Introduction to risk-averse optimisation

Pantelis Sopasakis

IMT Institute for Advanced Studies Lucca

February 17, 2016
I. Introductory

✓ Scenario trees
✓ Conditional risk measures
At every node $\alpha$ at stage $k$, we need to take a decision $u^\alpha_k$ in a non-anticipating way.

We assume that the disturbance is not measured at time $k$. 
Scenario trees

With every node $\alpha$ at stage $k < N$ we associate a set of children, $\text{child}(\alpha)$. 
Set $\text{child}(\alpha)$ can be identified with a subset of $\Omega_N$. Recall that this way we may construct a *filtration* on the scenario tree.
Filtration

\[ \mathcal{F}_N = 2^{\Omega_N} \]
Filtration
Filtration

\[ \mathcal{F}_{N-2} \]

\[ N \]
Filtration

$\mathcal{F}_{N-3}$

$N$
\[ \mathcal{F}_0 = \{ \emptyset, \Omega_N \} \]
Conditional probability

For every set $\text{child}(\alpha)$ we define a probability vector of conditional probabilities.

Now $\text{child}(\alpha)$ becomes a probability space...
The siblings probability space

On the probability space \( \text{child}(\alpha) \) we may define random variables

\[
Z : \text{child}(\alpha) \rightarrow \mathbb{IR}
\]
The siblings probability space

These random variables are identified by vectors

\[ Z \in \mathbb{R}^{\text{child}(\alpha)} \]
For every such random variable $Z$ we may define a trivial risk measure as

$$\rho^\alpha(Z) = \mathbb{E}_{p^\alpha}[Z]$$
Of course we can define other conditional risk measures

$$\rho^\alpha : \mathbb{R}^{|\text{child}(\alpha)|} \rightarrow \mathbb{R}.$$
For instance we may use

\[ \rho^\alpha[Z] = \inf_{t \in \mathbb{R}} \{ t + \lambda^{-1}_\alpha \mathbb{E}_p^\alpha [Z - t]^+ \}, \]

with \( \lambda_\alpha \in (0, 1) \).
Conditional risk measures

\[ \rho^\alpha[Z] = \mathbb{E}_{p^\alpha}[Z] + c\mathbb{E}_{p^\alpha}[Z - \mathbb{E}_{p^\alpha}[Z]^p]^{1/p}, \]

with \( c \in [0, 1]. \)
Conditional risk mappings

\[ \Omega_{k+1} = \bigcup_{\alpha \in \Omega_k} \text{child}(\alpha) \]

\[ \alpha_1 \rightarrow \alpha_2 \rightarrow \alpha_3 \]

\[ k \rightarrow k + 1 \]

\[ \rho_{\alpha_1} : \mathbb{R}^3 \rightarrow \mathbb{R} \]

\[ \rho_{\alpha_2} : \mathbb{R}^2 \rightarrow \mathbb{R} \]

\[ \rho_{\alpha_3} : \mathbb{R}^3 \rightarrow \mathbb{R} \]

\[ \rho_{k+1} = (\rho_{\alpha_1}, \rho_{\alpha_2}, \rho_{\alpha_3}) \]

\[ \rho_{k+1} : \mathbb{R}^{\Omega_{k+1}} \rightarrow \mathbb{R}^{\Omega_k} \]
Conditional risk mappings

$Z : \Omega_N \rightarrow \mathbb{R}$, $\mathcal{F}_N$ - meas.
Conditional risk mappings

\[ Z : \Omega_N \to \mathbb{R}, \ F_{N-1} - \text{meas.} \]

\[ \mathbb{R}^{\Omega_{N-1}} \cong \{ Z : \Omega_{N-1} \to \mathbb{R} \} \]

\[ \cong \{ Z : \Omega_N \to \mathbb{R}, F_{N-1} - \text{meas} \} \]
Conditional risk mappings

Define

\[ Z_k := \{ Z : \Omega_N \to \mathbb{R}, \mathcal{F}_k \text{ measurable} \} \]
\[ \cong \{ Z : \Omega_k \to \mathbb{R} \} \]
\[ \cong \mathbb{R}^{\Omega_k} \cong \mathbb{R}^{|\Omega_k|}. \]

Of course

\[ Z_0 \cong \mathbb{R}. \]

Then a conditional risk mapping \( \rho_{k+1} : \Omega_{k+1} \to \mathbb{R}^{\Omega_k} \) can be identified by a mapping

\[ \rho_{k+1} : Z_{k+1} \to Z_k. \]
Conditional risk mappings – Examples

Conditional expectation. \( \rho_{k+1} : \mathcal{Z}_{k+1} \rightarrow \mathcal{Z}_k \) with

\[
\rho_{k+1}[Z](\alpha) = \rho^\alpha[Z] = \mathbb{E}_{\rho^\alpha}[Z].
\]
Conditional risk mappings – Examples

Conditional expectation. \( \rho_{k+1} : \mathcal{Z}_{k+1} \rightarrow \mathcal{Z}_k \) with

\[
\rho_{k+1}[Z](\alpha) = \rho^\alpha[Z] = \mathbb{E}_{p^\alpha}[Z].
\]

Conditional AV@R\(_\alpha[\cdot]\). \( \rho_{k+1} : \mathcal{Z}_{k+1} \rightarrow \mathcal{Z}_k \) with

\[
\rho_{k+1}[Z](\alpha) = \rho^\alpha[Z] = \inf_t \{ t + \lambda_{\alpha}^{-1}\mathbb{E}_{p^\alpha}[Z - t]_+ \}. 
\]
Conditional expectation. \( \rho_{k+1} : \mathcal{Z}_{k+1} \rightarrow \mathcal{Z}_k \) with

\[
\rho_{k+1}[Z](\alpha) = \rho^\alpha[Z] = \mathbb{E}_{p^\alpha}[Z].
\]

Conditional AV@R\(_\alpha\). \( \rho_{k+1} : \mathcal{Z}_{k+1} \rightarrow \mathcal{Z}_k \) with

\[
\rho_{k+1}[Z](\alpha) = \rho^\alpha[Z] = \inf_t \{ t + \lambda^{-1}_\alpha \mathbb{E}_{p^\alpha}[Z - t]_+ \}.
\]

Conditional MUS. \( \rho_{k+1} : \mathcal{Z}_{k+1} \rightarrow \mathcal{Z}_k \) and for \( c_\alpha \geq 0 \)

\[
\rho_{k+1}[Z](\alpha) = \rho^\alpha[Z] = \mathbb{E}_{p^\alpha}[Z] + c_\alpha \mathbb{E}_{p^\alpha}[Z - \mathbb{E}_{p^\alpha}[Z]]_+.
\]
Regularity assumptions

Conditional risk mappings $\rho_{k+1} : \mathcal{Z}_{k+1} \rightarrow \mathcal{Z}_k$ must satisfy the following

1. Convexity: For $Z_1, Z_2 \in \mathcal{Z}_{k+1}$ and $\lambda \in (0, 1)$,

   $$\rho_{k+1}(\lambda Z_1 + (1 - \lambda) Z_2) \leq \lambda \rho_{k+1}(Z_1) + (1 - \lambda) \rho_{k+1}(Z_2)$$
Regularity assumptions

Conditional risk mappings $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ must satisfy the following:

1. **Convexity:** For $Z_1, Z_2 \in \mathcal{Z}_{k+1}$ and $\lambda \in (0, 1)$,
   
   $$\rho_{k+1}(\lambda Z_1 + (1 - \lambda) Z_2) \leq \lambda \rho_{k+1}(Z_1) + (1 - \lambda) \rho_{k+1}(Z_2)$$

2. **Monotonicity:** For $Z_1, Z_2 \in \mathcal{Z}_{k+1}$ with $Z_1 \leq Z_2$,
   
   $$\rho_{k+1}(Z_1) \leq \rho_{k+1}(Z_2)$$
Regularity assumptions

Conditional risk mappings $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ must satisfy the following:

1. Convexity: For $Z_1, Z_2 \in \mathcal{Z}_{k+1}$ and $\lambda \in (0, 1)$,

   $$\rho_{k+1}(\lambda Z_1 + (1 - \lambda) Z_2) \leq \lambda \rho_{k+1}(Z_1) + (1 - \lambda) \rho_{k+1}(Z_2)$$

2. Monotonicity: For $Z_1, Z_2 \in \mathcal{Z}_{k+1}$ with $Z_1 \leq Z_2$,

   $$\rho_{k+1}(Z_1) \leq \rho_{k+1}(Z_2)$$

3. Translation equivariance: For $Z \in \mathcal{Z}_{k+1}$ and $Y \in \mathcal{Z}_k$,

   $$\rho_{k+1}(Z + Y) = Y + \rho_{k+1}(Z)$$
Regularity assumptions

Conditional risk mappings $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ must satisfy the following:

1. **Convexity**: For $Z_1, Z_2 \in \mathcal{Z}_{k+1}$ and $\lambda \in (0, 1)$,
   \[
   \rho_{k+1}(\lambda Z_1 + (1 - \lambda)Z_2) \leq \lambda \rho_{k+1}(Z_1) + (1 - \lambda) \rho_{k+1}(Z_2)
   \]

2. **Monotonicity**: For $Z_1, Z_2 \in \mathcal{Z}_{k+1}$ with $Z_1 \leq Z_2$,
   \[
   \rho_{k+1}(Z_1) \leq \rho_{k+1}(Z_2)
   \]

3. **Translation equivariance**: For $Z \in \mathcal{Z}_{k+1}$ and $Y \in \mathcal{Z}_k$,
   \[
   \rho_{k+1}(Z + Y) = Y + \rho_{k+1}(Z)
   \]

4. **Positive homogeneity**: $\rho_{k+1}(\alpha Z) = \alpha \rho_{k+1}(Z)$ for all $\alpha \geq 0$. 
The mapping $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ is a **conditional risk mapping** if the underlying risk measures $\rho^\alpha$ are **coherent** for all $\alpha \in \Omega_k$. 
Let $\rho_k : \mathcal{Z}_k \to \mathcal{Z}_{k-1}$, for $k \in \mathbb{N}_{[0,N]}$, be coherent risk measures. Verify that $\rho := \rho_1 \circ \rho_2 \circ \cdots \circ \rho_N$, $\rho : \mathcal{Z}_N \to \mathcal{Z}_1$, is a coherent risk measure on $\mathcal{Z}_N$ (recall that $\rho_1 : \mathcal{Z}_1 \to \mathcal{Z}_0 \cong \mathbb{IR}$).
When all \( \rho^\alpha \) at stage \( k \) are coherent, there is a \( \mathcal{A}_{k+1}(\alpha) \) so that

\[
\rho^\alpha[Z] = \max_{\rho^\alpha \in \mathcal{A}_{k+1}(\alpha)} \mathbb{E}_{\rho^\alpha}[Z].
\]

Recall that \( Z \) is a function \( Z : \text{child}(\alpha) \rightarrow \mathbb{R} \), or, equivalently a vector \( Z \in \mathbb{R}^{\lvert \text{child}(\alpha) \rvert} \). We will now extend \( \rho^\alpha \) on \( \mathbb{R}^{\lvert \Omega_{k+1} \rvert} \), i.e.,

\[
\rho^\alpha : \mathbb{R}^{\lvert \Omega_{k+1} \rvert} \rightarrow \mathbb{R}.
\]

Let’s see how this happens...
Extension of $p^\alpha$

\[
\begin{align*}
\alpha_1 & \quad p_1^{\alpha_1} \\ & \quad p_2^{\alpha_1} \\
\alpha_2 & \quad 0 \\ & \quad 0
\end{align*}
\]

\{ \text{child}(\alpha_1) \}
A measure on $\Omega_{k+1}$

Having a probability measure $\nu$ on $\Omega_k$ and knowing the probability measures $p^\alpha$ for each $\alpha \in \Omega_k$ we may define a prob. measure on $\Omega_{k+1}$ as

$$\mu = \sum_{\alpha \in \Omega_k} \nu(\alpha) p^\alpha.$$
Understanding and using $\mu$

For every such measure $\mu$ on $\Omega_{k+1}$ notice that for $Z \in \mathcal{Z}_{k+1}$

$$\mathbb{E}_\mu[Z | \mathcal{F}_k](\alpha) = \mathbb{E}_{p^\alpha}[Z]$$

for $\alpha \in \Omega_k$. Note that

1. In the RHS above, $p^\alpha$ is a prob. vector on $\Omega_{k+1}$
2. Given a measure $\nu$ on $\Omega_k$, we created a measure $\mu$ on $\Omega_{k+1}$
3. $\mathcal{F}_k$ provides all necessary info up to time $k$, therefore
4. $\mathbb{E}_\mu[Z | \mathcal{F}_k]$ is a function of $\alpha$
5. $\mathbb{E}_\mu[Z | \mathcal{F}_k]$ is a random variable
We can now define the set of probability measures on $\Omega_{k+1}$

$$
\mathcal{C}_{k+1} = \left\{ \mu = \sum_{\alpha \in \Omega_k} \nu(\alpha) p^\alpha \Bigm| \nu : \text{prob. measure on } \Omega_k, p^\alpha \in \mathcal{A}_{k+1}(\alpha), \forall \alpha \in \Omega_k \right\}.
$$

Then

$$
\rho_{k+1}[Z] = \max_{\mu \in \mathcal{C}_{k+1}} \mathbb{E}_\mu[Z \mid \mathcal{F}_k]
$$
Interpretations of $\rho_{k+1}$

For all $\alpha \in \Omega_k$, $\rho_{k+1}$ is defined sibling-wise as

$$(\rho_{k+1}[Z])(\alpha) = \max_{p \in \mathcal{A}_{k+1}(\alpha)} \mathbb{E}_p[Z]$$

We also saw that $\rho_{k+1}[Z] = \rho_{k+1}[Z](\nu)$ is

$$\rho_{k+1}[Z] = \max_{\mu} \left\{ \mathbb{E}_\mu[Z] \mid \mathcal{F}_k \right\} \begin{cases} \nu : \text{prob. measure on } \Omega_k \\ \mu = \sum_{\alpha \in \Omega_k} \nu(\alpha)p^\alpha \\ p^\alpha \in \mathcal{A}_{k+1}(\alpha), \forall \alpha \in \Omega_k \end{cases}$$
✓ We revised the notion of a filtration
✓ RVs at stage $k$ can be written as $\mathcal{F}_k$-meas RVs on $\Omega_N$
✓ $Z_k \cong \mathbb{R}^{\Omega_k}$
✓ We introduced the very useful notion of risk mappings
✓ We derived a nice representation of conditional risk mappings
II. Risk-averse optimisation

✓ Problem statement
✓ Dynamic programming equations
Problem statement

Assume the system dynamics is given by

\[ x_{k+1} = f(x_k, u_k, w_k), \]

where \( w_k \) is a random variable. We introduce the risk-averse optimisation problem in nested form

\[
\inf_{u_0 \in \mathcal{U}_0(x_0)} \ell_0(x_0, u_0, \omega) + \rho_1 \left[ \inf_{u_1 \in \mathcal{U}_1(x_1, \omega)} \ell_1(x_1, u_1, \omega) + \ldots \\
+ \rho_{N-1} \left[ \inf_{u_{N-1} \in \mathcal{U}_{N-1}(x_{N-1}, \omega)} \ell_{N-1}(x_{N-1}, u_{N-1}, \omega) + \\
+ \rho_N [\ell_N(x_N, \omega)] \right] \right]
\]
Remark #1

The innermost term of the problem is $\rho_N[\ell_N(x_N, \omega)]$. We require that $\ell_N(x_N, \omega)$ is $\mathcal{F}_N$-measurable, i.e., $\ell_N \in \mathcal{Z}_N$ and $\rho_N$ is a mapping

$$\rho_N : \mathcal{Z}_N \to \mathcal{Z}_{N-1},$$

thus $\rho_N[\ell_N(x_N, \omega)] \in \mathcal{Z}_{N-1}$.
Remark #2

The first term of the optimisation problem is $\ell_0(x_0, u_0, \omega)$. We require that $\ell_0$ is $\mathcal{F}_0$-measurable, i.e., $\ell_0 \in \mathcal{Z}_0$. But, recall that

$$\mathcal{F}_0 = \{\Omega_N, \emptyset\},$$

$$\mathcal{Z}_0 \cong \mathbb{R},$$

therefore $\ell_0$ is deterministic.
Remark #3

Take a look at:

\[
\rho_{N-1}\left[\inf_{u_{N-1} \in \mathcal{U}_{N-1}(x_{N-1},\omega)} \ell_{N-1}(x_{N-1}, u_{N-1}, \omega) + \rho_{N}[\ell_{N}(x_{N}, \omega)]\right]
\]

Here \(x_{N}\) depends on \(u_{N-1}\): \(x_{N} = f(x_{N-1}, u_{N-1}, w_{N-1})\). This way, the above becomes a \(\mathcal{F}_{N-2}\)-measurable RV which depends on \(x_{N-1}\).
Remark #4

Overall, we require that functions $\ell_k(x, \cdot)$ are $\mathcal{F}_k$-measurable and control laws are causal ($\iff \mathcal{F}_k$-measurable).
Alternative formulations

Risk-averse problems can be written in the following ways

✓ Original nested formulation
✓ Dynamic programming
✓ Using the composite risk measure
Dynamic programming

Define

\[ V_0^*(x_N, \omega) := \rho_N[\ell_N(x_N, \omega)] \]

Here, \( \ell_N(x_N, \cdot) \in \mathcal{Z}_N \) and \( V_0^*(x_N, \cdot) \in \mathcal{Z}_{N-1} \). Then

\[ V_k^*(x_{N-k}, \omega) := \inf_{u_{N-k}} \left\{ \ell_{N-k}(x_{N-k}, u_{N-k}, \omega) \right. \]
\[ + \rho_{N-k+1}[V_{k+1}^*(f(x_{N-k}, u_{N-k}, \omega), \omega)] \}

and \( \ell_{N-k}(x, u, \cdot) \in \mathcal{Z}_{N-k} \) & \( V_{k+1}^*(x, \cdot) \in \mathcal{Z}_{N-k+1} \), thus \( V_k^*(x, \cdot) \in \mathcal{Z}_{N-k} \).
Dynamic programming overview

1. At every iteration $k$ of DP, $V_k(x, \cdot)$ is an $\mathcal{F}_{N-k}$-meas. RV
2. $f(x, u, \cdot)$ are $\mathcal{F}_k$-meas
3. $\ell_k(x, u, \cdot)$ are $\mathcal{F}_k$-meas and $\ell_N(x, \cdot)$ is $\mathcal{F}_N$-measurable
4. $\rho_k : \mathcal{Z}_k \to \mathcal{Z}_{k-1}$
5. The last iteration of DP involved $\rho_1 : \mathcal{Z}_1 \to \mathcal{Z}_0 \cong \mathbb{R}$
Dynamic programming

We showed that the DP iteration is defined by

\[ V^*_k(x, \omega) := \inf_u \left\{ \ell_{N-k}(x, u, \omega) + \rho_{N-k+1}[V^*_{k+1}(f(x, u, \omega), \omega)] \right\} \]

\[ = \inf_u \left\{ \ell_{N-k}(x, u, \omega) + \sup_{\mu \in \mathcal{C}_{N-k+1}} \mathbb{E}_\mu[V^*_{k+1}(f(x, u, \omega) \mid \mathcal{F}_{N-k})(\omega)] \right\} \]

and this is because for all \( Z \in \mathcal{Z}_{N-k+1} \)

\[ \rho_{N-k+1}[Z] = \sup_{\mu \in \mathcal{C}_{N-k+1}} \mathbb{E}_\mu[Z \mid \mathcal{F}_{N-k}] \]
Composite risk measure

This representation is the counterpart of the **product space** representation (see stochastic MPC slides).

We shall introduce a risk measure $\bar{\rho}$ so that the problem is written as

$$\min_{u_0, u_1, \ldots, u_{N-1}} \bar{\rho} \left[ \sum_{k=0}^{N-1} \ell_k(x_k, u_k, \omega) + \rho_N(x_N, \omega) \right]$$
Composite risk measure

Since all $\rho_k$ are conditional risk measures it holds that

$$\rho_{N-1}[Z_{N-1} + \rho_N[Z_N]] = \rho_{N-1}[\rho_N[Z_{N-1} + Z_N]]$$
Composite risk measure

Since all $\rho_k$ are conditional risk measures it holds that

$$\rho_{N-1}[Z_{N-1} + \rho_N[Z_N]] = \rho_{N-1}[\rho_N[Z_{N-1} + Z_N]]$$

$\triangleright Z_N \in \mathcal{Z}_N$
Composite risk measure

Since all $\rho_k$ are conditional risk measures it holds that

$$
\rho_{N-1}[Z_{N-1} + \rho_N[Z_N]] = \rho_{N-1}[\rho_N[Z_{N-1} + Z_N]]
$$

- $Z_N \in \mathcal{Z}_N$
- $\rho_N : \mathcal{Z}_N \rightarrow \mathcal{Z}_{N-1}$
Composite risk measure

Since all $\rho_k$ are conditional risk measures it holds that

$$\rho_{N-1}[Z_{N-1} + \rho_N[Z_N]] = \rho_{N-1}[\rho_N[Z_{N-1} + Z_N]]$$

- $Z_N \in \mathcal{Z}_N$
- $\rho_N : \mathcal{Z}_N \rightarrow \mathcal{Z}_{N-1}$
- $\rho_N[Z_N + Y] = Y + \rho_N[Z_N]$ for $Y \in \mathcal{Z}_{N-1}$
Composite risk measure

Then the original nested formulation

$$\inf_{u_0 \in \mathcal{U}_0(x_0)} \ell_0(x_0, u_0, \omega) + \rho_1 \left[ \inf_{u_1 \in \mathcal{U}_1(x_1, \omega)} \ell_1(x_1, u_1, \omega) + \ldots \right.$$ 

$$+ \rho_{N-1} \left[ \inf_{u_{N-1} \in \mathcal{U}_{N-1}(x_{N-1}, \omega)} \ell_{N-1}(x_{N-1}, u_{N-1}, \omega) + \rho_N \left[ \ell_N(x_N, \omega) \right] \right]$$

becomes...
Composite risk measure

...becomes

$$\inf_{u_0, u_1, \ldots, u_{N-1}} (\rho_1 \circ \rho_2 \circ \cdots \circ \rho_{N-1} \circ \rho_N)[\sum_{k=0}^{N-1} \ell_k(x_k, u_k, \omega) + \ell_N(x_N, \omega)],$$

subject to the system dynamics and the standard measurability assumption on $u_k$. 
Composite risk measure

...becomes

$$\inf_{u_0, u_1, \ldots, u_{N-1}} (\rho_1 \circ \rho_2 \circ \cdots \circ \rho_{N-1} \circ \rho_N)[\sum_{k=0}^{N-1} \ell_k(x_k, u_k, \omega) + \ell_N(x_N, \omega)],$$

subject to the system dynamics and the standard measurability assumption on $u_k$.

$\rho$ is a risk measure on $\mathcal{Z}_0 \times \cdots \times \mathcal{Z}_N$, that is

$$\rho(Z_0, Z_1, \ldots, Z_N) = Z_0 + \rho_1[Z_1 + \cdots + \rho_{N-1}[Z_{N-1} + \rho_N[Z_N]]].$$
Composite risk measure

Since

$$\sum_{k=0}^{N-1} \ell_k(x_k, u_k, \omega) + \ell_N(x_N, \omega)$$

as a function of $\omega$ is $\mathcal{F}_N$-measurable, we may naturally identify $\rho$ with the coherent risk measure

$$\bar{\rho}(Z_0 + Z_1 + \ldots + Z_N) = Z_0 + \rho_1[Z_1 + \ldots + \rho_{N-1}[Z_{N-1} + \rho_N[Z_N]]]$$

this is called composite risk measure.
The risk-averse optimisation problem becomes simply

\[
\inf_{u_0, \ldots, u_N} \bar{\rho} \left[ \sum_{k=0}^{N-1} \ell_k(x_k, u_k, \omega) + \ell_N(x_N, \omega) \right],
\]

subject to the system dynamics and measurability assumptions.

Note, however, that usually it is not possible to have a closed form for \( \bar{\rho} \).
Just like their stochastic counterpart, risk-averse problems admit three representations:

- Nested
- Dynamic programming
- Composite risk measure (product space)
- The last one can be cumbersome to work with
- We can also formulate the risk-averse problem using multi-period conditional risk mappings (we’ll discuss that later)
References
