Black-Box Control for Linear Dynamical Systems

Xinyi Chen[1 2], Elad Hazan[1 2],
[1]Google AI Princeton [2]Department of Computer Science, Princeton University

PRINCETON UNIVERSITY Google AI

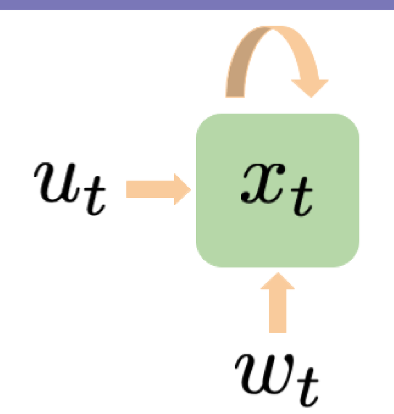
Robust and Optimal Control

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad x_t \in \mathbb{R}^{d_x}, \quad u_t \in \mathbb{R}^{d_u}$$

	Known Costs	Unknown Costs
Stochastic Noise	LQR (Known System), Adaptive Control (Unknown System)	Online LQR (Unknown System)
Adversarial Noise	H_∞-Control (Known System)	Nonstochastic Control (Unknown System)

Nonstochastic Control

- Adversarial noise: $\|w_t\| \leq 1$.
- Changing convex costs $c_t(x, u)$.
- Goal: Minimize Regret



$$\text{Regret}_T(\mathcal{A}) = \sum_{t=1}^{T} c_t(x_t^{\mathcal{A}}, u_t^{\mathcal{A}}) - \min_{\pi \in \Pi} \sum_{t=1}^{T} c_t(x_t^{\pi}, u_t^{\pi}).$$

In this work, we introduce an algorithm that achieves:

Sublinear regret in controlling an unknown system with black-box interactions.

Related Work

- **Black-box Adaptive Control with Quadratic Costs**: $\tilde{O}(\sqrt{T})$ regret by [3]. Our setting permits general convex cost functions.
- **Nonstochastic Control with Known Stabilizing Controller**: [4] gives sublinear regret. [2] considers partially observed systems.
- **Identification and Stabilization of Linear Systems:**
  - Stochastic Noise: [1] uses least squares; [6] finds a stabilizing controller in finite time.
  - Adversarial Noise: [5] for non-explosive systems.

The System Complexity

- The system (A, B) is (k, κ) strongly controllable for $\kappa \geq 1$.
- The system satisfies $\|A\|, \|B\| \leq \beta$ for some $\beta \geq 1$, and the cost functions are G-Lipschitz.
- Define the system complexity:

$$\mathcal{L} = kd_u + d_x + G + \beta + \kappa.$$

- Given a strongly stable controller K, the stabilized system $(A + BK, B)$ is (k, κ^*) strongly controllable.

Main Results

With high probability, our algorithm has regret

$$\text{Regret}_T \leq 2^{O(\mathcal{L} \log \mathcal{L})} + \tilde{O}(\text{poly}(\mathcal{L}, \kappa^*) T^{2/3})$$

Lower bound: any deterministic black-box control algorithm suffers regret $2^{\Omega(\mathcal{L})}$ in the worst case.

Lower Bound

Black-Box Control Algorithm: a deterministic algorithm $\mathcal{A}$ that outputs controls depending only on past states and cost functions.
Time-invariant Cost Functions: $c_t(x, u) = \|x\|^2 + \|u\|^2$

For any black-box control algorithm $\mathcal{A}$, there exists a noiseless linear dynamical system that is stabilizable and $(1, 1)$ controllable, and a sequence of time-invariant cost functions, such that

$$\text{Regret}(\mathcal{A}) = 2^{\Omega(\mathcal{L})}.$$

- We iteratively set the rows of the system as we receive controls from the controller $\mathcal{A}$.
- First finite-time bound which is exponential in dimension for any online control setting.

Main Algorithm

- **Phase 1: Black-box System Identification**
  - Coarse-grained identification of the system by using large controls, and obtain system estimates $(\hat{A}, \hat{B})$. Final state may be exponentially large.
- **Phase 2: Stable Controller Recovery:** Solve SDP proposed in [7] with estimates $(\hat{A}, \hat{B})$ and obtain strongly stable controller $\hat{K}$.
- **Phase 3: Nonstochastic Control:** Call Algorithm 1 in [4] with $\hat{K}$.

Phase 1: System Identification

- Define $C_{k-1} = [B \; AB \; \cdots A^{k-1}B]$, $C_k = [AB \; A^2B \; \cdots A^kB]$. By controllability, A is the unique solution to $XC_{k-1} = C_k$.
- We construct $C_0 \approx C_{k-1}$, $C_1 \approx C_k$ by estimating each A^jB. Output solution to $XC_0 = C_1$.

Estimating A^jB: **For** $t = 1, \ldots, (k+1)d_u$:

- Once every $k + 1$ iterations, play a control whose magnitude increases multiplicatively. The i-th control is the i-th standard basis vector scaled by ξ_i.
- Otherwise control with $u_t = 0$.

Let x_i^j be the state at j iterations after the ith nonzero-control. Construct estimate of A^jB as $\hat{M}_j = [\frac{x_1^j}{\xi_1} \; \frac{x_2^j}{\xi_2} \; \cdots \; \frac{x_{d_u}^j}{\xi_{d_u}}]$.

References (More in Paper)

[1] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis. Finite Time Identification in Unstable Linear Systems. Automatica, 2018..

[2] M. Simchowitz, K. Singh, and E. Hazan. Improper Learning for Non-Stochastic Control. Conference on Learning Theory, 2020.

[3] Y. Abbasi-Yadkori, and C. Szepesvári. Regret Bounds for the Adaptive Control of Linear Quadratic Systems. Conference on Learning Theory, 2011.

[4] E. Hazan, S. Kakade, and K. Singh. The Nonstochastic Control Problem. International Conference on Algorithmic Learning Theory, 2019.

[5] M. Simchowitz, R. Boczar, and B. Recht. Learning Linear Dynamical Systems with Semi-Parametric Least Squares. Conference on Learning Theory, 2019.

[6] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis. Finite-Time Adaptive Stabilization of Linear Systems. IEEE Transactions on Automatic Control, 2019.

[7] A. Cohen, A. Hassidim, T. Koren, N. Lazic, Y. Mansour, and K.Talwar. Online Linear Quadratic Control. International Conference on Machine Learning, 2018.