Assignment 2

Author: **Leonardo Colosi 1799057**

Contributors: **Bruno Francesco Nocera 1863075, Silverio Manganaro 1817504, Simone Tozzi, 1615930, Paolo Renzi 1887793, Jacopo Tedeschi 1882789, Amine Ahardane 2050689.**

MARR, RL

November 22, 2023

Contents

1 Theory

1.1 Exercise 1

Given the following table:

$$Q(s,a) \;=\; \begin{pmatrix} Q(1,1) & Q(1,2) \\ Q(2,1) & Q(2,2) \end{pmatrix} \;=\; \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Assuming $\alpha \;=\; 0.1$ and $\gamma \;=\; 0.5$, after the experience: $(s, a, r, s') = (1, 2, 3, 2)$ we can compute the Q-table update for:

1. Q-Learning
2. SARSA in the case $a' \;=\; \pi_\epsilon(s') \;=\; 2$

1. In the case of Q-Learing we can procede using the following update rule:

$$Q(s,a) = Q(s,a) \;+\; \alpha[\,r \;+\; \gamma(max_a Q(s',a')) \;-\; Q(s,a)]$$

So we would have:

$$\begin{aligned} & Q(s,a) = Q(1,2) \;+\; 0.1[\,3 \;+\; 0.5(max_a Q(2,a')) \;-\; Q(1,2)] \\ \rightarrow \quad & Q(1,2) \;+\; 0.1[\,3 \;+\; 0.5(max(Q(2,1),\; Q(2,2))) \;-\; Q(1,2)] \\ \rightarrow \quad & 2 \;+\; 0.1[\,3 \;+\; 0.5 \cdot 4 \;-\; 2] \\ \rightarrow \quad & 2 \;+\; 0.3 = 2.3 \end{aligned}$$

2. For SARSA we use as update rule:

$$Q(s,a) = Q(s,a) \;+\; \alpha[\,r \;+\; \gamma(Q(s',a')) \;-\; Q(s,a)]$$

in this case $a' = 2$, so we would have:

$$\begin{aligned} & Q(s,a) = Q(1,2) \;+\; 0.1[\,3 \;+\; 0.5(Q(2,a')) \;-\; Q(1,2)] \\ \rightarrow \quad & Q(1,2) \;+\; 0.1[\,3 \;+\; 0.5(Q(2,2)) \;-\; Q(1,2)] \\ \rightarrow \quad & 2 \;+\; 0.1[\,3 \;+\; 0.5 \cdot 4 \;-\; 2) \\ \rightarrow \quad & 2 \;+\; 0.3 = 2.3 \end{aligned}$$

1.2 Exercise 2

2 Code Implementation

2.1 Policy Iteration

2.2 iLQR