Assignment 1, Q5

Majid Ghasemi

September 2024

1 MDPs – Q5

Based on the formulae given in the question, let us answer the first part.

1.1 First Part

Let us re-write the given recursive one-step formula to solve the question. Prior to it, let us put the required information:

$\gamma = 0.7$, and Rewards starting from the first timestep to the last are: 2, 1, 4, -2, 1, and the episode is terminated at $T = 5$, which means there is no return and rewards after it.

We have this general equation:

$$G_t = R_{t+1} + \gamma G_{t+1} \tag{1}$$

Based on it, we can recursively calculate the return in each timestep as follows:

$$G_4 = R_5 + \gamma G_5 \tag{2}$$
$$G_3 = R_4 + \gamma G_4 \tag{3}$$
$$G_2 = R_3 + \gamma G_3 \tag{4}$$
$$G_1 = R_2 + \gamma G_2 \tag{5}$$
$$G_0 = R_1 + \gamma G_1 \tag{6}$$
$$\tag{7}$$

By substituting the numbers given, we have:

$$G_4 = 1 + 0.7(0) \tag{8}$$

Hence, $G_4 = 1$. Recursively, we can update all the returns by using prior G to that timestep:

$$G_3 = -2 + 0.7(1) \tag{9}$$

$G_3 = -1.3$.

$$G_2 = 4 + 0.7(-1.3) \tag{10}$$

$G_2 = 3.09$

$$G_1 = 1 + 0.7(3.09) \tag{11}$$

$G_1 = 3.163$.
And, lastly, $G_0 = 4.2141$.

$$G_0 = 2 + 0.7(3.163) \tag{12}$$

1.2 Second Part

To verify the results we got with the recursive approach, we can use the below formula to see whether we will find the same result or not:

$$G_t = \sum_{k=1}^{T-t} \gamma^{k-1} R_{t+k} \tag{13}$$

Where can be written as:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots + \gamma^{T-t-1} R_T \tag{14}$$

Now, we have:

$$G_0 = R_1 + \gamma R_2 + \gamma^2 R_3 + \gamma^3 R_4 + \gamma^4 R_5 \tag{15}$$

By substituting the numbers we have:

$$G_0 = 2 + 0.7(1) + 0.7^2(4) + 0.7^3(-2) + 0.7^4(1) \tag{16}$$

Where $G_0 = 4.2141$.
As it is evident and as expected, the both ways give the same answer for G_0.

SIGNED AS: MAJID GHASEMI , `September 25`