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

\[ x_{t+1} = Ax_t + Bu_t + w_t, x_0 \sim \mathcal{D}, w_t \sim \mathcal{N}(0, \sigma^2 I) \]
Setting: Linear Dynamical System and Convex Cost functions

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

\[
\max_{\pi} \mathbb{E}_\pi \left[ \sum_{h=0}^{H-1} c(x_h, \hat{a}_h) \right] \quad \frac{1}{2} x^T Q x + \frac{1}{2} u^T R u
\]

Convex Function!
Setting: Linear Dynamical System and Convex Cost functions

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

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

Convex Function!

We cannot guarantee the optimal policy is linear anymore…
Setting: Linear Dynamical System and Convex Cost functions

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

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

Convex Function!

We cannot guarantee the optimal policy is linear anymore…

But we consider a restricted policy class

\[ \Pi = \{ \pi = \{-K_t\}_{t=0}^{H-1} : K_t \in \mathcal{H}, \forall t \} \]
Setting: Linear Dynamical System and Convex Cost functions

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

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

Convex Function!

We cannot guarantee the optimal policy is linear anymore…

But we consider a restricted policy class

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

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

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 \( \pi := \{ -K_t \}_{t=0}^{H-1} \).
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 $\pi$ and the system is at $x_t$, we can compute all previous noises:

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

$$u_t = M_0 x_0 + M_1 w_0 + \cdots + M_t w_{t-1}$$
A New Convex Parameterization of Controllers

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

Assume we roll out \( \pi \) and the system is at \( x_t \), we can compute all previous noises:

\[
\begin{align*}
    w_{t-1} &= x_t - Ax_{t-1} - Bu_{t-1}, \\
    w_{t-2} &= x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots
\end{align*}
\]

\( u_t = -K_t x_t \)
A New Convex Parameterization of Controllers

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

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

\[ w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, \quad w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots \]

\[ u_t = -K_t x_t \]
\[ x_{t+1} = A x_{t+1} + B u_{t+1} + w_{t+1} \]
\[ u_{t-1} = -K_{t-1} x_{t-1} \]
\[ x_{t-1} = A x_{t-2} + B u_{t-2} + w_{t-2} \]
\[ u_{t-2} = -K_{t-2} x_{t-2} \]
A New Convex Parameterization of Controllers

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

Assume we roll out \( \pi \) and the system is at \( x_t \), we can compute all previous noises:

\[
w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, \quad w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots
\]

\[
u_t = -K_t x_t
\]

\[
= -K_t w_{t-1} - K_t(A - BK_{t-1})x_{t-1}
\]

\[
= -K_t w_{t-1} - K_t(A - BK_{t-1})(Ax_{t-2} - BK_{t-2}x_{t-2} + w_{t-2})
\]
A New Convex Parameterization of Controllers

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

Assume we roll out \( \pi \) and the system is at \( x_t \), we can compute all previous noises:

\[
w_{t-1} = x_t - Ax_{t-1} - Bu_{t-1}, \quad w_{t-2} = x_{t-1} - Ax_{t-2} - Bu_{t-2}, \ldots
\]

\[
u_t = -K_t x_t
\]

\[
= -K_t w_{t-1} - K_t (A - BK_{t-1}) x_{t-1}
\]

\[
= -K_t w_{t-1} - K_t (A - BK_{t-1}) (A x_{t-2} - BK_{t-2} x_{t-2} + w_{t-2})
\]

\[
= -K_t \left( \underbrace{w_{t-1}}_{:= M_{t-1,t}} \right) - K_t (A - BK_{t-1}) \left( \underbrace{w_{t-2}}_{:= M_{t-2,t}} \right) - K_t (A - BK_{t-1}) (A - BK_{t-2}) \left( \underbrace{x_{t-2}}_{:= M_{t-3,t}} \right)
\]

\[
-x_{t-2} = Ax_{t-3} + Bu_{t-3} + w_{t-3}
\]

\[-K_t (A - BK_{t-1}) (A - BK_{t-2}) (A - BK_{t-3}) : w_{t-3}\]
A New Convex Parameterization of Controllers

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

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

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

$$u_t = -K_t x_t$$

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

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

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

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

$$= \left[-K_t \left( \prod_{\tau=1}^{t} (A - BK_{t-\tau}) \right) \right] x_0 + \sum_{\tau=0}^{t-1} \left[-K_t \prod_{h=1}^{t-1-\tau} (A - BK_{t-h}) \right] w_{\tau}$$
A New Convex Parameterization of Controllers

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

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

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

$$u_t = -K_t x_t$$

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

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

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

$$:= M_{t-1, t}$$

$$:= M_{t-2, t}$$

$$:= M_{t-3, t}$$

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

$$M_t$$

$$M_{t, t}$$
A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization
\[
\{\{M_t, M_{t-1}; \ldots; M_0; t\}\}_{t=0}^{H-1},
\]
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 $w_0, \ldots, w_{H-1}$. 
A New Convex Parameterization of Controllers

[Claim] For any linear controllers \( \pi := \{ -K_t \}_{t=0}^{H-1} \), there exists a parameterization \( \{ \{ M_t, M_{t-1}; \ldots; M_0; t \} \}_{t=0}^{H-1} \), that generates the same sequence trajectory, given a any fixed \( x_0 \), and fixed noise \( w_0, \ldots, w_{H-1} \).

\[
\begin{align*}
x_0 & \rightarrow \quad u_0 = -K_0x_0 \\
x_0 & \rightarrow \quad u_0 = M_0x_0
\end{align*}
\]
A New Convex Parameterization of Controllers

[Claim] For any linear controllers \( \pi := \{-K_t\}_{t=0}^{H-1} \), there exists a parameterization \( \left\{ \{M_t, M_{t-1}, \ldots, M_0\}_t \right\}_{t=0}^{H-1} \), that generates the same sequence trajectory, given a any fixed \( x_0 \), and fixed noise \( w_0, \ldots, w_{H-1} \).

\[
\begin{align*}
  x_0 &\quad \rightarrow \quad u_0 = -K_0x_0 \quad \rightarrow \quad x_1 \\
  x_0 &\quad \rightarrow \quad u_0 = M_0x_0 \quad \rightarrow \quad x_1, w_0
\end{align*}
\]
A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization

$\{\{M_t, M_{t-1}; \ldots; M_0; t\}\}_{t=0}^{H-1}$, that generates the same sequence trajectory,

given a any fixed $x_0$, and fixed noise $w_0, \ldots, w_{H-1}$.

$x_0 \rightarrow u_0 = -K_0x_0 \rightarrow x_1 \rightarrow u_1 = -K_1x_1$

$x_0 \rightarrow u_0 = M_0x_0 \rightarrow x_1, w_0$
A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization

$$\left\{\{M_t, M_{t-1}; \ldots; M_0; t\}\right\}_{t=0}^{H-1},$$

that generates the same sequence trajectory, given a any fixed $x_0$, and fixed noise $w_0, \ldots, w_{H-1}$.

\[
x_0 \rightarrow u_0 = -K_0x_0 \rightarrow x_1 \rightarrow u_1 = -K_1x_1
\]

\[
x_0 \rightarrow u_0 = M_0x_0 \rightarrow x_1, w_0 \rightarrow u_1 = M_1x_0 + M_{0;1}w_0
\]
A New Convex Parameterization of Controllers

[Claim] For any linear controllers $\pi := \{-K_t\}_{t=0}^{H-1}$, there exists a parameterization

$\{\{M_t, M_{t-1}; t, \ldots, M_0; t\}\}_{t=0}^{H-1}$

that generates the same sequence trajectory,

given a any fixed $x_0$, and fixed noise $w_0, \ldots, w_{H-1}$.

\[
\begin{align*}
x_0 &\quad \rightarrow \quad u_0 = -K_0x_0 \quad \rightarrow \quad x_1 \quad \rightarrow \quad u_1 = -K_1x_1 \quad \rightarrow \quad x_2 \\
x_0 &\quad \rightarrow \quad u_0 = M_0x_0 \quad \rightarrow \quad x_1, w_0 \quad \rightarrow \quad u_1 = M_1x_0 + M_{0;1}w_0 \quad \rightarrow \quad x_2, w_1
\end{align*}
\]
A New Convex Parameterization of Controllers

**Claim** For any linear controllers \( \pi := \{-K_t\}_{t=0}^{H-1} \), there exists a parameterization \( \left\{ \{M_t, M_{t-1}; \ldots; M_0; t\} \right\}_{t=0}^{H-1} \), that generates the same sequence trajectory, given a any fixed \( x_0 \), and fixed noise \( w_0, \ldots, w_{H-1} \).

\[ x_0 \rightarrow u_0 = -K_0x_0 \rightarrow x_1 \rightarrow u_1 = -K_1x_1 \rightarrow x_2 \rightarrow u_2 = -K_2x_2 \]

\[ \Delta \]

\[ x_0 \rightarrow u_0 = M_0x_0 \rightarrow x_1, w_0 \rightarrow u_1 = M_1x_0 + M_{0;1}w_0 \rightarrow x_2, w_1 \]
[Claim] For any linear controllers \( \pi := \{-K_t\}_{t=0}^{H-1} \), there exists a parameterization 
\( \left\{ \{M_t, M_{t-1}, \ldots M_0\}\right\}_{t=0}^{H-1} \), that generates the same sequence trajectory,

given a any fixed \( x_0 \), and fixed noise \( w_0, \ldots, w_{H-1} \).

\[
egin{align*}
x_0 & \rightarrow u_0 = -K_0x_0 \rightarrow x_1 \rightarrow u_1 = -K_1x_1 \rightarrow x_2 \rightarrow u_2 = -K_2x_2 \\
\end{align*}
\]
A New Convex Parameterization of Controllers

[Claim] Given controller \( \bar{\pi} := \left\{ \{ M_t, M_{t+1}, \ldots, M_{t+H-1}\} \right\}_{t=0}^{H-1} \), \( u_t \) & \( x_t \) are all linear with respect to the parameters, \( \forall t \).
A New Convex Parameterization of Controllers

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

\[ x_0, u_0 = M_0 x_0 \]
A New Convex Parameterization of Controllers

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

\[ x_0, u_0 = M_0 x_0 \]
\[ x_1 = A x_0 + B M_0 x_0 + w_0 \]
A New Convex Parameterization of Controllers

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

\begin{align*}
x_0, u_0 & = M_0 x_0 \\
x_1 & = A x_0 + B M_0 x_0 \\
u_1 & = M_1 x_0 + M_{0:1} w_0 \quad \Leftarrow u_1 \text{ linear wrt } M_s
\end{align*}
A New Convex Parameterization of Controllers

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

\[
\begin{align*}
x_0, u_0 &= M_0x_0 \\
x_1 &= Ax_0 + BM_0x_0 \\
u_1 &= M_1x_0 + M_{0;1}w_0 \\
x_2 &= Ax_1 + BM_1x_0 + BM_{0;1}w_0
\end{align*}
\]
A New Convex Parameterization of Controllers

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

\[
\begin{align*}
x_0, u_0 &= M_0 x_0 \\
x_1 &= A x_0 + B M_0 x_0 \\
u_1 &= M_1 x_0 + M_{0;1} w_0 \\
x_2 &= A x_1 + B M_1 x_0 + B M_{0;1} w_0
\end{align*}
\]

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

$$\mathbb{E}_{\tilde{\pi}} \left[ \sum_{t=0}^{H-1} c(x_t, u_t) \right]$$

is convex with respect the parameters, $\forall t$
[Claim] Given controller \( \tilde{\pi} := \left\{ \{ M_t, M_0; t, \ldots M_{t-1}; t \} \right\}_{t=0}^{H-1} \),

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

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 $\tilde{\pi} := \left\{ \{M_t, M_{0:t}, \ldots M_{t-1:t}\} \right\}_{t=0}^{H-1}$, $E_{\tilde{\pi}} \left[ \sum_{t=0}^{H-1} c(x_t, u_t) \right]$ 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 \( \tilde{\pi} := \left\{ \{M_{t}, M_{0 \mid t}, \ldots M_{t-1 \mid t}\}\right\}_{t=0}^{H-1}, \)

\[
\mathbb{E}_{\tilde{\pi}} \left[ \sum_{t=0}^{H-1} c(x_t, u_t) \right]
\]

is convex with respect the parameters, \( \forall t \)

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

Over-Parameterized:

For \( \pi := \{-K_{t}\}_{t=0}^{H-1} \), we have \( H \times d \times k \) parameters
\( \Delta \)
[Claim] Given controller $\tilde{\pi} := \left\{ \{M_t, M_{0:t}, \ldots M_{t-1:t}\} \right\}_{t=0}^{H-1}$, 

\[
E_{\tilde{\pi}} \left[ \sum_{t=0}^{H-1} c(x_t, u_t) \right]
\]

is convex with respect the parameters, $\forall t$

Convexity allows to perform Gradient Descent directly on parameters

\[
k_t \leftrightarrow M_0, M_{0:t} \ldots M_{t-1:t}
\]

\[
\left\{ \{M_t, M_{0:t}, \ldots M_{t-1:t}\} \right\}_{t=0}^{H-1}
\]

Over-Parameterized:

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

For $\tilde{\pi} := \left\{ \{M_t, M_{0:t}, \ldots M_{t-1:t}\} \right\}_{t=0}^{H-1}$, 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 \( \{w^k_0, \ldots, w^k_{H-1}\} \) and (convex) cost function \( c^k(x, u) \),
2. Without knowing noises and cost, learner proposes a sequence of controllers \( \{\{M^k_t, M^k_{0:t}, \ldots, M^k_{t-1:t}\}\}_{t=0}^{H-1} \).
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 \( \{ \{ M_t^k, M_0^k; t, \ldots M_{t-1; t}^k \} \}_{t=0}^{H-1} \)
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 \( \{ w^k_0, \ldots, w^k_{H-1} \} \) and (convex) cost function \( c^k(x, u) \),
2. Without knowing noises and cost, learner proposes a sequence of controllers
\[ \{ \{ M^k_t, M^k_{0,t}, \ldots M^k_{t-1,t} \} \}_{t=0}^{H-1} \]
3. Learner executes controllers, and suffers total cost
\[ \sum_{h=0}^{H} c^k(x^k_h, a^k_h) \]

Goal: No-Regret
\[ \sum_{k=0}^{K-1} \sum_{h=0}^{H} c^k(x^k_h, a^k_h) - \min_{\{ -K^*_h \}_{h=0}^{H-1}} \sum_{k=0}^{K-1} J^k(\{ -K^*_h \}_{h=0}^{H-1}) = o(K), \]
What we did covered:

Online Control with Adversarial noises and costs

On the k-th day,
1. adversary decides sequence of noises \( \{ w^k_0, \ldots, w^k_{H-1} \} \) and (convex) cost function \( c^k(x, u) \),
2. Without knowing noises and cost, learner proposes a sequence of controllers \( \{ \{ M^k_t, M^k_{0:t}, \ldots, M^k_{t-1:t} \} \}_{t=0}^{H-1} \),
3. Learner executes controllers, and suffers total cost \( \sum_{h=0}^{H} c^k(x^k_h, a^k_h) \)

Goal: No-Regret

\[
\sum_{k=0}^{K-1} \sum_{h=0}^{H} c^k(x^k_h, a^k_h) - \min_{\{ -K^*_h \}_{h=0}^{H-1}} \sum_{k=0}^{K-1} J^k(\{ -K^*_h \}_{h=0}^{H-1}) = o(K),
\]

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

as \( K \to \infty \)
\[ u \sim N(0, B^2I) \]

Collect \((x, u, x')\)

\[
(\hat{A}, \hat{B}) = \arg \min_{A, B} \sum_{i} \left\| A x + B u - x'_{i} \right\|^2_{2} + \lambda \left\| A \right\|_{F}^2 + \lambda \left\| B \right\|_{F}^2
\]

\[
\hat{\pi} = \text{Run LQR}(\hat{A}, \hat{B})
\]