in \mathcal{M}$ such that

$$\varrho(\hat{X}) = \int_0^1 \frac{1}{p} F^{(-2)}(\hat{X}; p)\, \hat{\mu}(dt),$$

and for every $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$ satisfying the conditions

$$F^{(-2)}(Y;t) \le F^{(-2)}(\hat{X};t), \quad \text{for all} \quad t \in [0,1],$$
$$F^{(-2)}(Y;t) = F^{(-2)}(\hat{X};t), \quad \text{for all} \quad t \in \operatorname{supp}(\hat{\mu}),$$

the point $\hat{X}$ is also a solution of problem $(\mathscr{Q}')$ with $[\alpha, \beta] = [0,1]$.

Stochastic Dominance Constraints: The Nonconvex Case

$$(\mathscr{P}') \qquad \begin{aligned} &\min f(H(z)) \\ &\text{s.t. } F^{(2)}(G(z);\eta) \le F^{(2)}(Y;\eta) \quad \eta \in [a,b] \\ &\quad z \in Z \end{aligned}$$

Assumption

$H(\cdot)$ and $G(\cdot)$ are continuously Fréchet differentiable.

We define $I_0(z) = \{\eta \in [a,b] : F^{(2)}(G(z);\eta) = F^{(2)}(Y;\eta)\}$.

The differential constraint qualification condition at $z_0 \in Z$

A point $z_S \in Z$ and a constant $\delta > 0$ exist such that for all $\eta \in I_0(z_0)$

$$\int_{\{G(z_0)<\eta\}} [G'(z_0)(z_S - z_0)](\omega)\, P(d\omega)$$

$$\ge \int_{\{G(z_0)=\eta\}} \max\big(0, -[G'(z_0)(z_S - z_0)](\omega)\big)\, P(d\omega) + \delta.$$

The non-convex case: continued

The normal cone of Z at $\hat{z}$ is denoted by $N_Z(\hat{z})$

Theorem

If the point $\hat{z}$ is a local minimum of problem $\mathscr{P}'$ and the differential constraint qualification condition is satisfied at $\hat{z}$, then a function $\hat{u} \in \mathcal{U}_2$ exists such that

$$0 \in [H'(\hat{z})]^* \partial f(H(\hat{z})) - [G'(\hat{z})]^* \int_{\Omega} \partial u(G(\hat{z}))\, P(d\omega) + N_Z(\hat{z}),$$

$$\mathbb{E}\big[u(G(\hat{z}))\big] = \mathbb{E}\big[u(Y_0)\big].$$

The integral $\int_{\Omega} \partial u(Y)\, P(d\omega)$ is understood as a collection of integrals of all measurable selections of the multifunction $\omega \to \partial u(Y(\omega))$.

Key idea of the proof Local linearization of H and G around $\hat{z}$ and convex approximation of the objective function and the feasible set such that the tangent cones of both feasible sets at $\hat{z}$ are the same. (DD, A. Ruszczyński, Composite semi-infinite optimization, Control and Cybernetics, 36 (2007) 3, 633–646).

Control problem with order constraint and its risk-neutral equivalent

$$
(\mathfrak{C}) \qquad
\begin{aligned}
\max\ & \sum_{t=1}^{T} \mathbb{E}\big[G_t(s_t, v_t)\big] + \mathbb{E}\big[G_{T+1}(s_{T+1})\big] \\
\text{s.t. }\ & s_{t+1} = A_t s_t + B_t v_t + e_t, \quad t = 1, \dots, T, \\
& \big(G_1(s_1, v_1), \dots, G_1(s_T, v_T), G_{T+1}(s_{T+1})\big) \succeq_{(2)}^{\text{dis}} (Y_1, \dots, Y_T, Y_{T+1}) \\
& v_t \in V_t \text{ a.s.}, \quad t = 1, \dots, T.
\end{aligned}
$$

Theorem

If $(\hat{s}, \hat{v})$ is an optimal solution of problem $(\mathfrak{C})$, then a random discount sequence $\xi_t \in \mathcal{L}_\infty(\Omega, \mathcal{F}_t, P)$, $t = 1, \dots, T+1$, exists such that $(\hat{s}, \hat{v})$ is an optimal solution of the problem

$$
\begin{aligned}
\max\ & \sum_{t=1}^{T} \mathbb{E}\big[(1 + \xi_t)G_t(s_t, v_t)\big] + \mathbb{E}\big[(1 + \xi_{T+1})G_{T+1}(s_{T+1})\big] \\
\text{s.t. }\ & s_{t+1} = A_t s_t + B_t v_t + e_t, \quad t = 1, \dots, T, \\
& v_t \in V_t \quad \text{a.s.}, \quad t = 1, \dots, T.
\end{aligned}
$$

Maximum Principle

If $(\hat{s}, \hat{v})$ is an optimal solution of the control problem, then there exist: discount factors $\xi_t \in \mathcal{L}_\infty(\Omega, \mathcal{F}_t, P)$, $t = 1, \ldots, T+1$,

$$\xi_1 \geq \xi_2 \geq \cdots \geq \xi_T \geq \xi_{T+1} \geq 0 \quad \text{a.s.}$$

subgradients $(\sigma_t^s, \sigma_t^v) \in \mathcal{L}_q^{n_s}(\Omega, \mathcal{F}_t, P) \times \mathcal{L}_q^{n_v}(\Omega, \mathcal{F}_t, P)$ satisfying

$$(\sigma_t^s(\omega), \sigma_t^v(\omega)) \in \partial g_t(\hat{s}_t(\omega), \hat{v}_t(\omega))$$
$$\sigma_{T+1}^s(\omega) \in \partial g_{T+1}(\hat{s}_{T+1}(\omega))$$

and dual variables $\hat{y}_t \in \mathcal{L}_q^{n_s}(\Omega, \mathcal{F}_{t+1}, P)$, $t = 1, \ldots, T$, satisfying the adjoint equations

$$y_T = (1 + \xi_{T+1})\sigma_{T+1}^s$$
$$y_{t-1} = A_t' \mathbb{E}[y_t | \mathcal{F}_t] + (1 + \xi_t)\sigma_t^s, \quad t = T, \ldots, 2$$

such that for almost all $\omega \in \Omega$ the control $\hat{v}_t(\omega)$ is a solution of

$$\max_{c \in V_t} (1 + \xi_t(\omega)) g_t(\hat{s}_t(\omega), c) + \left\langle \mathbb{E}[y_t | \mathcal{F}_t](\omega), B_t(\omega) c \right\rangle$$

Challenge: The decisions are not time-consistent.

Two-stage stochastic optimization problems with order constraint on the recourse

First Stage Problem:

$$\begin{aligned} \min_x \ & f(x) + \mathbb{E}\big[\mathcal{Q}(x,\xi)\big] \\ \text{s.t.} \ & \mathcal{Q}(x,\xi) \preceq_{\text{ic}} Z, \\ & x \in \mathcal{D}. \end{aligned}$$

where $\mathcal{Q}(x,\xi)$ is the optimal value of the second stage problem

Second Stage Problem:

$$\mathcal{Q}(x,\xi) = \min_y \{q^\top y : \ Tx + Wy = h, \ y \in \mathcal{Y}\}.$$

$\mathcal{D} \subset \mathbb{R}^n$ and $\mathcal{Y} \subseteq \mathbb{R}^m$ are closed convex sets,
$f : \mathbb{R}^n \to \mathbb{R}$ is a convex function,
$\xi = (q, W, T, h)$; q, T, h are random.

R. Schultz, R. Gollmer, and U. Gotzes (2008)

Numerical Methods: Finite localizations

If $z \in \mathbb{R}^n$, then at most $n+2$ target values η_k, scalarizations c^k and shortfall levels $\mathbb{E}\big[(\eta_k - Y^\top c^k)_+\big]$, $k = 1, \dots, n+2$, exists such that problem $(\mathscr{P})$ is equivalent to

$$
\begin{aligned}
&\min f(z) \\
&\text{ s.t. } \mathbb{E}\big[(\eta_k - \langle c^k, G(z)\rangle)_+\big] \le \mathbb{E}\big[(\eta_k - \langle c^k, Y\rangle)_+\big], \\
&\qquad\qquad k = 1, \dots, n+2, \\
&\qquad z \in Z.
\end{aligned}
$$

Corollary

If $\hat{x}$ is an optimal solution of problem $(\mathscr{P})$, then a piecewise linear function $\hat{\varphi} \in \mathcal{U}^m(\mathfrak{M})$ exists with no more than $n+2$ pieces such that conditions (5)-(6) are satisfied.

Similar result is established for $(\mathscr{Q})$.

Homem de Mello– Merothra (2009), Noyan–Rudolf (2013) have shown that for a finite probability space (c^j, η_j) are vertices of a particular polyhedron.

Numerical Approaches

Primal Dominance Constraints

$$
(\mathscr{P}) \qquad \begin{aligned} &\min f(z) \\ &\text{s.t. } \mathbb{E}([\eta - G(z)]_+) \le \mathbb{E}([\eta - Y]_+) \quad \forall\, \eta \in [a, b], \\ &\quad z \in Z \end{aligned}
$$

Nonsmooth problem even for linear f and G.
Denote $v(\eta) = \mathbb{E}([\eta - Y]_+)$.

Equivalent formulation

$$
\begin{aligned} &\min f(z) \\ &\text{s.t. } \int_A (\eta - G(z))\, dP \le v(\eta) \quad \forall \eta \in [a, b],\ \forall A \in \mathcal{F} \\ &\quad z \in Z \end{aligned}
$$

If Y has finitely many realizations, then the dominance constraint reduces to finitely many inequalities.

Primal Outer Approximation Method

$\Omega = \{\omega_1, \ldots, \omega_N\}$ with probabilities $p_1, \ldots, p_N$.
The realizations of Y are y_j with shortfalls $v_j = \mathbb{E}([y_j - Y]_+)$ and $G^j(z) = [G(z)](\omega_j)$.

Step 1: At iteration k, solve the relaxation to obtain z^k:

$$\begin{aligned} &\min f(z) \\ &\text{s.t.} \sum_{i \in A^j} p_i\big(y_j - G^i(z)\big) \le v_j \quad j = 1, \ldots, k-1 \\ &\qquad z \in Z. \end{aligned}$$

Step 2: Check the order: calculate

$$\delta_k = \max_{1 \le j \le N} \mathbb{E}\big[(y_j - G(z^k))_+ - (y_j - Y)_+\big].$$

If $\delta_k \le 0$, stop; otherwise, continue.

Step 3: Determine j^*, for which $\mathbb{E}([y_{j^*} - G(z^k)]_+) > v_{j^*}$ and define a new event $A^k = \{1 \le i \le N : y_{j^*} > G^i(z^k)\}$ Increase k by one, and go to Step 1.

Multivariate version

Step 1: At iteration k, solve the master problem:

$$\min f(z)$$
$$\text{s.t. } \sum_{i \in A^j} p_i \big(\eta_j - \langle c^j, G^i(z)\rangle\big) \leq \mathbb{E}\big[\big(\eta_j - \langle c^j, Y\rangle\big)_+\big], \quad j \in J_k.$$
$$z \in Z.$$

Let z^k denote its solution and let $X^k = G(z^k)$.

Step 2: Calculate the quantity

$$\delta_k = \sup_{\eta, c} \big\{\mathbb{E}\big[(\eta - c^\top G(z^k))_+ - (\eta - c^\top Y)_+\big] : (\eta, c) \in [a, b] \times S\big\}.$$

If $\delta_k \leq 0$, stop; otherwise, continue.

Step 3: Determine (η_k, c^k) such that

$$\mathbb{E}\big[(\eta - c^\top G(z^k))_+ - (\eta - c^\top Y)_+\big] \geq \frac{\delta_k}{2}.$$

Set $J_{k+1} = J_k \cup \{(\eta_k, c^k)\}$; $k \leftarrow k + 1$, and go to Step 1.

Order-Verification Problem

$$\min\left\{\mathbb{E}\big[\eta - c^\top Y\big]_+ - \mathbb{E}\big[\eta - c^\top X^k\big]_+ \;:\; \eta\in[a,b],\ c\in S\right\} \geq 0?$$

- DC-optimization Linearization of the function $\mathbb{E}\big[\max(0, \eta - c^\top X^k)\big]$, $X^k = G(z^k)$, at each point $(c^i, \eta_i) \in J_k$ by subdifferentiation. Event
$$\mathcal{A}^{ik} = \{\omega \in \Omega : \langle c^i, X^k(\omega)\rangle \leq \eta_i\},$$
Subgradient of h^k at (c^i, η_i) by Strassen's Theorem:
$$\begin{pmatrix} \mathbb{P}(\mathcal{A}^{ik}) \\ -\mathbb{E}[X^k \mathbf{1}_{\mathcal{A}^{ik}}] \end{pmatrix} \in \partial h^k(\eta_i, c^i).$$
- Combinatorial methods for finite probability space.
Every event $\mathcal{A}$ is represented by $\alpha_i = \begin{cases} 1, & \text{if } i \in \mathcal{A}; \\ 0, & \text{if } i \notin \mathcal{A} \end{cases}$ for $i = 1, \dots, N$.
$$\min\ \mathbb{E}\big[(\eta - c^\top Y)_+\big] - \sum_{i=1}^N p_i \alpha_i (\eta - c^\top X^i)$$
$$\text{s. t. } (c, \eta) \in S \times [a,b],\ \alpha \in \{0,1\}^N.$$
We alternate between minimizing in α for a fixed (c, η) and minimizing in (c, η) keeping α fixed.

Inverse Stochastic Dominance Constraints

$$
(\mathscr{Q}) \qquad \begin{aligned} & \max\ f(z) \\ & \text{s.t. } F^{(-2)}(G(z);p) \geq F^{(-2)}(Y;p), \quad \forall\, p \in [\alpha,\beta], \\ & \qquad z \in Z \end{aligned}
$$

Z is a closed subset of a vector space $\mathcal{Z}$, $[\alpha,\beta] \subset (0,1)$
$G : \mathcal{Z} \to \mathcal{L}_1(\Omega,\mathcal{F},P)$ and $f : \mathcal{Z} \to \mathbb{R}$ are continuous.

Orders and conditional expectation

- For $X, Y \in \mathcal{L}_1(\Omega,\mathcal{F},P)$, $X \succeq_{(2)} Y$ if and only if

$$
\mathbb{E}[X|A] \geq \frac{1}{P(A)} F^{(-2)}(Y; P(A)) \quad \forall\, A \in \mathcal{F}:\ P(A) > 0.
$$

- For $X, Y \in \mathcal{L}_1(\Omega,\mathcal{F},P)$, $X \preceq_{\text{ic}} Y$ if and only if

$$
\mathbb{E}[X|A] + \frac{1}{P(A)} L(Y; -P(A)) \leq 0 \quad \forall\, A \in \mathcal{F}:\ P(A) > 0.
$$

Quantile Cutting Plane Methods: The Scaled Method

Step 1: Solve the problem to obtain z^k:

$$\begin{aligned} &\max f(z) \\ &\text{s.t. } \mathbb{E}[G(z)|B^j] \geq \frac{1}{P(B^j)} F^{(-2)}(Y; P(B^j)) \quad j = 1, \dots, k \\ &\qquad z \in Z. \end{aligned}$$

Step 2: Consider the sets $A_t^k = \{G(z^k) \leq t\}$ and let

$$\delta_k = \sup_t \left\{ \frac{1}{P(A_t^k)} F^{(-2)}(Y; P(A_t^k)) - \mathbb{E}[G(z^k)|A_t^k] : P(A_t^k) > 0 \right\}.$$

If $\delta_k \leq 0$, stop; otherwise, continue.

Step 3: Find t_k such that $P(G(z^k) \leq t_k) > 0$ and

$$\mathbb{E}[G(z^k)|A_{t_k}^k] - \frac{1}{P(A_{t_k}^k)} F^{(-2)}(Y; P(A_{t_k}^k)) \leq -\frac{\delta_k}{2}.$$

Set $B^{k+1} = A_{t_k}^k$, increase k by one, and go to Step 1.

Two-stage problems: Multicut decomposition method

For the finite probability space $\Omega = \{\omega_1, \dots, \omega_N\}$ with probabilities $P(\omega_i) = p_i$, $i = 1, \dots, N$, denote $T^i = T(\omega_i)$, $h^i = h(\omega_i)$, $y^i = y(\omega_i)$, and $q^i = q(\omega_i)$.

$$\begin{aligned} \min\ & f(x) + \sum_{i=1}^{N} p_i Q^i(x) \\ \text{s.t. } & \mathcal{Q}(x,\xi) \preceq_{\text{ic}} Z, \\ & x \in \mathcal{D}. \end{aligned}$$

$$Q^i(x) := \inf_y \{(q^i)^\top y : \ W^i y = h^i - T^i x,\ y \geq 0\}, \quad i = 1, \dots, N.$$

Idea: Approximate the functions $Q^i(\cdot)$ and their domains by cutting planes and use the approximate models to construct an approximation of the distribution of the recourse function $\mathcal{Q}(\cdot)$.

The functions $Q^i(\cdot)$ are convex.
The subdifferential of $Q^i(\cdot)$ is $\partial Q^i(x^k) = -(T^i)^\top \psi^i(x^k)$, where

$$\psi^i(x^k) = \operatorname{argmax}\{\pi^\top(h^i - T^i x^k) | (W^i)^T \pi \le q^i\}$$

is the set of optimal solutions of the dual to the second stage problem at $x = x^k$. For $\pi^{k,i} \in \psi^i(x^k)$, it holds

$$Q^i(x) \ge Q^i(x^k) - \langle (T^i)^T \pi^{k,i}, x - x^k \rangle.$$

Objective cuts

$$\begin{aligned} Q^s(x) &\ge \alpha^{k,i} + \langle g^{k,i}, x \rangle, \\ \text{where} \quad g^{k,i} &= -(T^i)^T \pi^{k,i}, \\ \alpha^{k,i} &= Q^i(x^k) + \langle \pi^{k,i}, T^i x^k \rangle. \end{aligned}$$

Two-stage problems: Feasibility Cuts

We derive an inequality that must be satisfied $x \in \text{dom } Q^i$. Denote by $U^i(x)$ the optimal value of

$$\begin{aligned} \min_{x,y}\ & \|u\| \\ \text{s.t.}\ & Wy + u = h^i - T^i x, \\ & y \geq 0. \end{aligned}$$

$U(\cdot)$ is convex and $U^i(x) = 0$ if and only if $Q^i(x) < \infty$. Let $r^{k,i} \in \partial U^i(x^k)$. Then,

$$0 \geq U^i(x^k) + \langle r^{k,i}, x - x^k \rangle = U^i(x^k) - \langle r^{k,i}, x^k \rangle + \langle r^{k,i}, x \rangle.$$

Feasibility cuts

$$\beta^{k,i} + (r^{k,i})^\top x \leq 0,$$

$$\text{where} \quad \begin{aligned} r^{k,i} &= (-T^i)^\top \Delta^i(x^k) \\ \beta^{k,i} &= U^i(x^k) - \langle r^{k,i}, x^k \rangle. \end{aligned}$$

Quantile function-based decomposition method for two-stage problems

The cuts constructed until iteration k are used in the following

Approximate problem

$$
\begin{aligned}
\min_{x,\upsilon}\ & f(x) + \sum_{i=1}^{N} p_i v^i \\
\text{s.t.}\ & \alpha^{j,i} + (g^{j,i})^\top x \le v^i \quad j \in J_{\mathrm{o}}^k(i),\ i = 1,\dots,N, \\
& \beta^{j,i} + (r^{j,i})^\top x \le 0 \quad j \in J_{\mathrm{f}}^k(i),\ i = 1,\dots,N, \\
& \frac{1}{P(A^j)}\Big[\sum_{i\in A^j} p_i v^i\Big] + \frac{1}{P(A^j)} L(Z; P(A^j) - 1) \le 0 \quad A^j \in J_{\mathrm{e}}^k, \qquad (7) \\
& x \in \mathcal{D}.
\end{aligned}
$$

Here $\upsilon = (v^1,\dots,v^N)^\top$ and V is a random variable with realizations $v^1,\dots,v^N$. Inequalities (7) define the event cuts which approximate the ordering constraint.

Quantile function-based decomposition method: Algorithm

Step 1. Solve the second stage problems with $x = x^k$.

(a) If $Q^i(x^k) < \infty$, then set $J_{\mathrm{f}}^k(i) = J_{\mathrm{f}}^{k-1}(i)$. If $Q^i(x^k) > v^{k,i}$, then construct an objective cut and set $J_{\mathrm{o}}^k(i) = J_{\mathrm{o}}^{k-1}(i) \cup \{k\}$; otherwise set $J_{\mathrm{o}}^k(i) = J_{\mathrm{o}}^{k-1}(i)$.

(b) If $Q^i(x^k) = \infty$, construct a feasibility cut and update $J_{\mathrm{f}}^k(i) = J_{\mathrm{f}}^{k-1}(i) \cup \{k\}$.

Step 2. If $Q^s(x^k) < \infty$ for all $i = 1, \ldots, N$, then construct an event cut. Set A^k to be the new event and $J_{\mathrm{e}}^k = J_{\mathrm{e}}^{k-1} \cup \{A^k\}$.

Step 3. If $Q(x^k) = \sum_{i=1}^N p_i v^{k,i}$ and $\delta \le 0$, then stop; otherwise continue.

Step 4. Solve the approximate problem. If it is infeasible, then stop. Otherwise, denote by (x^{k+1}, V^{k+1}) its solution; increase k by one, and go to Step 1.

Dual approach: Reduced Problems and Their Duals

Given a collection of point (η_j, c^j), $j \in J_k$ for some index set J_k, $k \in \mathbb{N}$, the reduced problem is

$$\begin{aligned} &\min f(z) \\ &\text{s.t. } \mathbb{E}\big[\big(\eta_j - (c^j)^\top G(z)\big)_+\big] \le \mathbb{E}\big[\big(\eta_j - (c^j)^\top Y\big)_+\big], \quad j \in J_k. \\ &\quad z \in Z \end{aligned}$$

The Lagrangian of the reduced problem is defined on $Z \times \mathbb{R}_+^{|J_k|}$ is

$$L_k(z, \mu^k) = f(z) + \mathbb{E}\Big[\sum_{j \in J_k} \mu_j^k \big(\eta_j - (c^j)^\top G(z)\big)_+ - \sum_{j \in J_k} \mu_j^k \big(\eta_j - (c^j)^\top Y\big)_+\Big].$$

Extension $\tilde{\mu}^k$ of μ^k to the entire set $[a, b] \times S$ by setting

$$\tilde{\mu}^k(\mathcal{A}) = \sum_{j \in J_k} \mu_j^k \mathbf{1}_{\mathcal{A} \cap \{(\eta_j, c^j)\}}.$$

The reduced dual function $D_k\colon \mathbb{R}^{|J_k|} \to \mathbb{R}$ at the point μ^k

$$D_k(\mu^k) = \min_{z \in Z} L_k(z, \mu^k) = D(\tilde{\mu}^k).$$

Subgradients of the Reduced Dual Function

The subgradients of $D_k(\mu)$

$$\Gamma_j^k(\mu) = \mathbb{E}\big[\big(\eta_j - (c^j)^\top G(z_\mu^k)\big)_+ - \big(\eta_j - (c^j)^\top Y\big)_+\big], \quad j = 1, \dots, |J_k|,$$

where z_μ^k is such that $D_k(\mu) = L_k(z_\mu^k, \mu)$.

For atomic measure μ^ℓ with atoms on the set $\{(\eta_j, c^j),\ j \in J_\ell\}$, we define an extension $\mu^{\ell,k} \in \mathbb{R}_+^{|J_k|}$ to the set $\{(\eta_j, c^j),\ j \in J_k \supset J_\ell\}$ by setting

$$\mu_j^{\ell,k} = \begin{cases} \mu_j^\ell & \text{if } j \in J_\ell \\ 0 & \text{if } j \notin J_\ell. \end{cases}$$

Note: the subgradients $\Gamma^{\ell,k}(\mu^{\ell,k})$ are <u>not</u> subgradients of the dual function.

Piecewise linear model $\mathcal{D}^k : \mathbb{R}^{|J_k|} \to \mathbb{R}$ of the reduced dual function

$$\mathcal{D}^k(\mu) = \min_{1 \le \ell \le k} \{D(\tilde{\mu}^\ell) + (\Gamma^{\ell,k})^\top (\mu - \mu^{\ell,k})\}.$$

Dual Approximative Bundle Method

Step 1: Solve $\min_{x\in\mathcal{X}} L_k(x,\mu^k)$ and calculate new subgradients $\Gamma^{\ell,k}$ at μ^k.

Step 2: If $k=1$ or if $\mathcal{D}(\mu^k) \geq (1-\gamma)\mathcal{D}(w^{k-1}) + \gamma\vartheta^{k-1}(\mu^k)$,
then set $w^k := \mu^k$; else set $w^k := w^{k-1,k}$.

Step 3: Calculate a solution (θ_{k+1},μ^{k+1}) of the master problem:

$$\max_{\mu\geq 0}\{\theta - \frac{\varrho}{2}\|\mu - w^k\|_2^2 \ : \ \mathcal{D}(\mu^\ell) + (\Gamma^{\ell,k})^\top(\mu-\mu^\ell) \geq \theta \ \ \ell = 1\dots k\}.$$

Set $\vartheta^k(\mu^{k+1}) = \theta_{k+1}$. Let π_ℓ be the optimal Lagrange multiplier of the ℓ-th constraint. Define $\tilde{x}^k = \sum_{\ell=1}^k \pi_\ell x^\ell$.

Step 4: Calculate the quantities

$$\delta_k = \sup_{(\eta,c)\in[a,b]\times S} \mathbb{E}\Big[\big(\eta - \langle c, G(\tilde{x}^k)\rangle\big)_+ - \big(\eta - \langle c, Y\rangle\big)_+\Big],$$

$$\delta_k' = \max_{j\in J_k} \mathbb{E}\Big[\big(\eta_j - \langle c^j, G(\tilde{x}^k)\rangle\big)_+ - \big(\eta_j - \langle c^j, Y\rangle\big)_+\Big].$$

Step 5: If $\mathcal{D}(w^k) \geq \theta^{k+1} - \varepsilon$ and $\delta_k \leq \varepsilon$, then stop; otherwise continue.

Step 6: If $\delta_k > \varepsilon$ and $\delta_k' \leq \frac{\delta_k}{4}$, then determine (η_*, c^*) such that

$$\mathbb{E}\big[(\eta_* - \langle c^*, G(\tilde{x}^k)\rangle)_+ - (\eta_* - \langle c^*, Y\rangle)_+\big] \geq \tfrac{1}{2}\delta_k$$

and set $J_{k+1} = J_k \cup \{(\eta_*, c^*)\}$; else set $J_{k+1} = J_k$.
Set $k \leftarrow k+1$ and go to Step 1.

Further Numerical Methods

- Subgradient-Based Approximation Methods for second-order dominance constraints with linear $G(\cdot)$ in primal form
  C. Fabian, G. Mitra, and D. Roman, 2008;
- Combinatorial methods for first-order dominance constraints
  G. Rudolf, N. Noyan, A. Ruszczyński 2006 based on second-order stochastic dominance relaxation
  $(\{X : X \succeq_{(2)} Y\} = \overline{\text{conv}}\{X : X \succeq_{(1)} Y\})$;
- Cutting Surface Method and Sample Average Approximation
  Homem de Mello-Mehrotra 2009, 2012;
- Exact Penalty Method and Sample Average Approximation
  Meskarian-Fliege-Xu 2014;
- Strassen theorem representation
  Luedtke 2012, Armbruster-Luedtke 2014;
- Augmented Lagrangian method
  Dentcheva-Martinez-Wolfhagen 2016.

Further research

- Non-convex problems Optimality conditions for problems with first- and higher-order dominance constraints with non-convex functions (DD–A.Ruszczyński 2004.)
- Stochastic dominance efficiency in multi-objective optimization and its relations to dominance constraints (G. Mitra, C. Fabian, K. Darby-Dowman, D. Roman, 2006, 2009)
- Two-stage problems with order constraints on the recourse function (R. Schultz, F. Neise, R. Gollmer, U. Gotzes, D. Drapkin, 2007, 2009, 2010, DD–E.Wolfhagen 2016)
- Stability and sensitivity analysis (DD-R. Henrion-A Ruszczyński, 2007; Y. Liu-H.Xu, 2010; DD-Römisch 2013; Klaus-Schultz 2015)
- Robust Dominance Relation and relations to robust preferences (DD-A. Ruszczyński, 2010)
- Shape optimization with stochastic-order constraints Gotzes-Schultz 2014
- Asymptotic distributions in Markov decision processes Haskell-Jane 2012, Haskell-Shanthikumar-Shen 2015

Portfolio Optimization

Assets $j = 1, \ldots, n$ with random return rates R_j
Reference return rate Y (e.g. index, existing portfolio, etc.)
Decision variables $z_j,\ j = 1, \ldots, n$, Z -polyhedral set
Portfolio return rate $R(z) = \sum_{j=1}^{n} z_j R_j$

$$\begin{aligned} \max\ & f(z) - \lambda\sigma \\ \text{s.t.}\ & \sum_{j=1}^{n} z_j R_j \succeq_{(2)} \langle c, Y \rangle - \sigma \quad \text{for all } c \in S \\ & z \in Z \end{aligned}$$

$f(x) = \mathbb{E}\big[R(x)\big]$ or $f(x) = -\varrho\big[R(x)\big]$: measure of risk.
S is the M-dimensional simplex, $\lambda > 0$.

All Statements are Equivalent

$$\sum_{j=1}^{n} z_j R_j \succeq_{(2)} c^\top Y$$

$$F^{(-2)}\Big(\sum_{j=1}^{n} z_j R_j; p\Big) \geq F^{(-2)}(c^\top; p) \text{ for all } p \in [0,1]$$

continuum of CVaR constraints for every risk level $p \in [0, 1]$

$$\mathbb{E}\Big[u\Big(\sum_{j=1}^{n} z_j R_j\Big)\Big] \geq \mathbb{E}\big[u(c^\top Y)\big]$$

for <u>all</u> concave nondecreasing functions u (von Neuman-Morgenstern utility)

$$\int_0^1 F^{(-1)}\Big(\sum_{j=1}^{n} z_j R_j; p\Big)\, dw(p) \geq \int_0^1 F^{(-1)}(c^\top Y; p)\, dw(p)$$

for <u>all</u> concave nondecreasing functions w (rank dependent utility)

Other Applications

- Electricity markets: portfolio of contracts and/or acceptability of contracts; risk-averse power generation planning;
- Inverse models, forecasting, learning: Compare the forecast errors via stochastic dominance and design data collection for model calibration;
- Network design: assign capacity to optimize network throughput;
- Robotics: control of position and communication of robots;
- Budget allocation in disaster prevention and relief planning;
- Supply-chain models: cost vs. market share;
- Medicine: radiation therapy design.