Theory of coherent measures

Miloš Kopa

Faculty of Mathematics and Physics
Charles University in Prague
Czech Republic
e-mail:kopa@karlin.mff.cuni.cz
Measure of risk assigns a real number to any random variable $L$ (loss).

Favorite risk measures:

- **variance**: $\text{var}(L) = \mathbb{E}(L - \mathbb{E}L)^2$
- **standard deviation**: $\text{sd}(L) = \left(\mathbb{E}(L - \mathbb{E}L)^2\right)^{\frac{1}{2}}$
- **semivariance**: $r_s(L) = \mathbb{E}\left[\max(0, L - \mathbb{E}L)^2\right]$
- **mean absolute deviation**: $r_a(L) = \mathbb{E}|L - \mathbb{E}L|$
- **mean absolute semideviation**: $r_{as}(L) = \mathbb{E}\left[\max(0, L - \mathbb{E}L)\right]$

**Value at Risk (VaR):**

$\text{VaR}_\alpha(L) = \inf \left\{ l \in \mathbb{R}, P(L > l) \leq 1 - \alpha \right\}$

**Conditional Value at Risk (CVaR):**

$\text{CVaR}_\alpha(L) = \inf \left\{ a \in \mathbb{R}, a + \frac{1}{1-\alpha} \mathbb{E}\left[\max(0, L - a)\right] \right\}$,

alternatively:

$\text{CVaR}_\alpha(L) = \beta \mathbb{E}(L|L > \text{VaR}_\alpha(L)) + (1 - \beta) \text{VaR}_\alpha(L)$ with some $\beta \in [0, 1]$
Questions for today

Risk measures:

- What are the “reasonable” properties that should have all “good” risk measures?
- Which of the considered measures has the properties?
- Is it possible to generalize a very well-known and popular standard deviation (variance)?
- What is the dual expression of measures with these properties?

Multiobjective optimization:

- How to formulate an optimization problem when multiple objective are considered?
- What are the best solutions of such problems?
- How to find all these best solutions?
Coherent risk measures

CRM: $\mathcal{R} : \mathcal{L}_2(\Omega) \rightarrow (-\infty, \infty]$ that satisfies

(R1) Translation equivariance: $\mathcal{R}(L + C) = \mathcal{R}(L) + C$ for all $X$ and constants $C$,

(R2) Positive homogeneity: $\mathcal{R}(0) = 0$, and $\mathcal{R}(\lambda L) = \lambda \mathcal{R}(L)$ for all $L$ and all $\lambda > 0$,

(R3) Subadditivity: $\mathcal{R}(L + M) \leq \mathcal{R}(L) + \mathcal{R}(M)$ for all $L$ and $M$,

(R4) Monotonicity: $\mathcal{R}(L) \geq \mathcal{R}(M)$ when $L \geq M$. 

Strictly expectation bounded risk measures satisfy (R1), (R2), (R3), and (R5) $\mathcal{R}(L) > E[L]$ for all nonconstant $L$, whereas $\mathcal{R}(L) = E[L]$ for constant $L$.

Note: (R2) & (R3) implies convexity of $\mathcal{R}$: for each $a \in [0, 1]$ we have:

$$\mathcal{R}(aL + (1-a)M) \leq a \mathcal{R}(L) + (1-a) \mathcal{R}(M)$$

Coherent risk measures

CRM: $\mathcal{R} : \mathcal{L}_2(\Omega) \to (-\infty, \infty]$ that satisfies

(R1) Translation equivariance: $\mathcal{R}(L + C) = \mathcal{R}(L) + C$ for all $X$ and constants $C$,

(R2) Positive homogeneity: $\mathcal{R}(0) = 0$, and $\mathcal{R}(\lambda L) = \lambda \mathcal{R}(L)$ for all $L$ and all $\lambda > 0$,

(R3) Subaditivity: $\mathcal{R}(L + M) \leq \mathcal{R}(L) + \mathcal{R}(M)$ for all $L$ and $M$,

(R4) Monotonicity: $\mathcal{R}(L) \geq \mathcal{R}(M)$ when $L \geq M$.

**Strictly expectation bounded risk measures** satisfy (R1), (R2), (R3), and

(R5) $\mathcal{R}(L) > \mathbb{E}[L]$ for all nonconstant $L$, whereas $\mathcal{R}(L) = \mathbb{E}[L]$ for constant $L$. 

Note: (R2)&(R3) implies convexity of $\mathcal{R}$: for each $a \in [0, 1]$ we have:

$$\mathcal{R}(aL + (1-a)M) \leq a \mathcal{R}(L) + (1-a) \mathcal{R}(M)$$
Coherent risk measures

CRM: $\mathcal{R} : L_2(\Omega) \to (-\infty, \infty]$ that satisfies

1. **Translation equivariance:** $\mathcal{R}(L + C) = \mathcal{R}(L) + C$ for all $X$ and constants $C$,
2. **Positive homogeneity:** $\mathcal{R}(0) = 0$, and $\mathcal{R}(\lambda L) = \lambda \mathcal{R}(L)$ for all $L$ and all $\lambda > 0$,
3. **Subadditivity:** $\mathcal{R}(L + M) \leq \mathcal{R}(L) + \mathcal{R}(M)$ for all $L$ and $M$,
4. **Monotonicity:** $\mathcal{R}(L) \geq \mathcal{R}(M)$ when $L \geq M$.

**Strictly expectation bounded risk measures** satisfy (R1), (R2), (R3), and

5. **(R5)** $\mathcal{R}(L) > \mathbb{E}[L]$ for all nonconstant $L$, whereas $\mathcal{R}(L) = \mathbb{E}[L]$ for constant $L$.

Note: (R2)&(R3) implies convexity of $\mathcal{R}$: for each $a \in [0, 1]$ we have:

$$\mathcal{R}(aL + (1-a)M) \leq \mathcal{R}(aL) + \mathcal{R}((1-a)M) = a\mathcal{R}(L) + (1-a)\mathcal{R}(M)$$

Coherent properties for the popular risk measures

- variance: none of (R1)-(R5)
- standard deviation: (R2)
- semivariance: none of (R1)-(R5)
- mean absolute deviation: (R2), (R3)
- mean absolute semideviation: (R2), (R3)
- Value at Risk (VaR): (R1),(R2), (R4)
- Conditional Value at Risk (CVaR): (R1)-(R5)
Consider a measurable space $(\Omega, \mathcal{A})$ and the set $\mathcal{P}$ of all probability measures on the space.

**Definition**

A set $Q \subset \mathcal{P}$ is called a risk envelope if for each $Q \in Q$ one has: $Q \geq 0$ and $\mathbb{E}Q = 1$.

**Theorem**

$\mathcal{R}$ is a coherent risk measure if and only if there exists a risk envelope $Q$ such that:

$$\mathcal{R}(L) = \max_{Q \in \mathcal{Q}} \mathbb{E}(QL)$$

and $Q$ can be chosen as a convex set.

Interpretation: A coherent risk measure can be understood as a worst-case expectation with respect to some class of probability distributions on $(\Omega, \mathcal{A})$. It means for some distribution $P'$. If the probability distribution of $L$ is $P$ then $Q = \frac{dP'}{dP}$.
To simplify the situation consider a measurable space with $M$ atoms (discrete distributions). Moreover let $L$ has a discrete uniform distribution on the space - atoms are equiprobable, i.e. discrete distribution with $M$ equiprobable scenarios $l_j$, $j = 1, 2, ..., M$. Assume that $M(1 - \alpha)$ is an integer number. Then:

$$CVaR_\alpha(L) = \min_{a, z_j} \ a + \frac{1}{(1 - \alpha) M} \sum_{j=1}^{M} z_j$$

s. t. $z_j \geq l_j - a, j = 1, .., M$

$z_j \geq 0, j = 1, .., M$
Example - Risk envelope for CVaR

And dual problem:

\[ CVaR_\alpha(L) = \max_{y_j} \sum_{j=1}^{M} y_j l_j \]

s. t. \[ \sum_{j=1}^{M} y_j = 1 \]

\[ y_j \leq \frac{1}{(1 - \alpha) M} \]

\[ y_j \geq 0, \ j = 1, \ldots, M \]

Note that optimal solution: \( y_j^* = 0 \) for \( j = 1, 2, \ldots, M\alpha \)

\[ y_j^* = \frac{1}{(1 - \alpha)M} \] for \( j = M\alpha + 1, \ldots, M \). Hence the risk envelope for CVaR is:

\[ Q = \{ Q : EQ = 1, 0 \leq Q \leq \frac{1}{(1 - \alpha)} \} \]
A return measure is defined as a functional $\mathcal{E}(L) = -\mathcal{R}(L)$ for a coherent risk measure $\mathcal{R}$. It is obvious that the expectation belongs to this class.
Rockafellar, Uryasev and Zabarankin (2006A, 2006B): GDM are introduced as an extension of *standard deviation* but they need not to be symmetric with respect to *upside* $X - \mathbb{E}X$ and *downside* $\mathbb{E}X - X$ of a random variable $X$. 

Any functional $D: L^2(\Omega) \rightarrow [0, \infty]$ is called a general deviation measure if it satisfies

1. $D(X + C) = D(X)$ for all $X$ and constants $C$,
2. $D(0) = 0$, and $D(\lambda X) = \lambda D(X)$ for all $X$ and all $\lambda > 0$,
3. $D(X + Y) \leq D(X) + D(Y)$ for all $X$ and $Y$,
4. $D(X) \geq 0$ for all $X$, with $D(X) > 0$ for nonconstant $X$.

$(D2) \& (D3) \Rightarrow$ convexity
Rockafellar, Uryasev and Zabarankin (2006A, 2006B): GDM are introduced as an extension of *standard deviation* but they need not to be symmetric with respect to *upside* $X - \mathbb{E}X$ and *downside* $\mathbb{E}X - X$ of a random variable $X$.

Any functional $D : \mathcal{L}_2(\Omega) \to [0, \infty]$ is called a general deviation measure if it satisfies

1. $D(X + C) = D(X)$ for all $X$ and constants $C$,
2. $D(0) = 0$, and $D(\lambda X) = \lambda D(X)$ for all $X$ and all $\lambda > 0$,
3. $D(X + Y) \leq D(X) + D(Y)$ for all $X$ and $Y$,
4. $D(X) \geq 0$ for all $X$, with $D(X) > 0$ for nonconstant $X$.

(D2) & (D3) $\Rightarrow$ convexity
Deviation measures

- **Standard deviation**

\[
D(X) = \sigma(X) = \sqrt{\mathbb{E} \| X - \mathbb{E}X \|_2}
\]

- **Mean absolute deviation**

\[
D(X) = \mathbb{E}[|X - \mathbb{E}X|].
\]

- **Mean absolute lower and upper semideviation**

\[
D_-(X) = \mathbb{E}\left[ \min(0, X - \mathbb{E}X) \right], \quad D_+(X) = \mathbb{E}\left[ \max(0, X - \mathbb{E}X) \right].
\]

- **Worst-case deviation**

\[
D(X) = \sup_{\omega \in \Omega} |X(\omega) - \mathbb{E}X|.
\]

For any $\alpha \in (0, 1)$ a finite, continuous, lower range dominated deviation measure

$$D_{\alpha}(X) = CVaR_{\alpha}(X - \mathbb{E}X).$$  \hspace{1cm} (1)

The deviation is also called **weighted mean absolute deviation from the** $(1 - \alpha)$-th quantile, see Ogryczak, Ruszczynski (2002), because it can be expressed as

$$D_{\alpha}(X) = \min_{\xi \in \mathbb{R}} \frac{1}{1 - \alpha} \mathbb{E}[\max\{(1 - \alpha)(X - \xi), \alpha(\xi - X)\}]$$ \hspace{1cm} (2)

with the minimum attained at any $(1 - \alpha)$-th quantile. In relation with CVaR minimization formula, see Pflug (2000), Rockafellar and Uryasev (2000, 2002).
According to Proposition 4 in Rockafellar et al (2006 A):

- if $\mathcal{D} = \lambda \mathcal{D}_0$ for $\lambda > 0$ and a deviation measure $\mathcal{D}_0$, then $\mathcal{D}$ is a deviation measure.

- if $\mathcal{D}_1, \ldots, \mathcal{D}_K$ are deviation measures, then
  - $\mathcal{D} = \max\{\mathcal{D}_1, \ldots, \mathcal{D}_K\}$ is also deviation measure.
  - $\mathcal{D} = \lambda_1 \mathcal{D}_1 + \cdots + \lambda_K \mathcal{D}_K$ is also deviation measure, if $\lambda_k > 0$ and $\sum_{k=1}^{K} \lambda_k = 1$.

We say that general deviation measure $\mathcal{D}$ is 

**lower semicontinuous (lsc)** if all the subsets of $\mathcal{L}_2(\Omega)$ having the form $\{X : \mathcal{D}(X) \leq c\}$ for $c \in \mathbb{R}$ (level sets) are closed;
We say that general deviation measure $\mathcal{D}$ is

(LSC) **lower semicontinuous** (lsc) if all the subsets of $\mathcal{L}_2(\Omega)$ having the form $\{X : \mathcal{D}(X) \leq c\}$ for $c \in \mathbb{R}$ (level sets) are closed;

(D5) **lower range dominated** if

$\mathcal{D}(X) \leq EX - \inf_{\omega \in \Omega} X(\omega)$ for all $X$. 

Miloš Kopa

Theory of coherent measures and multiobjective optimization
Theorem 2 in Rockafellar et al (2006 A):

**Theorem**

*Deviation measures correspond one-to-one with strictly expectation bounded risk measures under the relations*

- \( D(X) = R(X - E[X]) \)
- \( R(X) = E[-X] + D(X) \)

*In this correspondence, \( R \) is coherent if and only if \( D \) is lower range dominated.*


Convergence of approximate solutions in mean-risk models

Václav Kozmík

Faculty of Mathematics and Physics
Charles University in Prague

Supervisor: Miloš Kopa
Introduction

■ Mean-risk models
  □ aim to find optimal portfolio of assets
  □ analytical solutions for continuous distributions
  □ solutions using generated scenarios
    ■ predetermined, for instance by historical data
    ■ generated with the assumption of continuous distribution
    ■ generated according to few moment estimators

■ comparison of the approaches mentioned above
  □ convergence and its properties
  □ different continuous distributions
  □ different risk measures

■ computational part
  □ data processing and generating the scenarios
  □ optimization tasks in GAMS
Efficient portfolios

- we consider portfolio based on $N$ assets
- weights of the assets $w_i$, $\sum_{i=1}^{N} w_i = 1$
- expected returns $u_w$ (always using expectation)
- different risk measures $r_w$
- minimal required returns $u_e$

Definition

Portfolio of given $N$ assets with weights $w$ is (mean-risk) efficient, if there are no other weights $w_1^*, \ldots, w_N^*$ such that $\sum_{i=1}^{N} w_i^* = 1$ and $u_{w^*} \geq u_w$ and $r_{w^*} \leq r_w$. 
Classical optimization task

Efficient portfolios can be obtained while solving following task:

\[
\begin{align*}
\min_{\mathbf{w}} & \quad r_{\mathbf{w}} \\
\text{s. t.} & \quad u_{\mathbf{w}} \geq u_e \\
& \quad \sum_{i=1}^{N} w_i = 1 \\
& \quad w_i \in \mathbb{R}, \ i = 1, \ldots, N.
\end{align*}
\]

Non-negativity condition:

\[w_1, \ldots, w_N \geq 0.\]
Risk measures

- variance
- VaR
- cVaR
- absolute deviation
- semivariance

**Definition**

Let \( \alpha \in (0, 1) \) be the threshold and \( L \) random variable which represents the loss from holding the portfolio. Then we define \( \text{VaR}_\alpha \) as:

\[
\text{VaR}_\alpha(L) = \inf \{ l \in \mathbb{R}, P(L > l) \leq 1 - \alpha \}
\]
Risk measures

Definition

cVaR$_{\alpha}$ is defined as:

\[
cVaR_{\alpha}(L) = \inf \left\{ a \in \mathbb{R}, a + \frac{1}{1 - \alpha} \mathbb{E} \left[ \max (0, L - a) \right] \right\}.
\]

Absolute deviation can be calculated as:

\[
r_a(L) = \mathbb{E} |L - \mathbb{E}L|.
\]

Semivariance can be calculated as:

\[
r_s(L) = \mathbb{E} \left[ \max (0, L - \mathbb{E}L)^2 \right]
\]
Elliptical distributions

- generalization of normal distribution
- include normal distribution, Student distribution, logistic elliptical distribution and others
- symmetrical around the mean
- simple analysis of linear combinations

Theorem

Let $X \sim E(\mu, \Sigma, \psi)$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Then it holds:

$$AX + b \sim E(A\mu + b, A\Sigma A^T, \psi).$$
Variance

We get classical optimization task which can be used for all elliptical distributions:

$$\begin{align*}
\min_{\mathbf{w}} & \quad \mathbf{w}^T \mathbf{V} \mathbf{w} \\
\text{s. t.} & \quad \mathbf{w}^T \boldsymbol{\mu} \geq u_e \\
& \quad \sum_{i=1}^{N} w_i = 1 \\
& \quad w_i \in \mathbb{R}, \ i = 1, \ldots, N
\end{align*}$$
Normal distribution

- VaR
  \[ VaR_\alpha(L) = -w^T \mu + q_\alpha \sqrt{w^T V w} \]

- cVaR
  \[ cVaR_\alpha(L) = -w^T \mu + \frac{\exp \left\{ -\frac{q_\alpha^2}{2} \right\}}{(1 - \alpha) \sqrt{2\pi}} \sqrt{w^T V w} \]

- absolute deviation
  \[ r_a(L) = \sqrt{\frac{2}{\pi}} \sqrt{w^T V w} \]

- semivariance
  \[ r_s(L) = \frac{1}{2} \text{E} \left[ (L - \text{E}L)^2 \right] \]
Student distribution

- VaR

\[
\text{VaR}_\alpha(L) = \text{VaR}_\alpha(L) = -w^T \mu + t_{\alpha,\nu} \sqrt{w^T \Sigma w}
\]

- cVaR

\[
c\text{VaR}_\alpha(L) = -w^T \mu + \frac{\Gamma \left(\frac{\nu-1}{2}\right) \sqrt{\nu} \left(1 + \frac{t_{\alpha,\nu}^2}{\nu}\right)^{-\frac{\nu-1}{2}}}{\Gamma \left(\frac{\nu-2}{2}\right) (1 - \alpha)(\nu - 2) \sqrt{\pi}} \sqrt{w^T \Sigma w}
\]

- absolute deviation

\[
r_a(L) = \frac{2\sqrt{\nu} \Gamma \left(\frac{\nu+1}{2}\right)}{(\nu - 1) \sqrt{\pi} \Gamma \left(\frac{\nu}{2}\right)} \sqrt{w^T \Sigma w}
\]

- semivariance

\[
r_s(L) = \frac{1}{2} E \left[ (L - EL)^2 \right]
\]
Variance - scenarios

- suppose we have $M$ scenarios of possible stock prices
- we can use mean and variance-covariance estimators $\hat{\mu}$ and $\hat{\Sigma}$ to minimize variance
- allows us to process estimators before running the optimization task and therefore is quick

$$
\min_w \mathbf{w}^T \hat{\Sigma} \mathbf{w} \\
\text{s. t. } \mathbf{w}^T \hat{\mu} \geq u_e \\
\sum_{i=1}^{N} w_i = 1 \\
w_i \in \mathbb{R}, \ i = 1, \ldots, N.
$$
VaR - scenarios

- general case could be nonconvex
- we reformulate the task using integer programming
- still hardly computable - $M$ binary variables

$$\begin{align*}
\min_{\nu, \mathbf{w}, \delta^j} & \quad \nu \\
\text{s. t.} & \quad -\mathbf{w}^T \mathbf{\psi}^j \leq \nu + K \delta^j, \ j = 1, \ldots, M \\
& \quad \sum_{j=1}^{M} \delta^j = \lfloor (1 - \alpha) M \rfloor \\
& \quad \delta^j \in \{0, 1\}, \ j = 1, \ldots, M \\
& \quad \frac{1}{M} \sum_{j=1}^{M} \mathbf{w}^T \mathbf{\psi}^j \geq u_e \ldots
\end{align*}$$
cVaR - scenarios

- linear programming task, can be solved quickly

\[
\begin{align*}
\min_{a,w,z} & \quad a + \frac{1}{(1 - \alpha) M} \sum_{j=1}^{M} z^j \\
\text{s. t.} \quad & z^j \geq -w^T l^j - a, \ j = 1, \ldots, M \\
& z^j \geq 0, \ j = 1, \ldots, M \\
& \frac{1}{M} \sum_{j=1}^{M} w^T l^j \geq u_e \\
& \sum_{i=1}^{N} w_i = 1 \\
& w_i \in \mathbb{R}, \ i = 1, \ldots, N.
\end{align*}
\]
Absolute deviation - scenarios

- linear programming task

\[
\begin{align*}
\min_{w,z^j} & \quad \frac{1}{M} \sum_{j=1}^{M} z^j \\
\text{s. t.} & \quad w^T l^i - \frac{1}{M} \sum_{i=1}^{M} w^T l^i \leq z^j, \quad j = 1, \ldots, M \\
& \quad - w^T l^j + \frac{1}{M} \sum_{i=1}^{M} w^T l^i \leq z^j, \quad j = 1, \ldots, M \\
& \quad \frac{1}{M} \sum_{j=1}^{M} w^T l^j \geq u_e
\end{align*}
\]
Semivariance - scenarios

- quadratic programming task

\[
\min_{w, z^j} \frac{1}{M} \sum_{j=1}^{M} (z^j)^2 \\
\text{s. t. } z^j \geq -w^T \psi + \frac{1}{M} \sum_{i=1}^{M} w^T l^i, \quad j = 1, \ldots, M \\
z^j \geq 0, \quad j = 1, \ldots, M \\
\frac{1}{M} \sum_{j=1}^{M} w^T \psi \geq u_e \\
\ldots
\]
Computational part

- **own software in C++ & GAMS**
- **software configuration**
  - 50 iterations of generating scenarios, average of the solutions found
  - scenarios are always generated from scratch
  - non-negativity condition
  - ML estimators of distribution parameters
  - threshold 95%
  - different minimal expected returns - small, medium, high
- **maximal number of scenarios**
  - for VaR approx. 1 000 scenarios
  - other risk measures up to 50 000 scenarios
- **data used**
  - stock market indices
  - Japan, USA, Great Britain, Czech Rep. and Germany
  - from 15.9.2008 to 18.9.2009
Convergence issues

- we can experience difficulties with large number of scenarios

- solutions from repeated iterations form clusters
Convergence issues - normal distribution

Semivariance - cluster analysis

Weight of the index in the portfolio

Japan USA Britain Czech R. Germany

Semivariance - cluster analysis

cluster 1

cluster 2
Convergence issues - Student distribution

Semivariance - cluster analysis

Weight of the index in the portfolio

Japan  USA  Britain  Czech R.  Germany

cluster 1
cluster 2
cluster 3
cluster 4
Cluster analysis

- we use the k-means procedure
- given the number of clusters we choose the mean of largest one as the optimal solution
- how to choose the optimal number of clusters
  - we use Bayes Information Criterion (BIC) modified for multidimensional case
  - if the information criterion decreases while adding more clusters we use the last largest cluster which had more than half of the observations included
- needed mostly for elliptical distributions and semivariance
Variance

Distance from the optimal solution vs. Number of scenarios.
cVaR

![Graph showing cVaR vs Number of scenarios for different scenarios and returns.](image-url)
Absolute deviation

![Graph showing absolute deviation vs number of scenarios](image-url)
Semivariance

![Semivariance Graph]

- Normal (small returns)
- Normal (medium returns)
- Normal (high returns)
- Student 7 (small returns)
- Student 7 (medium returns)
- Student 7 (high returns)
- Student 5 (small returns)
- Student 5 (medium returns)
- Student 5 (high returns)

Distance from the optimal solution vs. Number of scenarios
Normal distribution

Normal distribution (medium returns)

- Variance
- CVaR
- Absolute deviation
- Semivariance

Number of scenarios vs. Distance from the optimal solution
Moment generated scenarios

Distance from the optimal solution

Number of scenarios

- variance
- cVaR
- abs. deviation
- semivariance

Number of scenarios
Conclusion

- the convergence could not be achieved without proper analysis of repeated experiments
- using cluster analysis we achieved convergence for all risk measures
- the convergence properties depend on the minimal expected returns
  - even though we have smaller set of feasible solutions the convergence properties could be worse
- moment generated scenarios should be used only for a small sample of initial scenarios
- best convergence properties
  - small differences between distributions
  - variance or cVaR - fast computation
Conclusion

Thank you for your attention!

Miloš Kopa
e-mail:
kopa@karlin.mff.cuni.cz