CS4641 HW5

Kung-hsiang, Huang

2016-11-29

1 K-Means

(a)

Iteration 2:
Centers: 2, 8.3
Data assignment:
Cluster 1: 1, 2, 3, 4
Cluster 2: 9, 12, 6, 10, 9

Iteration 3:
Centers: 2.5, 9.2
Data assignment:
Cluster 1: 1, 2, 3, 4
Cluster 2: 9, 12, 6, 10, 9

(b)

Yes, because the last two assignments are the same, which means centers will not change after this iteration.

2 K-Means and Variance

(a)

Since the algorithm will assign a set of points with smaller variance to be of same clusters as K increases, the variance of solution will decrease.

(b)

By setting K to be equal to the number of data points can guarantee that the variance is 0 as each cluster only contain one instance.

3 Reinforcement Learning I

No, it does not effectively communicate the goal to the agent as time spent in the maze is not taken into consideration. A better approach would be setting the reward of states other than goal state to be -1 so that the longer the agent stays in the maze, the lower the point will be.

4 Reinforcement Learning II

(a)

Only the intervals between them are important as adding a constant C to all the rewards is essentially equal to adding another K to the values of all the states, which will be proven in (b). Therefore, we can add all the reward by the absolute value of the most negative reward value, and the solution will still remains the same.

(b)

$$
\begin{aligned}
V^\pi(s) &= E_\pi[R_t | s_t = s] \\
V^\pi(s) &= E_\pi[\sum_k \gamma^k r_{t+k+1} | s_t = s]
\end{aligned}
$$

Let $\hat{r} = r + C$

$$
\begin{aligned}
\hat{V}^\pi(s) &= E_\pi[\sum_k \gamma^k \hat{r}_{t+k+1} | s_t = s] \\
&= E_\pi[\sum_k \gamma^k r_{t+k+1} | s_t = s] + E_\pi[\sum_k \gamma^k C | s_t = s] \\
&= V^\pi(s) + \frac{C}{1-\gamma} \\
&= V^\pi(s) + K
\end{aligned}
$$

(c)

Based on (b), we know that

$$K = \frac{C}{1-\gamma}$$