Machine Learning for Systems & Control
5SC28
2023-2024
Exercise Set for Lecture 9:
Model Internalization-Based RL

1 Introduction

You can use the following resource to solve this model internalization exercise using PILCO:

- Matlab PILCO toolbox, available at `http://github.com/UCL-SML/pilco-matlab`

It is recommended to read the first chapter of the PILCO manual to gain a better understanding of PILCO, which provides an overview of the key concepts, techniques, and functionalities. For further information, you can refer to the original PILCO paper.

Additionally, there is a Python implementation available at `https://github.com/nrontsis/PILCO`. While it does not work as well as the MATLAB implementation, it might be interesting to experiment with (note that running the Python version requires installing TensorFlow).

2 Installing PILCO in MATLAB

1. Download the PILCO toolbox from the provided website or Canvas folder.
2. Add the PILCO toolbox to your path. In the MATLAB command window, type the following command:

   ```
   addpath(genpath('the_pilco_directory'))
   ```

   Ensure you provide the correct directory path to the PILCO toolbox in the `addpath` command so that MATLAB can locate and use the toolbox effectively.
3. Optional: Run an example from the `scenarios` folder to verify the installation.

Afterward, check out the examples in the `pilcoV09/scenarios` directory.

3 Exercise: The Cart-Pole Problem

The goal is to learn a policy (state feedback control map) that can swing up a pendulum attached to a cart and keep the cart around its original position. The cart has mass m_1 and the attached pendulum has mass m_2 and length l, swinging freely in the plane. The pendulum angle θ_2 is measured anti-clockwise from hanging down. The cart can move horizontally with an applied external force u and is affected by viscous friction b between the cart and the ground. The parameters of the system are:

- $m_1 = 0.5\,kg$
- $m_2 = 0.5\,kg$
- $l = 0.6\,m$
- $b = 0.1\,N/m/s$

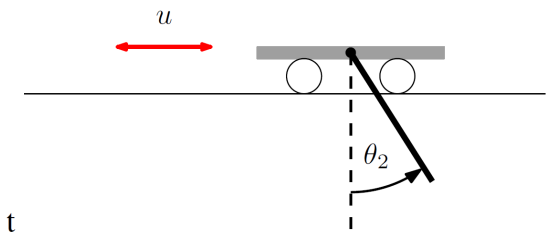

Figure 1: Cart-pole system

Use the PILCO toolbox to solve the problem. An example script, `Sol_1.m`, has been prepared with extensive documentation to demonstrate how this can be achieved. Observe how many episodes are required (compared to the previous model-free Q-learning) to swing up the pendulum.