Introduction to risk-averse optimisation

Pantelis Sopasakis

IMT Institute for Advanced Studies Lucca

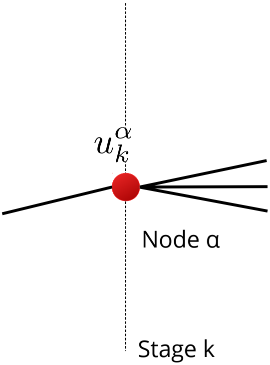
February 17, 2016

Outline

I. Introductory

- ✓ Scenario trees
- ✓ Conditional risk measures

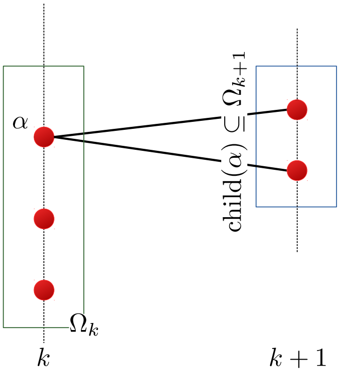
Scenario trees



At every node α at stage k we need to take a decision u_k^α in a non-anticipating way.

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

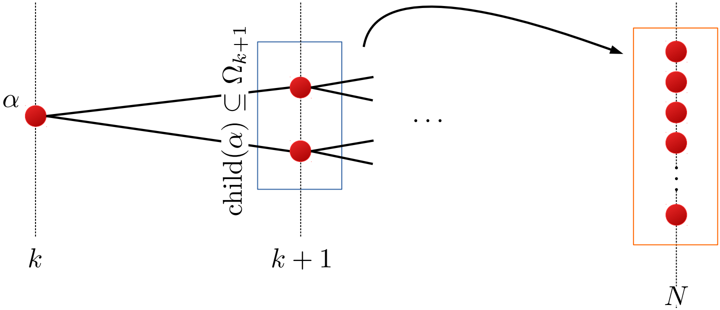
Scenario trees



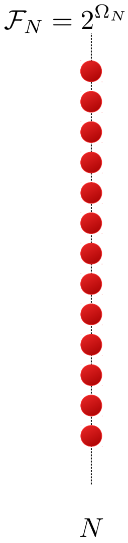
With every node α at stage $k < N$ we associate a set of children, $\mathrm{child}(\alpha)$.

Filtration

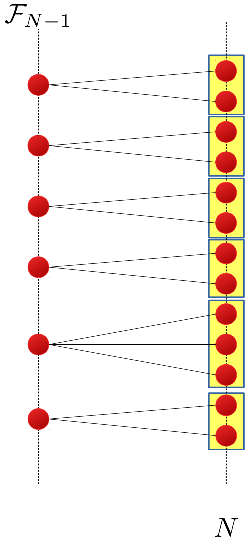
Set $\mathrm{child}(\alpha)$ can be identified with a subset of Ω_N. Recall that this way we may construct a *filtration* on the scenario tree.



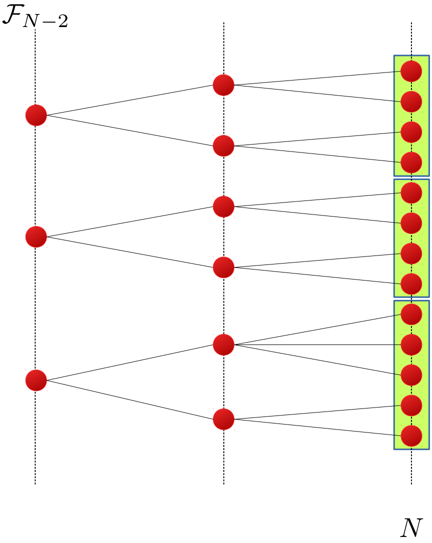
Filtration



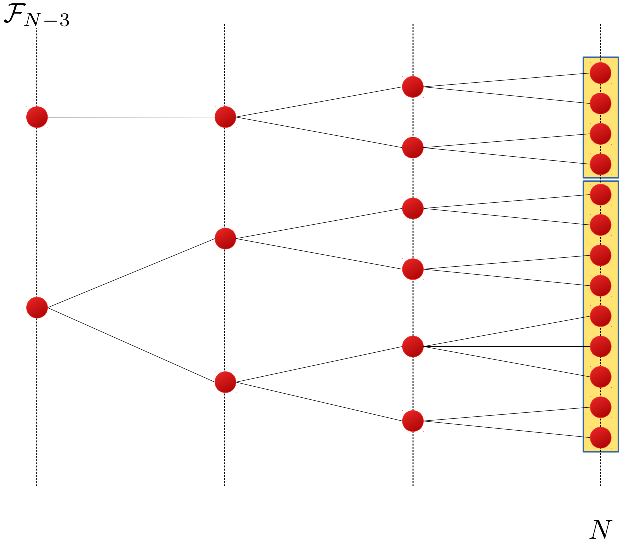
Filtration



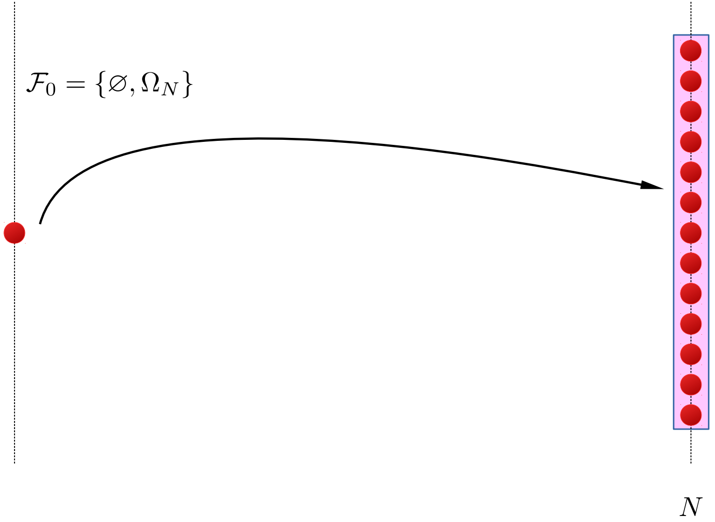
Filtration



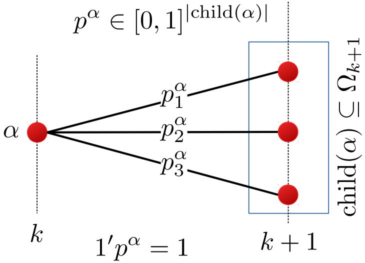
Filtration



Filtration



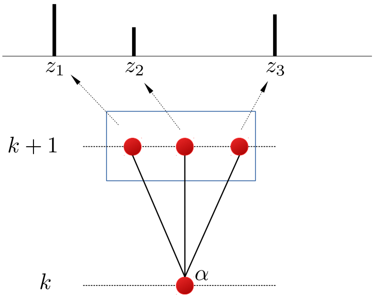
Conditional probability



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

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

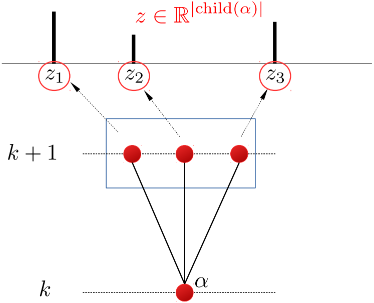
The siblings probability space



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

$$Z : \mathrm{child}(\alpha) \to \mathbb{R}$$

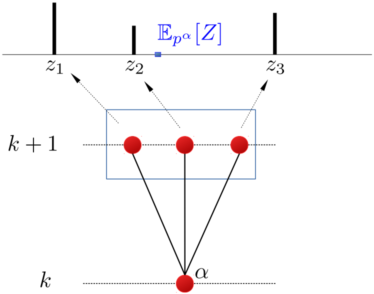
The siblings probability space



These random variables are identified by vectors

$$Z \in \mathbb{R}^{|\,\mathrm{child}(\alpha)|}$$

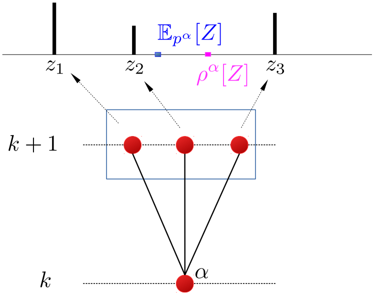
Conditional risk measures



For every such random variable Z we may define a trivial risk measure as

$$\rho^\alpha(Z) = \mathbb{E}_{p^\alpha}[Z]$$

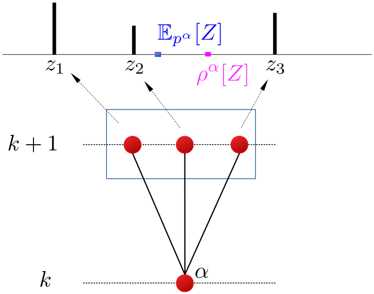
Conditional risk measures



Of course we can define other *conditional* risk measures

$$\rho^\alpha : \mathbb{R}^{|\,\mathrm{child}(\alpha)|} \to \mathbb{R}.$$

Conditional risk measures

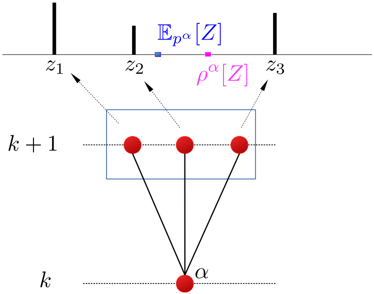


For instance we may use

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

with $\lambda_\alpha \in (0,1)$.

Conditional risk measures

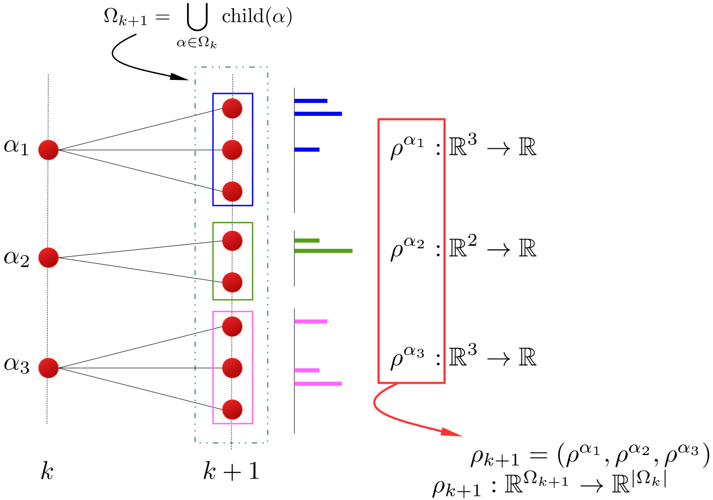


... or

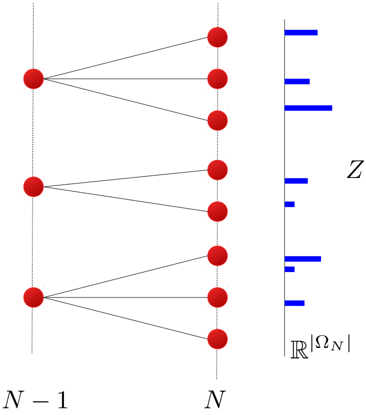
$$\rho^\alpha[Z] = \mathbb{E}_{p^\alpha} + c\mathbb{E}_{p^\alpha}[[Z - \mathbb{E}_{p^\alpha}[Z]]^p]^{1/p},$$

with $c \in [0, 1]$.

Conditional risk mappings

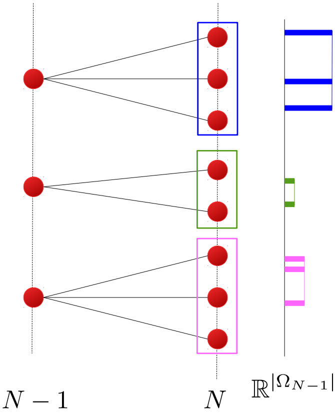


Conditional risk mappings



$$Z : \Omega_N \to \mathbb{R},\ \mathcal{F}_N - \text{meas.}$$

Conditional risk mappings



$$Z : \Omega_N \to \mathbb{R},\ \mathcal{F}_{N-1} - \text{meas.}$$

$$\mathbb{R}^{|\Omega_{N-1}|} \cong \{Z : \Omega_{N-1} \to \mathbb{R}\}$$
$$\cong \{Z : \Omega_N \to \mathbb{R}, \mathcal{F}_{N-1} - \text{meas}\}$$

Conditional risk mappings

Define

$$\begin{aligned}\mathcal{Z}_k &:= \{Z : \Omega_N \to \mathbb{R}, \mathcal{F}_k - \text{measurable}\}\\ &\cong \{Z : \Omega_k \to \mathbb{R}\}\\ &\cong \mathbb{R}^{\Omega_k} \cong \mathbb{R}^{|\Omega_k|}.\end{aligned}$$

Of course

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

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

$$\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k.$$

Conditional risk mappings – Examples

Conditional expectation. $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ with

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

Conditional risk mappings – Examples

Conditional expectation. $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ with

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

Conditional $\mathrm{AV@R}_{\alpha}[\cdot]$**.** $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ with

$$\rho_{k+1}[Z](\alpha) = \rho^{\alpha}[Z] = \inf_{t} \left\{ t + \lambda_{\alpha}^{-1} \mathbb{E}_{p^{\alpha}}[Z - t]_{+} \right\}.$$

Conditional risk mappings – Examples

Conditional expectation. $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ with

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

Conditional $\mathrm{AV@R}_{\alpha}[\cdot]$**.** $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ with

$$\rho_{k+1}[Z](\alpha) = \rho^{\alpha}[Z] = \inf_{t} \left\{ t + \lambda_{\alpha}^{-1} \mathbb{E}_{p^{\alpha}}[Z - t]_{+} \right\}.$$

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

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

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

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 $\alpha \geq 0$.

Regularity assumptions

The mapping $\rho_{k+1} : \mathcal{Z}_{k+1} \to \mathcal{Z}_k$ is a **conditional risk mapping** if the underlying risk measures ρ^α are **coherent** for all $\alpha \in \Omega_k$.

Exercise

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 $\rho_1 : \mathcal{Z}_1 \to \mathcal{Z}_0 \cong \mathbb{R}$).

Representation of conditional risk mappings

When all ρ^α at stage k are coherent, there is a $\mathfrak{A}_{k+1}(\alpha)$ so that

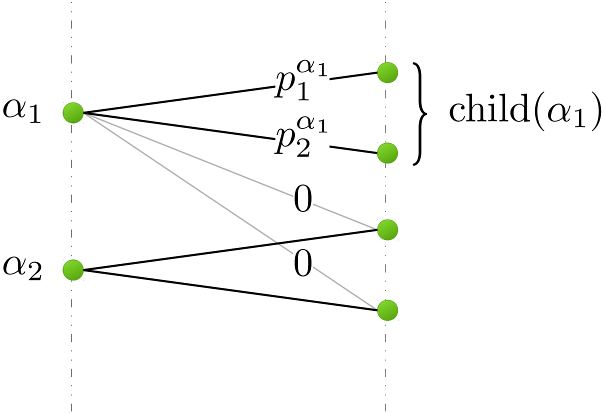
$$\rho^\alpha[Z] = \max_{p^\alpha \in \mathfrak{A}_{k+1}(\alpha)} \mathbb{E}_{p^\alpha}[Z].$$

Recall that Z is a function $Z : \mathrm{child}(\alpha) \to \mathbb{R}$, or, equivalently a vector $Z \in \mathbb{R}^{|\,\mathrm{child}(\alpha)|}$. We will now extend ρ^α on $\mathbb{R}^{|\Omega_{k+1}|}$, i.e.,

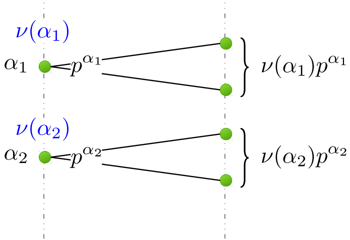
$$\rho^\alpha : \mathbb{R}^{|\Omega_{k+1}|} \to \mathbb{R}.$$

Let's see how this happens...

Extension of p^α



A measure on Ω_{k+1}



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

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

Understanding and using μ

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

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

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

1. In the RHS above, p^α is a prob. vector on Ω_{k+1}
2. Given a measure ν on Ω_k, we created a measure μ on Ω_{k+1}
3. $\mathcal{F}_k$ provides all necessary info up to time k, therefore
4. $\mathbb{E}_\mu[Z \mid \mathcal{F}_k]$ is a function of α
5. $\mathbb{E}_\mu[Z \mid \mathcal{F}_k]$ is a random variable

Representation of conditional risk mapprings

We can now define the **set of probability measures** on Ω_{k+1}

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

Then

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

Interpretations of ρ_{k+1}

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

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

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

$$\rho_{k+1}[Z] = \max_{\mu} \left\{ \mathbb{E}_\mu[Z \mid \mathcal{F}_k] \;\middle|\; \begin{array}{l} \nu : \text{prob. measure on } \Omega_k \\ \mu = \sum_{\alpha \in \Omega_k} \nu(\alpha) p^\alpha \\ p^\alpha \in \mathfrak{A}_{k+1}(\alpha), \forall \alpha \in \Omega_k \end{array} \right\}$$

End of first section

- ✓ We revised the notion of a filtration
- ✓ RVs at stage k can be written as $\mathcal{F}_k$-meas RVs on Ω_N
- ✓ $\mathcal{Z}_k \cong \mathbb{R}^{|\Omega_k|}$
- ✓ 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

$$\begin{aligned}\inf_{u_0 \in \mathcal{U}_0(x_0)} \ell_0(x_0, u_0, \omega) + \rho_1\Big[\inf_{u_1 \in \mathcal{U}_1(x_1, \omega)} \ell_1(x_1, u_1, \omega) + \ldots \\ + \rho_{N-1}\Big[\inf_{u_{N-1} \in \mathcal{U}_{N-1}(x_{N-1}, \omega)} \ell_{N-1}(x_{N-1}, u_{N-1}, \omega) + \\ + \rho_N[\ell_N(x_N, \omega)]\Big]\Big]\end{aligned}$$

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 ρ_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 ℓ_0 is $\mathcal{F}_0$-measurable, i.e., $\ell_0 \in \mathcal{Z}_0$. But, recall that

$$\mathcal{F}_0 = \{\Omega_N, \varnothing\},$$
$$\mathcal{Z}_0 \cong \mathbb{R},$$

therefore ℓ_0 is deterministic.

Remark #3

Take a look at:

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

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** ($\Leftrightarrow \mathcal{F}_k$-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

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

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

$$\begin{aligned} V_k^\star(x_{N-k}, \omega) := \inf_{u_{N-k}} \Big\{ & \ell_{N-k}(x_{N-k}, u_{N-k}, \omega) \\ & + \rho_{N-k+1}[V_{k+1}^\star(f(x_{N-k}, u_{N-k}, \omega), \omega)] \Big\} \end{aligned}$$

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

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

$$\begin{aligned} V_k^\star(x,\omega) &:= \inf_u \left\{\ell_{N-k}(x,u,\omega) + \rho_{N-k+1}[V_{k+1}^\star(f(x,u,\omega),\omega)]\right\} \\ &= \inf_u \left\{\ell_{N-k}(x,u,\omega) + \sup_{\mu\in\mathfrak{C}_{N-k+1}} \mathbb{E}_\mu[V_{k+1}^\star(f(x,u,\omega) \mid \mathcal{F}_{N-k}](\omega)\right\} \end{aligned}$$

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

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

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 ρ_k are conditional risk measures it holds that

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

Composite risk measure

Since all ρ_k are conditional risk measures it holds that

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

- $Z_N \in \mathcal{Z}_N$

Composite risk measure

Since all ρ_k are conditional risk measures it holds that

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

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

Composite risk measure

Since all ρ_k are conditional risk measures it holds that

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

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

Composite risk measure

Then the original nested formulation

$$\inf_{u_0 \in \mathcal{U}_0(x_0)} \ell_0(x_0, u_0, \omega) + \rho_1 \Big[\inf_{u_1 \in \mathcal{U}_1(x_1, \omega)} \ell_1(x_1, u_1, \omega) + \dots$$
$$+ \rho_{N-1} \big[\inf_{u_{N-1} \in \mathcal{U}_{N-1}(x_{N-1}, \omega)} \ell_{N-1}(x_{N-1}, u_{N-1}, \omega) +$$
$$+ \rho_N [\ell_N(x_N, \omega)] \big] \Big]$$

becomes...

Composite risk measure

...becomes

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

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

Composite risk measure

...becomes

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

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

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

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

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 ω is $\mathcal{F}_N$-measurable, we may naturally identify ρ 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**.

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 $\bar{\rho}$.

End of second section

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

1. A. Shapiro, D. Dentcheva, and A.P. Ruszczyński, *Lectures on stochastic programming: modeling and theory*, SIAM 2009.
2. T. Asamov and A. Ruszczyński, *Time-consistent approximations of risk-averse multistage stochastic optimization problems*, Mathematical Programming 153(2), Springer editions, 2014.