Alex Beyer

ENME743

Final Project Deliverable 1: Preliminary Data and Visualization

This deliverable is written about Standard project 4: Acrobot Control. The system is constructed as in [Sutton and Barto Ch 11.3 - The Acrobot](http://incompleteideas.net/book/11/node4.html), which is linked in the OpenAI gym page for the Acrobot.

**Data Collection and Processing**

In order to collect data a working simulation of the Acrobot dynamics was built within MATLAB. Considering this is a bang-bang optimal control problem we don’t need data *per se*, but being able to test simple behaviors via the dynamics gives us insight into how the system works, as well as providing a testbed that’s a bit less of a black box than the OpenAI implementation of the dynamcis. My implementation is largely constrained by the OpenAI gym model: the only passable action in is a 3-state toggle for the motor attached between joints 1 and 2 and the only outputs are the sines and cosines of joint angles along with the angular velocity of each joint. The problem is asking us to drive the Acrobot to a single state, namely [π/2,0] which is the unstable vertical equilibria. It’s worth noting here that the simulation end condition in OpenAI gym is such that the instability of the equilibrium doesn’t matter, the simulation stops the instant the arm is pointed straight up.

All together, an optimal solution to this problem requires a control law which is able to use a bang-bang control law (discrete inputs, -1, 0 and +1 in this case) to rotate the lower member in such a way that the inertia of the motion flings both members upright. To give the model the best chance of implementing this control policy, we should aim to give it some sense of the models dynamics. In principle the model should be able to find a close-to-optimal policy without any understanding of the dynamics, but it could help the model converge to a better policy faster. As an example, since our system is inertia driven we won’t be able to get the swing up behavior we want with a constant input in one direction (proved in the Data Visualization section, although because our model terminates the instant the arm is upright we could in theory achieve this for specific combinations of mass and link length) so we’d want to disincentivize the model from applying constant control inputs for long periods of time.