Polyhedral risk measures in stochastic programming

A. Eichhorn and W. Römisch

Humboldt-University Berlin
Institute of Mathematics
10099 Berlin, Germany

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

GAMS Workshop, Heidelberg, Sept. 1-3, 2003
1 Introduction

When wishing to replace the usual expectation-based objective by some functional measuring risk, at least the following three issues have to be addressed:

- What is an appropriate risk measure for the underlying practical model?
- Does this exchange lead to serious changes in structure and stability?
- Does the exchange cause serious computational problems?

References: Schultz/Tiedemann 02, Schultz 03

Of course, the stochastic programming user wishes that his choice of a risk measure leads to the answer no on the last two questions.

We will discuss a subclass of convex risk measures having this enjoyable property.

Our motivation stems from stochastic programming applications in electricity portfolio management, i.e., from solving large scale mixed-integer stochastic programs.
2 Polyhedral risk measures

Let $\mathcal{Z}$ denote a linear space of real random variables on some probability space $(\Omega, \mathcal{F}, IP)$. We assume that $\mathcal{Z}$ contains the constants. A functional $\rho : \mathcal{Z} \to \mathbb{R}$ is called a risk measure if it satisfies the following two conditions for all $z, \tilde{z} \in \mathcal{Z}$:

(i) If $z \leq \tilde{z}$, then $\rho(z) \geq \rho(\tilde{z})$ (monotonicity).

(ii) For each $r \in \mathbb{R}$ we have $\rho(z + r) = \rho(z) - r$ (translation invariance).

A risk measure $\rho$ is called convex if it satisfies the condition

$$\rho(\lambda z + (1 - \lambda) \tilde{z}) \leq \lambda \rho(z) + (1 - \lambda) \rho(\tilde{z})$$

for all $z, \tilde{z} \in \mathcal{Z}$ and $\lambda \in [0, 1]$.

A convex risk measure is called coherent if it is positively homogeneous, i.e., $\rho(\lambda z) = \lambda \rho(z)$ for all $\lambda \geq 0$ and $z \in \mathcal{Z}$.

References: Artzner/Delbaen/Eber/Heath 99, Föllmer/Schied 02
**Definition:** A risk measure \( \rho \) on \( Z \) will be called polyhedral if there exist \( k, l \in \mathbb{N}, a, c \in \mathbb{R}^k, q, w \in \mathbb{R}^l \), a polyhedral set \( X \subseteq \mathbb{R}^k \) and a polyhedral cone \( Y \subseteq \mathbb{R}^l \) such that

\[
\rho(z) = \inf \{ \langle c, x \rangle + \mathbb{E}[\langle q, y \rangle] : \langle a, x \rangle + \langle w, y \rangle = z, x \in X, y \in Y \}
\]

for each \( z \in Z \). Here, \( \mathbb{E} \) denotes the expectation on \( (\Omega, \mathcal{F}, \mathbb{P}) \) and \( \langle \cdot, \cdot \rangle \) the scalar product on \( \mathbb{R}^k \).

The notion polyhedral risk measure is motivated by the polyhedrality of \( \rho(z) \) as a function of the scenarios of \( z \) if \( z \) is discrete.

Assume that \( \rho \) is a polyhedral risk measure on the space \( Z = L_1(\Omega, \mathcal{F}, \mathbb{P}) \), that \( \langle w, Y \rangle = \mathbb{R} \) and \( \{ u \in \mathbb{R} : uw - q \in Y^* \} \neq \emptyset \), where \( Y^* \) is the polar cone of \( Y \). Then there exist two real numbers \( u_\ell, \ell = 1, 2 \), such that

\[
\rho(z) = \inf_{x \in X} \{ \langle c, x \rangle + \mathbb{E}[\max_{\ell=1,2} u_\ell (z - \langle a, x \rangle)] \}.
\]

In particular, \( \rho \) is a convex risk measure. It is coherent if \( X \) is a cone.
**Proposition:** Let $k, l \in \mathbb{IN}$, $a, c \in \mathbb{IR}^k$, $q, w \in \mathbb{IR}^l$, a polyhedral set $X \subseteq \mathbb{IR}^k$ and a polyhedral cone $Y \subseteq \mathbb{IR}^l$ be given such that

$$
\rho(z) = \inf \{ \langle c, x \rangle + \mathbb{E}[\langle q, y \rangle] : \langle a, x \rangle + \langle w, y \rangle = z, x \in X, y \in Y \}
$$

for each $z \in L_1(\Omega, \mathcal{F}, \mathbb{IP})$.

Let $\langle w, Y \rangle = \mathbb{IR}$ and $\emptyset \neq \{ u \in \mathbb{IR} : uw - q \in Y^* \} \subset \mathbb{IR}^-$ and $a, c$ and $X$ have the form $a = (\hat{a}, -1)$, $c = (\hat{c}, 1)$ and $X = \hat{X} \times \mathbb{IR}$, where $\hat{a}, \hat{c} \in \mathbb{IR}^{k-1}$ and $\hat{X} \subseteq \mathbb{IR}^{k-1}$.

Then $\rho$ is a convex risk measure on $L_1(\Omega, \mathcal{F}, \mathbb{IP})$ if it is finite.

Furthermore, if $\rho$ is finite it admits the following dual representation

$$
\rho(z) = \sup \left\{ -\mathbb{E}[\lambda z] + \inf_{\hat{x} \in \hat{X}} \langle \hat{c} + \hat{a}, \hat{x} \rangle : \lambda \in L_{p'}(\Omega, \mathcal{F}, \mathbb{IP}), \right. \\
\left. \mathbb{E}[\lambda] = 1, -(q + \lambda w) \in Y^* \right\}
$$

for each $z \in L_p(\Omega, \mathcal{F}, \mathbb{IP})$ with $1 < p < +\infty$ and

$$
\frac{1}{p} + \frac{1}{p'} = 1.
$$

Proof: by relying on stochastic programming methodology; the Lagrangian dual function has the form

$$
D(\lambda) = \inf_{x \in X} \{ (c + \mathbb{E}[\lambda]a, x) + \inf_{y \in L_p, y \in Y} \mathbb{E}[(q + \lambda w, y)] - \mathbb{E}[\lambda z] \}
$$

(Reference: Wets 70).
**Example 1:** Let \( k = l = 1, a = -1, c = 1 \) and \( X = \mathbb{R} \). Then the conditions in the Proposition imply \( Y = \mathbb{R} \), \( w \neq 0 \) and \( \frac{q}{w} < 0 \). We obtain that \( \rho \) has the form

\[
\rho(z) = \mathbb{E}\left[\frac{q}{w}z\right] + \inf\{(1 + \frac{q}{w})x : x \in \mathbb{R}\}.
\]

Hence, \( \rho \) is finite iff \( \frac{q}{w} = -1 \) iff \( \rho(z) = \mathbb{E}[-z] \).

**Example 2:** Conditional, Tail or Average Value at Risk

We consider the Average Value at Risk \( \text{AVaR}_\alpha \) defined by

\[
\text{AVaR}_\alpha(z) := \frac{1}{\alpha} \int_0^\alpha \text{VaR}_\gamma(z) d\gamma
\]

\[
= \inf_{r \in \mathbb{R}} \left\{ r + \frac{1}{\alpha} \mathbb{E}[\max\{0, -r - z\}] \right\},
\]

where \( \text{VaR}_\alpha(z) := \inf\{r \in \mathbb{R} : \mathbb{P}(z + r < 0) \leq \alpha\} \) is the Value at Risk at level \( \alpha \in (0, 1) \).

\( \text{AVaR}_\alpha \) is polyhedral by setting \( k = 1, l = 2, a = -1, c = 1, q = (\frac{1}{\alpha}, 0), w = (-1, 1), X = \mathbb{R} \) and \( Y = \mathbb{R}^2_+ \).

The condition \(- (q + \lambda w) \in Y^* \) in the dual representation is equivalent to \( \lambda \in [0, \frac{1}{\alpha}] \).
3 Multiperiod polyhedral risk measures

While the notion of a polyhedral risk measure is appropriate for two-stage stochastic programming models, we now consider a multiperiod extension in case that instead of the real random variable $z$ a real stochastic process $z = \{z_t, \mathcal{F}_t\}_{t=2}^T$ with a filtration $\mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_t \subseteq \cdots \subseteq \mathcal{F}_T = \mathcal{F}$ is given. It is assumed that $z_\tau$, $\tau = 2, \ldots, t$, is measurable with respect to $\mathcal{F}_t$, $t = 2, \ldots, T$.

As natural candidates for such an extension we consider functionals $\rho$ on the linear space $\times_{t=2}^T L_1(\Omega, \mathcal{F}_t, \mathbb{P})$ that are defined as optimal values of specific multi-stage stochastic programs. Namely, we assume that there are $k_t \in \mathbb{N}$, $c_t \in \mathbb{R}^{k_t}$, $t = 1, \ldots, T$, $a_t \in \mathbb{R}^{k_{t-1}}$, $w_t \in \mathbb{R}^{k_t}$, $t = 2, \ldots, T$, a polyhedral set $Y_1 \subseteq \mathbb{R}^{k_1}$ and polyhedral cones $Y_t \subseteq \mathbb{R}^{k_t}$, $t = 2, \ldots, T$, such that

$$\rho(z) = \inf \{ IE\left[ \sum_{t=1}^T \langle c_t, y_t \rangle \right] : y_1 \in Y_1, y_t \text{ is } \mathcal{F}_t\text{-measurable}, y_t \in Y_t,$$

$$\langle a_t, y_{t-1} \rangle + \langle w_t, y_t \rangle = z_t, t = 2, \ldots, T \}$$
Dynamic programming formulation of $\rho$:

$$
\rho(z) = \inf \{ \langle c_1, y_1 \rangle + \mathbb{E}[v_2(y_1, z)] : y_1 \in Y_1 \}
$$

$$
v_t(y_{t-1}, z_t, \ldots, z_T) := \inf \{ \langle c_t, y_t \rangle + \mathbb{E}[v_{t+1}(y_t, z_{t+1}, \ldots, z_T)|\mathcal{F}_t] : y_t \in Y_t, \langle a_t, y_{t-1} \rangle + \langle w_t, y_t \rangle = z_t \},
$$

$$
t = T, \ldots, 2,
$$

$$
v_{T+1}(y_T) := 0.
$$

References: Rockafellar/Wets 76, Evstigneev 76

Dual formulation for $\rho$:

$$
\rho(z) = \sup \{ \inf_{y_1 \in Y_1} \langle c_1 + \mathbb{E}[\lambda_2] a_2, y_1 \rangle - \mathbb{E} \left[ \sum_{t=2}^{T} \lambda_t z_t \right] : \lambda_t \in L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}), t = 2, \ldots, T, -(c_T + \lambda_T w_T) \in Y_T^*,
$$

$$
-(c_t + w_t \lambda_t + a_{t+1} \mathbb{E}[\lambda_{t+1}|\mathcal{F}_t]) \in Y_t^*, t = T-1, \ldots, 2
$$

holds whenever $z_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}), t = 2, \ldots, T, p > 1,$

$$
\frac{1}{p} + \frac{1}{p'} = 1,$

and the right-hand side is finite.

References: Eisner/Olsen 75, Rockafellar/Wets 76, 78
**Definition:** (Artzner/Delbaen/Eber/Heath/Ku 02)

A functional \( \rho \) on (a subset of) \( \times_{t=2}^{T} L_p^p(\Omega, \mathcal{F}_t, \mathbb{P}) \) \((p > 1)\) is called a multiperiod coherent risk measure if it satisfies the Fatou property and if there exist positive reals \( \eta_t > 0 \), \( t = 2, \ldots, T \), \( \sum_{t=2}^{T} \eta_t = 1 \), and a closed convex set \( \Lambda \) of nonnegative functions contained in \( \times_{t=2}^{T} L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) \), with \( \frac{1}{p} + \frac{1}{p'} = 1 \), such that

\[
\rho(z) = \sup \{- \sum_{t=2}^{T} \eta_t \mathbb{E}[\lambda_t z_t] : \lambda \in \Lambda, \sum_{t=2}^{T} \eta_t \mathbb{E}[\lambda_t] = 1 \}.
\]

\[\Lambda(\eta) := \{ \lambda \in \times_{t=2}^{T} L_{p'}(\Omega, \mathcal{F}_t, \mathbb{P}) : -(c_T + \lambda_T \frac{w_T}{\eta_T}) \in Y_T^*, \]

\[-(c_t + \frac{w_t}{\eta_t} \lambda_t + \frac{a_{t+1}}{\eta_{t+1}} \mathbb{E}[\lambda_{t+1} | \mathcal{F}_t]) \in Y_t^*, t = 2, \ldots, T - 1 \}\]

**Proposition:**

A multiperiod polyhedral risk measure is multiperiod coherent if

(i) \( \eta_t = \frac{1}{T-1} \), \( t = 2, \ldots, T \),

(ii) \( \lambda \in \Lambda(\eta) \) implies \( \lambda_t \geq 0 \), \( \mathbb{P} \)-a.s., \( t = 2, \ldots, T \), and

(iii) \( \lambda \in \Lambda(\eta) \) and \( \inf_{y_1 \in Y_1} \langle c_1 + \mathbb{E}[\lambda_2] a_2, y_1 \rangle = 0 \) imply \( \mathbb{E}[\lambda_t] = 1 \), \( t = 2, \ldots, T \).
4 Multistage SP models with minimal risk

We consider a stochastic process \( \{ \xi^t, \mathcal{F}_t \}_{t=2}^T \) with distribution \( P \) modelling the uncertain future and the multi-stage stochastic program

Minimize \[
\sum_{i=1}^I C_{i1}x_i^1 + \mathbb{E} \left[ \sum_{t=2}^T \sum_{i=1}^I C_{it}(\xi_t)x_i^t \right] \quad \text{(expectation)}
\]
or, alternatively,

Minimize \[
\sum_{i=1}^I C_{i1}x_i^1 + \rho \left( \left\{ \sum_{i=1}^I C_{it}(\xi_t)x_i^t \right\}_{t=2}^T \right) \quad \text{(risk measure)}
\]
such that

\[x^t \text{ is } \mathcal{F}_t - \text{measurable }, x_i^t \in X_{it}, t = 1, \ldots, T,\]
\[A_{it,t}x_i^t + A_{it,t-1}(\xi^t)x_i^{t-1} \geq g_{it}(\xi^t), t = 2, \ldots, T, i = 1, \ldots, I,\]
\[\sum_{i=1}^I B_{it}(\xi^t)x_i^t \geq d_t(\xi^t), t = 1, \ldots, T.\]

Here, \( x^t = (x_1^t, \ldots, x_I^t), t = 1, \ldots, T, \mathcal{F}_1 = \{\emptyset, \Omega\}, \)
\( \mathcal{F}_T = \mathcal{F}, A_{it,\tau}, \tau = t - 1, t, B_{it}, g_{it} \) and \( d_t \) are matrices and vectors possibly depending on \( \xi^t, t = 1, \ldots, T, \) and \( X_{it} \) subsets of Euclidean spaces.
Incorporating the multiperiod risk measure into the original program leads to a multistage model exhibiting the same structure and having \((x, y)\) as decision.

\[
\min_{(x,y)} \sum_{i=1}^{I} C_{i1} x^1_i + \mathbb{IE}\left[\sum_{t=1}^{T} \langle c_t, y_t \rangle\right]
\]

such that

\[
y_1 \in Y_1, y_t \text{ is nonanticipative, } y_t \in Y_t, \]

\[
\langle a_t, y_{t-1}\rangle + \langle w_t, y_t \rangle = \sum_{i=1}^{I} C_{it}(\xi^t)x^t_i, t = 2, \ldots, T,
\]

\[
x^t \text{ is nonanticipative, } x^t_i \in X^t_i, t = 2, \ldots, T,
\]

\[
A_{it,t}x^t_i + A_{it,t-1}(\xi^t)x^{t-1}_i \geq g_{it}(\xi^t), t = 2, \ldots, T, i = 1, \ldots, I,
\]

\[
\sum_{i=1}^{I} B_{it}(\xi^t)x^t_i \geq d_t(\xi^t), t = 1, \ldots, T.
\]

The stability behaviour and the metrics \(\mu_c\) as well as the decomposition structure (e.g. for scenario, node and geographic decomposition) do (almost) not change!