# Risk-Averse Optimization

Andrzej Ruszczyński

*Rutgers University*

# Optimization models with random data

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

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

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

$Y$ - a random vector in $\mathbb{R}^s$

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

## Questions

- What is a “better” value of the objective?
- What is a “feasible” solution?

## Expected value model

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

## Probabilistic / chance constraints

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

## Penalty for violating the constraints

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

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

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

# Risk Modeling

Random outcome (e.g., cost):

$$X_z(\omega) = f(z,\omega), \qquad f : \mathcal{Z} \times \Omega \to \mathbb{R}$$

$(\Omega, \mathcal{F}, P)$ - probability space, $\mathcal{Z}$ - some vector space

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

**Expected Utility Models** (von Neumann and Morgenstern, 1944)

$$\min_{z \in Z} \ \mathbb{E}\Big[u(X_z)\Big] \qquad \Big( = \int_{\Omega} u\big(f(z,\omega)\big)\, dP(\omega) \Big)$$

$u : \mathbb{R} \to \mathbb{R}$ is a (dis)utility function:

- non-decreasing for all rational decision makers
- convex for all risk-averse decision makers

Existence of utility functions derived from systems of axioms,
but they are very difficult to elicit in practice

# Elementary Mean–Risk Models

**Objective:** Choose decision $z$ to

- minimize the expected outcome, the mean $\mathbb{E}[X_z]$,
- minimize a scalar measure of uncertainty of $X_z$, the risk $r[X_z]$

**Examples** of risk measures:

$$r_1[X] = \mathbb{V}\mathrm{ar}[X] \qquad \text{(Markowitz' model)}$$

$$r_2[X] = \Big(\mathbb{E}[(X - \mathbb{E}X)_+^q]\Big)^{1/q} \qquad \text{(semideviation)}$$

$$r_3[X] = \min_{\xi} \mathbb{E}\Big[\max\Big(\frac{1-p}{p}(\xi - X), X - \xi\Big)\Big] \qquad \text{(deviation from quantile)}$$

**Optimization models**

$$\min_{z \in Z} \; \mathbb{E}[X_z] + \kappa\, r[X_z], \qquad 0 \le \kappa \le \kappa_{\text{max}}$$

Interesting application of parametric optimization

**Note**: Risk measures depend on the entire distribution,
$r[X_z]$ is nonlinear w.r.t. probability and possibly nonconvex in $z$.

## Application to Portfolio Problems

Assets $j = 1, \ldots, n$ with random returns $R_j$.
Decision variables $z_j,\ j = 1, \ldots, n$.
Portfolio return $R(z) = \sum_{j=1}^{n} z_j R_j$.

Discrete distribution: returns $r_{jt}$ with probabilities $p_t$.

Mean–Semideviation Model

$$\max_{z} \mathbb{E}R(z) - \kappa \mathbb{E}\Big[\mathbb{E}R(z) - R(z)\Big]_+$$

$$\text{subject to } \sum_{j=1}^{n} z_j = 1,\ z \geq 0.$$

## Linear programming formulation

$$\max_{z} \; \mu - \kappa \sum_{t=1}^{T} p_t v_t$$

subject to

$$\sum_{t=1}^{T} \sum_{j=1}^{n} p_t r_{jt} z_j = \mu$$

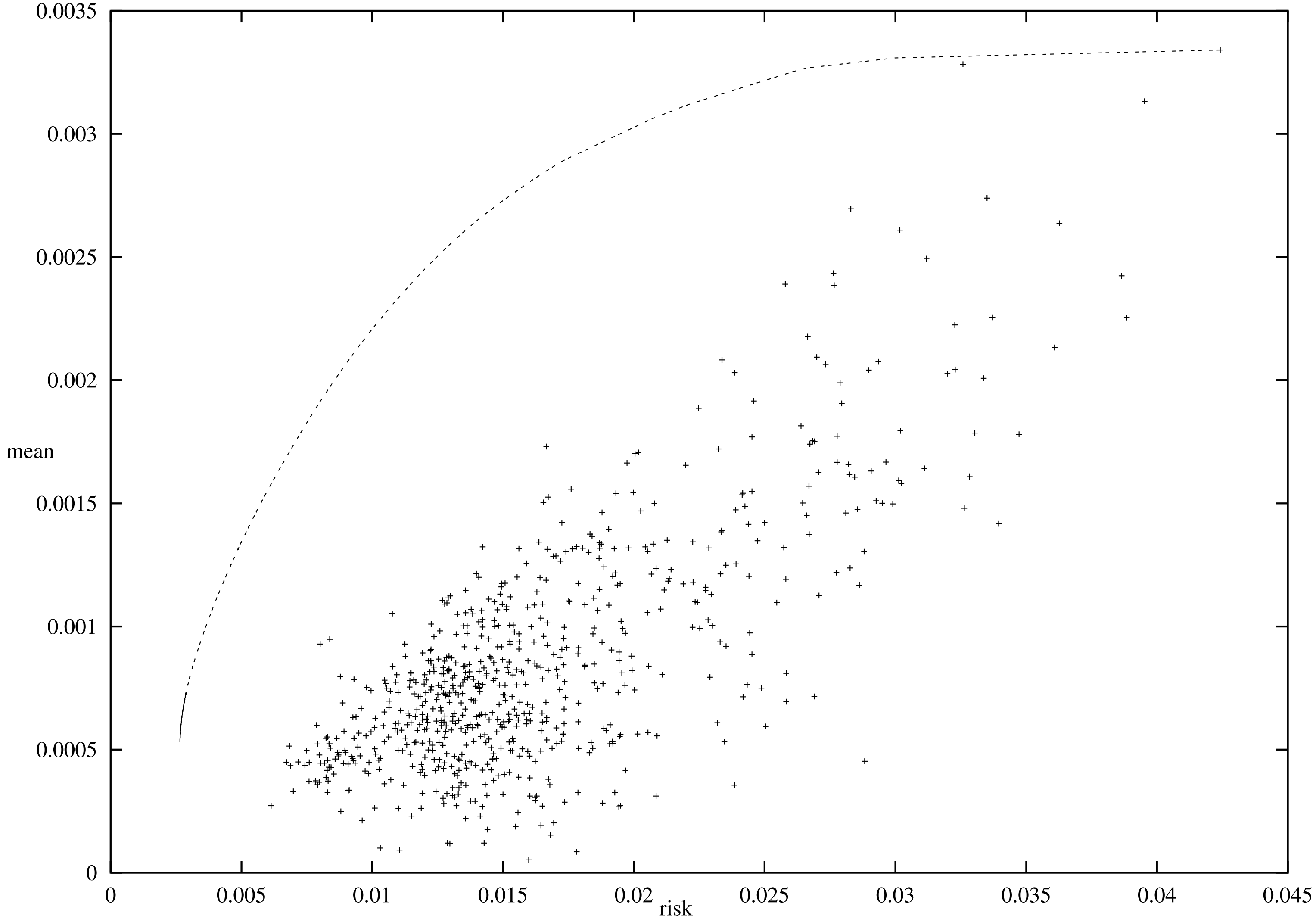
$$\sum_{j=1}^{n} r_{jt} z_j - \mu = u_t - v_t$$

$$\sum_{j=1}^{n} z_j = 1, \;\; z \geq 0.$$

At the optimal solution $u_t v_t = 0$, $t = 1, \ldots, T$.

For $\kappa = 0$ the solution is trivial:
*invest all in the security with the highest expected return*

Idea: Use a parametric simplex method to move from $\kappa = 0$ to $\kappa = 1$ and recover the entire efficient frontier.



The efficient frontier for the mean-semideviation model ($0 \leq \kappa \leq 1$) in a portfolio problem with 719 securities and 3080 realizations. Parametric optimization by Ruszczyński–Vanderbei (2003)

# Consistency with Stochastic Ordering

(Ogryczak–Ruszczyński, 1997–)

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

$$X_1 \preceq_s X_2 \quad \text{iff} \quad \mathbb{E}[u(X_1)] \le \mathbb{E}[u(X_2)] \quad \forall\, u \in \mathscr{U}$$

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

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

$$X_1 \preceq_s X_2 \;\Rightarrow\; \mathbb{E}[X_1] + \kappa\, r[X_1] \le \mathbb{E}[X_2] + \kappa\, r[X_2]$$

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

Popular risk measures consistent with the increasing convex order (second order stochastic dominance):

- semideviation ($\kappa_{\max} = 1$)
- deviation from quantile ($\kappa_{\max} = 1$)

but not variance.

Optimal solutions of mean–risk optimization models cannot be dominated

# General Theory of Risk Functionals

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

(A1) **Convexity**:

$\rho(\alpha X_1 + (1-\alpha) X_2) \leq \alpha \rho(X_1) + (1-\alpha)\rho(X_2)$ for all $X_1, X_2 \in \mathcal{X}$ and $\alpha \in [0, 1]$

(A2) **Monotonicity**: If $X_1, X_2 \in \mathcal{X}$ and $X_1(\cdot) \leq X_2(\cdot)$, then $\rho(X_1) \leq \rho(X_2)$

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

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

Kijima-Ohnishi (1993), Artzner-Delbaen-Eber-Heath (1999),
Föllmer-Schied (2002), Rockafellar-Uryasev-Zabarankin (2005)

# Conjugate Duality of Risk Functionals

Pairing of a linear topological space $\mathcal{X}$ with a linear topological space $\mathcal{Y}$ of regular (signed) measures on $\Omega$ :

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

**Theorem:** If $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ is proper, lower semicontinuous* and convex, then

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

with $\mathcal{A} := \mathrm{dom}(\rho^*) \subset \mathcal{Y}$. Moreover:
(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 $\mu \in \mathcal{A}$

Delbaen, Föllmer–Schied (2002), Rockafellar–Uryasev–Zabarankin (2005), Cheridito–Delbaen–Kupper (2005), Ruszczyński–Shapiro (2003)

**Lower semicontinuity follows from monotonicity if $\rho$ is finite and $\mathcal{X}$ is a Banach lattice*

# Law-Invariant Risk Functionals

**Definition:** $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ is law invariant with respect to $P$, if

$$P[X_1 \le t] = P[X_2 \le t] \text{ for all } t \in \mathbb{R} \quad \Rightarrow \quad \rho(X_1) = \rho(X_2)$$

$$(X_1 \overset{\mathscr{D}}{\sim} X_2)$$

**Definition:** A law invariant risk functional $\rho$ is consistent with stochastic order "$\preceq_s$", if $X_1 \preceq_s X_2 \quad \Rightarrow \quad \rho(X_1) \le \rho(X_2) \quad \forall\, X_1, X_2 \in \mathcal{X}$.

**Theorem:** Suppose that the space $\Omega$ allows a continuous random variable. If $\rho$ is law invariant and convex, then it is consistent with the increasing convex order iff it is monotone (in the sense of (A2)).

Idea of the proof: If $X_1 \preceq_{\text{icx}} X_2$ then by Dentcheva–Ruszczyński (2004)

$$X_1 \le Y \in \overline{\text{conv}}\left\{\hat{X}_2 : \hat{X}_2 \overset{\mathscr{D}}{\sim} X_2\right\}$$

As $\rho$ is monotone, convex and law invariant, $\rho(X_1) \le \rho(Y) \le \rho(X_2)$.

# Optimization of Risk Functionals

Consider $X_z(\omega) = f(z, \omega)$, $\omega \in \Omega$.

The composite optimization problem:

$$\min_{z \in Z} \rho(X_z) \qquad \text{(P)}$$

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

$$\min_{z \in Z} \left\{ \mathbb{E}_{\widehat{\mu}}[X_z] - \rho^*(\widehat{\mu}) \right\} = \min_{z \in Z} \max_{\mu \in \mathscr{P}} \left\{ \mathbb{E}_{\mu}[X_z] - \rho^*(\mu) \right\}$$

If $\mathcal{X}$ Banach, $\mathcal{Y} = \mathcal{X}^*$, $X_z$ l.s.c, then duality:

$$\min_{z \in Z} \rho(X_z) = \max_{\mu \in \mathscr{P}} \inf_{z \in Z} \left\{ \mathbb{E}_{\mu}[X_z] - \rho^*(\mu) \right\}$$

# Risk Value of Perfect Information

Define the operation "inf" on the family $\{X_z : z \in Z\}$ of elements of $\mathcal{X}$ as the pointwise infimum:

$$\Big[\inf_{z\in Z} X_z\Big](\omega) := \inf\Big\{f(z,\omega) : z \in Z\Big\}, \quad \omega \in \Omega.$$

Suppose that $\inf_{z\in Z} X_z \in \mathcal{X}$.

We define the risk value of perfect information as

$$\mathrm{RVPI}_\rho := \inf_{z\in Z} \rho(X_z) - \rho\Big(\inf_{z\in Z} X_z\Big)$$

The *expected value of perfect information* with respect to measure $\mu$:

$$\mathrm{EVPI}_\mu := \inf_{z\in Z} \mathbb{E}_\mu[X_z] - \mathbb{E}_\mu\Big[\inf_{z\in Z} X_z\Big].$$

**Theorem**: Suppose that $\rho(\cdot)$ is lower semicontinuous and assumptions (A1)–(A4) are satisfied. Let $\mathcal{A} \subset \mathcal{P}$ be the convex set for which the dual representation of $\rho$ holds true. Then

$$\inf_{\mu\in\mathcal{A}} \mathrm{EVPI}_\mu \le \mathrm{RVPI}_\rho \le \sup_{\mu\in\mathcal{A}} \mathrm{EVPI}_\mu$$

# Conditional Risk Mappings

Let $\mathcal{F}_1 \subset \mathcal{F}_2$ be sigma algebras of subsets of a set $\Omega$, and $\mathcal{X}_1 \subset \mathcal{X}_2$ be linear spaces of real valued functions on $\Omega$ measurable with respect to $\mathcal{F}_1$ and $\mathcal{F}_2$.

**Definition**: A mapping $\rho : \mathcal{X}_2 \to \mathcal{X}_1$ is a conditional risk mapping if the following properties hold:

(A1) **Convexity**: If $\alpha \in [0,1]$ and $X, Y \in \mathcal{X}_2$, then

$$\rho\Big(\alpha X + (1-\alpha)Y\Big) \le \alpha\rho(X) + (1-\alpha)\rho(Y)$$

(A2) **Monotonicity**: If $X \le Y$, then $\rho(X) \le \rho(Y)$

(A3) **Predictable Translation Equivariance**: If $X_1 \in \mathcal{X}_1$ and $X_2 \in \mathcal{X}_2$, then

$$\rho(X_1 + X_2) = X_1 + \rho(X_2)$$

Riedel (2003), Ruszczyński–Shapiro (2004)

# Conjugate Duality of Conditional Risk Mappings

For each $\omega \in \Omega$ we define a set of probability measures $\mathscr{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega) \subset \mathscr{P}_{\mathcal{Y}_2}$ defined as the set of all $\nu \in \mathscr{P}_{\mathcal{Y}_2}$ such that for every $B \in \mathcal{F}_1$

$$\nu(B) = \begin{cases} 1, & \text{if } \omega \in B, \\ 0, & \text{if } \omega \notin B. \end{cases}$$

Note that $\omega$ is fixed here and $B$ varies in $\mathcal{F}_1$.

**Theorem**: Let $\rho = \rho_{\mathcal{X}_2|\mathcal{X}_1}$ be a lower semicontinuous conditional risk mapping satisfying assumptions (A1)–(A3). Then

$$\rho_\omega(X) = \sup_{\mu \in \mathscr{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)} \Big\{ \langle \mu, X \rangle - \rho^*_\omega(\mu) \Big\}, \quad \omega \in \Omega,\ X \in \mathcal{X}_2,$$

and $\rho^*_\omega(\mu)$ is the convex conjugate of $\rho_\omega(X)$. Conversely, suppose that a mapping $\rho : \mathcal{X}_2 \to \mathcal{X}_1$ can be represented in the above form for some (proper) function $\rho^* : \mathcal{Y}_2 \times \Omega \to \overline{\mathbb{R}}$. Then $\rho$ is lower semicontinuous and satisfies conditions (A1)–(A3).

For separable $\mathcal{X}_2$, all weakly* measurable selections of $\mathscr{P}_{\mathcal{Y}_2|\mathcal{F}_1} : \Omega \rightrightarrows \mathscr{P}_{\mathcal{Y}_2}$ are conditional probabilities

# Recursive Risk Models

Consider a sequence of sigma algebras $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_T$, with $\mathcal{F}_1 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$, and let $\mathcal{X}_1 \subset \cdots \subset \mathcal{X}_T$ be a corresponding sequence of linear spaces of $\mathcal{F}_t$-measurable functions, $t = 1, ..., T$.
Let $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}} : \mathcal{X}_t \to \mathcal{X}_{t-1}$ be conditional risk mappings. Denote $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_T$ and $X := (X_1, X_2, \ldots, X_T)$, where $X_t \in \mathcal{X}_t$, $t = 1, \ldots, T$.
Define a risk function $\tilde{\rho} : \mathcal{X} \to \mathbb{R}$:

$$\tilde{\rho}(X) := X_1 + \rho_{\mathcal{X}_2|\mathcal{X}_1}\Big[X_2 + \rho_{\mathcal{X}_3|\mathcal{X}_2}\Big(X_3 + \ldots \\ \cdots + \rho_{\mathcal{X}_{T-1}|\mathcal{X}_{T-2}}\big[X_{T-1} + \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}\big(X_T\big)\big]\Big)\Big]$$

Since $\mathcal{F}_1 = \{\emptyset, \Omega\}$, the space $\mathcal{X}_1$ can be identified with $\mathbb{R}$, and hence $\tilde{\rho}(X)$ is real valued. By assumption (A3) we have

$$X_{T-1} + \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}(X_T) = \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}(X_{T-1} + X_T).$$

Applying this formula for $t = T, T-1, \ldots, 2$ we obtain the equation:

$$\tilde{\rho}(X) = \rho_T(X_1 + \ldots + X_T),$$

where

$$\rho_t := \rho_{\mathcal{X}_2|\mathcal{X}_1} \circ \cdots \circ \rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}, \quad t = 2, ..., T.$$

Since each conditional risk mapping $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$ satisfies (A1)–(A3), it follows that the function $\rho_T$ satisfies (A1)–(A3) as well. Moreover, if conditional risk mappings $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$ are positively homogeneous, then $\rho$ is positively homogeneous. Assuming further that the spaces $\mathcal{X}_t$ are separable and $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$ are lower semicontinuous, we obtain

$$\tilde{\rho}(X) = \sup_{\mu \in \mathcal{A}} \mathbb{E}_\mu[X_1 + ... + X_T],$$

where the set $\mathcal{A} := \mathcal{A}_1 \circ ... \circ \mathcal{A}_{T-1}$ is given by the composition of the multifunctions $\mathcal{A}_\tau : \Omega \rightrightarrows \mathscr{P}_{\mathcal{Y}_{\tau+1}}$, $\tau = 1, \dots, T-1$.

The composition $\mathcal{A}_1 \circ \mathcal{A}_2$ is the set $\mathcal{A}(\omega)$ formed by *all* measures $\mu \in \mathcal{Y}_3$ representable in the form

$$\mu(S) = \int_\Omega \Big[\mu_2(\tilde{\omega})\Big](S)\, d\mu_1(\tilde{\omega}), \;\; S \in \mathcal{F}_3,$$

where $\mu_2(\cdot) \in \mathcal{A}_2(\cdot)$ is a weakly* $\mathcal{F}_2$-measurable selection and $\mu_1 \in \mathcal{A}_1(\omega)$.

# Dynamic Risk Optimization

Suppose random outcomes $X_t \in \mathcal{X}_t$ result from decisions $z_t$ in some stochastic system. We introduce linear spaces $\mathcal{Z}_t$ of $\mathcal{F}_t$-measurable functions $Z_t : \Omega \to \mathbb{R}^{n_t}$ and consider functions $f_t : \mathbb{R}^{n_t} \times \Omega \to \mathbb{R}$, $t = 1, \ldots, T$.

With functions $f_t$ we associate mappings $F_t : \mathcal{Z}_t \to \mathcal{X}_t$ defined as follows

$$\Big[F_t(Z_t)\Big](\omega) := f_t(Z_t(\omega), \omega), \quad Z_t \in \mathcal{Z}_t,\ \omega \in \Omega.$$

We assume that the functions $f_t(z_t, \omega)$ are *random lower semicontinuous* and that the mappings $F_t$ are well defined: for every $Z_t \in \mathcal{Z}_t$ the function $f_t(Z_t(\cdot), \cdot)$ belongs to the space $\mathcal{X}_t$, $t = 1, ..., T$.

We say that the mapping $F_t$ is convex if $\Big[F_t(\cdot)\Big](\omega)$ is convex for all $\omega \in \Omega$. Then for every conditional risk mapping $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$, satisfying (A1)–(A3), the function $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}(F_t(\cdot))$ is convex.

Let $\mathcal{Z} = \mathcal{Z}_1 \times \cdots \times \mathcal{Z}_T$, and let $F : \mathcal{Z} \to \mathcal{X}$ be defined as

$$F(Z) := (F_1(Z_1), \dots, F_T(Z_T)).$$

With the risk function $\tilde{\rho}$ and the mapping $F$ we can associate the (convex) function

$$\vartheta(Z) := \tilde{\rho}(F(Z)) = F_1(Z_1) + \rho_{\mathcal{X}_2|\mathcal{X}_1}\Big[F_2(Z_2) + \rho_{\mathcal{X}_3|\mathcal{X}_2}\Big(F_3(Z_3) + \cdots$$
$$\cdots + \rho_{\mathcal{X}_{T-1}|\mathcal{X}_{T-2}}\big[F_{T-1}(Z_{T-1}) + \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}\big(F_T(Z_T)\big)\big]\Big)\Big].$$

Suppose that we are given $\mathcal{F}_t$-measurable, closed-valued multifunctions

$$\mathcal{G}_t : \mathbb{R}^{n_{t-1}} \times \Omega \rightrightarrows \mathbb{R}^{n_t}, \quad t = 2, \dots, T,$$

with $\mathcal{G}_1 \subset \mathbb{R}^{n_1}$ being a fixed (deterministic) set. We define the set

$$\mathfrak{G} := \Big\{Z \in \mathcal{Z} : Z_t(\omega) \in \mathcal{G}_t(Z_{t-1}(\omega), \omega),\ \omega \in \Omega,\ t = 1, \dots, T\Big\},$$

and consider the problem

$$\min_{Z \in \mathfrak{G}} \vartheta(Z).$$

We refer to it as the nested formulation
of a multistage risk optimization problem.

## Interchangeability

We assume that the spaces $\mathcal{Z}_t$ are solid in the sense that for every two elements $\underline{X}, \overline{X} \in \mathcal{X}_t$ and every $\mathcal{F}_t$-measurable function $X_t$

$$\underline{X}(\cdot) \le X_t(\cdot) \le \overline{X}(\cdot) \;\; \Rightarrow \;\; X_t \in \mathcal{X}_t$$

Furthermore, we assume that there exist elements $\underline{X}_t \in \mathcal{X}_t$ such that for all $Z \in \mathfrak{S}$ we have $F_t(Z_t) \ge \underline{X}_t$, $t = 1, \dots, T$.

The spaces $\mathcal{Z}_t$ are assumed to be decomposable: for every $Z_t \in \mathcal{Z}_t$ and $B \in \mathcal{F}_t$, and every bounded and $\mathcal{F}_t$-measurable function $W : \Omega \to \mathbb{R}^n$, the space $\mathcal{Z}_t$ also contains the function $Y_t(\cdot) := \mathbb{1}_{\Omega \setminus B}(\cdot) Z_t(\cdot) + \mathbb{1}_B(\cdot) W(\cdot)$.

Now let $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ be a risk function on a decomposable $\mathcal{X}$.
By using monotonicity of $\rho$, we obtain
the interchangeability formula for risk functions:

$$\rho\left(\inf_{z \in \mathbb{R}^n} h(z, \cdot)\right) = \inf_{Z \in \mathcal{Z}} \rho(H_Z),$$

where $H_Z(\omega) := h(Z(\omega), \omega)$.

# Dynamic Programming Equations

Under all previous assumptions we obtain the following relations.

Define the value function

$$V_t(z_{t-1}) = \inf_{Z_t,\dots,Z_T} \Big[ F_t(Z_t) + \rho_{\mathcal{X}_{t+1}|\mathcal{X}_t}\Big( F_{t+1}(Z_{t+1}) + \cdots \\ \cdots + \rho_{\mathcal{X}_{T-1}|\mathcal{X}_{T-2}}\Big[ F_{T-1}(Z_{T-1}) + \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}\Big( F_T(Z_T)\Big)\Big]\Big)\Big]$$

The optimal policy and the value function satisfy the Bellman equations

$$\Big[V_t(z_{t-1})\Big](\omega) = \inf_{z_t \in \mathcal{G}_t(z_{t-1},\omega)} \Big\{ f_t(z_t,\omega) + \Big[\rho_{\mathcal{X}_{t+1}|\mathcal{X}_t}\Big(V_{t+1}(z_t)\Big)\Big](\omega)\Big\}$$

Suppose now that the conditional risk mappings $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$ are lower semicontinuous and positively homogeneous. Then there exist convex closed sets $\mathcal{A}_t(\omega) \subset \mathscr{P}_{\mathcal{Y}_t|\mathcal{F}_{t-1}}(\omega)$ such that

$$\Big[\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}(X_t)\Big](\omega) = \sup_{\mu_t \in \mathcal{A}_t(\omega)} \mathbb{E}_{\mu_t}\Big[X_t\Big] = \sup_{\nu \in \mathcal{D}_t} \mathbb{E}_{\nu}\Big[X_t\Big|\mathcal{F}_{t-1}\Big](\omega) \quad \text{(separability)}$$

Under convexity and mild regularity assumptions the 'inf' and 'sup' operators can be interchanged (duality).

(Ruszczyński–Shapiro, 2004)

# Risk Aversion by Stochastic Dominance Constraints

(Dentcheva–Ruszczyński, 2003–)

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

New Model:

$$\begin{aligned} \max\ & \mathbb{E}[X_z] && \text{(or some other objective)} \\ \text{subject to } & X_z \succeq_s Y && \text{(stochastic ordering constraint)} \\ & z \in Z \end{aligned}$$

$X_z$ is preferred over $Y$ by all decision makers
having utility functions in the generator $\mathscr{U}$:

$$\mathbb{E}[u(X_z)] \geq \mathbb{E}[u(Y)] \quad \forall\ u \in \mathscr{U}$$

- if $\mathscr{U}$ is the set of all nondecreasing functions
  – first order dominance "$\succeq_{(1)}$"

- if $\mathscr{U}$ is the set of all concave nondecreasing functions
  – second order dominance "$\succeq_{(2)}$"