Homework 3

E. Todorov, AMATH/CSE 579

Due June 12 noon

Here we will construct a global control policy for the cart-pole system. Let y denote the horizontal position of the cart and θ the angle of the pole/pendulum which is attached to the cart, where $\theta = 0$ corresponds to the upright position. The state vector is $x = \left(\theta, y, \dot{\theta}, \dot{y}\right)$. The system is under-actuated: there is a scalar control u pushing the cart in the horizontal direction. The goal is to push the cart in such a way that the pendulum swings to the upright position and stays there.

The dynamics are:

$$\ddot{\theta} = \frac{9.8 \sin\theta - \left(u + \dot{\theta}^2 \sin\theta\right)\cos\theta - 0.1\dot{\theta}}{8/3 - \cos^2\theta}$$
$$\ddot{y} = u + \dot{\theta}^2 \sin\theta - \ddot{\theta}\cos\theta$$

Note the unusual form of the horizontal acceleration $\ddot{y}$: it is expressed as a function of the angular acceleration $\ddot{\theta}$, which needs to be computed first.

You should design a cost function that captures this task. In addition to the task cost (with minimum at $\theta = 0$) you should include the usual quadratic control cost, with some weight that produces sensible controls.

Recall that for optimal control problems in the form

$$\dot{x} = a(x) + B(x)u$$
$$\ell(x, u) = q(x) + \frac{1}{2}u^T R u$$

the optimal value function for the infinite-horizon average-cost setting is the solution to the HJB equation

$$c + v^*(x) = \ell(x, u^*(x)) + (a(x) + B(x)u^*(x))^T v_x^*(x)$$

where the optimal control law is given by

$$u^*(x) = -R^{-1}B(x)^T v_x^*(x)$$

Let us now replace the (unknown) function $v^*(x)$ with a parametric function approximator $v(x, w)$. The approach here is to fund the w for which the above HJB equation is satisfied as closely as possible. The average cost c is also an unknown parameter that needs to be computed along the way.

We can fit w using a collocation method as follows. Define a set of collocation states $\{x_n\}$ which should be larger than the number of free parameters, and should cover the region of state space where we want to obtain a good solution (since our state space is only 4D, we can cover all of it here). Define the Bellman residual error

$$E(w, c) = \frac{1}{N}\sum_{n=1}^{N}\Big(\ell(x_n, u(x_n, w)) + (a(x_n) + B(x_n)u(x_n, w))^T v_x(x_n, w) - c - v(x_n, w)\Big)^2$$

where the control law $u(x, w)$ is given by

$$u(x, w) = -R^{-1}B(x)^T v_x(x, w)$$

The quantity $E(w, c)$ can now be minimized numerically via gradient descent. You should be able to compute the gradient of E analytically.

For the function approximator, use a mixture of Gaussians whose centers are scattered around the state space. For simplicity fix the centers and covariances of the Gaussians, and only adjust their scalar weights. Thus the function approximator is

$$v(x, w) = \sum_{i=1}^{K} w_i \exp\Big(-(x - r_i)^T S(x - r_i)\Big)$$

where $\{r_i\}$ are the centers of the Gaussians, and S is half the inverse covariance matrix (assumed to be the same for all Gaussians). Again, we fix $\{r_i\}$ and S in advance and adjust $\{w_i\}$ and c by optimizing the quantity E.

You should experiment with the number of placement of the Gaussian bases and collocation states, so that the resulting control law can indeed control the cart-pole system and make the pole go to the upright position.

Submit your code along with figures illustrating the behavior of the resulting control law, and some text summarizing your observations.