Convex Parameterization for Linear Dynamical Systems

Sham Kakade and Wen Sun

CS 6789: Foundations of Reinforcement Learning

Setting: Linear Dynamical System and Convex Cost functions

$$x_{t+1} = Ax_t + Bu_t + w_t, x_0 \sim \mathscr{D}, w_t \sim \mathscr{N}(0,\sigma^2 I)$$

Setting: Linear Dynamical System and Convex Cost functions

$$x_{t+1} = Ax_t + Bu_t + w_t, x_0 \sim \mathscr{D}, w_t \sim \mathcal{N}(0, \sigma^2 I)$$

$$\frac{1}{2} x^T Q x + \frac{1}{2} u^T R u$$

$$\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{h=0}^{H-1} c(x_h, \overset{u}{\cancel{a}}_h) \right]$$

Convex Function!

Setting: Linear Dynamical System and Convex Cost functions

$$x_{t+1} = Ax_t + Bu_t + w_t, x_0 \sim \mathscr{D}, w_t \sim \mathscr{N}(0,\sigma^2 I)$$

$$\max_{\pi} \mathbb{E}_{\pi}\left[\sum_{h=0}^{H-1} c(x_h, a_h)\right]$$

Convex Function!

We cannot guarantee the optimal policy is linear anymore…

Setting: Linear Dynamical System and Convex Cost functions

$$x_{t+1} = Ax_t + Bu_t + w_t, x_0 \sim \mathscr{D}, w_t \sim \mathscr{N}(0,\sigma^2 I)$$

$$\max_{\pi} \mathbb{E}_{\pi}\left[\sum_{h=0}^{H-1} c(x_h, a_h)\right]$$

Convex Function!

We cannot guarantee the optimal policy is linear anymore…

But we consider a restricted policy class

$$\Pi = \{\pi = \{-K_t\}_{t=0}^{H-1} : K_t \in \mathscr{K}, \forall t\}$$

Setting: Linear Dynamical System and Convex Cost functions

$$x_{t+1} = Ax_t + Bu_t + w_t, x_0 \sim \mathcal{D}, w_t \sim \mathcal{N}(0, \sigma^2 I)$$

$\begin{pmatrix} x_{t+1} \\ u_t \end{pmatrix}$

$$\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{h=0}^{H-1} c(x_h, a_h) \right]$$

Convex Function!

We cannot guarantee the optimal policy is linear anymore…

But we consider a restricted policy class

$$\Pi = \{\pi = \{-K_t\}_{t=0}^{H-1} : K_t \in \mathcal{K}, \forall t\}$$

Goal: find the best linear controllers: $\{-K_t^\star\}_{t=0}^{H-1} := \arg\min_{\pi \in \Pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^{H-1} c(x_t, u_t) \right]$

A New Convex Parameterization of Controllers

$\frac{1}{2}x^T Q x + \frac{1}{2}u^T R u$

Unfortunately, even with quadratic cost, $\max_{\pi} \mathbb{E}_{\pi}\left[\sum_{h=0}^{H-1} c(x_h, a_h)\right]$ is not convex

with respect to $\pi := \{-K_t\}_{t=0}^{H-1}$

Δ

A New Convex Parameterization of Controllers

Unfortunately, even with quadratic cost, $\max_{\pi} \mathbb{E}_{\pi}\left[\sum_{h=0}^{H-1} c(x_h, a_h)\right]$ is not convex with respect to $\pi := \{-K_t\}_{t=0}^{H-1}$

Solution: Over-Parameterization

A New Convex Parameterization of Controllers

Let us consider an arbitrary linear controller $\pi := \{-K_t\}_{t=0}^{H-1}$

A New Convex Parameterization of Controllers

Let us consider an arbitrary linear controller $\pi := \{-K_t\}_{t=0}^{H-1}$

Assume we roll out π and the system is at x_t, we can compute all previous noises:

$$w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots$$

x_t

$$u_t = M_0 x_0 + M_1 w_0 + \cdots + M_t w_{t-1}$$

A New Convex Parameterization of Controllers

Let us consider an arbitrary linear controller $\pi := \{-K_t\}_{t=0}^{H-1}$

Assume we roll out π and the system is at x_t, we can compute all previous noises:

$$w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots$$

$u_t = -K_t x_t$

A New Convex Parameterization of Controllers

Let us consider an arbitrary linear controller $\pi := \{-K_t\}_{t=0}^{H-1}$

Assume we roll out π and the system is at x_t, we can compute all previous noises:

$$w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots$$

$u_t = -K_t x_t$

$x_t = Ax_{t-1} + Bu_{t-1} + w_{t-1}$

$= -K_t w_{t-1} - K_t(A - BK_{t-1})x_{t-1}$

$u_{t-1} = -K_{t-1} x_{t-1}$

Δ

$x_{t-1} = Ax_{t-2} + Bu_{t-2} + w_{t-2}$

$u_{t-2} = -K_{t-2} \cdot x_{t-2}$

A New Convex Parameterization of Controllers

Let us consider an arbitrary linear controller $\pi := \{-K_t\}_{t=0}^{H-1}$

Assume we roll out π and the system is at x_t, we can compute all previous noises:

$$w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots$$

$$u_t = -K_t x_t$$

$$= -K_t w_{t-1} - K_t(A - BK_{t-1})x_{t-1}$$

$$= -K_t w_{t-1} - K_t(A - BK_{t-1})\underbrace{(Ax_{t-2} - BK_{t-2}x_{t-2} + w_{t-2})}_{= x_{t-1}}$$

A New Convex Parameterization of Controllers

Let us consider an arbitrary linear controller $\pi := \{-K_t\}_{t=0}^{H-1}$

Assume we roll out π and the system is at x_t, we can compute all previous noises:

$$w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots$$

$$u_t = -K_t x_t$$

$$= -K_t w_{t-1} - K_t(A - BK_{t-1})x_{t-1}$$

$$= -K_t w_{t-1} - K_t(A - BK_{t-1})(Ax_{t-2} - BK_{t-2}x_{t-2} + w_{t-2})$$

$$= \underbrace{-K_t}_{:=M_{t-1;t}} w_{t-1} \underbrace{-K_t(A - BK_{t-1})}_{:=M_{t-2;t}} w_{t-2} \underbrace{-K_t(A - BK_{t-1})(A - BK_{t-2})}_{:=M_{t-3;t}} x_{t-2}$$

A

$x_{t-2} = Ax_{t-3} + Bu_{t-3} + w_{t-3}$

$-K_{t-3} \cdot x_{t-3}$

$$\underbrace{-K_t(A - BK_{t-1})(A - BK_{t-2})(A - BK_{t-3}) \cdot w_{t-3}}$$

A New Convex Parameterization of Controllers

Let us consider an arbitrary linear controller $\pi := \{-K_t\}_{t=0}^{H-1}$

Assume we roll out π and the system is at x_t, we can compute all previous noises:

$$w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots$$

$$u_t = -K_t x_t$$

$$= -K_t w_{t-1} - K_t(A - BK_{t-1})x_{t-1}$$

$$= -K_t w_{t-1} - K_t(A - BK_{t-1})(Ax_{t-2} - BK_{t-2}x_{t-2} + w_{t-2})$$

$$= \underbrace{-K_t}_{:=M_{t-1;t}} w_{t-1} \underbrace{-K_t(A - BK_{t-1})}_{:=M_{t-2;t}} w_{t-2} \underbrace{-K_t(A - BK_{t-1})(A - BK_{t-2})}_{:=M_{t-3;t}} x_{t-2}$$

$$= \underbrace{\left[-K_t \left(\prod_{\tau=1}^{t} (A - BK_{t-\tau}) \right) \right]}_{M_t} x_0 + \sum_{\tau=0}^{t-1} \underbrace{\left[-K_t \prod_{h=1}^{t-1-\tau} (A - BK_{t-h}) \right]}_{M_{\tau;t}} w_\tau$$

A New Convex Parameterization of Controllers

Let us consider an arbitrary linear controller $\pi := \{-K_t\}_{t=0}^{H-1}$

Assume we roll out π and the system is at x_t, we can compute all previous noises:

$$w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots$$

$$u_t = -K_t x_t \quad \Longleftrightarrow \quad u_t = M_t x_0 + M_{0;t} w_0 + M_{1;t} w_1 + \ldots + M_{t-1;t} w_{t-1}$$

$$= -K_t w_{t-1} - K_t(A - BK_{t-1})x_{t-1}$$

$$= -K_t w_{t-1} - K_t(A - BK_{t-1})(Ax_{t-2} - BK_{t-2}x_{t-2} + w_{t-2})$$

$$= \underbrace{-K_t}_{:=M_{t-1;t}} w_{t-1} \underbrace{-K_t(A - BK_{t-1})}_{:=M_{t-2;t}} w_{t-2} \underbrace{-K_t(A - BK_{t-1})(A - BK_{t-2})}_{:=M_{t-3;t}} x_{t-2}$$

$$= \underbrace{\left[-K_t\left(\prod_{\tau=1}^{t}(A - BK_{t-\tau})\right)\right]}_{M_t} x_0 + \sum_{\tau=0}^{t-1} \underbrace{\left[-K_t \prod_{h=1}^{t-1-\tau}(A - BK_{t-h})\right]}_{M_{\tau;t}} w_\tau$$

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{\{M_t, M_{t-1;t}, \ldots M_{0;t}\}\right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{\{M_t, M_{t-1;t}, \ldots M_{0;t}\}\right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

x_0
✓

x_0
✓

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{\{M_t, M_{t-1;t}, \ldots M_{0;t}\}\right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

$$x_0 \longrightarrow \underbrace{u_0 = -K_0 x_0}$$

$$x_0 \longrightarrow u_0 = M_0 x_0$$

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{\{M_t, M_{t-1;t}, \ldots M_{0;t}\}\right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

$$x_0 \longrightarrow u_0 = -K_0 x_0 \longrightarrow x_1$$

$$x_0 \longrightarrow u_0 = M_0 x_0 \longrightarrow x_1, w_0$$

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{\{M_t, M_{t-1;t}, \ldots M_{0;t}\}\right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

$$x_0 \longrightarrow u_0 = -K_0 x_0 \longrightarrow x_1 \longrightarrow u_1 = -K_1 x_1$$

$$x_0 \longrightarrow u_0 = M_0 x_0 \longrightarrow x_1, w_0$$

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{\{M_t, M_{t-1;t}, \ldots M_{0;t}\}\right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

$$x_0 \longrightarrow u_0 = -K_0 x_0 \longrightarrow x_1 \longrightarrow u_1 = -K_1 x_1$$

$$x_0 \longrightarrow u_0 = M_0 x_0 \longrightarrow x_1, w_0 \longrightarrow u_1 = M_1 x_0 + M_{0;1} w_0$$

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{\{M_t, M_{t-1;t}, \ldots M_{0;t}\}\right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

$$x_0 \longrightarrow u_0 = -K_0 x_0 \longrightarrow x_1 \longrightarrow u_1 = -K_1 x_1 \longrightarrow \underline{x_2}$$

$$x_0 \longrightarrow u_0 = M_0 x_0 \longrightarrow x_1, w_0 \longrightarrow u_1 = M_1 x_0 + M_{0;1} w_0 \longrightarrow x_2, w_1$$

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{ \{M_t, M_{t-1;t}, \ldots M_{0;t}\} \right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

$$x_0 \longrightarrow u_0 = -K_0 x_0 \longrightarrow x_1 \longrightarrow u_1 = -K_1 x_1 \longrightarrow x_2 \longrightarrow u_2 = -K_2 x_2$$

$$x_0 \longrightarrow u_0 = M_0 x_0 \longrightarrow x_1, w_0 \longrightarrow u_1 = M_1 x_0 + M_{0;1} w_0 \longrightarrow x_2, w_1$$

A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization $\left\{\{M_t, M_{t-1;t}, \ldots M_{0;t}\}\right\}_{t=0}^{H-1}$, that generates the same sequence trajectory, given a any fixed x_0, and fixed noise $w_0, \ldots, w_{H-1}$.

$$x_0 \longrightarrow u_0 = -K_0 x_0 \longrightarrow x_1 \longrightarrow u_1 = -K_1 x_1 \longrightarrow x_2 \longrightarrow u_2 = -K_2 x_2$$

$$x_0 \longrightarrow u_0 = M_0 x_0 \longrightarrow x_1, w_0 \longrightarrow u_1 = M_1 x_0 + M_{0;1} w_0 \longrightarrow x_2, w_1 \longrightarrow u_2 = M_2 x_0 + M_{0;2} w_0 + M_{1;2} w_1$$

A New Convex Parameterization of Controllers

[Claim] Given controller $\widetilde{\pi} := \left\{ \{M_t, M_{0;t}, \ldots M_{t-1;t}\} \right\}_{t=0}^{H-1}$, u_t & x_t are all linear with respect to the parameters, $\forall t$

A New Convex Parameterization of Controllers

[Claim] Given controller $\widetilde{\pi} := \left\{ \{ M_t, M_{0;t}, \ldots M_{t-1;t} \} \right\}_{t=0}^{H-1}$, u_t & x_t are all linear with respect to the parameters, $\forall t$

$$x_0, u_0 = M_0 x_0 \quad \checkmark$$

A New Convex Parameterization of Controllers

[Claim] Given controller $\widetilde{\pi} := \left\{ \{M_t, M_{0;t}, \ldots M_{t-1;t}\} \right\}_{t=0}^{H-1}$, u_t & x_t are all linear with respect to the parameters, $\forall t$

$$x_0, u_0 = M_0 x_0$$

linear

$$x_1 = Ax_0 + BM_0 x_0 + w_0$$

A New Convex Parameterization of Controllers

[Claim] Given controller $\widetilde{\pi} := \left\{ \{M_t, M_{0;t}, \ldots M_{t-1;t}\} \right\}_{t=0}^{H-1}$, u_t & x_t are all linear with respect to the parameters, $\forall t$

$$x_0, u_0 = M_0 x_0$$

$$x_1 = Ax_0 + BM_0 x_0$$

$$u_1 = M_1 x_0 + M_{0;1} w_0$$

$\leftarrow$ u_1 linear wrt M_s

A

A New Convex Parameterization of Controllers

[Claim] Given controller $\widetilde{\pi} := \left\{ \{M_t, M_{0;t}, \ldots M_{t-1;t}\} \right\}_{t=0}^{H-1}$, u_t & x_t are all linear with respect to the parameters, $\forall t$

$$x_0, u_0 = M_0 x_0$$

$$x_1 = Ax_0 + BM_0 x_0$$

$$u_1 = M_1 x_0 + M_{0;1} w_0$$

$$x_2 = Ax_1 + BM_1 x_0 + BM_{0;1} w_0$$

x_2 is linear wrt M_s

A New Convex Parameterization of Controllers

[Claim] Given controller $\widetilde{\pi} := \left\{ \{ M_t, M_{0;t}, \ldots M_{t-1;t} \} \right\}_{t=0}^{H-1}$, $u_t \;\&\; x_t$ are all linear with respect to the parameters, $\forall t$

$$x_0, u_0 = M_0 x_0$$

$$x_1 = Ax_0 + BM_0 x_0$$

$$u_1 = M_1 x_0 + M_{0;1} w_0$$

convex

$$x_2 = Ax_1 + BM_1 x_0 + BM_{0;1} w_0$$

[Claim] Given controller $\widetilde{\pi} := \left\{ \{ M_t, M_{0;t}, \ldots M_{t-1;t} \} \right\}_{t=0}^{H-1}$, $\sum_{t=0}^{H-1} c(x_t, u_t)$ is convex with respect the parameters, $\forall t$

[Claim] Given controller $\widetilde{\pi} := \left\{ \{M_t, M_{0;t}, \ldots M_{t-1;t}\} \right\}_{t=0}^{H-1}$**,**

$\mathbb{E}_{\widetilde{\pi}} \left[\sum_{t=0}^{H-1} c(x_t, u_t) \right]$ **is convex with respect the parameters,** $\forall t$

[Claim] Given controller $\widetilde{\pi} := \left\{ \{M_t, M_{0;t}, \ldots M_{t-1;t}\} \right\}_{t=0}^{H-1}$,

$\mathbb{E}_{\widetilde{\pi}}\left[\sum_{t=0}^{H-1} c(x_t, u_t) \right]$ **is convex with respect the parameters, $\forall t$**

Convexity allows to perform Gradient Descent directly on parameters $\left\{ \{M_t, M_{0;t}, \ldots M_{t-1;t}\} \right\}_{t=0}^{H-1}$

[Claim] Given controller $\widetilde{\pi} := \left\{ \{ M_t, M_{0;t}, \ldots M_{t-1;t} \} \right\}_{t=0}^{H-1}$,

$\mathbb{E}_{\widetilde{\pi}} \left[\sum_{t=0}^{H-1} c(x_t, u_t) \right]$ is convex with respect the parameters, $\forall t$

Convexity allows to perform Gradient Descent directly on parameters

$$\left\{ \{ M_t, M_{0;t}, \ldots M_{t-1;t} \} \right\}_{t=0}^{H-1}$$

Over-Parameterized:

[Claim] Given controller $\widetilde{\pi} := \left\{ \{ M_t, M_{0;t}, \ldots M_{t-1;t} \} \right\}_{t=0}^{H-1}$,

$\mathbb{E}_{\widetilde{\pi}} \left[\sum_{t=0}^{H-1} c(x_t, u_t) \right]$ is convex with respect the parameters, $\forall t$

Convexity allows to perform Gradient Descent directly on parameters

$$\left\{ \{ M_t, M_{0;t}, \ldots M_{t-1;t} \} \right\}_{t=0}^{H-1}$$

$K_t \in \mathbb{R}^{k \times d}$

Over-Parameterized:

For $\pi := \{-K_t\}_{t=0}^{H-1}$, we have $\underset{\Delta}{H} \times \underline{d \times k}$ parameters

[Claim] Given controller $\widetilde{\pi} := \left\{\{M_t, M_{0;t}, \ldots M_{t-1;t}\}\right\}_{t=0}^{H-1}$,

$$\mathbb{E}_{\widetilde{\pi}}\left[\sum_{t=0}^{H-1} c(x_t, u_t)\right]$$ **is convex with respect the parameters, $\forall t$**

Convexity allows to perform Gradient Descent directly on parameters

$K_t \leftrightarrow M_0 \; M_{0;t} \cdots M_{t-1;t}$

$$\left\{\{M_t, M_{0;t}, \ldots M_{t-1;t}\}\right\}_{t=0}^{H-1}$$

Over-Parameterized:

For $\pi := \{-K_t\}_{t=0}^{H-1}$, we have $H \times d \times k$ parameters

For $\widetilde{\pi} := \left\{\{M_t, M_{0;t}, \ldots M_{t-1;t}\}\right\}_{t=0}^{H-1}$, we have $\approx H^2 \times d \times k$ parameters

What we covered:

1. LQR formulation and DP for LQR (Riccati Equation)

2. You don't really need to remember the exact quadratic formulation and linear control, I always derive them from scratch when I use LQR (but it could be time consuming..)

3. Another form of over-parameterized controllers which leads to a convex parameterization (hence we can do gradient descent).

What we did covered:

Online Control with Adversarial noises and costs

What we did covered:

Online Control with Adversarial noises and costs

On the k-th day,

What we did covered:

Online Control with Adversarial noises and costs

On the k-th day,

1. adversary decides sequence of noises $\{w_0^k, \ldots, w_{H-1}^k\}$ and (convex) cost function $c^k(x, u)$,

What we did covered:

Online Control with Adversarial noises and costs

On the k-th day,

1. adversary decides sequence of noises $\{w_0^k, \ldots, w_{H-1}^k\}$ and (convex) cost function $c^k(x, u)$,
2. Without knowing noises and cost, learner proposes a sequence of controllers

$$\{\{M_t^k, M_{0;t}^k, \ldots M_{t-1;t}^k\}\}_{t=0}^{H-1}$$

What we did covered:

Online Control with Adversarial noises and costs

On the k-th day,

1. adversary decides sequence of noises $\{w_0^k, \ldots, w_{H-1}^k\}$ and (convex) cost function $c^k(x, u)$,
2. Without knowing noises and cost, learner proposes a sequence of controllers

$$\{\{M_t^k, M_{0;t}^k, \ldots M_{t-1;t}^k\}\}_{t=0}^{H-1}$$

3. Learner executes controllers, and suffers total cost $\sum_{h=0}^{H} c^k(x_h^k, a_h^k)$

What we did covered:

Online Control with Adversarial noises and costs

On the k-th day,

1. adversary decides sequence of noises $\{w_0^k, \ldots, w_{H-1}^k\}$ and (convex) cost function $c^k(x, u)$,
2. Without knowing noises and cost, learner proposes a sequence of controllers $\{\{M_t^k, M_{0;t}^k, \ldots M_{t-1;t}^k\}\}_{t=0}^{H-1}$
3. Learner executes controllers, and suffers total cost $\sum_{h=0}^{H} c^k(x_h^k, a_h^k)$

Goal: No-Regret $\sum_{k=0}^{K-1} \sum_{h=0}^{H} c^k(x_h^k, a_h^k) - \min_{\{-K_h^\star\}_{h=0}^{H-1}} \sum_{k=0}^{K-1} J^k(\{-K_h^\star\}_{h=0}) = o(K),$

What we did covered:

Online Control with Adversarial noises and costs

On the k-th day,

1. adversary decides sequence of noises $\{w_0^k, \ldots, w_{H-1}^k\}$ and (convex) cost function $c^k(x, u)$,
2. Without knowing noises and cost, learner proposes a sequence of controllers $\{\{M_t^k, M_{0;t}^k, \ldots M_{t-1;t}^k\}\}_{t=0}^{H-1}$
3. Learner executes controllers, and suffers total cost $\sum_{h=0}^{H} c^k(x_h^k, a_h^k)$

Goal: No-Regret $\frac{1}{K}\left(\sum_{k=0}^{K-1}\sum_{h=0}^{H} c^k(x_h^k, a_h^k) - \min_{\{-K_h^\star\}_{h=0}^{H-1}} \sum_{k=0}^{K-1} J^k(\{-K_h^\star\}_{h=0})\right) = o(K)$,

$\frac{\sqrt{K}}{K} \to 0$ as $K \to \infty$

$$\left(J^k(\pi) = \sum_{h=0}^{H-1} c^k(x_h, u_h), \text{where } x_{h+1} = Ax_h + Bu_t + w_t^k, u_t = \pi(x_t) \right)$$

$u \sim N(0, \beta^2 I)$

Collect (x, u, x')

$$(\hat{A}, \hat{B}) = \underset{A, B}{\arg\min} \sum \| Ax + Bu - x' \|_2^2 + \lambda \|A\|_F^2 + \lambda \|B\|_F^2$$

$$\hat{\pi} = \text{RunLQR}(\hat{A}, \hat{B})$$