Risk-Averse Control of Partially Observable Systems

Andrzej Ruszczyński

Rutgers

Multi-Stage Stochastic Optimization for Clean Energy Transition
Oaxaca, September, 2019

Risk Models with Variable Probability Measures

$\mathcal{X}$ - Polish space with Borel σ-algebra $\mathcal{B}(\mathcal{X})$
$\mathcal{P}(\mathcal{X})$ - the set of probability measures on $\big(\mathcal{X}, \mathcal{B}(\mathcal{X})\big)$.
$\mathbb{B}(\mathcal{X})$ - the space of all real-valued bounded measurable functions on $\mathcal{X}$.

Probabilistic Model: a pair $[Z, P] \in \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$.

A measurable functional $\rho : \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ is called a *risk form*

(i) It is monotonic, if $V \le W$ implies $\rho[V, P] \le \rho[W, P]$ for all $P \in \mathcal{P}(\mathcal{X})$;

(ii) It is normalized if $\rho[0, P] = 0$ for all $P \in \mathcal{P}(\mathcal{X})$;

(iii) It is translation equivariant if $\rho[a\mathbb{1} + V, P] = a + \rho[V, P]$ for all $a \in \mathbb{R}$;

(iv) It is positively homogeneous, if $\rho[\beta V, P] = \beta\rho[V, P]$ for all $\beta \in \mathbb{R}_+$;

(v) It has the support property, if $\rho\big[\mathbb{1}_{\mathrm{supp}(P)} V, P\big] = \rho[V, P]$.

Examples

Bilinear form (the expected value)

$$\mathbb{E}[Z,P] = \int_{\mathcal{X}} Z(x)\; P(dx) \quad = \mathbb{E}_P[Z]$$

Mean-semideviation

$$\rho[Z,P] = \mathbb{E}_P[Z] + \kappa\Big[\mathbb{E}_P\big[\big(\max(0, Z - \mathbb{E}_P[Z])\big)^p\big]\Big]^{\frac{1}{p}}$$

where $p \geq 1$, $\kappa \in [0,1]$.

Inverse risk measure

$$\rho[Z,P] = \min_{\eta\in\mathbb{R}} \Big\{\eta + \kappa\Big[\mathbb{E}_P\big[\big(\max(0, Z - \eta)\big)^p\big]\Big]^{\frac{1}{p}}\Big\}$$

where $p \geq 1$, $\kappa > 1$.

All law invariant risk measures may be cast as risk forms

Preliminaries 1

State Consistency

If a risk form $\rho[Z, P]$ has the the normalization, translation equivariance, and support properties then for every $Z \in \mathbb{B}(\mathcal{X})$ and every $x \in \mathcal{X}$

$$\rho\big[Z, \delta_x\big] = Z(x)$$

A probabilistic model $[Z, P]$ is smaller than a probabilistic model $[Z', P']$ in the increasing convex order, written $[Z, P] \preceq [Z', P']$, if for all $\eta \in \mathbb{R}$

$$\int_{\mathcal{X}} \big[Z(x) - \eta\big]_+ \, P(dx) \leq \int_{\mathcal{X}} \big[Z'(x) - \eta\big]_+ \, P'(dx).$$

A risk form $\rho[Z, P]$ is consistent with the increasing convex order, if

$$[Z, P] \preceq [Z', P'] \implies \rho[Z, P] \leq \rho[Z', P'].$$

Preliminaries 2

A risk form $\rho[Z, P]$ is comonotonically convex, if for all comonotonic functions $Z, V \in \mathcal{B}(\mathcal{X})$, all $P \in \mathcal{P}(\mathcal{X})$, and all $\lambda \in [0, 1]$,

$$\rho[\lambda Z + (1 - \lambda) V, P] \leq \lambda \rho[Z, P] + (1 - \lambda) \rho[V, P].$$

With every stochastic model $[Z, P]$ we associate its distribution function,

$$F[Z, P](t) = P[Z \leq t], \quad t \in \mathbb{R},$$

and its quantile function

$$\Phi[Z, P](p) = \inf \left\{ \eta : P[Z \leq \eta] \geq p \right\}, \quad p \in (0, 1].$$

Duality

$\mathcal{M}$ – the set of countably additive finite measures on $(0,1]$

The conjugate functional $\rho^* : \mathcal{M} \to \mathbb{R} \cup \{+\infty\}$

$$\rho^*(\mu) = \sup_{[Z,P]\in\mathbb{B}(\mathcal{X})\times\mathcal{P}(\mathcal{X})} \left\{ \int_0^1 \Phi[Z,P](p)\ \mu(dp) - \rho[Z,P] \right\}.$$

Suppose $\mathcal{X}$ is uncountable. If a risk form $\rho : \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ is normalized, translation equivariant, comonotonically convex, and consistent with the increasing convex order, then a uniquely defined closed convex set

$$\mathcal{D}_\rho \subseteq \big\{\mu \in \mathcal{M} : \mu(0,\cdot] \textit{ is nondecreasing and convex on } (0,1],\ \ \mu(0,1] = 1\big\}$$

exists, such that for all $[Z,P] \in \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$

$$\rho[Z,P] = \sup_{\mu\in\mathcal{D}_\rho} \left\{ \int_0^1 \Phi[Z,P](p)\ \mu(dp) - \rho^*(\mu) \right\}$$

If the risk form is positively homogeneous, then $\rho^*(\mu) \equiv 0$.

Kusuoka Representation

The Average Value at Risk at level $\alpha \in [0,1]$ of a probabilistic model $[Z,P]$:

$$\mathrm{AVaR}_\alpha[Z,P] = \begin{cases} \frac{1}{\alpha}\int_{1-\alpha}^{1} \varPhi[Z,P](p)\; dp & \text{if } \alpha \in (0,1), \\ \varPhi[Z,P](1) & \text{if } \alpha = 0 \\ \mathbb{E}[Z,P] & \text{if } \alpha = 1 \end{cases}$$

$$= \inf_\eta \left\{ \eta + \frac{1}{\alpha}\mathbb{E}_P\big[(Z-\eta)_+\big] \right\} \qquad (\text{for } \alpha > 0)$$

Suppose the conditions of the Duality Theorem are satisfied and the risk form $\rho[\cdot,\cdot]$ is positively homogeneous. Then a convex subset $\varLambda_\rho$ of the set of probability measures on $[0,1]$ exists, such that for all $[Z,P]$

$$\rho[Z,P] = \sup_{\lambda\in\varLambda_\rho} \int_0^1 \mathrm{AVaR}_s[Z,P]\; \lambda(ds).$$

Conditional Risk Operator

Two Polish spaces $\mathcal{X}$ and $\mathcal{Y}$ and their Borel σ-algebras $\mathcal{B}(\mathcal{X})$ and $\mathcal{B}(\mathcal{Y})$

Every $P \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ can be disintegrated into its marginal $P_{\mathcal{X}} \in \mathcal{P}(\mathcal{X})$ and a transition kernel $P_{\mathcal{Y}|\mathcal{X}} : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ as follows: $P(dx, dy) = P_{\mathcal{X}}(dx)\, P_{\mathcal{Y}|\mathcal{X}}(dy|x)$.

Let $\mathcal{Q}(\mathcal{Y}|\mathcal{X})$ be the space of all kernels $Q : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$. For any $\lambda \in \mathcal{P}(\mathcal{X})$ and and any $Q \in \mathcal{Q}(\mathcal{Y}|\mathcal{X})$, the composition $P = \lambda \otimes Q$ defined as $P(dx, dy) = \lambda(dx) Q(dy|x)$ is an element of $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$

Suppose the risk form $\rho : \mathbb{B}(\mathcal{X} \times \mathcal{Y}) \times \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}$ is monotonic, translation equivariant, and normalized. Then it induces a conditional risk operator $\rho_{\mathcal{Y}|\mathcal{X}} : \mathbb{B}(\mathcal{X} \times \mathcal{Y}) \times \mathcal{Q}(\mathcal{Y}|\mathcal{X}) \to \mathbb{B}(\mathcal{X})$ defined as follows:

$$\rho_{\mathcal{Y}|\mathcal{X}}[Z, Q](x) = \rho[Z, \delta_x \otimes Q], \quad x \in \mathcal{X}$$

Conditional Risk Forms

Conditional risk operator

$$\rho_{\mathcal{Y}|\mathcal{X}}[Z, Q](x) = \rho[Z, \delta_x \otimes Q], \quad x \in \mathcal{X}$$

If the risk form ρ has the support property, we can define the conditional risk forms $\rho_{\mathcal{Y}|x} : \mathbb{B}(\mathcal{Y}) \times \mathcal{P}(\mathcal{Y}) \to \mathbb{R}$, $x \in \mathcal{X}$, as follows:

$$\rho_{\mathcal{Y}|x}[Z(x, \cdot), Q(x)] = \rho_{\mathcal{Y}|\mathcal{X}}[Z, Q](x), \quad x \in \mathcal{X}.$$

If the risk form $\rho[\cdot, \cdot]$ is monotonic (normalized, translation equivariant), then, for every $x \in \mathcal{X}$, the conditional risk form $\rho_{\mathcal{Y}|x}$ is monotonic (normalized, translation equivariant).

Conditional Consistency and Risk Disintegration

A risk form $\rho : \mathbb{B}(\mathcal{X} \times \mathcal{Y}) \times \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}$ is conditionally consistent if for all $Z, Z' \in \mathcal{B}(\mathcal{X} \times \mathcal{Y})$ and all $Q, Q' \in \mathcal{Q}(\mathcal{Y}|\mathcal{X})$ the inequality

$$\rho_{\mathcal{Y}|\mathcal{X}}[Z, Q] \leq \rho_{\mathcal{Y}|\mathcal{X}}[Z', Q']$$

implies that $\quad \rho[Z, \lambda \otimes Q] \leq \rho[Z', \lambda \otimes Q'], \ \forall \, \lambda \in \mathcal{P}(\mathcal{X}).$

Marginal Risk Form

Suppose $\rho : \mathbb{B}(\mathcal{X} \times \mathcal{Y}) \times \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}$ is monotonic, normalized, translation equivariant, has the support property, and is conditionally consistent. Then a marginal risk form $\rho_{\mathcal{X}} : \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ exists, such that for all $[Z, P] \in \mathbb{B}(\mathcal{X} \times \mathcal{Y}) \times \mathcal{P}(\mathcal{X} \times \mathcal{Y})$:

$$\rho[Z, P] = \rho_{\mathcal{X}}\big[\rho_{\mathcal{Y}|\mathcal{X}}[Z, P_{\mathcal{Y}|\mathcal{X}}], P_{\mathcal{X}}\big]$$

It is monotonic, normalized, translation equivariant, and has the support property.

Controlled Two-Stage System. Functional Formulation

Control Spaces - $\mathcal{U}_1$ (stage 1) and $\mathcal{U}_2$ (stage 2).
Random Data - X observed after first stage, Y - never observed.

After choosing $u_1 \in \mathcal{U}_1$, observation of X is made, and we choose control $u_2 \in U_2(X, u_1) \subset \mathcal{U}_2$ to minimize the risk of $c(X, Y, u_1, u_2)$.
The risk is measured by the form $\rho[\cdot, \cdot]$.

Functional Perspective

We represent u_2 it as a decision rule : $u_2 = \pi(x)$, $x \in \mathcal{X}$. The overall cost is:

$$Z^{u_1,\pi}(x, y) = c(x, y, u_1, \pi(x)), \quad (x, y) \in \mathcal{X} \times \mathcal{Y}.$$

The problem takes on the form

$$\begin{aligned} &\min_{u_1,\pi} \rho\big[Z^{u_1,\pi}, P\big] \\ &\text{s.t. } u_1 \in U_1, \\ &\qquad \pi(\cdot) \lessdot U_2(\cdot, u_1) \quad (\pi \text{ is a selection of } U_2) \end{aligned}$$

Controlled Two-Stage System. Hierarchical Formulation

Let the following assumptions be satisfied:

(i) The risk form ρ is monotonic, normalized, translation equivariant, has the support property, and is conditionally consistent;

(ii) The multifunction U_2 is upper-semicontinuous and has nonempty and compact values;

(iii) The function c is uniformly bounded, measurable, and lower-semicontinuous with respect to its second argument.

Then the functional problem is equivalent to the two-stage problem:

$$\min_{u_1 \in U_1} \rho_{\mathcal{X}}\big[V(\cdot, u_1), P_{\mathcal{X}}\big],$$

where $V(\cdot, \cdot)$ is the optimal value of the second stage problem:

$$V(x, u_1) = \min_{u_2 \in U_2(x,u_1)} \rho_{\mathcal{Y}|x}\big[c(x, \cdot, u_1, u_2), P_{\mathcal{Y}|\mathcal{X}}(x)\big], \quad x \in \mathcal{X}, \quad u_1 \in U_1.$$

Controlled Observation Distribution

After a control $u_1 \in U_1 \subset \mathcal{U}_1$ is chosen, the distribution of the observation X depends on Y and u_1 via a controlled kernel $K : \mathcal{Y} \times \mathcal{U}_1 \to \mathcal{P}(\mathcal{X})$.

Let P_Y be the marginal distribution of Y. After the first decision u_1 will be chosen, the joint distribution of (Y, X) will become

$$M(u_1) = P_{\mathcal{Y}} \otimes K(\cdot, u_1),$$

that is, $M(dy, dx|u_1) = P_{\mathcal{Y}}(dy)K(dx|y, u_1)$. Therefore, denoting the second stage decision by $u_2 = \pi(x)$, our problem is

$$\begin{aligned} \min_{u_1, \pi} \ & \rho\big[Z^{u_1,\pi}, M(u_1)\big], \\ \text{s.t. } & u_1 \in U_1, \\ & \pi(\cdot) \lessdot U_2(\cdot, u_1). \end{aligned}$$

Two-Stage Formulation

Marginal distribution of the observation: $M_{\mathcal{X}}(u_1) = \int_{\mathcal{Y}} K(y, u_1)\; P_{\mathcal{Y}}(dy)$

Disintegration: $M(u_1) = M_{\mathcal{X}}(u_1) \otimes \Gamma(u_1)$

The transition kernel Γ is the Bayes operator.

Under the same assumptions as in the uncontrolled observation case, the problem is equivalent to the two-stage problem:

$$\min_{u_1 \in U_1} \rho_{\mathcal{X}}\big[V(\cdot, u_1), M_{\mathcal{X}}(u_1)\big],$$

where $V(\cdot,\cdot)$ is the optimal value of the second stage problem:

$$V(x, u_1) = \min_{u_2 \in U_2(x,u_1)} \rho_{\mathcal{Y}|x}\big[c(x, \cdot, u_1, u_2), \Gamma(x, u_1)\big], \quad x \in \mathcal{X}, \quad u_1 \in U_1.$$

Partially Observable Discrete-Time Models

- Markov Process: $\{X_t, Y_t\}_{t=1,\dots,T}$ on the Borel state space $\mathcal{X} \times \mathcal{Y}$
- The process $\{X_t\}$ is observable, while $\{Y_t\}$ is not observable
- Control sets: $U_t : \mathcal{X} \rightrightarrows \mathcal{U}$, $t = 1, \dots, T$
- Transition kernel: $\mathbb{P}[(X_{t+1}, Y_{t+1}) \in C \mid x_t, y_t, u_t] = Q_t(x_t, y_t, u_t)(C)$
- Costs: $Z_t = c_t(X_t, Y_t, U_t)$, $t = 1, \dots, T$

Two relevant filtrations

- $\{\mathcal{F}_t^{X,Y}\}$ defined by the full state process $\{X_t, Y_t\}$
- $\{\mathcal{F}_t^{X}\}$ defined by the observed process $\{X_t\}$

Space of costs: $\mathcal{Z}_t = \left\{ Z : \Omega \to \mathbb{R} \,\middle|\, Z \text{ is } \mathcal{F}_t^{X,Y}\text{-measurable and bounded} \right\}$

Classical Problem:

$$\min \ \mathbb{E}\left\{c_1(X_1, Y_1, U_1) + c_2(X_2, Y_2, U_2) + \cdots + c_T(X_T, Y_T, U_T)\right\}$$

Partially Observable Discrete-Time Models

- Markov Process: $\{X_t, Y_t\}_{t=1,\dots,T}$ on the Borel state space $\mathcal{X} \times \mathcal{Y}$
- The process $\{X_t\}$ is observable, while $\{Y_t\}$ is not observable
- Control sets: $U_t : \mathcal{X} \rightrightarrows \mathcal{U}$, $t = 1, \dots, T$
- Transition kernel: $\mathbb{P}[(X_{t+1}, Y_{t+1}) \in C \mid x_t, y_t, u_t] = Q_t(x_t, y_t, u_t)(C)$
- Costs: $Z_t = c_t(X_t, Y_t, U_t)$, $t = 1, \dots, T$

Two relevant filtrations

- $\{\mathcal{F}_t^{X,Y}\}$ defined by the full state process $\{X_t, Y_t\}$
- $\{\mathcal{F}_t^{X}\}$ defined by the observed process $\{X_t\}$

Space of costs: $\mathcal{Z}_t = \Big\{ Z : \Omega \to \mathbb{R} \;\Big|\; Z \text{ is } \mathcal{F}_t^{X,Y}\text{-measurable and bounded} \Big\}$

Risk-Averse Problem:

$$\min \; \rho_{1,T}\big\{c_1(X_1, Y_1, U_1), c_2(X_2, Y_2, U_2), \dots, c_T(X_T, Y_T, U_T)\big\}$$

Dynamic Risk Measures

Probability space $(\Omega, \mathcal{F}, P)$ with filtration $\mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T \subset \mathcal{F}$
Adapted sequence of random variables (costs) $Z_1, Z_2, \ldots, Z_T$
Spaces: $\mathcal{Z}_t$ of $\mathcal{F}_t$-measurable functions and $\mathcal{Z}_{t,T} = \mathcal{Z}_t \times \cdots \times \mathcal{Z}_T$

Dynamic Risk Measure

A sequence of conditional risk measures $\rho_{t,T} : \mathcal{Z}_{t,T} \to \mathcal{Z}_t$, $t = 1, \ldots, T$.
Monotonicity condition:

$$\rho_{t,T}(Z) \leq \rho_{t,T}(W) \text{ for all } Z, W \in \mathcal{Z}_{t,T} \text{ such that } Z \leq W$$

Local property: For all $A \in \mathcal{F}_t$

$$\rho_{t,T}(\mathbb{1}_A Z) = \mathbb{1}_A \rho_{t,T}(Z)$$

Conditional Risk Evaluators

Space of observable random variables:

$$\mathcal{S}_t = \Big\{ S : \Omega \to \mathbb{R} \,\Big|\, S \text{ is } \mathcal{F}_t^X\text{-measurable and bounded} \Big\}, \quad t = 1, \dots, T$$

A mapping $\rho_{t,T} : \mathcal{Z}_t \times \cdots \times \mathcal{Z}_T \to \mathcal{S}_t$ is a conditional risk evaluator

(i) It is monotonic if $Z_s \le W_s$ for all $s = t, \dots, T$, implies that

$$\rho_{t,T}(Z_t, \dots, Z_T) \le \rho_{t,T}(W_t, \dots, W_T)$$

(ii) It is normalized if $\rho_{t,T}(0, \dots, 0) = 0$;

(iii) It is translation equivariant if $\forall (Z_t, \dots, Z_T) \in \mathcal{S}_t \times \mathcal{Z}_{t+1} \times \cdots \times \mathcal{Z}_T$,
$\rho_{t,T}(Z_t, \dots, Z_T) = Z_t + \rho_{t,T}(0, Z_{t+1}, \dots, Z_T)$;

(iv) It is decomposable if a mapping $\rho_t : \mathcal{Z}_t \to \mathcal{S}_t$ exists such that:

$$\rho_t(Z_t) = Z_t, \quad \forall Z_t \in \mathcal{S}_t,$$
$$\rho_{t,T}(Z_t, \dots, Z_T) = \rho_t(Z_t) + \rho_{t,T}(0, Z_{t+1}, \dots, Z_T), \quad \forall\, Z \in \mathcal{Z}_{t,T}$$

Risk Filters and their Time Consistency

A risk filter $\{\rho_{t,T}\}_{t=1,\dots,T}$ is a sequence of conditional risk evaluators $\rho_{t,T} : \mathcal{Z}_{t,T} \to \mathcal{S}_t$.

We have index risk filters by policy π, because π affects the measure P^π

History: $H_t^\pi = (X_1, X_2^\pi, \dots, X_t^\pi)$, $h_t = (x_1, x_2, \dots, x_t)$

A family of risk filters $\{\rho_{t,T}^\pi\}_{t=1,\dots,T}^{\pi\in\Pi}$ is stochastically conditionally time consistent if for any $\pi, \pi' \in \Pi$, for any $1 \le t < T$, for all $h_t \in \mathcal{X}^t$, all $(Z_t, \dots, Z_T) \in \mathcal{Z}_{t,T}$ and all $(W_t, \dots, W_T) \in \mathcal{Z}_{t,T}$, the conditions

$$Z_t = W_t$$

$$\left(\rho_{t+1,T}^\pi(Z_{t+1}, \dots, Z_T) \mid H_t^\pi = h_t\right) \preceq_{\text{st}} \left(\rho_{t+1,T}^{\pi'}(W_{t+1}, \dots, W_T) \mid H_t^{\pi'} = h_t\right)$$

imply

$$\rho_{t,T}^\pi(Z_t, Z_{t+1}, \dots, Z_T)(h_t) \le \rho_{t,T}^{\pi'}(W_t, W_{t+1}, \dots, W_T)(h_t)$$

The relation $\preceq_{\text{st}}$ is the conditional stochastic order

Bayes Operator

Belief State: Conditional distribution of Y_t given initial distribution ξ_1 and history $g_t = (\xi_1, x_1, u_1, x_2, \dots, u_{t-1}, x_t)$

$$\big[\varXi_t(g_t)\big](A) = \mathbb{P}[Y_t \in A \mid g_t], \quad \forall\, A \in \mathcal{B}(\mathcal{Y}), \quad t = 1, \dots, T$$

Conditional distribution of the observable part:

$$\mathbb{P}\left[X_{t+1} \in B \mid g_t, u_t\right] = \int_{\mathcal{Y}} \big[Q_t^X(x_t, \cdot, u_t)\big](B)\; d\varXi_t(g_t),$$

where $Q_t^X(x_t, y_t, u_t)$ is the marginal of $Q_t(x_t, y_t, u_t)$ on the space $\mathcal{X}$

Transition of the belief state - Bayes operator

$$\varXi_{t+1}(g_{t+1}) = \varGamma_t(x_t, \varXi_t(g_t), u_t, x_{t+1})$$

Example: $\mathcal{Y} = \{y^1, \dots, y^n\}$ and $Q_t(x, y, u)$ has density $q_t(x', y' | x, y, u)$

$$\big[\varGamma_t(x, \xi, u, x')\big](\{y^k\}) = \frac{\sum_{i=1}^n q_t(x', y^k \mid x, y^i, u)\, \xi^i}{\sum_{\ell=1}^n \sum_{i=1}^n q_t(x', y^\ell \mid x, y^i, u)\, \xi^i}$$

Markov Risk Filters

Policies $\pi = (\pi_1, \ldots, \pi_T)$ with decision rules $\pi_t(h_t) \in U_t(x_t)$

Markov Policy

For all $h_t, h'_t \in \mathcal{X}^t$, if $x_t = x'_t$ and $\xi_t = \xi'_t$, then
$\pi_t(h_t) = \pi_t(h'_t) = \pi_t(x_t, \xi_t)$

Policy value function:

$$v_t^\pi(h_t) = \rho_{t,T}^\pi\big(c_t(X_t, Y_t, \pi_t(H_t)), \ldots, c_T(X_T, Y_T, \pi_T(H_T))\big)(h_t)$$

A family of risk filters $\big\{\rho_{t,T}^\pi\big\}_{t=1,\ldots,T}^{\pi \in \Pi}$ is Markov if for all Markov policies $\pi \in \Pi$, for all $h_t = (x_1, \ldots, x_t)$ and $h'_t = (x'_1, \ldots, x'_t)$ in $\mathcal{X}^t$ such that $x_t = x'_t$ and $\xi_t = \xi'_t$, we have

$$v_t^\pi(h_t) = v_t^\pi(h'_t) = v_t^\pi(x_t, \xi_t)$$

Notation: $\rho_t\big(c_t(X_t, Y_t, u_t) = r_t(X_t, \xi_t, u_t)$

Structure of Markov Risk Filters

A family of risk filters $\{\rho^{\pi}_{t,T}\}^{\pi\in\Pi}_{t=1,\dots,T}$ is normalized, translation-invariant, stochastically conditionally time consistent, decomposable, and Markov if and only if transition risk mappings exist:

$$\sigma_t : \left\{\left(x_t, \xi_t, Q^{\pi}_t(h_t)\right) : \pi \in \Pi,\ h_t \in \mathcal{X}^t\right\} \times \mathcal{V} \to \mathbb{R}, \quad t = 1 \dots T-1,$$

(i) $\sigma_t(x, \xi, \cdot, \cdot)$ is normalized and strongly monotonic with respect to stochastic dominance

(ii) for all $\pi \in \Pi$, for all $t = 1, \dots, T-1$, and for all $h_t \in \mathcal{X}^t$,

$$v^{\pi}_t(h_t) = r_t(x_t, \xi_t, \pi_t(h_t)) + \sigma_t\left(x_t, \xi_t, Q^{\pi}_t(h_t), v^{\pi}_{t+1}(h_t, \cdot)\right)$$

Evaluation of a Markov policy π:

$$v^{\pi}_t(x_t, \xi_t) = r_t(x_t, \xi_t, \pi_t(x_t, \xi_t)) + \sigma_t\Big(x_t, \xi_t, Q^{\pi}_t(x_t, \xi_t), x' \mapsto v^{\pi}_{t+1}(x', \overbrace{\Gamma_t(x_t, \xi_t, \pi_t(x_t, \xi_t), x')}^{\xi'})\Big)$$

Examples of Transition Risk Mappings

Average Value at Risk

$$\sigma(x,\xi,m,v) = \min_{\eta\in\mathbb{R}} \left\{ \eta + \frac{1}{\alpha(x,\xi)} \int_{\mathcal{X}} \big(v(x') - \eta\big)_+ \; m(dx') \right\}$$

where $\alpha(x,\xi) \in [\alpha_{\min}, \alpha_{\max}] \subset (0,1]$.

Mean–Semideviation of Order p

$$\sigma(x,\xi,m,v) = \underbrace{\int_{\mathcal{X}} v(x') \; m(dx')}_{\mathbb{E}_m[v]} + \kappa(x,\xi) \left(\int_{\mathcal{X}} \big(v(x') - \mathbb{E}_m[v]\big)_+^p \; m(dx') \right)^{\frac{1}{p}}$$

where $\kappa(x,\xi) \in [0,1]$.

Entropic Mapping

$$\sigma(x,\xi,m,v) = \frac{1}{\gamma(x,\xi)} \ln\left(\mathbb{E}_m\left[e^{\gamma(x,\xi)\, v(x')} \right] \right), \quad \gamma(x,\xi) > 0$$

Dynamic Programming

Risk-averse optimal control problem:

$$\min_{\pi} \; \rho^{\pi}_{1,T}\big\{c_1(X_1, Y_1, U_1), c_2(X_2, Y_2, U_2), \dots, c_T(X_T, Y_T, U_T)\big\}$$

Theorem

If the risk measure is Markovian (+ general conditions), then the optimal solution is given by the dynamic programming equations:

$$v^*_T(x,\xi) = \min_{u \in \mathcal{U}_T(x)} r_T(x,\xi,u), \quad x \in \mathcal{X}, \quad \xi \in \mathcal{P}(\mathcal{X})$$

$$v^*_t(x,\xi) = \min_{u \in \mathcal{U}_t(x)} \bigg\{ r_t(x,\xi,u) + \\ \sigma_t\Big(x, \xi, \int_{\mathcal{Y}} K^X_t(x,y,u)\,\xi(dy), x' \mapsto v^*_{t+1}\big(x', \Gamma_t(x,\xi,u,x')\big)\Big)\bigg\}, \\ x \in \mathcal{X}, \quad \xi \in \mathcal{P}(\mathcal{Y}), \quad t = T-1, \dots, 1$$

Optimal Markov policy $\hat{\Pi} = \{\hat{\pi}_1, \dots, \hat{\pi}_T\}$ - the minimizers above

Risk-Averse Clinical Trials (Darinka Dentcheva and Curtis McGinity)

- In stages $t = 1, \dots, T$ successive patients are given drugs (cytotoxic agents), to which severe toxic response (even death) is possible
- Probability of toxic response ($x_{t+1} = 1$) depends on the unknown optimal dose η^* and the administered dose (control) u_t:

$$F(u_t, \eta) = \frac{1}{1 + e^{-\varphi(u_t,\eta)}}$$

- The "belief state" ξ_t, the conditional probability distribution of the unknown optimal dose, is the current state of the system
- The state evolves according to Bayes operator, depending on the response of the patient: for $\eta \in \mathcal{Y}$ (the range of doses)

$$\xi_{t+1}(\eta) \sim \begin{cases} F(u_t, \eta)\, \xi_t(\eta) & \text{if toxic } \ (x_{t+1} = 1) \\ \big(1 - F(u_t, \eta)\big)\, \xi_t(\eta) & \text{if not toxic } \ (x_{t+1} = 0) \end{cases}$$

- Cost per stage: $c_t(\eta, u_t) = \gamma_t |u_t - \eta|$ (other forms possible)

Medical ethics naturally motivates risk-averse control

Total Cost Models

Find the best policy $\pi = (\pi_1, \dots, \pi_T)$ to determine doses $u_t = \pi_t(\xi_t)$

Expected Value Model

$$\min_{\pi \in \Pi} \mathbb{E}^{\pi} \left[\sum_{t=1}^{T+1} \gamma_t |u_t - \eta^*| \right]$$

γ_{T+1} is the weight of the final recommendation u_{T+1}

Risk-Averse Model

$$\min_{\pi \in \Pi} \rho^{\pi}_{1,T+1} \left[\left\{ \gamma_t |u_t - \eta^*| \right\}_{t=1,\dots,T+1} \right]$$

Two sources of risk

- Unknown state η^* (only belief state ξ_t available at time t)
- Unknown evolution of $\{\xi_t\}$ due to random responses of patients

Dynamic Programming Equations

- All memory is carried by the belief state ξ_t
- For each ξ_t and u_t, only two next states are possible, corresponding to $x_{t+1} = 0$ or 1

Simplified equation

$$v_t(\xi) = \min_u \left\{ r_t(\xi, u) + \sigma\Big(\xi, \int_{\mathcal{Y}} \mathbb{P}[x' = 1|y, u]\, \xi(dy), v^*_{t+1}\big(\Gamma_t(x, \xi, u, \cdot)\big)\Big) \right\}$$

Examples:

$$r_t(\xi, u) = \mathbb{E}_\xi\big[|u - \eta|\big]$$

$$\sigma\big(\xi, p, \varphi(\cdot)\big) = \mathbb{E}_\xi\Big[\max_{x' \in \{0,1\}} \varphi(x')\Big]$$

Any law invariant risk measure on the space of functions on U (for r_t) or on $U \times \{0, 1\}$ (in the case of σ_t) can be used here.

Limited Lookahead Policies

At each time t, assume that this is the last test before the final recommendation, and solve the two-stage problem

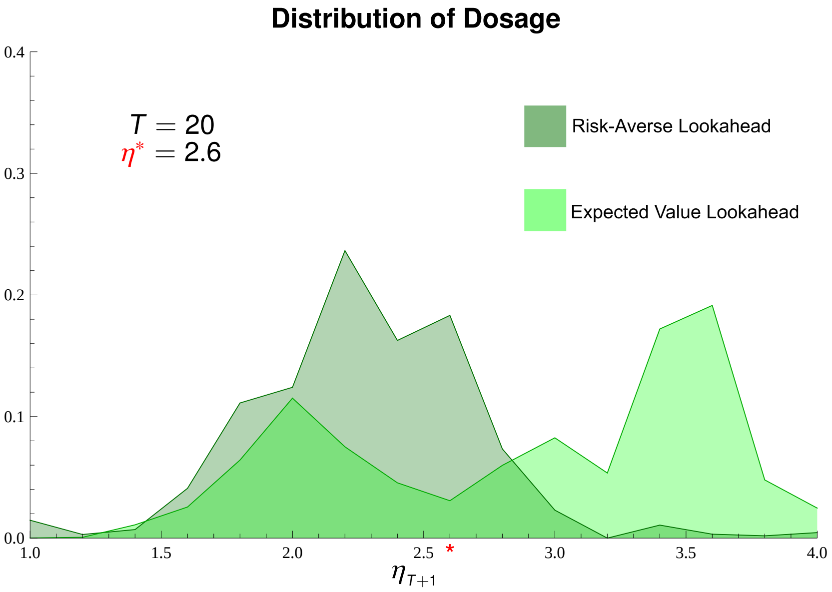
Risk-Neutral

$$\min_{u_t} \mathbb{E}_{\xi_t}\left\{\gamma_t|u_t-\eta| + \overline{\gamma}_{t+1}\ \mathbb{E}_{\text{response}}\left[\min_{u_{t+1}} \mathbb{E}_{\xi_{t+1}}|u_{t+1}-\eta|\right]\right\}$$

Risk-Averse

$$\min_{u_t} \mathbb{E}_{\xi_t}\left\{\gamma_t|u_t-\eta| + \overline{\gamma}_{t+1}\ \max_{\text{response}}\left[\min_{u_{t+1}} \mathbb{E}_{\xi_{t+1}}|u_{t+1}-\eta|\right]\right\}$$

$$\overline{\gamma}_{t+1} = \sum_{\tau=t+1}^{T+1} \gamma_\tau \qquad \text{(weight of the future)}$$


Distribution of Dosage
0.4
0.3
0.2
0.1
0.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
T = 20
η* = 2.6
Risk-Averse Lookahead
Expected Value Lookahead
*
η_{T+1}

Machine Deterioration (with Jingnan Fan)

We consider the problem of minimizing costs of a machine in T periods.

Unobserved state: $y_t \in \{1,2\}$, with 1 being the "good" and 2 the "bad" state
Observed state: x_t - cost incurred in the previous period
Control: $u_t \in \{0,1\}$, with 0 meaning "continue", and 1 meaning "replace"

The dynamics of Y is Markovian, with the transition matrices $K^{[u]}$:

$$K^{[0]} = \begin{pmatrix} 1-p & p \\ 0 & 1 \end{pmatrix} \quad K^{[1]} = \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix}$$

Distribution of costs:

$$\mathbb{P}\big[x_{t+1} \le C \,\big|\, y_t = i, u_t = 0\big] = \int_{-\infty}^{C} f_i(x)\,dx, \quad i = 1,2$$

$$\mathbb{P}\big[x_{t+1} \le C \,\big|\, y_t = i, u_t = 1\big] = \int_{-\infty}^{C} f_1(x)\,dx, \quad i = 1,2$$

Value and Policy Monotonicity

Belief state: $\xi_i \in [0,1]$ - conditional probability of the "good" state
The optimal value functions: $v_t^*(x,\xi) = x + w_t^*(\xi),\ t = 1,\dots,T+1$

Dynamic programming equations

$$w_t^*(\xi) = \min \Big\{ R + \sigma\big(f_1, x' \mapsto x' + w_{t+1}^*(1-p)\big); \\ \sigma\big(\xi f_1 + (1-\xi) f_2, x' \mapsto x' + w_{t+1}^*(\Gamma(\xi, x'))\big) \Big\},$$

with the final stage value $w_{T+1}^*(\cdot) = 0$.

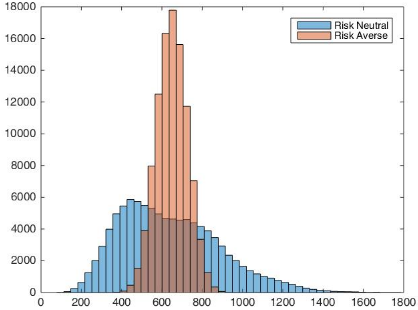
If $\frac{f_1}{f_2}$ is non-increasing, then the functions $w_t^*(\cdot)$ are non-increasing and thresholds $\xi_t^* \in [0,1],\ t = 1,\dots,T$ exist, such that the policy

$$u_t^* = \begin{cases} 0 & \text{if } \xi_t > \xi_t^*, \\ 1 & \text{if } \xi_t \le \xi_t^*, \end{cases}$$

is optimal

Numerical Illustration

Cost distributions f_1 and f_2: uniform with $\int_0^\eta f_1(x)\,dx \preceq \int_0^\eta f_2(x)\,dx$
Transition risk mapping: mean–semideviation



Empirical distribution of the total cost for the risk-neutral model (blue) and the risk-averse model (orange)

References

- D. Dentcheva, A. Ruszczyński: Risk Forms: Representation, Disintegration, and Application to Partially Observable Two-Stage Systems, *Mathematical Programming, Series B*, 2019, *OnLine First*.
- D. Dentcheva and A. Ruszczyński, Risk preferences on the space of quantile functions, *Mathematical Programming, Series B* 148 (2014), No. 1-2, 181–200.
- J. Fan and A. Ruszczyński, Risk measurement and risk-averse control of partially observable discrete-time Markov systems, *Mathematical Methods of Operations Research*, 2018, 1–24.
- C. McGinity, D. Dentcheva and A. Ruszczyński, Risk-averse approximate dynamic programming for partially observable systems with application to clinical trials, *in preparation*
- A. Ruszczyński, Risk-averse dynamic programming for Markov decision processes, *Mathematical Programming*, Series B 125 (2010) 235–261
- Ö. Çavuş and A. Ruszczyński, Computational methods for risk-averse undiscounted transient Markov models, *Operations Research*, 62 (2), 2014, 401–417.
- Ö. Çavuş and A. Ruszczyński, Risk-averse control of undiscounted transient Markov models, *SIAM J. on Control and Optimization*, 52(6), 2014, 3935–3966.