Reinforcement learning and Optimal control for Robotics (ROB-GY 6323)

New York University, Fall 2024

Due on 12/08, 11:59 PM

Name: Raman Kumar Jha
NYU ID: N13866145

The implemented algorithm is Q-learning with function approximation for the inverted pendulum control problem. The key components can be described as:

$$Q(s,a) \approx f_\theta(s,a)$$

where f_θ is a neural network with parameters θ. The update rule follows:

$$\theta \leftarrow \theta - \alpha \nabla_\theta \left(Q(s,a) - \left(g(s,a) + \gamma \min_{a'} Q(s',a') \right) \right)^2$$

where:

- $s = [\theta, \dot{\theta}]$ is the state (angle and angular velocity)
- $a \in \{-5, 0, 5\}$ are the possible control actions
- $g(s,a)$ is the immediate cost:

$$g(s,a) = 0.01(1 - \cos(\theta - \pi))^2 + 0.001\dot{\theta}^2 + 0.00001u^2$$

- $\gamma = 0.7$ is the discount factor
- α is the learning rate

The algorithm uses ε-greedy exploration with decay and SGD optimization with momentum and gradient clipping to prevent instability.

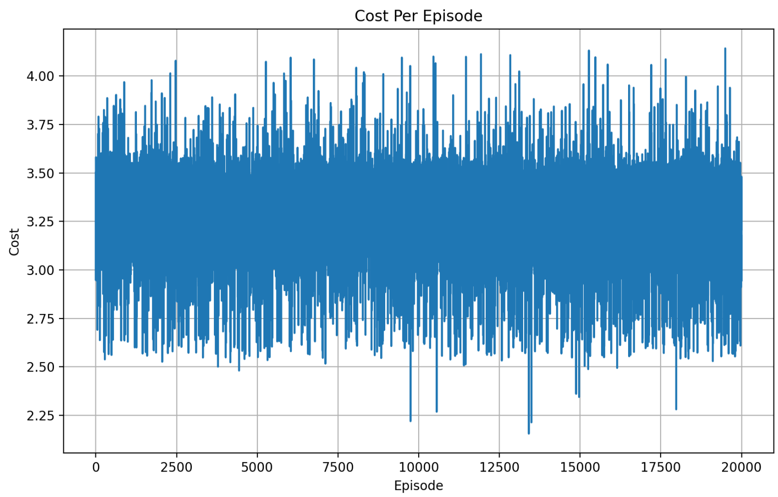

Figure 1: Cost per episode of the Pendulum

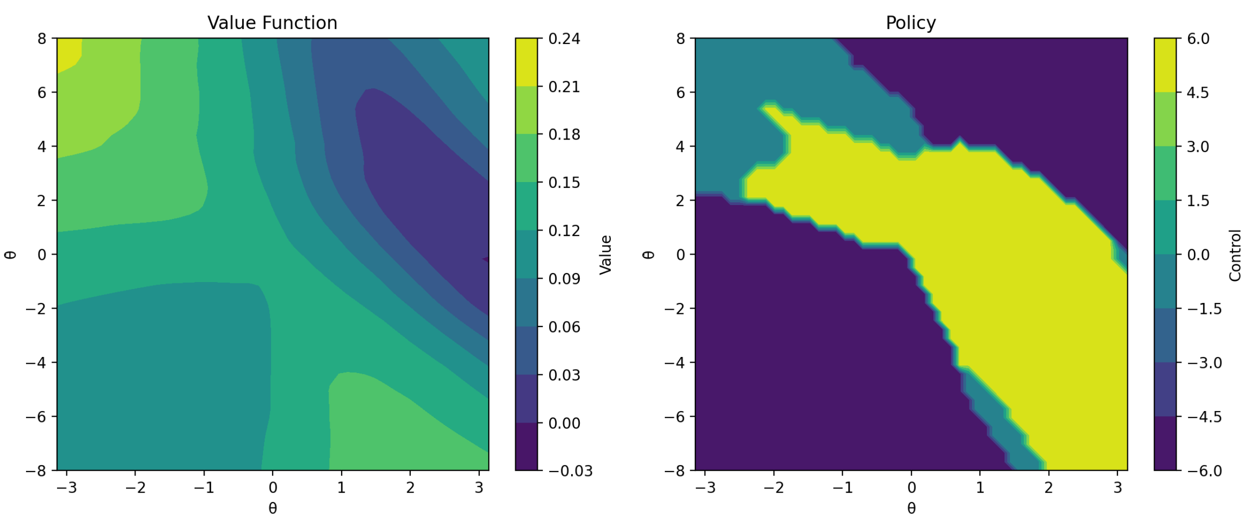

Figure 2: Learned value Function and the Policy of the Pendulum