# Lab 3 LSPI

Eduardo Santoro Morgan
PUC-Rio
ELE2761 - Reinforcement Learning
Prof. Wouter Caarls

May 2, 2019

## 1

Figures 1 and 2 show the resulting policy and convergence plot from running LSPI with parameters $\gamma = 0.9 and N = 10000$. We can see that the resulting policy looks somewhat similar to what we had previously achieved using SARSA. The convergence takes many less iterations of the full algorithm, however each iteration performs

## 2

### 2.1

### 2.2

### 2.3

### 2.4

### 2.5

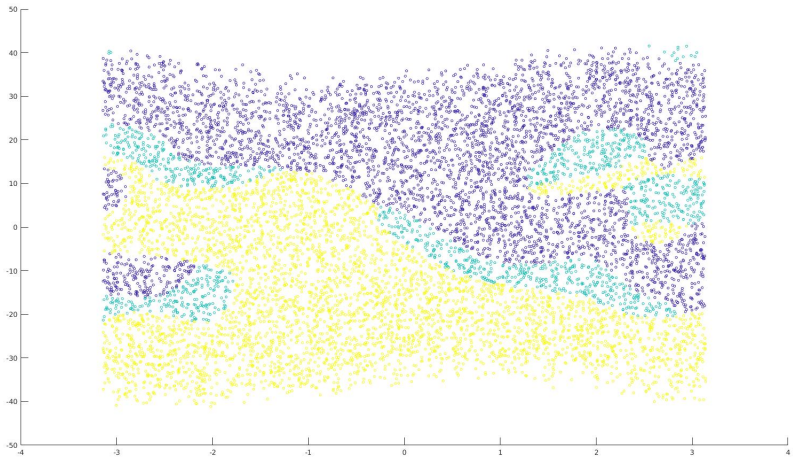


Figure 1: Policy Exercise 4.1

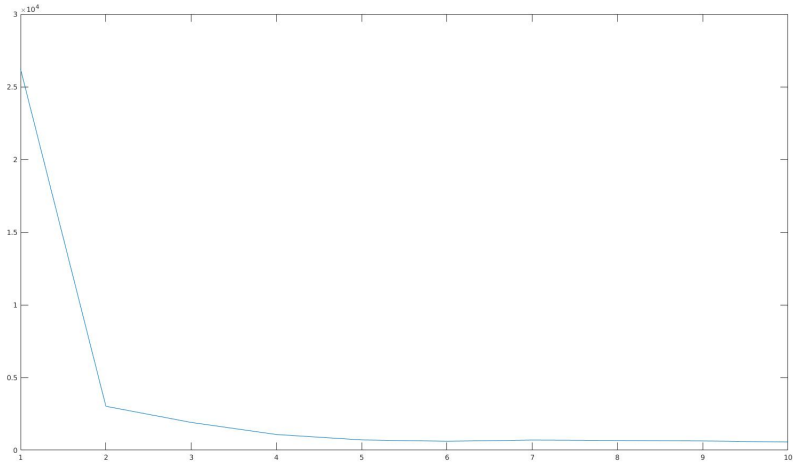


Figure 2: Convergence Exercise 4.1. Computed using L1 distance between policies of subsequent iterations

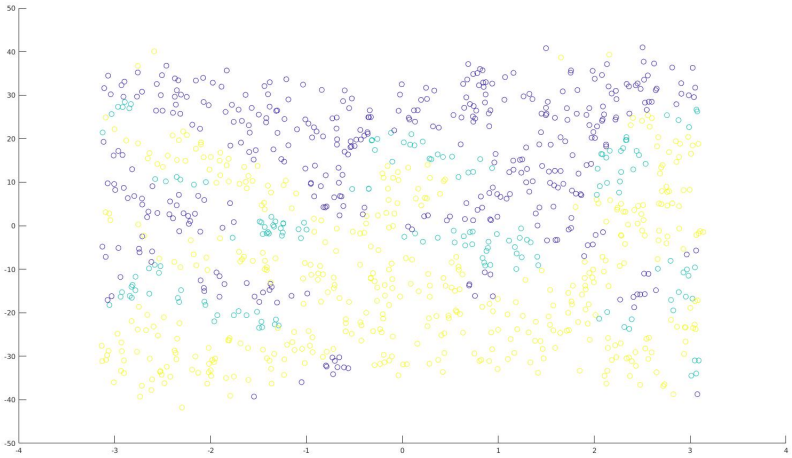

Figure 3: Policy Exercise 4.2

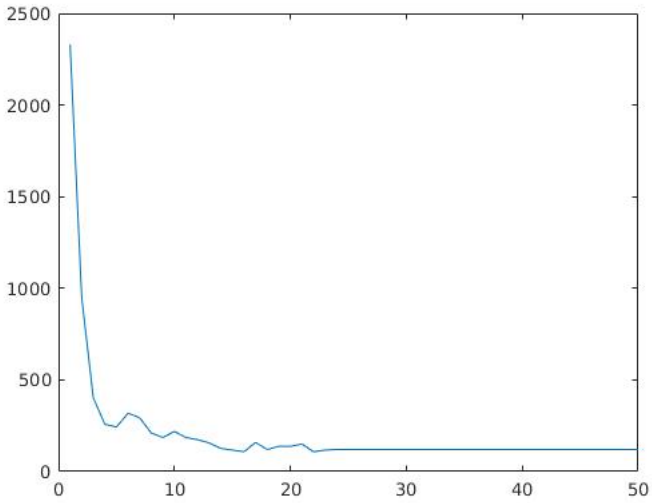


Figure 4: Convergence Exercise 4.2.