Risk-Averse Control of Markov Systems

Andrzej Ruszczyński

University of Florida

November 2015

Controlled Markov Models

- State space $\mathcal{X}$ (Borel)
- Control space $\mathcal{U}$ (Borel)
- Feasible control set $U : \mathcal{X} \rightrightarrows \mathcal{U}$, $t = 1, 2, \dots$
- Controlled transition kernel $Q : \text{graph}(U) \to \mathcal{P}(\mathcal{X})$, $t = 1, 2, \dots$
  $\mathcal{P}(\mathcal{X})$ - set of probability measures on $\mathcal{X}$
- Cost functions $c : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$, $t = 1, 2, \dots$
- State history $h_t = (x_1, \dots, x_t) \in \mathcal{X}^t$ (up to time $t = 1, 2, \dots$)
- Policy $\pi_t : \mathcal{X}^t \to \mathcal{U}$, $t = 1, 2, \dots$ (always supported in $U(x_t)$)
- Markov policy $\pi_t : \mathcal{X} \to \mathcal{U}$, $t = 1, 2, \dots$
  (stationary if $\pi_t = \pi_1$ for all t)

$$x_t \longrightarrow u_t = \pi_t(x_t)$$

$$(x_t, u_t) \longrightarrow x_{t+1} \sim Q(x_t, u_t)$$

Risk-Neutral Total Cost Problem

Infinite horizon expected cost problem:

$$\min_{\pi_1,\pi_2,\dots} \mathbb{E}^{\Pi}\left[\sum_{t=1}^{\infty} \alpha^{t-1} c_t(x_t,u_t)\right], \quad \alpha \in (0,1]$$

with controls $u_t = \pi_t(x_1,\dots,x_t)$

Two Cases:

Discounted models (with $\alpha < 1$) and transient models (with $\alpha = 1$)

Standard Results:

- A deterministic Markov policy is optimal
- Optimal policy can be found by dynamic programming equations

Our Intention

Introduce risk aversion to the problem by replacing
the expected value by dynamic risk measures

Using Dynamic Risk Measures for Markov Decision Processes

- Controlled Markov process x_t^{Π}, $t = 1, \dots, T$
- Policy $\Pi = \{\pi_1, \pi_2, \dots, \pi_T\}$ with $u_t = \pi_t(x_t)$ implies measure P^{Π}
- Cost sequence $Z_t^{\Pi} = c(x_t^{\Pi}, \pi_t(x_t^{\Pi}))$ (bounded), $t = 1, \dots, T$,
- Dynamic time-consistent risk measure

$$J_T(\Pi) = Z_1^{\Pi} + \rho_1^{\Pi}\Big(Z_2^{\Pi} + \cdots + \rho_{T-1}^{\Pi}(Z_T^{\Pi})\cdots\Big)$$

- Risk-averse optimal control problem: $\min\limits_{\Pi} \lim\limits_{T\to\infty} J_T(\Pi)$

Difficulties

- Probability measure P^{Π}, processes x_t^{Π} and Z_t^{Π} depend on policy Π
- The one-step risk measures $\rho_t^{\Pi}(\cdot)$ depend on Π and may depend on history $\Rightarrow$ no Markov policies

Idea

We only need to measure risk of random sequences that may occur

History $h_t = (x_1, \dots, x_t)$. Process $Z_t^{\Pi}(h_t) = c(x_t, \pi_t(h_t))$, $t = 1, \dots, T$

A family of conditional risk measures $\{\rho_{t,T}^{\Pi}\}_{t=1,\dots,T}^{\Pi \in \varPi}$ is stochastically conditionally time-consistent if for all feasible policies Π, Π', all $1 \le t \le T-1$, and for all histories $h_t \in \mathcal{X}^t$, the relations

$$Z_t^{\Pi}(h_t) = Z_t^{\Pi'}(h_t)$$

$$\left(\rho_{t+1,T}^{\Pi}(Z_{t+1}^{\Pi}, \dots, Z_T^{\Pi}) \,\middle|\, H_t^{\Pi} = h_t\right) \preceq_{\text{st}} \left(\rho_{t+1,T}^{\Pi'}(Z_{t+1}^{\Pi'}, \dots, Z_T^{\Pi'}) \,\middle|\, H_t^{\Pi'} = h_t\right)$$

imply

$$\rho_{t,T}^{\Pi}(Z_t^{\Pi}, \dots, Z_T^{\Pi})(h_t) \le \rho_{t,T}^{\Pi'}(Z_t^{\Pi'}, \dots, Z_T^{\Pi'})(h_t)$$

The conditional stochastic order $\preceq_{\text{st}}$:

$$\begin{aligned} Q_t^{\Pi}(h_t)\big(\{y : Z_t^{\Pi}(h_t) + \rho_{t+1,T}^{\Pi}(Z_{t+1}^{\Pi}, \dots, Z_T^{\Pi})(h_t, y) > \eta\}\big) \\ \le Q_t^{\Pi'}(h_t)\big(\{y : Z_t^{\Pi'}(h_t) + \rho_{t+1,T}^{\Pi'}(Z_{t+1}^{\Pi'}, \dots, Z_T^{\Pi'})(h_t, y) > \eta\}\big) \end{aligned}$$

A family of process-based dynamic risk measures $\{\rho_{t,T}^{\Pi}\}_{t=1,\dots,T}^{\Pi\in\mathbb{\Pi}}$ for a Markov decision problem is Markovian if for all Markov policies $\Pi \in \mathbb{\Pi}$, for any measurable and bounded $c_1, \dots, c_T : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$, and for all $h_t = (x_1, \dots, x_t)$ and $h_t' = (x_1', \dots, x_t')$ such that $x_t = x_t'$, we have

$$\rho_{t,T}^{\Pi}\big(c_t(X_t, \pi_t(X_t)), \dots, c_T(X_T, \pi_T(X_T))\big)(h_t) \\ = \rho_{t,T}^{\Pi}\big(c_t(X_t, \pi_t(X_t)), \dots, c_T(X_T, \pi_T(X_T))\big)(h_t').$$

If the current state x_t is the same, and the same Markov policy Π is used, then the risk is the same. The risk measure can be written as a function of the state:

$$\rho_{t,T}^{\Pi}\big(c_t(X_t, \pi_t(X_t)), \dots, c_T(X_T, \pi_T(X_T))\big)(x_t)$$

For a fixed history-dependent policy Π and every $h_t \in \mathcal{X}^t$, we write

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

If a family of process-based dynamic risk measures $\left\{\rho_{t,T}^{\Pi}\right\}_{t=1,\dots,T}^{\Pi\in\varPi}$ is Markovian, translation-invariant, and stochastically conditionally time-consistent, then there exist transition risk mappings

$$\sigma_t : \big\{\big(x, Q_t(x,u)\big) \,:\, u \in U(x),\ x \in \mathcal{X}\big\} \times \mathcal{V} \to \mathbb{R}, \quad t = 1,\ldots,T-1$$

($\mathcal{V}$ - space of measurable bounded functions on $\mathcal{X}$)
such that for all $\Pi \in \varPi$, for all $t = 1,\ldots,T-1$, and all $h_t \in \mathcal{X}^t$, the functional $\sigma_t\big(x_t, Q_t(x_t,\pi_t(h_t),\cdot)\big)$ is a law-invariant risk measure on $\big(\mathcal{X},\mathcal{B}(\mathcal{X}),Q_t\big)$ and for any $c = \{c_t\}_{t=1\dots T}$, we have

$$v_t^{c,\Pi}(h_t) = c_t(x_t,\pi_t(h_t)) + \sigma_t\big(x_t, Q_t(x_t,\pi_t(h_t)), v_{t+1}^{c,\Pi}(h_t,\cdot)\big),\ t = 1\ldots T-1$$

For a fixed history-dependent policy Π and every $h_t \in \mathcal{X}^t$, we write

$$v_t^{c,\Pi}(h_t) = \rho_{t,T}^{\Pi}\big(c_t(X_t, \pi_t(H_t)), \dots, c_T(X_T, \pi_T(H_T))\big)(h_t)$$

If a family of process-based dynamic risk measures $\{\rho_{t,T}^{\Pi}\}_{t=1,\dots,T}^{\Pi\in\mathbb{\Pi}}$ is Markovian, translation-invariant, and stochastically conditionally time-consistent, then there exist transition risk mappings

$$\sigma_t : \big\{\big(x, Q_t(x,u)\big) \,:\, u \in U(x),\ x \in \mathcal{X}\big\} \times \mathcal{V} \to \mathbb{R}, \quad t = 1, \dots, T-1$$

($\mathcal{V}$ - space of measurable bounded functions on $\mathcal{X}$)
such that for all $\Pi \in \mathbb{\Pi}$, for all $t = 1, \dots, T-1$, and all $h_t \in \mathcal{X}^t$, the functional $\sigma_t\big(x_t, Q_t(x_t, \pi_t(h_t), \cdot)\big)$ is a law-invariant risk measure on $\big(\mathcal{X}, \mathcal{B}(\mathcal{X}), Q_t\big)$ and for any $c = \{c_t\}_{t=1\dots T}$, we have (for Markovian Π)

$$v_t^{c,\Pi}(x_t) = c_t(x_t, \pi_t(x_t)) + \sigma_t\big(x_t, Q_t(x_t, \pi_t(x_t)), v_{t+1}^{c,\Pi}(\cdot)\big), \ t = 1 \dots T-1$$

Finite Horizon Risk-Averse Control Problem

Consider a controlled Markov process $\{X_t\}$ with $u_t = \pi_t(X_1, \dots, X_t)$.

Risk-averse optimal control problem:

$$\min_{\Pi} J_T(\Pi, x_1) = c_1(x_1, u_1) + \rho_1^{\Pi}\Big(c_2(X_2, u_2) + \cdots + \rho_{T-1}^{\Pi}\big(c_T(X_T, u_T) + \rho_T\big(c_{T+1}(X_{T+1})\big) \cdots \big)\Big)$$

Theorem

If the conditional measures ρ_t^{Π} are Markovian (+ general conditions), then the optimal solution is given by the dynamic programming equations:

$$v_{T+1}(x) = c_{T+1}(x), \quad x \in \mathcal{X}$$

$$v_t(x) = \min_{u \in U(x)} \Big\{ c_t(x, u) + \sigma_t\big(x, Q_t(x, u), v_{t+1}\big) \Big\}, \ x \in \mathcal{X}, \ t = T, \dots, 1$$

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

Finite Horizon Risk-Averse Control Problem

Consider a controlled Markov process $\{X_t\}$ with $u_t = \pi_t(X_1, \dots, X_t)$.

Risk-averse optimal control problem:

$$\min_{\Pi} \; J_T(\Pi, x_1) = c_1(x_1, u_1) + \rho_1^{\Pi}\Big(c_2(X_2, u_2) + \cdots + \rho_{T-1}^{\Pi}\Big(c_T(X_T, u_T) + \rho_T\big(c_{T+1}(X_{T+1})\big) \cdots \Big)\Big)$$

Theorem

If the conditional measures ρ_t^{Π} are Markovian (+ general conditions), then the optimal solution is given by the dynamic programming equations:

$$v_{T+1}(x) = c_{T+1}(x), \quad x \in \mathcal{X}$$

$$v_t(x) = \min_{u \in U(x)} \Big\{ c_t(x,u) + \max_{\mu \in \mathcal{A}_t(x, Q_t(x,u))} \mathbb{E}_\mu\big[v_{t+1}\big] \Big\}, \; x \in \mathcal{X}, \; t = T, \dots, 1$$

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

Discounted risk measure $(0 < \alpha < 1)$

$$J_T^{\alpha}(\Pi, x) = Z_1^{\Pi} + \rho_1^{\Pi}\Big(\alpha Z_2^{\Pi} + \cdots + \rho_{T-1}^{\Pi}(\alpha^{T-1} Z_T^{\Pi}) \cdots\Big)$$

Optimal cost: $J^*(x) = \inf_{\Pi} \lim_{T\to\infty} J_T^{\alpha}(\Pi, x)$

Assume that the model is stationary, the conditional risk measures ρ_t, $t = 1, \dots, T$, are Markovian (+ technical conditions). Then a bounded function $v : \mathcal{X} \to \mathbb{R}$ satisfies the dynamic programming equations

$$v(x) = \min_{u \in U(x)} \Big\{ c(x, u) + \alpha\sigma\big(x, Q(x, u), v\big) \Big\}, \quad x \in \mathcal{X},$$

if and only if $v(\cdot) \equiv J^*(\cdot)$. Moreover, the minimizer $\pi^*(x)$, $x \in \mathcal{X}$, on the right hand side exists and defines an optimal Markov policy $\Pi^* = \{\pi^*, \pi^*, \dots\}$.

If $\alpha = 1$ additional conditions of risk transient models

Infinite Horizon Risk (for stationary and coherent models)

Discounted risk measure ($0 < \alpha < 1$)

$$J_T^{\alpha}(\Pi, x) = Z_1^{\Pi} + \rho_1^{\Pi}\Big(\alpha Z_2^{\Pi} + \cdots + \rho_{T-1}^{\Pi}\big(\alpha^{T-1} Z_T^{\Pi}\big)\cdots\Big)$$

Optimal cost: $J^*(x) = \inf_{\Pi} \lim_{T\to\infty} J_T^{\alpha}(\Pi, x)$

Assume that the model is stationary, the conditional risk measures ρ_t, $t = 1, \dots, T$, are Markovian (+ technical conditions). Then a bounded function $v : \mathcal{X} \to \mathbb{R}$ satisfies the dynamic programming equations

$$v(x) = \min_{u \in U(x)} \Big\{ c(x, u) + \alpha \max_{\mu \in \mathcal{A}(x, Q(x,u))} \mathbb{E}_{\mu}[v] \Big\}, \quad x \in \mathcal{X},$$

if and only if $v(\cdot) \equiv J^*(\cdot)$. Moreover, the minimizer $\pi^*(x)$, $x \in \mathcal{X}$, on the right hand side exists and defines an optimal Markov policy $\Pi^* = \{\pi^*, \pi^*, \dots\}$.

If $\alpha = 1$ additional conditions of risk transient models

Continuous-Time Markov Chains

For a finite state space $\mathcal{X}$, we consider a continuous-time Markov chain $\{X_t\}_{0 \le t \le T}$ with the transition function

$$Q_{t,r}(y|x) = P(X_r = y \,|\, X_t = x),$$

where $x, y \in \mathcal{X}$ and $0 \le t < r \le T$. We assume that the transition rates

$$G_t(y|x) = \lim_{\tau \downarrow 0} \frac{1}{\tau}\big[Q_{t,t+\tau}(y|x) - \delta_x(y)\big], \quad x, y \in \mathcal{X},$$

are uniformly bounded for all $0 \le t \le T$. Here,

$$\delta_x(y) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

The rates constitute the generator $G_t : \mathcal{X} \to \mathcal{M}(\mathcal{X})$, where $\mathcal{M}(\mathcal{X})$ is the set of signed measures on $\mathcal{X}$

$\xi_{[0,t]}$ - history (path) of the process X up to time t
$\Xi^{\xi_t}_{t,r}$ - space of paths on $[t, r]$ starting from ξ_t; $P^{\xi_t}_{t,r}$ - corresponding measure

A dynamic risk measure $\varrho = \{\varrho_{t,T}\}_{t\in[0,T]}$ is stochastically conditionally time-consistent, if for all $0 \le t \le r \le T$, all $\xi_{[0,t]} \in \Xi_{[0,t]}$, if

$$\varrho_{r,T}(Z_T)\,|\,\xi_{[0,t]} \preceq_{\rm st} \varrho_{r,T}(W_T)\,|\,\xi_{[0,t]}$$

then

$$\varrho_{t,T}(Z_T)(\xi_{[0,t]}) \le \varrho_{t,T}(W_T)(\xi_{[0,t]}) \qquad (\star)$$

It is strongly stochastically conditionally time-consistent, if for any two times $r_1, r_2 \in [t, T]$, the inequality

$$\varrho_{r_1,T}(Z_T)\,|\,\xi_{[0,t]} \preceq_{\rm st} \varrho_{r_2,T}(W_T)\,|\,\xi_{[0,t]} \qquad \text{implies} \qquad (\star)$$

The conditional stochastic order "$\preceq_{\rm st}$": for all $\eta \in \mathbb{R}$

$$P^{\xi_t}_{t,r_1}\{\varrho_{r_1,T}(Z_T)\,|\,\xi_{[0,t]} > \eta\} \le P^{\xi_t}_{t,r_2}\{\varrho_{r_2,T}(W_T)\,|\,\xi_{[0,t]} > \eta\}$$

Cost of the process starting from ξ_t at time t:

$$Z_{t,T}^{\xi_t} = \int_t^T c_s\big(X_s^{t,\xi_t}\big)\, ds + f\big(X_T^{t,\xi_t}\big)$$

A dynamic risk measure $\big\{\varrho_{t,T}\big\}_{t\in[0,T]}$ is Markovian, if for all $0 \le t < T$, all paths $\xi_{[0,t]}, \xi'_{[0,t]}$, the equality $\xi_t = \xi'_t$ implies that for all bounded measurable functions $c : [t, T] \times \mathcal{X} \to \mathbb{R}$ and $f : \mathcal{X} \to \mathbb{R}$ we have

$$\varrho_{t,T}\big(Z_{t,T}^{\xi_t}\big)(\xi_{[0,t]}) = \varrho_{t,T}\big(Z_{t,T}^{\xi'_t})\big)(\xi'_{[0,t]}).$$

The risk of the future costs $Z_{t,T}^{\xi_t}$ is a function of the last observed state ξ_t.

For Markovian risk measures having the local property we write

$$v_t(\xi_t) = \varrho_{t,T}\big(Z_{t,T}^{\xi_t}\big)(\xi_t)$$

Cost accumulated on the interval $[t, r]$, given state ξ_t:

$$I^{\xi_t}_{t,r}(c) = \int_t^r c_s(X^{t,\xi_t}_s)\, ds$$

$\Xi^{\xi_t}_{t,r}$ - space of paths on $[t, r]$ starting from ξ_t; $P^{\xi_t}_{t,r}$ - corresponding measure

If $\{\varrho_{t,T}\}_{t\in[0,T]}$ is stochastically conditionally time-consistent, translation invariant, and Markovian, then for every $0 \le t \le r \le T$ and every $\xi_t \in \mathcal{X}$ a functional $\varsigma^{\xi_t}_{t,r} : \mathcal{L}_\infty(\Xi^{\xi_t}_{t,r}, P^{\xi_t}_{t,r}) \to \mathbb{R}$ exists such that

$$v_t(\xi_t) = \varsigma^{\xi_t}_{t,r}\Big(I^{\xi_t}_{t,r}(c) + v_r(X^{t,\xi_t}_r)\Big)$$

Moreover, the functional $\varsigma^{\xi_t}_{t,r}(\cdot)$ is law invariant with respect to the probability measure $P^{\xi_t}_{t,r}$.

If ϱ is coherent, then $\varsigma^{\xi_t}_{t,r}(\cdot)$ is a coherent measure of risk

Assumption: $\varsigma^{\xi_t}_{t,r}(\cdot)$ is Lipschitz continuous in $\mathcal{L}_p\big(\Xi^{\xi_t}_{t,r}, P^{\xi_t}_{t,r}\big)$, $p \in [1,\infty)$.

If a dynamic risk measure $\{\varrho_{t,T}\}_{t\in[0,T]}$ is strongly stochastically conditionally time-consistent, translation invariant, and Markovian, then for every $t \in [0, T]$ a functional $\sigma_t : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \mathcal{L}(\mathcal{X}) \to \mathbb{R}$ exists such that for every Z_T, for all $\xi_t \in \mathcal{X}$, and all $r \in [t, T]$ we have

$$v_t(\xi_t) = \int_t^r c_s(\xi_t)\, ds + \sigma_t\big(\xi_t, Q_{t,r}(\,\cdot\,|\xi_t), v_r\big) + o(r-t),$$

(i) $\sigma_t(\cdot,\cdot,\cdot)$ is law invariant with respect to the second argument

(ii) If ϱ is coherent, then $\sigma_t(\xi_t,\cdot,\cdot)$ is a coherent measure of risk

(iii) For all $x \in \mathcal{X}$ and all $v \in \mathcal{L}(\mathcal{X})$, we have $\sigma_t(x, \delta_x, v) = v(x)$, where δ_x is the Dirac measure at x [state consistency]

$\mathcal{Q}$ - set of stochastic kernels $Q : \mathcal{X} \to \mathcal{P}(\mathcal{X})$

If $\sigma_t(x, m, \cdot)$ is coherent, then the following dual representation is true:

$$\sigma_t\big(x, m, v\big) = \max_{\mu \in \mathcal{A}_t(x,m)} \sum_{y \in \mathcal{X}} v(y)\mu(y), \quad v \in \mathcal{L}(\mathcal{X}),$$

where $\mathcal{A}_t(x, m) \subset \mathcal{P}(\mathcal{X})$ is a nonempty, convex, closed, and bounded set. We define the multikernel $\mathfrak{M} : \mathcal{Q} \rightrightarrows \mathcal{Q}$:

$$\mathfrak{M}(Q) = \big\{M \in \mathcal{Q} : M(x) \in \mathcal{A}(x, Q(x)),\ \forall\, x \in \mathcal{X}\big\}.$$

A multifunction $\mathfrak{M}$ is semi-differentiable at the point I in the direction $K \in \mathcal{T}_{\mathcal{Q}}(I)$ if a nonempty set $\mathfrak{D}(K) \subset \mathfrak{S}$ exists, such that for every sequence $\varepsilon_n \downarrow 0$ and every sequence $K_n \to K$, $K_n \in \mathcal{T}_{\mathcal{Q}}(I)$,

$$\lim_{n\to\infty} \frac{1}{\varepsilon_n}\big[\mathfrak{M}(I + \varepsilon_n K_n) - I\big] = \mathfrak{D}(K),$$

- Semiderivatives $\mathfrak{D}(K)$ of many transition risk mappings (semideviations, average value at risk, etc.) exist and can be calculated, for every tangent direction K
- In our case, $K = G_t$ (the generator of the system)
- For small time increments δ, we can derive the "chain rule"

$$\mathfrak{M}(Q_{t,t+\delta}) \approx I + \delta\mathfrak{D}(G_t)$$

- Using the support functions $s_x(v) = \sup_{\lambda \in \mathfrak{D}(G_t)(x)} \sum_{y \in \mathcal{X}} \lambda(y)v(y)$, we derive the value function representation by ODEs:

$$\frac{\partial v_t(x)}{\partial t} = -c_t(x) - s_x(v_t), \quad t \in [0, T], \quad x \in \mathcal{X},$$
$$v_T(x) = f(x), \quad x \in \mathcal{X}.$$

- Close approximations by discrete-time models can be constructed, and the discrete-time theory and methods apply

Filtered probability space $(\Omega, \mathscr{F}, P, \mathbb{F})$
Filtration $\mathbb{F}$ is generated by n-dimensional Brownian motion $\{W_t\}_{t\in[0,T]}$
Controlled diffusion process with initial value $\zeta \in \mathcal{L}_2(\Omega, \mathscr{F}_t, \mathbb{P}; \mathbb{R}^n)$:

$$dX_s^{t,\zeta;u} = b(s, X_s^{t,\zeta;u}, u_s)\, ds + \sigma(s, X_s^{t,\zeta;u}, u_s)\, dW_s, \quad s \in [t, T],$$
$$X_t^{t,\zeta;u} = \zeta,$$

with functions $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times d}$.
Cost rate $c : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$; Final cost $\Psi : \mathbb{R}^n \to \mathbb{R}$.
Cost accumulated in the interval $[t, T]$

$$\xi_{t,T}(u, \zeta) := \int_t^T c(s, X_s^{t,\zeta;u}, u_s)\, ds + \Psi(X_T^{t,\zeta;u}), \quad \text{a.s.}.$$

All functions are assumed to be sufficiently regular (Lipschitz or bounded).

$$\min_{u(\cdot)\in\mathcal{U}} \varrho_{0,T}\left(\int_0^T c(s, X_s^{0,x_0;u}, u_s)\, ds + \Psi(X_T^{0,x_0;u})\right)$$

$$dX_s^{0,x_0;u} = b(s, X_s^{0,x_0;u}, u_s)\, ds + \sigma(s, X_s^{0,x_0;u}, u_s)\, dW_s, \quad s \in [0,T]$$

where $\{\varrho_{t,r}\}_{0\le t\le r\le T}$ is a dynamic risk measure on the space of square-integrable adapted processes on $[0,T]\times\Omega$

Time consistency: $\varrho_{t,r}(Y_r) = \varrho_{t,s}\big(\varrho_{s,r}(Y_r)\big)$, for all $t \le s \le r$
Local property: $\varrho_{t,r}(\mathbb{1}_A Y_r) = \mathbb{1}_A \varrho_{t,r}(Y_r)$, for all events $A \in \mathcal{F}_t$.

Structure of $\varrho_{t,r}(\cdot)$ [Coquet, Hu, Mémin, Peng (2002)]

Under mild conditions, a generator $g : [0,T]\times\mathbb{R}\times\mathbb{R}^n$ exists, such that $\varrho_{t,r}(\xi) = Y_t$, where (Y,Z) solve backward stochastic differential equation

$$-dY_s = g(s, Y_s, Z_s)\, ds - Z_s'\, dW_s, \quad s \in [t,r], \quad Y_r = \xi.$$

If g is convex, pos.-homogeneous, independent of y, then ϱ is coherent

Value function

$$V(t,x) = \inf_{u(\cdot)\in\mathcal{U}} \varrho_{t,T}\left(\int_t^T c(s, X_s^{t,x;u}, u_s)\, ds + \Psi(X_T^{t,x;u})\right)$$

Dynamic Programming Equation

For any $(t,x) \in [0,T) \times \mathbb{R}^n$ and all $r \in [t,T]$, we have

$$V(t,x) = \inf_{u(\cdot)\in\mathcal{U}} \varrho_{t,r}\left[\int_t^r c\big(s, X_s^{t,x;u}, u_s\big)\, ds + V\big(r, X_r^{t,x;u}\big)\right].$$

Related decoupled forward–backward system:

$$\begin{aligned}
dX_s^{t,x;u} &= b(s, X_s^{t,x;u}, u_s)\, ds + \sigma(s, X_s^{t,x;u}, u_s)\, dW_s, \quad s \in [t,r]\\
X_t^{t,x;u} &= x\\
-dY_s^{t,x;u} &= \big[c(s, X_s^{t,x;u}, u_s) + g(s, Z_s^{t,x;u})\big]\, ds - Z_s^{t,x;u}\, dW_s, \quad s \in [t,r]\\
Y_r^{t,\zeta;u} &= V\big(r, X_r^{t,x;u}\big)
\end{aligned}$$

Laplacian operator:

$$\big[\mathcal{L}^{\alpha} w\big](t,x) = \partial_t w(t,x) +$$

$$+ \sum_{i,j=1}^{n} \frac{1}{2}\big(\sigma(t,x,\alpha)\sigma(t,x,\alpha)^{\top}\big)_{ij}\partial_{x_i x_j} w(t,x) + \sum_{i=1}^{n} b_i(t,x,\alpha)\partial_{x_i} w(t,x).$$

Risk-Averse HJB Equation

On the space $\mathcal{C}_b^{1,2}([0,T]\times\mathbb{R}^n)$, we consider the following equation

$$\min_{\alpha\in U}\Big\{c(t,x,\alpha) + \big[\mathcal{L}^{\alpha} v\big](t,x) + g\big(t, [\mathcal{D}_x v\cdot\sigma^{\alpha}](t,x)\big)\Big\} = 0 \quad \forall (t,x)$$

$$v(T,x) = \Psi(T,x), \quad x\in\mathbb{R}^n.$$

If the functions b and σ are bounded, then the value function $V(t,x)$ is a viscosity solution of the risk-averse HJB equation.
Conversely, if the HJB equation has a solution, it is equal to $V(t,x)$.

Extensions and Current Research on Risk-Averse Control

- Partially Observable Markov Processes (with Jingnan Fan)
  - process-based risk measures
  - transition risk mappings on the observable part
  - dynamic programming equations
- Risk-Averse Control of Clinical Trials
  (with Darinka Dentcheva and Curtis McGinity)
  - new dynamic models of clinical trials
  - approximate dynamic programming methods
- Risk-Averse Control of Diffusion Processes (with Jianing Yao)
  - approximation by risk-averse Markov chains

References

- A. Ruszczyński, Risk-averse dynamic programming for Markov decision processes, *Mathematical Programming*, Series B 125 (2010) 235–261
- Ö. Çavuș and A. Ruszczyński, Computational methods for risk-averse undiscounted transient Markov models, *Operations Research*, 62 (2), 2014, 401–417.
- Ö. Çavuș and A. Ruszczyński, Risk-averse control of undiscounted transient Markov models, *SIAM J. on Control and Optimization*, 52(6), 2014, 3935–3966.
- J. Fan and A. Ruszczyński, Process-based risk measures for observable and partially observable discrete-time controlled systems, *submitted for publication*.
- D. Dentcheva and A. Ruszczyński, Risk measures for continuous-time Markov chains, *submitted for publication*.
- A. Ruszczyński and J. Yao, Risk-averse control of diffusion processes, *submitted for publication*.