Rienforcement learning: Assignment 2

Oluwatomilayo, Adegbite 500569283xercise 3.19
Nikolas, Maier 500461990

February 21, 2019

1

1.1 Question 4

$$V^{\pi}(s) = R_s^{\pi(s)} + \gamma \sum_{s' \in S} P_{ss'}^{\pi(s)} V_{\pi}(s') \tag{1}$$

Let $\pi' = greedy(V_{apr})$

$$\pi'(s) = arg\max_{a \in A} \left\{ R_s^a + \gamma \sum_{s' \in S} P_{ss'}^a V_{apr}(s') \right\} \tag{2}$$

$$L^{\pi}(s) = V^*(s) - V^{\pi'}(s) \tag{3}$$

$$\max_{s \in S} \left\{ L^{\pi'}(s) \right\} \leq \frac{2\gamma\epsilon}{1-\gamma} \tag{4}$$

Must show that $\mid V^*(s) - V_{apr}(s) \mid \leq \epsilon$ for every state s