On The Hardness of Learning to Stabilize Linear Systems

Xiong Zeng[1], Zexiang Liu[1], Zhe Du[1], Necmiye Ozay[1], Mario Sznaier[2]

1: University of Michigan, Ann Arbor 2: Northeastern University, Boston

Introduction:

Learning an initial stabilizing controller is an important step in many learning-based control tasks. This poster studies the statistical hardness of learning to stabilize linear time-invariant systems. Our analysis shows a larger class of systems that are hard to learn to stabilize compared with earlier work [2]. This work is published in [1].

Problem Setup:

Given

- a discrete-time linear time-invariant (LTI) system S from **a class $\mathcal{C}_n$** :

$$x_{t+1} = Ax_t + Bu_t + w_t,$$

where $x_t \in \mathbb{R}^n, w_t \sim N(0, \sigma_w^2 \boldsymbol{I}),\ \sigma_w^2 > 0,$

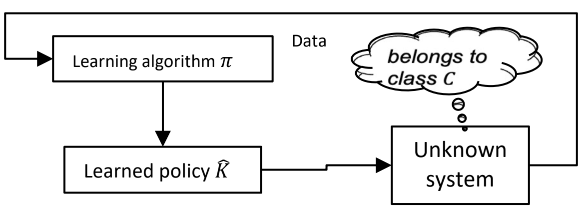
- a time budget N and an input budget $\mathbb{E}[\|u_t\|^2] \leq \sigma_u^2$.

Consider a learning algorithm π that

i. interacts with the system for N units of time, and

ii. outputs a linear static state-feedback controller $\widehat{K}$ at time N.

We want $\widehat{K}$ to stabilize the (unknown) system S.



(Use your favorite algorithm π: model-free, model-based, etc.)

Question: Given a class C_n of linear systems, how does the **sample complexity** for learning stabilizing static state-feedback controllers depend on the **system dimension $\boldsymbol{n}$**?

Main Result: Consider two special LTI systems with

$$A_1 = A_2 = \begin{bmatrix} r & v & 0 & \cdots & 0 \\ 0 & 0 & v & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & \cdots & v \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ v \end{bmatrix}, B_2 = \begin{bmatrix} m \\ 0 \\ \vdots \\ v \end{bmatrix}.$$

Key insight: if the class $\mathcal{C}_n$ contains a pair of systems (A_1, B_1) and (A_2, B_2) that are **hard to distinguish**, and **not co-stabilizable,** then **learning to stabilize is hard.**

Proposition 1: There exists a state-feedback controller K that stabilizes both (A_1, B_1) and (A_2, B_2), only if

$$m \leq \mathcal{O}\left(\frac{1}{\exp(n)}\right).$$

Theorem 1: For any class C_n of systems containing such a pair with $m = 2\left(\frac{2v}{r-1}\right)^n$ and any learning algorithm, the least amount of data for learning to stabilize arbitrary systems in C_n **grows exponentially** with the state dimension n. Specifically,

$$\inf_{S \in C_n} P_{S,\pi}^N \left(\rho\left(A + B\,\widehat{K}\,\right) < 1\right) \geq 1 - \delta,$$

is satisfied only if

$$N \geq \frac{\delta_w^2}{2\delta_u^2}\left(\frac{r-1}{2v}\right)^{2n} log\,\frac{1}{3\delta},$$

where $n \geq 2, r > 1, and\ 0 < v < \frac{r-1}{2}$.

The key ideas in the proof:

i. Use **Ackermen's formula** to show that (A_1, B_1) and (A_2, B_2) are not co-stabilizable (Proposition 1);

ii. Show that the KL divergence between (A_1, B_1) and (A_2, B_2) decays exponentially with n (indistinguishability).

i + ii + Birge's Inequality ➔ at least exp(n) samples to learn a stabilizing feedback gain for an unknown system being either (A_1, B_1) and (A_2, B_2).

Comparison with Existing Result:

Tsiamis et al. [2]	Stable A	Degenerate noise
Our Work	Unstable A	Non-degenerate noise

Compared with [2], the system classes in our work are **not necessarily hard to identify**, but still **hard to stabilize**.

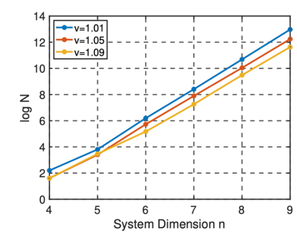
Experiments:

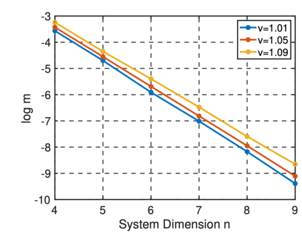
Certainty equivalence LQR

$$\min_{\mathbf{u}_0,\mathbf{u}_1,\cdots} \lim_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\sum_{t=0}^{T}\left(\mathbf{x}_t^\top \mathbf{Q}\mathbf{x}_t + \mathbf{u}_t^\top \mathbf{R}\mathbf{u}_t\right)\right]$$
$$\text{s.t.} \quad \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t$$

LMI-based sufficient condition for co-stabilizability

$$\begin{aligned} \text{find} \quad & \mathbf{K}, \mathbf{P} \\ \text{s.t.} \quad & (\mathbf{A} + \mathbf{B}_1\mathbf{K})^\top \mathbf{P} (\mathbf{A} + \mathbf{B}_1\mathbf{K}) \prec \mathbf{P} \\ & (\mathbf{A} + \mathbf{B}_2(m)\mathbf{K})^\top \mathbf{P} (\mathbf{A} + \mathbf{B}_2(m)\mathbf{K}) \prec \mathbf{P} \\ & \mathbf{P} \succ 0 \end{aligned}$$





Current work:

Reduce sample complexity by using **dynamic time-varying** controllers.

References:

[1] X. Zeng, Z. Liu, Z. Du, N. Ozay, and M. Sznaier, On the Hardness of Learning to Stabilize Linear Systems, CDC, 2023

[2] Tsiamis, Anastasios, et al. "Learning to control linear systems can be hard." Conference on Learning Theory. PMLR, 2022.