Robust-Adaptive Control of Linear Systems: beyond Quadratic Costs

Edouard Leurent[1,2],
Odalric-Ambrym Maillard[1],
Denis Efimov[1]

[1]Univ. Lille, Inria, CNRS,
Centrale Lille, UMR 9189 – CRIStAL,
[2]Renault Group

Most RL algorithms rely on trial-and-error

- random exploration
- optimism in the face of uncertainty

Our work: ensure safety at all time

- pessimism in the face of uncertainty
- learn from observations

Dynamical priors with structured uncertainty

State x, control u, bounded disturbance ω

$$\dot{x}(t) = A(\theta)x(t) + Bu(t) + \omega(t),$$

$$A(\theta) = A + \sum_{i=1}^{d} \theta_i \phi_i,$$

where $A, B, \phi_0, \dots, \phi_d$ are known.

Performance measure

- Discounted return $\sum_n \gamma^n R(x(t_n))$.
- Reward R only assumed to be bounded.

Given N observations,

1. Build a confidence region for the dynamics

$$\mathbb{P}\left(\theta \in \mathcal{C}_{N,\delta}\right) \geq 1 - \delta,$$

2. Maximize the worst-case performance V^r

$$\sup_{\mathbf{u}\in(\mathbb{R}^q)^{\mathbb{N}}} \underbrace{\inf_{\substack{\theta\in\mathcal{C}_{N,\delta}\\ \omega\in[\underline{\omega},\overline{\omega}]^{\mathbb{R}}}} \left[\sum_{n=N+1}^{\infty} \gamma^n R(x(t_n)) \middle| \mathbf{u}, \theta, \omega\right]}_{V^r(\mathbf{u})},$$

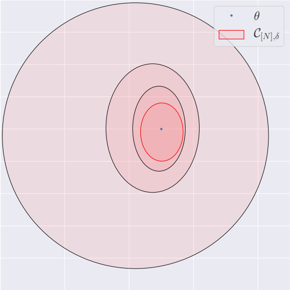
Related work

- Robust Dynamic Programming (finite states only)
- Linear-Quadratic setting (quadratic costs only)

Model Estimation

Confidence ellipsoid from non-asymptotic linear regression.

$$\|\theta_{N,\lambda} - \theta\|_{G_{N,\lambda}} \leq \beta_N(\delta)$$

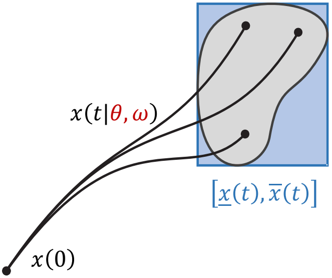


To optimise V^r $\rightarrow$ approximate it by a tractable surrogate $\hat{V}^r$

Interval Prediction

Propagate uncertainty $\mathcal{C}_{N,\delta}$ through time to bound the state

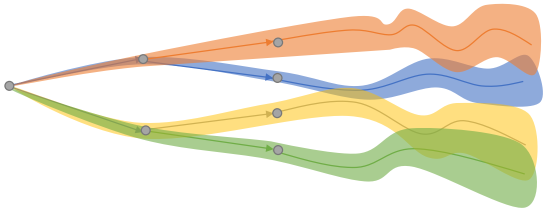
$$\underline{x}(t) \leq x(t) \leq \overline{x}(t), \quad \forall t \geq t_N$$



Pessimistic Planning

Surrogate pessimistic return, maximized by tree based-planning

$$\hat{V}^r(\mathbf{u}) \overset{\text{def}}{=} \sum_{n=N+1}^{\infty} \gamma^n \min_{x \in [\underline{x}(t_n), \overline{x}(t_n)]} R(x).$$



Theorem (Lower bound)

$$\hat{V}^r(\mathbf{u}) \leq V^r(\mathbf{u}) \leq V(u)$$

Theorem (Suboptimality bound)

Under two conditions:

1. *Lipschitz rewards R;*
2. *a stability condition over $A(\theta_N)$, for $N > N_0$*

we can bound the suboptimality with probability $1-\delta$ as:

$$\underbrace{V(u_\star) - V(u_K)}_{\text{suboptimality}} \leq \underbrace{\Delta_\omega}_{\text{robustness to disturbances}} + \underbrace{\mathcal{O}\left(\frac{\beta_N(\delta)^2}{\lambda_{\min}(G_{N,\lambda})}\right)}_{\text{estimation error}} + \underbrace{\mathcal{O}\left(K^{-\frac{\log 1/\gamma}{\log \kappa}}\right)}_{\text{planning error}}$$

Corollary (Asymptotic Near-Optimality)

If the features $\phi_i x_n$ are persistently excited,

- *the stability condition can be relaxed to $A(\theta)$ only*
- *the suboptimality bound takes the more explicit form*

$$V(u_\star) - V(u_K) \leq \Delta_\omega + \mathcal{O}\left(\frac{\log(N^{d/2}/\delta)}{N}\right) + \mathcal{O}\left(K^{-\frac{\log 1/\gamma}{\log \kappa}}\right)$$

which ensures near-optimality as $N \to \infty$ and $K \to \infty$.

Thank You!

I am looking for a postdoctoral position.