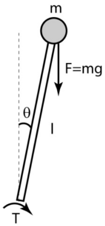
CMPE 252, Artificial Intelligence and Data Engineering, Spring 2023

Home Assignment 3, Part 2, LQR

Sequential Decision Making in Linear Models with Quadratic Reward

due: April 7, 2023, 11:59pm.

Setting



Equation-Of-Motion of the pendulum: $\begin{bmatrix} \dot{\theta}_1(t) \\ \dot{\theta}_2(t) \end{bmatrix} = \begin{bmatrix} \dot{\theta}_2(t) \\ -\frac{g}{\ell} \sin \theta_1(t) \end{bmatrix}$

where $\theta_1(t)$ and $\theta_2(t)$ are the angle and the angular velocity.

- EOM in discrete time, t, with the control gain, $B_{dX \times dU}$ and control u:

$$\text{Dynamics model: } x[t+1] = A_D x[t] + Bu[t]$$

Parameters

- $A_D = \begin{pmatrix} 1.00049 & 0.0100016 \\ 0.098016 & 1.00049 \end{pmatrix}$
- $B = \begin{pmatrix} 0.01 \\ 0.0005 \end{pmatrix}$

Questions

1. Starting from $x[0] = [\theta_0, \dot{\theta}_0] = [+0.1; 0.0]$, simulate the pendulum dynamics for $T = 499$ steps with $u[t] = 0$.
2. Starting from $x[0] = [\theta_0, \dot{\theta}_0] = [-0.1; 0.0]$, simulate the pendulum dynamics for $T = 499$ steps with $u[t] = 0$.
3. Explain your observations. How do θ and $\dot{\theta}$ change? Does the pendulum stays near the top?
4. Plot on the same figure the angle, $x[1, :]$, the angular velocity, $x[2, :]$, for 1&2 above as a function of time. Add the legend.
5. Repeat 1-4 above with the optimal state-feedback control signal, $u^{lqr}[t] = K[t]x[t]$, with $R = 0.1$, and with three different values for $\alpha \in [0.001, \text{your value}, \text{your value}]$ in $Q = \alpha * IdentityMatrix_{2\times 2}$. You can use the same matrix Q for Q_f. The summary of LQR appears at the next slide.
6. (Bonus, 5 points) Visualize the value function, $V(x) = x' P_0 x$, on a 2D plot with the axes: $\theta = -0.3 : 0.001 : 0.3$ and $\dot{\theta} = -0.3 : 0.001 : 0.3$. Explain your observation and the shape of V.

Summary of LQR solution via DP

1. set $P_N := Q_f$

2. for $t = N, \ldots, 1,$

$$P_{t-1} := Q + A^T P_t A - A^T P_t B (R + B^T P_t B)^{-1} B^T P_t A$$

3. for $t = 0, \ldots, N-1$, define $K_t := -(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A$

4. for $t = 0, \ldots, N-1$, optimal u is given by $u_t^{\mathrm{lqr}} = K_t x_t$

- optimal u is a linear function of the state (called *linear state feedback*)
- recursion for min cost-to-go runs backward in time

What to submit

- Your code in ipynb, which enables to reconstruct your plots.
- A pdf file with the plots, the explanations, and your full code.
- Submit these two files separately (not in a zip file) in Canvas.