Online state tracking in presence of adversarial disturbances (r1)

Farnaz Adib Yaghmaie (farnaz.adib.yaghmaie@liu.se)

Introduction

We propose an online Reinforcement Learning (RL) algorithm to achieve tracking in presence of adversarial disturbances.

- The disturbance is arbitrary, bounded, and unknown.
- The reference is generated by a linear system with unknown dynamics.
- The controller is learned to optimize any convex cost function.

Problem Formulation

Consider the following linear dynamical system

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{1}$$

where $w_k \in \mathbb{R}^n$ denotes bounded disturbance which is selected in adversarial setting. The reference is generated by

$$\begin{aligned} z_{k+1} &= Sz_k, \\ r_k &= Fz_k, \end{aligned} \tag{2}$$

We aim to generate $u_k = \mathcal{A}(x_{1:k}, r_{1:k})$ to minimize the following cost function

$$J_T(\mathcal{A}) = \sum_{k=1}^{T} c_k(e_k, u_k), \tag{3}$$

where e_k is the state tracking error

$$e_k = x_k - r_k. \tag{4}$$

Optimal tracking problem

Consider (1)-(2). Generate $u_k = \mathcal{A}(x_{1:k}, r_{1:k})$ to minimize (3).

Assumptions

Assumption 1: (A, B) is known and stabilizable. (No loss of generality, see Algorithm 2 in **[r2]**.)
Assumption 2: (S, F) is unknown but observable. z_k is not measurable but the output r_k is measurable. r_k is bounded.
Assumption 3: The cost $c_k(e_k, u_k)$ is convex.

The linear history-based policy

We parameterize the controller as

$$u_k^\pi = Kx_k + \sum_{t=1}^{m_w} M_w^{[t-1]} w_{k-t} + \sum_{s=0}^{m_r-1} M_r^{[s]} r_{k-s}, \tag{5}$$

where K is stabilizing. In Lemma 2 of **[r1]**, we prove that (5) can approximate $u_k^{\text{lin}} = K_{fb}x_k + K_{ff}z_k$ and as such, it can solve the state tracking problem in absence of disturbance.

Online state tracking algorithm

1: **Initialize:** Select a stabilizing K and set M_w, M_r arbitrarily.
2: **for** $k = 1, .., T$ **do**
3: Record r_k and execute u_k^π.
4: Observe x_{k+1} and record $w_k = k_{k+1} - Ax_k - Bu_k$.
5: Suffer $c_k(e_k, u_k)$ and compute $\tilde{c}_k$ as it follows: Let $\tilde{x}_{k-H} = 0$. Compute for H steps

$$\begin{aligned} \tilde{u}_{k-H}^\pi &= K\tilde{x}_{k-H} + \sum_{t=1}^{m_w} M_w^{[t-1]} w_{k-H-t} + \sum_{s=0}^{m_r-1} M_r^{[s]} r_{k-H-s} \\ \tilde{x}_{k-H+1} &= \tilde{x}_{k-H} + B\tilde{u}_{k-H} + w_{k-H}, \end{aligned} \tag{6}$$

to have $\tilde{u}_k, \tilde{x}_k$. Let

$$\tilde{c}_k = c_k(\tilde{x}_k - r_k, \tilde{u}_k) \tag{7}$$

6: Update M_w, M_r

$$M_w = M_w - \eta \nabla_{M_w} \tilde{c}_k, \; M_r = M_r - \eta \nabla_{M_r} \tilde{c}_k. \tag{8}$$

Simulation results

The dynamical system:

$$x_{k+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} u_k + w_k.$$

The reference:

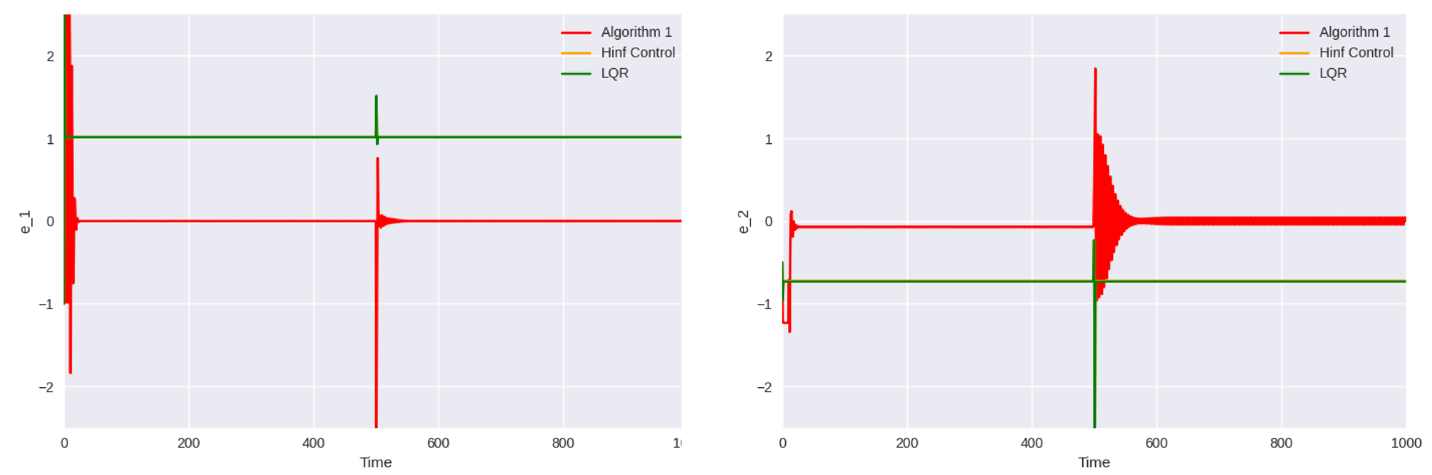
$$z_{k+1} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} z_k, \; r_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} z_k, \; z_0 = \begin{bmatrix} 1 \\ -2 \\ 0.5 \end{bmatrix} \; 0 \le k < 500,$$

$$z_{k+1} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} z_k, \; r_k = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix} z_k, \; z_{500} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \; k \ge 500.$$

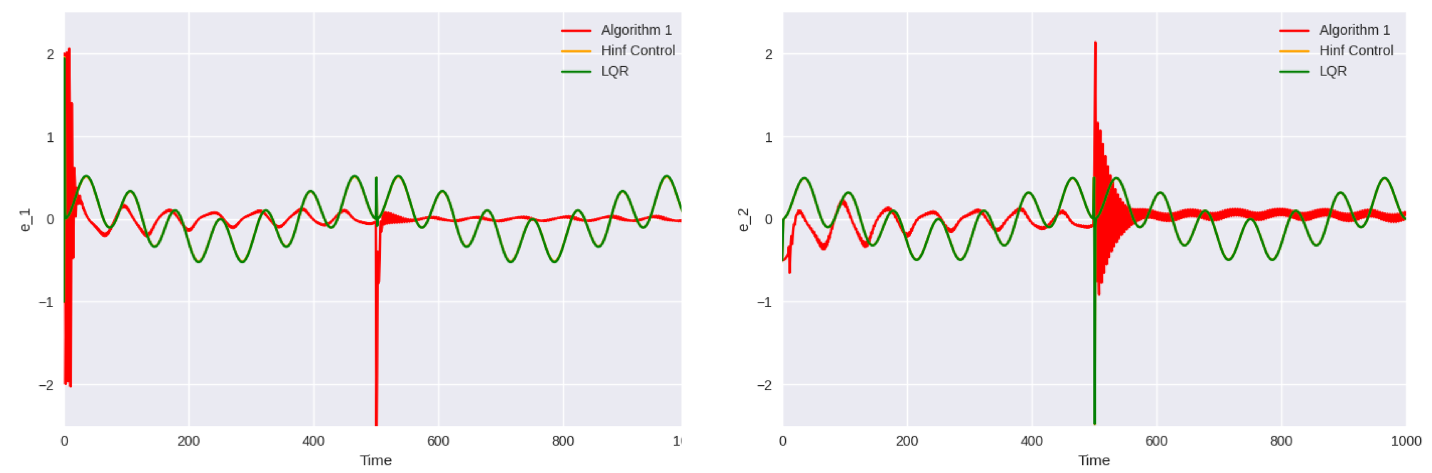
A quadratic cost with $Q = 20I_2$, $R = I_2$ is considered. Three disturbances are examined

- **Case 1)** $w_{1k} = 1$, $w_{2k} = -0.7$,
- **Case 2)** $w_{1k} = w_{2k} = 0.5\sin(6\pi k/T)\sin(8\pi k/T)$, and
- **Case 3)** $w_{1k} = w_{2k} = 0.5\sin(8\pi k/T)$.

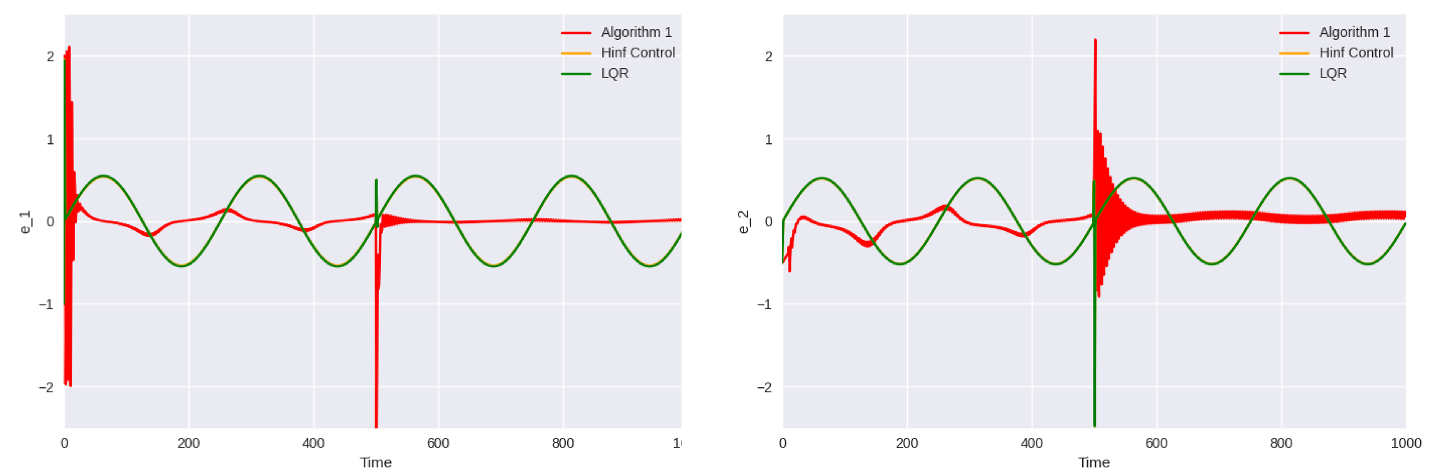
The results of the the online state tracking algorithm are shown in red.



Tracking error for Case 1



Tracking error for Case 2



Tracking error for Case 3

References

[r1] F. Adib Yaghmaie "Online state tracking in presence of adversarial disturbances", Submitted to ACC 2022.
[r2] E. Hazan, S. M. Kakade, and K. Singh. "The Nonstochastic Control Problem". In Algorithmic Learning Theory, pages 408—-421, 2020.

LINKÖPING UNIVERSITY
LINK-SIC

LINKÖPING UNIVERSITY
Division of Automatic Control