How Should a Robot Assess Risk? Towards an Axiomatic Theory of Risk in Robotics

Marco Pavone
Stanford University, Department of Aeronautics and Astronautics

Joint work with Y.-L. Chow, A. Majumdar, S. Mannor, S. Singh, A. Tamar

Autonomous Systems Lab

Planning under uncertainty

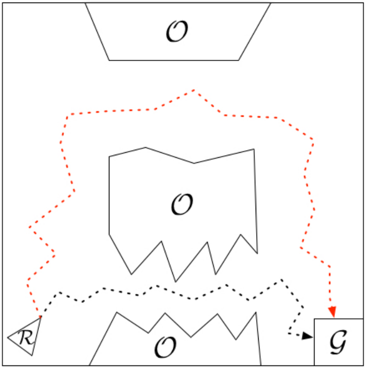
MDP formulation

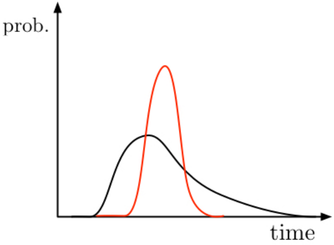
- States x, actions u
- Transitions $P(\cdot|x,u)$
- Cost $c(x_k, u_k)$
- Policy $u_k \sim \pi(x_k)$
- Standard MDP objective:

$$\min_{\pi} \mathbb{E}_P^{\pi}\left[\sum_{k=0}^{N-1} c(x_k, u_k) + c(x_N)\right]$$

- Many algorithms...

Why should a robot be risk-aware?





Risk-aware decision making

- Risk-aware decision making:

$$\min_{\pi} \rho_P^{\pi}\left[\sum_{k=0}^{N-1} c(x_k, u_k) + c(x_N)\right],$$

  where ρ is a *risk measure*

- Some possible choices:
  - Worst-case assessment: $\rho(X) = \max_i X_i$
  - Expected utility theory: $\rho(X) = \mathbb{E}\left(u(X)\right)$, where u is a *disutility function*
  - Mean-risk models: $\rho(X) = \mathbb{E}\left(X\right) + \gamma\mathbb{D}[X]$, where $\mathbb{D}[X]$ is a measure of uncertainty (e.g., variance)
  - Value at risk: $\rho(X) = \min\{x \mid \mathbb{P}[X > x] \leq \alpha\} =: \mathrm{VaR}_\alpha(X)$

Potential pitfalls

- Mean-variance risk metric: $\mathbb{E}[X] + \mathrm{Variance}[X]$

Outcome	ω_1	ω_2	ω_3	ω_4
Probability	0.25	0.25	0.25	0.25
X	1	2	3	4
X'	2	2	3	4

Potential pitfalls

- Mean-variance risk metric: $\mathbb{E}[X] + \mathrm{Variance}[X]$

Outcome	ω_1	ω_2	ω_3	ω_4
Probability	0.25	0.25	0.25	0.25
X	1	2	3	4
X'	2	2	3	4

The robot would strictly prefer X'!

Towards an axiomatic theory of risk in robotics

[Majudmar, Pavone, ISRR '17 (submitted)]

- Parallels recent efforts in the finance and OR communities

Towards an axiomatic theory of risk in robotics

[Majudmar, Pavone, ISRR ’17 (submitted)]

- Parallels recent efforts in the finance and OR communities
- Rationality axioms:

Towards an axiomatic theory of risk in robotics

[Majudmar, Pavone, ISRR '17 (submitted)]

- Parallels recent efforts in the finance and OR communities
- Rationality axioms:
  1. Monotonicity: $X \leq X' \Rightarrow \rho(X) \leq \rho(X')$

Towards an axiomatic theory of risk in robotics

[Majudmar, Pavone, ISRR '17 (submitted)]

- Parallels recent efforts in the finance and OR communities
- Rationality axioms:
  1. Monotonicity: $X \leq X' \Rightarrow \rho(X) \leq \rho(X')$
  2. Translational invariance: $\rho(X + a) = \rho(X) + a, \quad \forall a \in \mathbb{R}$

Towards an axiomatic theory of risk in robotics

[Majudmar, Pavone, ISRR '17 (submitted)]

- Parallels recent efforts in the finance and OR communities
- Rationality axioms:
  1. Monotonicity: $X \leq X' \Rightarrow \rho(X) \leq \rho(X')$
  2. Translational invariance: $\rho(X + a) = \rho(X) + a, \quad \forall a \in \mathbb{R}$
  3. Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X), \quad \forall \lambda \geq 0$

Towards an axiomatic theory of risk in robotics

[Majudmar, Pavone, ISRR '17 (submitted)]

- Parallels recent efforts in the finance and OR communities
- Rationality axioms:
  1. Monotonicity: $X \leq X' \Rightarrow \rho(X) \leq \rho(X')$
  2. Translational invariance: $\rho(X + a) = \rho(X) + a, \quad \forall a \in \mathbb{R}$
  3. Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X), \quad \forall \lambda \geq 0$
  4. Sub-additivity: $\rho(X + X') \leq \rho(X) + \rho(X')$

Towards an axiomatic theory of risk in robotics

[Majudmar, Pavone, ISRR '17 (submitted)]

- Parallels recent efforts in the finance and OR communities
- Rationality axioms:
  1. Monotonicity: $X \leq X' \Rightarrow \rho(X) \leq \rho(X')$
  2. Translational invariance: $\rho(X + a) = \rho(X) + a, \quad \forall a \in \mathbb{R}$
  3. Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X), \quad \forall \lambda \geq 0$
  4. Sub-additivity: $\rho(X + X') \leq \rho(X) + \rho(X')$

Metrics satisfying axioms A1-A4 are called *coherent risk metrics*

Static risk: risk-sensitive MDPs

[Chow, Tamar, Pavone, Mannor, NIPS '15 and TAC '17]

Representation theorem for coherent risk [(Artzner et al., MF "99)]

- ρ is coherent iff there exists a *convex* set of probability distributions $\mathcal{Q}$ such that

$$\rho(X) = \max_{Q \in \mathcal{Q}} \mathbb{E}_Q(X)$$

Static risk: risk-sensitive MDPs

[Chow, Tamar, Pavone, Mannor, NIPS '15 and TAC '17]

Representation theorem for coherent risk [(Artzner et al., MF "99)]

- ρ is coherent iff there exists a *convex* set of probability distributions $\mathcal{Q}$ such that

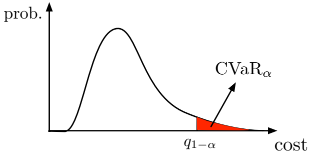
$$\rho(X) = \max_{Q \in \mathcal{Q}} \mathbb{E}_Q(X)$$

CVaR definition

- $q_\alpha(X)$ – α quantile
- α-CVaR

$$\text{CVaR}_\alpha(X) := \mathbb{E}\left(X | X \geq q_{1-\alpha}(X)\right)$$

- Sensitive to rare, disastrous cases



$$\mathcal{Q}_{\text{CVaR}} = \left\{ Q \in \mathbb{R}^L \,\middle|\, 0 \leq Q(j) \leq \frac{P(j)}{\alpha}, \forall j, \text{ and } \sum_j Q(j) = 1 \right\}$$

- Risk-sensitive MDP:

$$\min_{\pi \in \Pi_H} \quad \text{CVaR}_\alpha \left[\sum_{k=0}^{\infty} \gamma^k c(x_k, u_k) \right]$$

Solution algo and robustness

Algorithm:

- Key idea: state augmentation, α becomes a state variable
- On augmented state space, Markovian policies are optimal
- Approximate DP algorithm via linear interpolation with provable error bounds

Solution algo and robustness

Algorithm:

- Key idea: state augmentation, α becomes a state variable
- On augmented state space, Markovian policies are optimal
- Approximate DP algorithm via linear interpolation with provable error bounds

Connection to model uncertainty:

Temporally-budgeted perturbations

- Multiplicative perturbations: $\hat{P}(x_{k+1}|x_k) = P(x_{k+1}|x_k) \cdot \delta_k(x_{k+1}|x_k)$
- Budget: $\delta_1(x_1|x_0)\delta_2(x_2|x_1)\cdots\delta_N(x_N|x_{N-1}) \leq \eta$
- Set of possible perturbations Δ_η

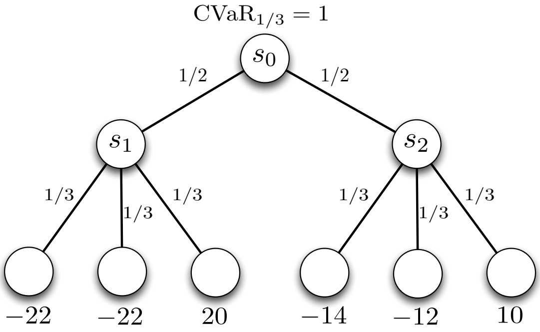
CVaR as robustness measure

$$\text{CVaR}_{\frac{1}{\eta}}\left[\sum_{k=1}^N c(x_k)\right] = \sup_{(\delta_1,\ldots,\delta_N)\in\Delta_\eta} \mathbb{E}_{\hat{P}}\left[\sum_{k=1}^N c(x_k)\right]$$

Dynamic risk: risk-sensitive MPC

[Chow, Singh, Majumdar, Pavone, TAC '17 (submitted)]

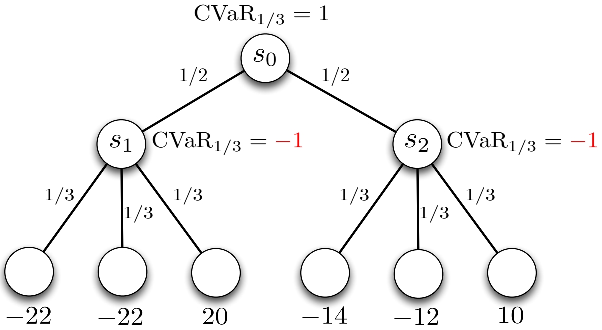
- Time inconsistency in risk evaluation



Dynamic risk: risk-sensitive MPC

[Chow, Singh, Majumdar, Pavone, TAC '17 (submitted)]

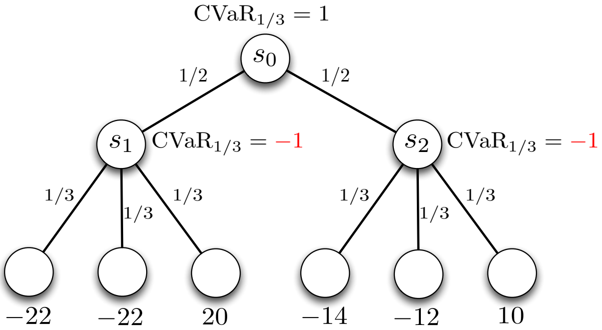
- Time inconsistency in risk evaluation



Dynamic risk: risk-sensitive MPC

[Chow, Singh, Majumdar, Pavone, TAC '17 (submitted)]

- Time inconsistency in risk evaluation



⇒ multi-stage decision making requires a refinement of the notion of risk

Time-Consistent, Dynamic Risk Measures

- Key result: any time-consistent, dynamic risk measure can be written as [Ruszczynski, JMP '10]

$$\rho_{k,N}(X_k, \ldots, X_N) = X_k + \rho_k(X_{k+1} + \rho_{k+1}(X_{k+2} + \ldots + \rho_{N-2}(X_{N-1} + \rho_{N-1}(X_N)) \ldots))$$

Time-Consistent, Dynamic Risk Measures

- Key result: any time-consistent, dynamic risk measure can be written as [Ruszczynski, JMP '10]

$$\rho_{k,N}(X_k, \ldots, X_N) = X_k + \rho_k(X_{k+1} + \rho_{k+1}(X_{k+2} + \ldots + \rho_{N-2}(X_{N-1} + \rho_{N-1}(X_N)) \ldots))$$

- Example before:

$$\rho_{0,2}(c(x_2)) = \mathrm{CVaR}_{1/3}(\mathrm{CVaR}_{1/3}(c(x_2))) = -1$$

Risk-sensitive MPC

[Chow, Singh, Majumdar, Pavone, TAC '17 (submitted)]

Optimization problem $\mathcal{MPC}$ — Given initial state $x_{k|k}$ and a prediction horizon $N \geq 1$, solve

$$\begin{aligned} \min_{\pi_{k|k},\ldots,\pi_{k+N-1|k}} \quad & J\left(x_{k|k}, \pi_{k|k}, \ldots, \pi_{k+N-1|k}, F\right) \\ \text{such that} \quad & x_{k+h+1|k} = A(w_{k+h})x_{k+h|k} + B(w_{k+h})\pi_{k+h|k}(x_{k+h|k}), \\ & \qquad \text{for } h \in \{0, \ldots, N-1\} \end{aligned}$$

where J is a time-consistent, dynamic risk measure

Risk-sensitive MPC

[Chow, Singh, Majumdar, Pavone, TAC '17 (submitted)]

Optimization problem $\mathcal{MPC}$ — Given initial state $x_{k|k}$ and a prediction horizon $N \geq 1$, solve

$$\begin{aligned}
\min_{\pi_{k|k},\ldots,\pi_{k+N-1|k}} \quad & J\left(x_{k|k}, \pi_{k|k}, \ldots, \pi_{k+N-1|k}, F\right) \\
\text{such that} \quad & x_{k+h+1|k} = A(w_{k+h})x_{k+h|k} + B(w_{k+h})\pi_{k+h|k}(x_{k+h|k}), \\
& \text{for } h \in \{0, \ldots, N-1\}
\end{aligned}$$

where J is a time-consistent, dynamic risk measure

Key results:

1. Compositional structure of risk leads to a DP-style solution
2. Alternatively, dual characterization of risk leads to a convex programming solution

Conclusions

Risk-sensitive and resilient decision making:

- Requires a tractable and intuitive theory that is compatible with the modern theory of risk
- Coherent risk metrics (and refinements) appear as a promising tool

Points of discussion:

1. Different or additional axioms?
2. Need of time consistency?
3. How to choose a risk metric?
4. Risk-based legal frameworks for AI?

Acknowledgement: This work is supported by ONR, Science of Autonomy Program and YIP Program.