Risk-Averse Optimal Path Problems for Markov Models

Andrzej Ruszczyński and Özlem Çavuş

RUTGERS

Georg Pflug Anniversary Workshop

Vienna 2011

Research supported by the NSF award CMMI-0965689

How to Measure Risk of Sequences?

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: $\mathcal{Z}_t = \mathcal{L}_p(\Omega, \mathcal{F}_t, P)$, $p \in [1, \infty]$, and $\mathcal{Z}_{t,T} = \mathcal{Z}_t \times \cdots \times \mathcal{Z}_T$

Conditional Risk Measure

A mapping $\rho_{t,T} : \mathcal{Z}_{t,T} \to \mathcal{Z}_t$ satisfying the monotonicity condition:

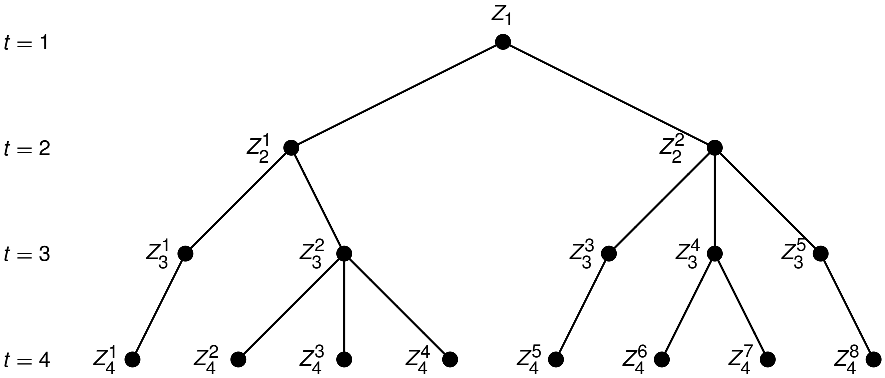
$$\rho_{t,T}(Z) \leq \rho_{t,T}(W) \text{ for all } Z, W \in \mathcal{Z}_{t,T} \text{ such that } Z \leq W$$

Dynamic Risk Measure

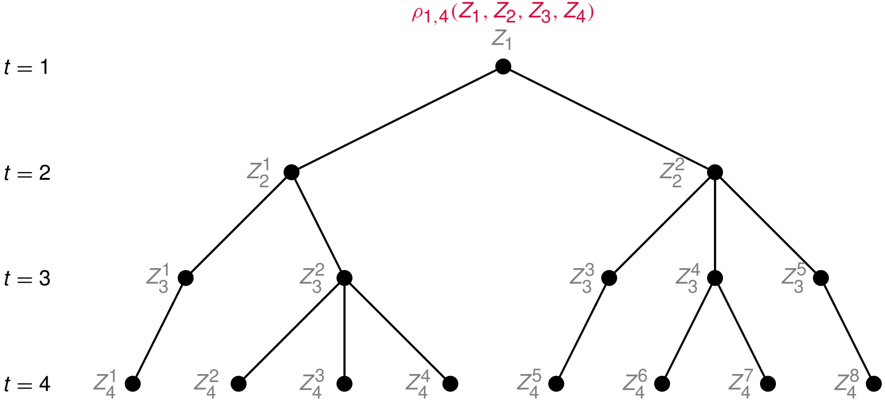
A sequence of conditional risk measures $\rho_{t,T} : \mathcal{Z}_{t,T} \to \mathcal{Z}_t$, $t = 1, \ldots, T$

$$\rho_{1,T}(Z_1, Z_2, Z_3, \ldots, Z_T) \in \mathcal{Z}_1 = \mathbb{R}$$
$$\rho_{2,T}(Z_2, Z_3, \ldots, Z_T) \in \mathcal{Z}_2$$
$$\rho_{3,T}(Z_3, \ldots, Z_T) \in \mathcal{Z}_3$$
$$\vdots$$

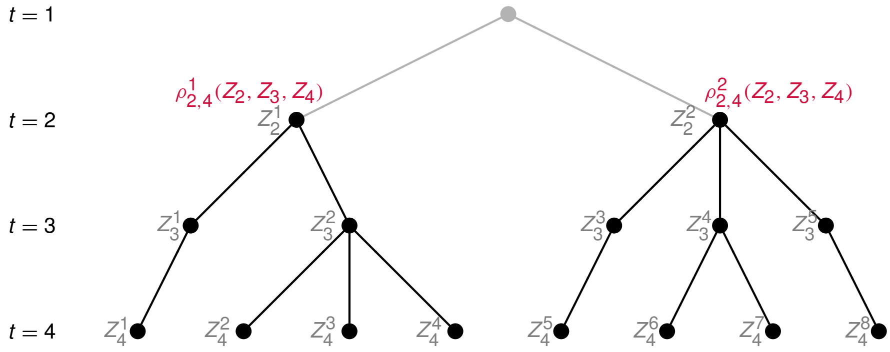
Evaluating Risk on a Scenario Tree



Evaluating Risk on a Scenario Tree



Evaluating Risk on a Scenario Tree



Time Consistency of Dynamic Risk Measures

A dynamic risk measure $\{\rho_{t,T}\}_{t=1}^{T}$ is time-consistent if for all $\tau < \theta$

$$Z_k = W_k,\ k = \tau, \dots, \theta - 1 \quad \text{and} \quad \rho_{\theta,T}(Z_\theta, \dots, Z_T) \le \rho_{\theta,T}(W_\theta, \dots, W_T)$$

imply that $\quad \rho_{\tau,T}(Z_\tau, \dots, Z_T) \le \rho_{\tau,T}(W_\tau, \dots, W_T)$

Define $\rho_t(Z_{t+1}) = \rho_{t,T}(0, Z_{t+1}, 0, \dots, 0)$

Suppose a dynamic risk measure $\{\rho_{t,T}\}_{t=1}^{T}$ is time-consistent and

$$\rho_{t,T}(Z_t, Z_{t+1}, \dots, Z_T) = Z_t + \rho_{t,T}(0, Z_{t+1}, \dots, Z_T)$$
$$\rho_{t,T}(0, \dots, 0) = 0$$

Then for all t we have the nested representation

$$\rho_{t,T}(Z_t, \dots, Z_T) = Z_t + \rho_t\Big(Z_{t+1} + \rho_{t+1}\big(Z_{t+2} + \dots + \rho_{T-1}(Z_T)\cdots\big)\Big)$$

Coherent One-Step Conditional Risk Measures

Stronger assumptions about one-step measures $\rho_t : \mathcal{Z}_{t+1} \to \mathcal{Z}_t$:

- Convexity: $\rho_t(\lambda Z + (1-\lambda)W) \le \lambda\rho_t(Z) + (1-\lambda)\rho_t(W)$
  $\forall\ \lambda \in (0,1),\ Z, W \in \mathcal{Z}_{t+1}$
- Monotonicity: If $Z \le W$ then $\rho_t(Z) \le \rho_t(W),\ \forall\ Z, W \in \mathcal{Z}_{t+1}$
- Predictable Translation Equivariance:
  $\rho_t(Z + W) = Z + \rho_t(W),\ \forall\ Z \in \mathcal{Z}_t,\ \ W \in \mathcal{Z}_{t+1}$
- Positive Homogeneity: $\rho_t(\tau Z) = \tau\rho_t(Z), \forall\ Z \in \mathcal{Z}_{t+1},\ \ \tau \ge 0$

Scandolo ('03), Riedel ('04), R.-Shapiro ('06), Cheridito-Delbaen-Kupper ('06), Föllmer-Penner ('06), Artzner-Delbaen-Eber-Heath-Ku ('07), Pflug-Römisch ('07)

Example: Conditional Mean–Semideviation

$$\rho_t(Z_{t+1}) = \mathbb{E}[Z_{t+1}|\mathcal{F}_t] + \kappa\mathbb{E}\Big[\big(Z_{t+1} - \mathbb{E}[Z_{t+1}|\mathcal{F}_t]\big)_+^s\Big|\mathcal{F}_t\Big]^{\frac{1}{s}}$$

Here $s \in [1,p]$ and $\kappa \in [0,1]$ may be $\mathcal{F}_t$-measurable

Coherent One-Step Conditional Risk Measures

Stronger assumptions about one-step measures $\rho_t : \mathcal{Z}_{t+1} \to \mathcal{Z}_t$:

- Convexity: $\rho_t(\lambda Z + (1-\lambda)W) \le \lambda \rho_t(Z) + (1-\lambda)\rho_t(W)$
  $\forall\, \lambda \in (0,1),\ Z, W \in \mathcal{Z}_{t+1}$
- Monotonicity: If $Z \le W$ then $\rho_t(Z) \le \rho_t(W),\ \forall\, Z, W \in \mathcal{Z}_{t+1}$
- Predictable Translation Equivariance:
  $\rho_t(Z + W) = Z + \rho_t(W),\ \forall\, Z \in \mathcal{Z}_t,\ \ W \in \mathcal{Z}_{t+1}$
- Positive Homogeneity: $\rho_t(\tau Z) = \tau \rho_t(Z), \forall\, Z \in \mathcal{Z}_{t+1},\ \ \tau \ge 0$

Scandolo ('03), Riedel ('04), R.-Shapiro ('06), Cheridito-Delbaen-Kupper ('06), Föllmer-Penner ('06), Artzner-Delbaen-Eber-Heath-Ku ('07), Pflug-Römisch ('07)

Example: Conditional Mean–Semideviation

$$\rho_t(Z_{t+1}) = \mathbb{E}[Z_{t+1}|\mathcal{F}_t] + \kappa \mathbb{E}\Big[\big(Z_{t+1} - \mathbb{E}[Z_{t+1}|\mathcal{F}_t]\big)_+^s \Big| \mathcal{F}_t\Big]^{\frac{1}{s}}$$

Here $s \in [1, p]$ and $\kappa \in [0, 1]$ may be $\mathcal{F}_t$-measurable

Coherent One-Step Conditional Risk Measures

Stronger assumptions about one-step measures $\rho_t : \mathcal{Z}_{t+1} \to \mathcal{Z}_t$:

- Convexity: $\rho_t(\lambda Z + (1-\lambda)W) \le \lambda\rho_t(Z) + (1-\lambda)\rho_t(W)$
  $\forall\, \lambda \in (0,1),\;\; Z, W \in \mathcal{Z}_{t+1}$
- Monotonicity: If $Z \le W$ then $\rho_t(Z) \le \rho_t(W),\;\; \forall\, Z, W \in \mathcal{Z}_{t+1}$
- Predictable Translation Equivariance:
  $\rho_t(Z + W) = Z + \rho_t(W),\;\; \forall\, Z \in \mathcal{Z}_t,\;\; W \in \mathcal{Z}_{t+1}$
- Positive Homogeneity: $\rho_t(\tau Z) = \tau\rho_t(Z), \forall\, Z \in \mathcal{Z}_{t+1},\;\; \tau \ge 0$

Scandolo ('03), Riedel ('04), R.-Shapiro ('06), Cheridito-Delbaen-Kupper ('06), Föllmer-Penner ('06), Artzner-Delbaen-Eber-Heath-Ku ('07), Pflug-Römisch ('07)

Example: Conditional Mean–Semideviation

$$\rho_t(Z_{t+1}) = \mathbb{E}[Z_{t+1}|\mathcal{F}_t] + \kappa\mathbb{E}\Big[\big(Z_{t+1} - \mathbb{E}[Z_{t+1}|\mathcal{F}_t]\big)_+^s \Big| \mathcal{F}_t\Big]^{\frac{1}{s}}$$

Here $s \in [1,p]$ and $\kappa \in [0,1]$ may be $\mathcal{F}_t$-measurable

Coherent One-Step Conditional Risk Measures

Stronger assumptions about one-step measures $\rho_t : \mathcal{Z}_{t+1} \to \mathcal{Z}_t$:

- **Convexity:** $\rho_t(\lambda Z + (1-\lambda)W) \leq \lambda\rho_t(Z) + (1-\lambda)\rho_t(W)$
  $\forall\ \lambda \in (0,1),\ Z, W \in \mathcal{Z}_{t+1}$
- **Monotonicity:** If $Z \leq W$ then $\rho_t(Z) \leq \rho_t(W)$, $\forall\ Z, W \in \mathcal{Z}_{t+1}$
- **Predictable Translation Equivariance:**
  $\rho_t(Z + W) = Z + \rho_t(W)$, $\forall\ Z \in \mathcal{Z}_t,\ W \in \mathcal{Z}_{t+1}$
- **Positive Homogeneity:** $\rho_t(\tau Z) = \tau\rho_t(Z)$, $\forall\ Z \in \mathcal{Z}_{t+1},\ \tau \geq 0$

Scandolo ('03), Riedel ('04), R.-Shapiro ('06), Cheridito-Delbaen-Kupper ('06), Föllmer-Penner ('06), Artzner-Delbaen-Eber-Heath-Ku ('07), Pflug-Römisch ('07)

Example: Conditional Mean–Semideviation

$$\rho_t(Z_{t+1}) = \mathbb{E}[Z_{t+1}|\mathcal{F}_t] + \kappa\mathbb{E}\Big[\big(Z_{t+1} - \mathbb{E}[Z_{t+1}|\mathcal{F}_t]\big)_+^s\Big|\mathcal{F}_t\Big]^{\frac{1}{s}}$$

Here $s \in [1, p]$ and $\kappa \in [0, 1]$ may be $\mathcal{F}_t$-measurable

Coherent One-Step Conditional Risk Measures

Stronger assumptions about one-step measures $\rho_t : \mathcal{Z}_{t+1} \to \mathcal{Z}_t$:

- Convexity: $\rho_t(\lambda Z + (1-\lambda)W) \le \lambda \rho_t(Z) + (1-\lambda)\rho_t(W)$
  $\forall\, \lambda \in (0,1),\ Z, W \in \mathcal{Z}_{t+1}$
- Monotonicity: If $Z \le W$ then $\rho_t(Z) \le \rho_t(W),\ \forall\, Z, W \in \mathcal{Z}_{t+1}$
- Predictable Translation Equivariance:
  $\rho_t(Z + W) = Z + \rho_t(W),\ \forall\, Z \in \mathcal{Z}_t,\ W \in \mathcal{Z}_{t+1}$
- Positive Homogeneity: $\rho_t(\tau Z) = \tau \rho_t(Z), \forall\, Z \in \mathcal{Z}_{t+1},\ \tau \ge 0$

Scandolo ('03), Riedel ('04), R.-Shapiro ('06), Cheridito-Delbaen-Kupper ('06), Föllmer-Penner ('06), Artzner-Delbaen-Eber-Heath-Ku ('07), Pflug-Römisch ('07)

Example: Conditional Mean–Semideviation

$$\rho_t(Z_{t+1}) = \mathbb{E}[Z_{t+1}|\mathcal{F}_t] + \kappa \mathbb{E}\Big[\big(Z_{t+1} - \mathbb{E}[Z_{t+1}|\mathcal{F}_t]\big)_+^s \Big| \mathcal{F}_t\Big]^{\frac{1}{s}}$$

Here $s \in [1, p]$ and $\kappa \in [0,1]$ may be $\mathcal{F}_t$-measurable

Coherent One-Step Conditional Risk Measures

Stronger assumptions about one-step measures $\rho_t : \mathcal{Z}_{t+1} \to \mathcal{Z}_t$:

- **Convexity:** $\rho_t(\lambda Z + (1-\lambda)W) \le \lambda\rho_t(Z) + (1-\lambda)\rho_t(W)$ $\forall\, \lambda \in (0,1),\; Z, W \in \mathcal{Z}_{t+1}$
- **Monotonicity:** If $Z \le W$ then $\rho_t(Z) \le \rho_t(W), \;\forall\, Z, W \in \mathcal{Z}_{t+1}$
- **Predictable Translation Equivariance:** $\rho_t(Z + W) = Z + \rho_t(W), \;\forall\, Z \in \mathcal{Z}_t,\;\; W \in \mathcal{Z}_{t+1}$
- **Positive Homogeneity:** $\rho_t(\tau Z) = \tau\rho_t(Z), \forall\, Z \in \mathcal{Z}_{t+1},\;\; \tau \ge 0$

Scandolo ('03), Riedel ('04), R.-Shapiro ('06), Cheridito-Delbaen-Kupper ('06), Föllmer-Penner ('06), Artzner-Delbaen-Eber-Heath-Ku ('07), Pflug-Römisch ('07)

Example: Conditional Mean–Semideviation

$$\rho_t(Z_{t+1}) = \mathbb{E}[Z_{t+1}|\mathcal{F}_t] + \kappa\mathbb{E}\Big[\big(Z_{t+1} - \mathbb{E}[Z_{t+1}|\mathcal{F}_t]\big)_+^s \Big| \mathcal{F}_t\Big]^{\frac{1}{s}}$$

Here $s \in [1, p]$ and $\kappa \in [0, 1]$ may be $\mathcal{F}_t$-measurable

Controlled Markov Models

- Finite state space $\mathcal{X}$
- Finite control space $\mathcal{U}$
- Feasible control sets $U : \mathcal{X} \rightrightarrows \mathcal{U}$
- Controlled transition matrices $Q(u)$
- Cost function $c : \mathcal{X} \times \mathcal{U} \times \mathcal{X} \to \mathbb{R}$
- State history $\mathcal{X}^t$ (up to time $t = 1, 2, \dots$)
- Policy $\pi_t : \mathcal{X}^t \to \mathcal{P}(\mathcal{U})$, $t = 1, 2, \dots$
  (with distributions supported on $U(x_t)$)
- Markov policy $\pi_t : \mathcal{X} \to \mathcal{P}(\mathcal{U})$, $t = 1, 2, \dots$
  (stationary if $\pi_t = \pi_1$ for all t)

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

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

Assumption: For every Markov policy the chain is absorbing with some absorbing state x_{A}

The Risk-Neutral Optimal Path Problem

$$\min_{\pi_1,\pi_2,\ldots} \mathbb{E}\left[\sum_{t=1}^{\infty} c(x_t, u_t, x_{t+1})\right]$$

with controls $u_t \sim \pi_t(x_1, \ldots, x_t)$

- The problem has an optimal solutions in form of a deterministic Markov policy
- The optimal policy can be found by dynamic programming equations
- It is sufficient to consider cost functions of form $\bar{c}(x_t, u_t)$.

Our Intention

Introduce risk aversion to both problems by replacing the expected value by dynamic risk measures

Using Dynamic Risk Measures for Markov Decision Processes

- Controlled Markov process x_t, $t = 1, \dots, T, T+1$
- Policy $\Pi = \{\pi_1, \pi_2, \dots, \pi_T\}$ defines $u_t \sim \pi_t(x_t)$
- Cost sequence $Z_t = c(x_{t-1}, u_{t-1}, x_t)$, $t = 2, \dots, T+1$
- Dynamic time-consistent risk measure

$$J(\Pi) = \rho_1\Big(c(x_1, u_1, x_2) + \rho_2\big(c(x_2, u_2, x_3) + \cdots + \rho_T(c(x_T, u_T, x_{T+1}))\cdots\big)\Big)$$

- Risk-averse optimal control problem

$$\min_{\Pi} J(\Pi)$$

Difficulty

The value of $\rho_t(\cdot)$ is $\mathcal{F}_t$-measurable and is allowed to depend on the entire history of the process. We cannot expect a Markov optimal policy if our attitude to risk depends on the whole past

New Construction of a Conditional Risk Measure

- Consider functions of the pair control–next state, in a fixed space $\mathcal{V} = \mathcal{U} \times \mathcal{X}$:

$$c_x(u, y) = c(x, u, y), \quad u \in \mathcal{U}, \quad y \in \mathcal{X}$$

- Additional argument: distribution of the control–next state pair

$$[\lambda \circ Q_x](u, y) = \lambda(u) q_{x,y}(u), \quad u \in \mathcal{U}, \quad y \in \mathcal{X}$$

The set of probability measures on $\mathcal{U} \times \mathcal{X}$:

$$\mathbb{P} = \left\{ p \in \mathcal{V} : \sum_{u \in \mathcal{U}} \sum_{y \in \mathcal{X}} p(u, y) = 1, \; p \geq 0 \right\}.$$

A measurable function $\sigma : \mathcal{V} \times \mathcal{X} \times \mathbb{P} \to \mathbb{R}$ is a *risk transition mapping* if for every $x \in \mathcal{X}$ and every $p \in \mathbb{P}$, the function $\varphi \mapsto \sigma(\varphi, x, p)$ is a coherent measure of risk on $\mathcal{V}$

Markov Risk Measures

A one-step conditional risk measure $\rho_t : \mathcal{Z}_{t+1} \to \mathcal{Z}_t$ is a Markov risk measure with respect to $\{x_t\}$, if there exists a risk transition mapping $\sigma_t : \mathcal{V} \times \mathcal{X} \times \mathbb{P} \to \mathbb{R}$ such that for all functions $w : \mathcal{X} \times \mathcal{U} \times \mathcal{X} \to \mathbb{R}$ and for all randomized controls λ on $U(x_t)$ we have

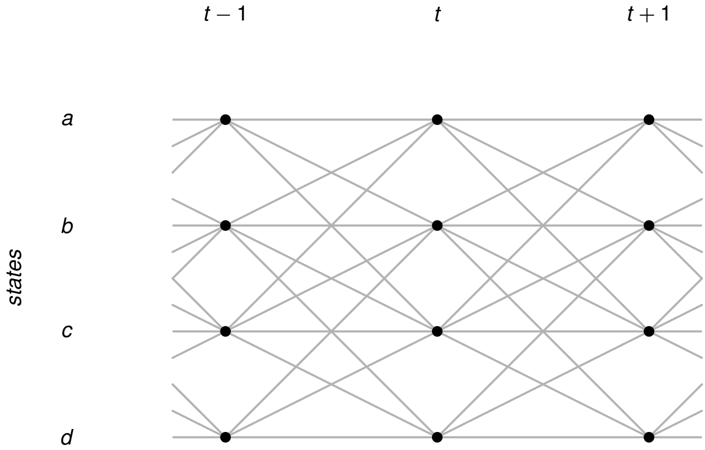
$$\rho_t\big(w(x_t, u_t, x_{t+1})\big) = \sigma_t\big(w_{x_t}, x_t, \lambda \circ Q_{x_t}\big)$$

A risk transition mapping $\sigma : \mathcal{V} \times \mathcal{X} \times \mathbb{P} \to \mathbb{R}$ is consistent with the first order stochastic dominance if for all φ, ψ in $\mathcal{V}$ and all $p, q \in \mathbb{P}$ such that $F_\varphi^p(\cdot) \le F_\psi^q(\cdot)$ we have $\sigma(\varphi, x, p) \ge \sigma(\psi, x, q)$ for all $x \in \mathcal{X}$

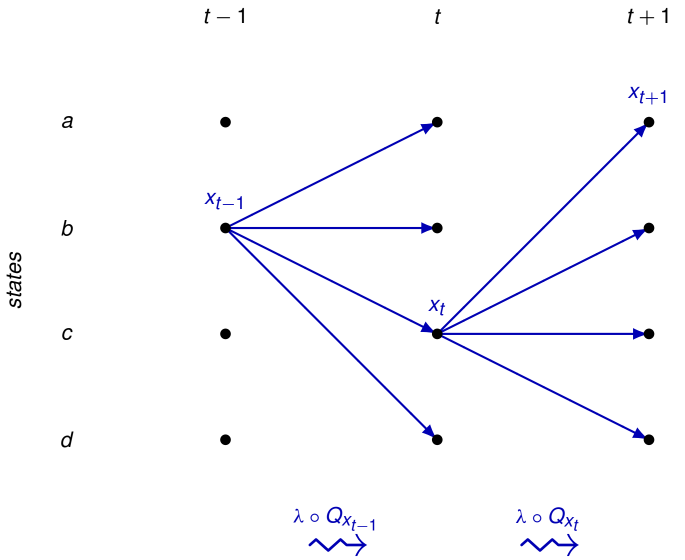
A risk transition mapping $\sigma : \mathcal{V} \times \mathcal{X} \times \mathbb{P} \to \mathbb{R}$ is proper if for all $\varphi \in \mathcal{V}$, all $x \in \mathcal{X}$, and all nontrivial $p \in \mathbb{P}$ we have

$$\sigma(\varphi, x, p) < \max\big\{\varphi(u, y) : p(u, y) > 0\big\}$$

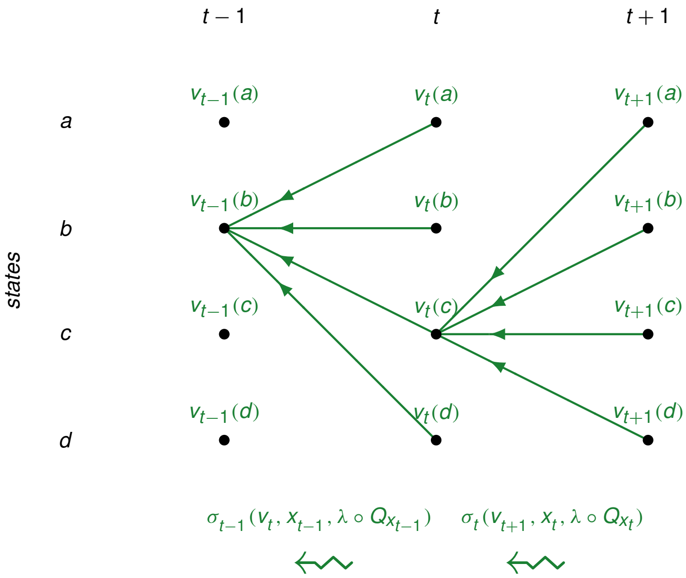
Markov Risk Evaluation



Markov Risk Evaluation



Markov Risk Evaluation



Risk of Infinite Sequences

Finite horizon risk

$$J_T(\Pi, x_1) = \rho_1\bigg(c(x_1, u_1, x_2) + \rho_2\Big(c(x_2, u_2, x_3) + \cdots \\ \cdots + \rho_T\big(c(x_T, u_T, x_{T+1})\big)\cdots\Big)\bigg)$$

Infinite horizon risk

$$J_\infty(\Pi, x_1) = \lim_{T\to\infty} J_T(\Pi, x_1)$$

Suppose the conditional risk measures ρ_t, $t = 1, 2, \ldots$ are Markov and share the same risk transition mapping $\sigma(\cdot,\cdot,\cdot)$, which is consistent with first order stochastic dominance and proper. Then for every stationary policy $\Pi = \{\pi, \pi, \ldots\}$ the limit $J_\infty(\Pi, x_1)$ is finite.

Dynamic Programming Equations

Suppose the risk transition mapping is consistent with first order dominance and proper. For a stationary policy $\Pi = \{\pi\}$ a function $v(x) \equiv J_\infty(\Pi, x)$ if and only if $v(x_\mathrm{A}) = 0$ and

$$v(x) = \sigma\big(c_x + v, x, \pi(x) \circ Q_x\big), \quad x \in \mathcal{X}$$

The optimal value function

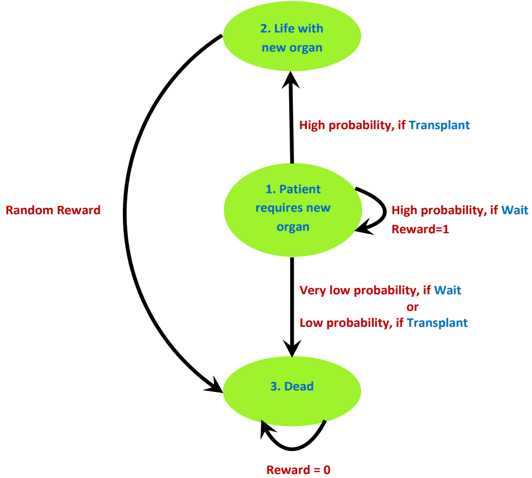
$$J^*(x) = \inf_\Pi J_\infty(\Pi, x), \quad x \in \mathcal{X}$$

In addition, suppose the risk transition mapping is continuous with respect to the third argument. Then $v(x) \equiv J^*(x)$ iff $v(x_\mathrm{A}) = 0$ and

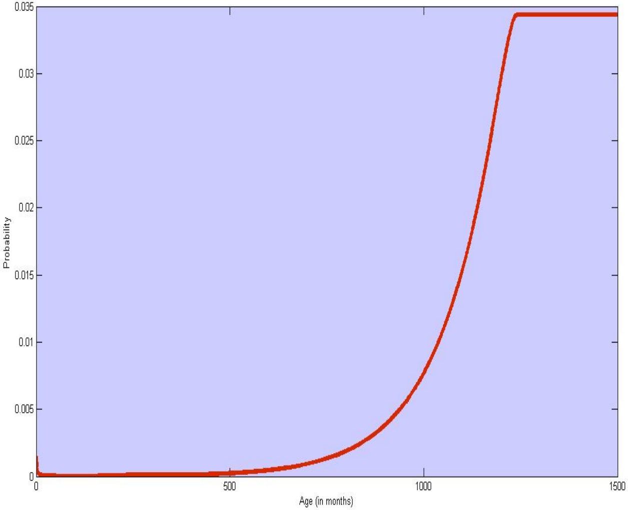
$$v(x) = \min_{\pi \in \mathcal{P}(U(x))} \sigma\big(c_x + v, x, \pi \circ Q_x\big), \quad x \in \mathcal{X}$$

The minimizer $\pi^*(x)$, $x \in \mathcal{X}$, defines an optimal randomized policy $\Pi^* = \{\pi^*\}$.

Example: Organ Transplant



Probability of Death, Jasiulewicz ('97)



- **Expected Total Reward:**
  The optimal policy is to wait

- **Mean Semi-Deviation with Deterministic Policies:**
  The optimal policy is to transplant

- **Mean Semi-Deviation with Randomized Policies:**
  Wait with probability 0.993983 and transplant with probability 0.006017

Summary

- Risk-averse stochastic shortest (longest) path problems are considered
- Markov conditional risk measures are introduced
- Finiteness of the overall risk is proved
- Dynamic programming equations are derived
- Randomized policies may be optimal
- Interesting applications follow