Multiperiod Risk Functionals in stochastic programming

G. Ch. Pflug and W. Römisch

University of Vienna
Institute of Statistics and Decision Support Systems
and
Humboldt-University Berlin
Institute of Mathematics

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

EURO XXI, Reykjavik (Iceland), July 2–5, 2006

WWTF – Wiener Wissenschafts-, Forschungs- und Technologiefonds

DFG Research Center MATHEON
Mathematics for key technologies

Multistage stochastic programs

Let $\{\xi_t\}_{t=1}^T$ be a discrete-time stochastic data process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with ξ_1 deterministic. The stochastic decision x_t at period t is assumed to be measurable with respect to $\mathcal{F}_t := \sigma(\xi_1, \ldots, \xi_t)$ (nonanticipativity).

Multistage stochastic optimization model:

$$\min \left\{ \mathbb{E}\left[\sum_{t=1}^{T} \langle b_t(\xi_t), x_t \rangle\right] \middle| \begin{array}{l} x_t \in X_t, t = 1, \ldots, T, A_{1,0}x_1 = h_1(\xi_1), \\ x_t \text{ is } \mathcal{F}_t - \text{measurable}, t = 1, \ldots, T, \\ A_{t,0}x_t + A_{t,1}x_{t-1} = h_t(\xi_t), t = 2, ., T \end{array} \right\}$$

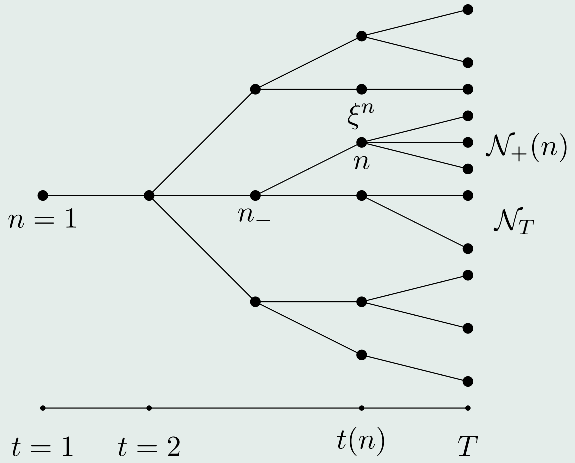
where the sets X_t, $t = 1, \ldots, T$, are polyhedral cones, the vectors $b_t(\cdot)$ and $h_t(\cdot)$ depend affine linearly on ξ_t.

If the process $\{\xi_t\}_{t=1}^T$ has a finite number of scenarios, they exhibit a scenario tree structure.

Typical applications: Power production planning, revenue and portfolio management models.

Data process approximation by scenario trees

The process $\{\xi_t\}_{t=1}^T$ is approximated by a process forming a scenario tree being based on a finite set $\mathcal{N} \subset \mathbb{N}$ of nodes.



Scenario tree with $T=5$, $N=22$ and 11 leaves

$n=1$ root node, n_- unique predecessor of node n, $\mathrm{path}(n) = \{1,\ldots,n_-,n\}$, $t(n) := |\mathrm{path}(n)|$, $\mathcal{N}_+(n)$ set of successors to n, $\mathcal{N}_T := \{n \in \mathcal{N} : \mathcal{N}_+(n) = \emptyset\}$ set of leaves, $\mathrm{path}(n)$, $n \in \mathcal{N}_T$, scenario with (given) probability π^n, $\pi^n := \sum_{\nu \in \mathcal{N}_+(n)} \pi^\nu$ probability of node n, ξ^n realization of $\xi_{t(n)}$.

Tree representation of the optimization model

$$\min \left\{ \sum_{n \in \mathcal{N}} \pi^n \langle b_{t(n)}(\xi^n), x^n \rangle \,\middle|\, \begin{array}{l} x^n \in X_{t(n)}, n \in \mathcal{N}, A_{1,0} x^1 = h_1(\xi^1) \\ A_{t(n),0} x^n + A_{t(n),1} x^{n-} = h_{t(n)}(\xi^n), n \in \mathcal{N} \end{array} \right\}$$

How to solve the optimization model ?

- Standard software (e.g., CPLEX)

- Decomposition methods for (very) large scale models
  (Ruszczynski/Shapiro (Eds.): Stochastic Programming, Handbook, 2003)

Open question:
How to model and incorporate risk into multiperiod models ?

Risk functionals

Let z be a real random variable on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that $z = z(x)$ is the revenue depending on a decision x in some stochastic optimization model. The traditional objective of such models consists in maximizing the expected revenue, i.e.,

$$\max_x \ \mathbb{E}[z(x)].$$

However, the revenue $z(x)$ of some or many decisions x might have fat tails, in particular, to the left. Looking only at the expectation of z hides any tail information.

Examples of risk functionals:

Upper semivariance:

$$sV_+(z) := \mathbb{E}[[\mathbb{E}[z] - z]_+^2] = \mathbb{E}[\max\{\mathbb{E}[z] - z, 0\}^2]$$

Value-at-Risk:

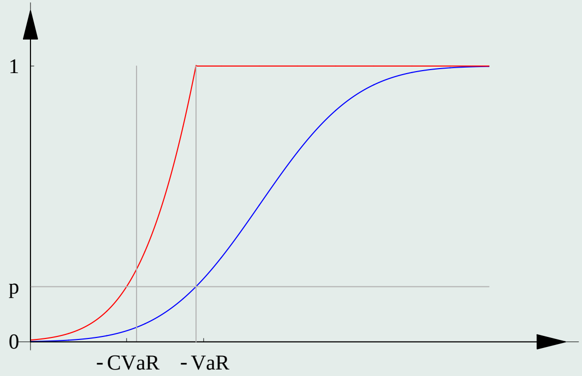
$$VaR_p(z) := -\inf\{r \in \mathbb{R} : \mathbb{P}(z \le r) \ge p\} \quad (p \in (0,1))$$

Conditional Value-at-risk:

$$CVaR_p(z) := \text{mean of the tail distribution function } F_p$$

where $F_p(t) := \begin{cases} 1 & t \geq -VaR_p(z), \\ \frac{F(t)}{p} & t < -VaR_p(z) \end{cases}$ and

$F(t) := \mathbb{P}(\{z \leq t\})$ is the distribution function of z.



$VaR_p(z)$ and $CVaR_p(z)$ for a continuously distributed z

Home Page
Title Page
Contents
Go Back
Full Screen
Close
Quit

Axiomatic characterization of risk:

Let $\mathcal{Z} = L_p(\Omega, \mathcal{F}, I\!\!P)$ for $1 \le p \le +\infty$. A mapping $\mathcal{A} : \mathcal{Z} \to I\!\!R$ is called acceptability functional if it is concave on $\mathcal{Z}$ and satisfies the following two conditions for all z, $\tilde{z} \in \mathcal{Z}$:

(i) If $z \le \tilde{z}$, then $\mathcal{A}(z) \le \mathcal{A}(\tilde{z})$ (monotonicity).

(ii) For each $r \in I\!\!R$ and $z \in \mathcal{Z}$ we have $\mathcal{A}(z + r) = \mathcal{A}(z) + r$ (translation equivariance).

An acceptability functional $\mathcal{A}$ is called positively homogeneous if $\mathcal{A}(\lambda z) = \lambda \mathcal{A}(z)$ holds for all $\lambda \ge 0$ and $z \in \mathcal{Z}$.
$\mathcal{A}$ is called strict if $\mathcal{A}(z) \le I\!\!E[z]$ for each $z \in \mathcal{Z}$.

Given an acceptability functional $\mathcal{A}$, the mapping $\rho := -\mathcal{A}$ is called a convex risk functional. ρ is called a coherent risk functional if $\mathcal{A}$ is positively homogeneous.

References: Artzner/Delbaen/Eber/Heath 99, Föllmer/Schied 02

Conditional risk mappings

Let $\mathcal{G} \subset \mathcal{F}$ be σ-fields and $\mathcal{Y} := L_p(\Omega, \mathcal{G}, I\!P)$ be the corresponding subspace of $\mathcal{Z}$.

A mapping $\mathcal{A} : \mathcal{Z} \to \mathcal{Y}$ is called conditional acceptability mapping or acceptability mapping with observable information $\mathcal{G}$ if it satisfies the following conditions for all $z, \tilde{z} \in \mathcal{Z}$:

(i) $\mathcal{A}(\lambda z + (1-\lambda)\tilde{z}) \geq \lambda \mathcal{A}(z) + (1-\lambda)\mathcal{A}(\tilde{z})$ for all $\lambda \in [0,1]$ (((pointwise) concavity)

(ii) If $z \leq \tilde{z}$, then $\mathcal{A}(z) \leq \mathcal{A}(\tilde{z})$ (monotonicity).

(iii) If $\tilde{z} \in \mathcal{Y}$, we have $\mathcal{A}(z + \tilde{z}) = \mathcal{A}(z) + \tilde{z}$ ((predictable) translation equivariance).

Notation: $\mathcal{A}(\cdot, \mathcal{G})$ or $\mathcal{A}_{\mathcal{F}|\mathcal{G}}$.

The mapping $\rho = \rho_{\mathcal{F}|\mathcal{G}} := -\mathcal{A}_{\mathcal{F}|\mathcal{G}}$ is called conditional convex risk mapping.

References: Detlefsen/Scandolo, Finance Stochast. 05, Ruszczynski/Shapiro, Math. OR 06

Multiperiod risk functionals

Let a filtration of σ-fields $\mathcal{F}_t$, $t = 1, \ldots, T$, and (real) random variables $\{z_t\}_{t=1}^T$ with $z_t \in L_p(\Omega, \mathcal{F}_t, I\!\!P)$, $1 \le p \le +\infty$, be given. Then it may become necessary to measure their risk by multiperiod functionals. We assume $\mathcal{F}_t \subseteq \mathcal{F}_{t+1} \subseteq \mathcal{F}$ and $\mathcal{F}_1 = \{\emptyset, \Omega\}$, i.e. z_1 is deterministic.

A functional $\mathcal{A} : \mathcal{Z} = \times_{t=1}^T L_p(\Omega, \mathcal{F}_t, I\!\!P) \to \overline{I\!\!R}$ is called multiperiod acceptability functional if it is concave and satisfies the following two conditions for all $z, \tilde{z} \in \times_{t=1}^T L_p(\Omega, \mathcal{F}_t, I\!\!P)$:
(i) If $z_t \le \tilde{z}_t$, $t = 1, \ldots, T$, then $\mathcal{A}(z_1, \ldots, z_T) \le \mathcal{A}(\tilde{z}_1, \ldots, \tilde{z}_T)$ (monotonicity),
(ii) If $\tilde{z}_t \in L_p(\Omega, \mathcal{F}_{t-1}, I\!\!P)$, then $\mathcal{A}(z_1, \ldots, z_t + \tilde{z}_t, \ldots, z_T) = I\!\!E[\tilde{z}_t] + \mathcal{A}(z_1, \ldots, z_T)$ ((predictable) translation equivariance).
Notation: $\mathcal{A}(z_1, \ldots, z_T; \mathcal{F}_1, \ldots, \mathcal{F}_T)$.

The mapping $\rho := -\mathcal{A}$ is called a multiperiod convex risk functional.

Dual representations and properties

Let $\mathcal{Z}^*$ denote the topological dual of $\mathcal{Z}$ for $p \in [1, +\infty)$, i.e., $\mathcal{Z}^* := \times_{t=1}^T L_{p'}(\Omega, \mathcal{F}_t, I\!P)$ with $\frac{1}{p} + \frac{1}{p'} = 1$, and let

$$\langle z^*, z \rangle = \sum_{t=1}^{T} I\!E[z_t^* z_t]$$

be the dual pairing between $\mathcal{Z}^*$ and $\mathcal{Z}$.

An acceptability functional $\mathcal{A}$ is called proper if $\mathcal{A}(z) < +\infty$ for all $z \in \mathcal{Z}$ and its domain dom$(\mathcal{A}) := \{Z \in \mathcal{Z} : \mathcal{A}(z) > -\infty\}$ is nonempty. If $\mathcal{A}$ is proper and upper semicontinuous, its domain is closed and convex.

The conjugate $\mathcal{A}^* : \mathcal{Z}^* \to \overline{I\!R}$ of $\mathcal{A}$ is given by

$$\mathcal{A}^*(z^*) := \inf_{z \in \mathcal{Z}} \{\langle z^*, z \rangle - \mathcal{A}(Y)\}.$$

The Fenchel-Moreau theorem of convex analysis then implies the representation

$$\mathcal{A}(z) = \inf_{z^* \in \mathcal{Z}^*} \{\langle z^*, z \rangle - \mathcal{A}^*(z^*)\}$$

if $\mathcal{A}$ is proper and upper semicontinuous.

Theorem:

Let $\mathcal{A} : \mathcal{Z} \to \overline{\mathbb{R}}$ be a proper multiperiod acceptability functional. Then the representation

$$\mathcal{A}(z) = \inf_{z^* \in \mathcal{Z}^*} \left\{ \sum_{t=1}^{T} \mathbb{E}[z_t^* z_t] - \mathcal{A}^*(z^*) \ : \ z_t^* \geq 0, \ \mathbb{E}[z_t^* | \mathcal{F}_{t-1}] = 1, \right.$$
$$\left. t = 2, \ldots, T \right\}$$

is valid for every $z \in \mathcal{Z}$ if $\mathcal{A}$ is upper semicontinuous.

Conversely, if $\mathcal{A}$ can be represented in the above form for some function $\mathcal{A}^* : \mathcal{Z}^* \to \overline{\mathbb{R}}$, then $\mathcal{A}$ is an upper semicontinuous multiperiod acceptability functional.

Moreover, $\mathcal{A}$ is locally Lipschitz continuous, superdifferentiable and Hadamard directionally differentiable on $\operatorname{int} \operatorname{dom}(\mathcal{A})$. Its directional derivative at $\bar{z} \in \operatorname{int} \operatorname{dom}(\mathcal{A})$ satisfies

$$\mathcal{A}'(\bar{z}, z) = \inf_{z^* \in \partial \mathcal{A}(\bar{z})} \langle z^*, z \rangle, \ \forall z \in \mathcal{Z},$$
$$\partial \mathcal{A}(\bar{z}) = \{ z^* \in \mathcal{Z}^* : \mathcal{A}(z) \leq \mathcal{A}(\bar{z}) + \langle z^*, z - \bar{z} \rangle, \ \forall z \in \mathcal{Z} \}.$$

Reference: Ruszczynski/Shapiro, Math. OR 06 (to appear)

Multiperiod polyhedral risk functionals

It is a natural idea to introduce acceptability and risk functionals as optimal values of certain stochastic programs.

Definition:

A multiperiod acceptability functional $\mathcal{A}$ on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ is called polyhedral if there are $k_t \in \mathbb{N}$, $c_t \in \mathbb{R}^{k_t}$, $t = 1, \ldots, T$, $w_{t\tau} \in \mathbb{R}^{k_{t-\tau}}$, $t = 1, \ldots, T$, $\tau = 0, \ldots, t-1$, a polyhedral set Y_1 and polyhedral cones $Y_t \subset \mathbb{R}^{k_t}$, $t = 2, \ldots, T$, such that

$$\mathcal{A}(z) = -\inf \left\{ \mathbb{E}\Big[\sum_{t=1}^{T} \langle c_t, y_t \rangle\Big] \;\middle|\; \begin{array}{c} y_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{k_t}),\ y_t \in Y_t \\ \sum_{\tau=0}^{t-1} \langle w_{t,\tau}, y_{t-\tau} \rangle = z_t,\ \ t = 1, \ldots, T \end{array} \right\}.$$

A mapping $\rho := -\mathcal{A}$ is called multiperiod polyhedral risk functional.

Remark: A convex combination of the expectation and of a multiperiod polyhedral acceptability functional is again a multiperiod polyhedral acceptability functional.

Polyhedral risk functionals preserve linearity and decomposition structures of optimization models. (Eichhorn/Römisch, SIAM J. Optim. 05)

Theorem:

Let $\mathcal{A}$ be a functional on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, I\!P)$ ($p \in [1, +\infty)$) having the form in the previous definition. Assume

(i) complete recourse: $\langle w_{t,0}, Y_t \rangle = I\!R$, $t = 2, \ldots, T$,

(ii) dual feasibility: $\left\{ u \in I\!R^T : c_t + \sum_{\nu=t}^{T} u_\nu w_{\nu,\nu-t} \in -Y_t^* \right\} \neq \emptyset$,

where the sets Y_t^* are the (polyhedral) polar cones of Y_t.

Then $\mathcal{A}$ is finite, continuous and concave on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, I\!P)$ and the following dual representation holds whenever $\frac{1}{p} + \frac{1}{p'} = 1$:

$$\mathcal{A}(z) = \inf \left\{ \begin{array}{l} I\!E\left[\sum_{t=1}^{T} z_t^* z_t\right] \\ -\inf_{y_1 \in Y_1} \left\langle c_1 + \sum_{\nu=1}^{T} I\!E\left[z_\nu^*\right] w_{\nu,\nu-1}, y_1 \right\rangle \end{array} \middle| \begin{array}{l} z_t^* \in L_{p'}(\Omega, \mathcal{F}_t, I\!P) \\ c_t + \sum_{\nu=t+1}^{T} I\!E\left[z_\nu^* | \mathcal{F}_t\right] w_{\nu,\nu-t} \in -Y_t^* \end{array} \right\}.$$

Idea: Determine the parameters k_t, c_t, $w_{t\tau}$ and Y_t such that

$$c_1 + \sum_{t=2}^{T} w_{t,t-1} \mathbb{E}[z_t^*] \in -Y_1^* \;\Rightarrow\; \mathbb{E}[z_2^*] = 1,$$

$$c_t + \sum_{\tau=t+1}^{T} w_{\tau,\tau-t} \mathbb{E}[z_\tau^* | \mathcal{F}_t] \in -Y_t^* \;\Rightarrow\; \mathbb{E}[z_{t+1}^* | \mathcal{F}_t] = 1 \text{ and } z_t^* \geq 0$$

$$c_T + w_{T,0} z_T^* \in -Y_T^* \;\Rightarrow\; z_T^* \geq 0.$$

We assume $k_1 \geq 2$, $k_t \geq 3$, $t = 2, \ldots, T-1$, and $k_T \geq 2$.

Furthermore, let the sets Y_t be of the form

$Y_1 = \mathbb{R} \times \hat{Y}_1, Y_t = \mathbb{R} \times \mathbb{R}_+ \times \hat{Y}_t,\ t = 2, \ldots, T-1, Y_T = \mathbb{R}_+ \times \hat{Y}_T,$

where $\hat{Y}_1$ is polyhedral and the sets $\hat{Y}_t$, $t = 2, \ldots, T$, are polyhedral cones. Finally, we set

$c_1 = (-1, \hat{c}_1)$, $c_t = (-1, 0, \hat{c}_t)$, $t = 2, \ldots, T-1$, $c_T = (0, \hat{c}_T)$, $w_{1,1} = (1, \hat{w}_{1,1})$, $w_{t,t} = (0, \hat{w}_{t,t})$, $t = 2, \ldots, T$, $w_{T,1} = (1, \hat{w}_{T,1})$, $w_{t,1} = (0, 1, \hat{w}_{t,1})$, $w_{t,1} = (1, 0, \hat{w}_{t,1})$, $w_{t,\tau} = (0, 0, \hat{w}_{t,\tau})$, $\tau = 2, \ldots, T-t$, $t = 1, \ldots, T-1$.

Corollary:

Let $\mathcal{A}$ be a functional on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, I\!P)$ $(p \in [1, +\infty))$ with parameters chosen as above. Assume complete recourse and dual feasibility. Then $\mathcal{A}$ is a finite, continuous and multiperiod acceptability functional having the representation

$$\mathcal{A}(z) = \inf \Big\{ \sum_{t=1}^{T} I\!E[z_t^* z_t] - \inf_{\hat{y}_1 \in \hat{Y}_1} \langle \bar{c}_1, \hat{y}_1 \rangle | z_t^* \in L_{p'}(\Omega, \mathcal{F}_t, I\!P),\ z_t^* \geq 0,$$
$$I\!E[z_t^* | \mathcal{F}_{t-1}] = 1,\ \bar{c}_t + \hat{w}_{t,1} z_t^* + \in -\hat{Y}_t^*,\ t = 2, \ldots, T \Big\},$$

where $\bar{c}_1 := \hat{c}_1 + \sum_{t=2}^{T} \hat{w}_{t,t-1} \in Y_1$, $\bar{c}_t := \hat{c}_t + \sum_{\tau=t+1}^{T} \hat{w}_{\tau,\tau-t}$, $t = 2, \ldots, T-1$, $\bar{c}_T := \hat{c}_T$.

Example: (value-of-information approach)

Let $k_t = 3$ for $t = 1, \ldots, T$, $\bar{c}_1 = (0,0)$, $\hat{Y}_1 = I\!R_+ \times I\!R_+$, $\bar{c}_t = \frac{1}{\alpha_t}$, $\hat{w}_{t,0} = -1$ and $\hat{Y}_t = I\!R_+$, $t = 1, \ldots, T-1$, $\bar{c}_T = (0, \frac{1}{\alpha_T})$, $\hat{Y}_T = I\!R_+ \times I\!R_+$ and $\hat{w}_{T,1} = (0,-1)$, where $\alpha_t \in (0,1)$. Then we obtain the following acceptability functional

$$\mathcal{A}(z) = \inf \Big\{ \sum_{t=2}^{T} I\!E[z_t^* z_t] \;:\; z_t^* \in L_{p'}(\Omega, \mathcal{F}_t, I\!P), I\!E[z_t^* | \mathcal{F}_{t-1}] = 1, \\ z_t^* \in [0, \frac{1}{\alpha_t}], t = 2, \ldots, T \Big\}.$$

$$\mathcal{A}(z) = I\!E\Big[\sum_{t=2}^{T} \inf \Big\{ z_t^* z_t : I\!E[z_t^* | \mathcal{F}_{t-1}] = 1, z_t^* \in [0, \frac{1}{\alpha_t}] \Big\} \Big]$$

$$\rho(z) = I\!E\Big[\sum_{t=2}^{T} AVaR_{\alpha_t}(z_t, \mathcal{F}_{t-1}) \Big]$$

Reference: Pflug/Ruszczynski 04, 05

Iterated conditional risk mappings

Let $\mathcal{A}_{\mathcal{F}_t|\mathcal{F}_{t-1}}$, $t = 2, \ldots, T$, be conditional acceptability mappings and we define an acceptability functional $\mathcal{A}$ on $\times_{t=1}^{T} L_p(\Omega, \mathcal{F}_t, I\!\!P)$

$$\mathcal{A}(z) := z_1 + \mathcal{A}_{\mathcal{F}_2|\mathcal{F}_1}\Big[z_2 + \cdots + \mathcal{A}_{\mathcal{F}_{T-1}|\mathcal{F}_{T-2}}[z_{T-1} + \mathcal{A}_{\mathcal{F}_T|\mathcal{F}_{T-1}}(z_T)]\Big]$$
$$= \mathcal{A}_{\mathcal{F}_2|\mathcal{F}_1} \circ \cdots \circ \mathcal{A}_{\mathcal{F}_{T-1}|\mathcal{F}_{T-2}} \circ \mathcal{A}_{\mathcal{F}_T|\mathcal{F}_{T-1}}(z_1 + \cdots + z_T),$$

where the latter representation is a consequence of the (predictable) translation equivariance.

Example: (polyhedral conditional acceptability mappings)

$$\mathcal{A}_{\mathcal{F}_t|\mathcal{F}_{t-1}}(z) = -\inf\Big\{\langle c_1, y_1\rangle + I\!\!E[\langle c_2, y_2\rangle|\mathcal{F}_{t-1}] : y_1 \in L_p(\Omega, \mathcal{F}_{t-1}, I\!\!P),$$
$$y_1 \in Y_1, y_2 \in L_p(\Omega, \mathcal{F}_t, I\!\!P), y_2 \in Y_2, \langle w_1, y_1\rangle + \langle w_2, y_2\rangle = z\Big\}$$

and select the parameters such that $\mathcal{A}_{\mathcal{F}_t|\mathcal{F}_{t-1}}$ is a conditional acceptability mapping.

Reference: Ruszczynski/Shapiro, Math. OR 06 (to appear)