Audit Access Expires May 11, 2020
You lose all access to this course, including your progress, on May 11, 2020.

6. Bellman Equations
Bellman Equations

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Recall from lecture the **Bellman Equations** are

$$\begin{aligned} V^*(s) &= \max_a Q^*(s,a) \\ Q^*(s,a) &= \sum_{s'} T(s,a,s') \left(R(s,a,s') + \gamma V^*(s')\right) \end{aligned}$$

where

- the **value function** $V^*(s)$ is the expected reward from starting at state s and acting optimally.
- the **Q-function** $Q^*(s,a)$ is the expected reward from starting at state s, then acting with action a, and acting optimally afterwards.

Value Function in Terms of Q Function

1/1 point (graded)

Let us work through a numerical example to understand the Bellman equations.

Let there be 4 possible actions, $a_1, a_2, a_3, a_4,$ from a given state s, and let the Q^* values be as follows:

$$\begin{aligned} Q^*(s,a_1) &= 10 \\ Q^*(s,a_2) &= -1 \\ Q^*(s,a_3) &= 0 \\ Q^*(s,a_4) &= 11. \end{aligned}$$

Enter the value of $V^*(s)$ below:

11 ✔ **Answer:** 11

Solution:

Note that $V^*(s)$ is given by:

$$V^*(s) = \max_a Q^*(s,a)$$
$$V^*(s) = \max(10,-1,0,11) = 11.$$

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Bellman Equation for Q Function

0/1 point (graded)

As above, let there be 4 possible actions, $a_1, a_2, a_3, a_4,$ from a given state s wth Q^* values given below:

$$Q^*(s,a_1) = 10$$
$$Q^*(s,a_2) = -1$$
$$Q^*(s,a_3) = 0$$
$$Q^*(s,a_4) = 11.$$

Let s' be a state that can be reached from s by taking the action a_1. Let

$$T(s,a_1,s') = 1$$
$$R(s,a_1,s') = 5$$
$$\gamma = 0.5.$$

Enter the value of $V^*(s')$ below:

21/2 ✖ **Answer:** 10

Solution:

Note that since T denotes probabilities, the following must be true:

$$\sum_{s'} T(s, a, s') = 1$$

. Also,

$$Q^*(s, a) = \sum_{s'} T(s, a, s') \left(R(s, a, s') + \gamma V^*(s')\right)$$

Since, $T(s, a_1, s') = 1$ and $\sum_{s'} T(s, a, s') = 1$, we would have $T(s, a_1, s\text{ ''}) = 0 \quad \forall s\text{ ''} \neq s'$.
The above equation would then reduce as follows

$$Q^*(s, a_1) = T(s, a_1, s') \left(R(s, a_1, s') + \gamma V^*(s')\right)$$

$$10 = 1 * (5 + 0.5 * V^*(s'))$$

$$V^*(s') = 5/0.5 = 10$$

Submit You have used 3 of 3 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 5 Reinforcement Learning (2 weeks) :Lecture 17.

Reinforcement Learning 1 / 6. Bellman Equations

Add a Post

Show all posts by recent activity

💬	reward vs. utility It would be better if you reserved the term "reward" for something that you receive at a singl...	1
💬	Staff, I first marked the correct anwser, then I thought I was on the second question and marked the wrong and sent. Could you please check I first marked the right one and reverse the second click....thanks in a...	3
💬	Textbook: Reinforcement Learning: An Introduction by Richard S. Sutton and Andrew G. Barto For those who are interested in Reinforcement Learning, check out this textbook for further c...	6
?	Q* and V* If Q* is the expected reward after taking some action and then acting optimally, how can V* f...	2
💬	[Staff] Entered correct answer, then wrong one accidentally For question 1, I entered the correct answer and got the point. Then, I entered the answer for...	2
💬	Incorrect Equation The last equation on the board at the end of the lecture is incorrect. An opening parentheses...	2
?	Any hints for Q2. Bellman Equation I tried naively to substitute the values in the Bellman equation but am not getting the answer...	6
💬	Yahtzee Optimal Player This lecture reminds me of a past game played: [Solitaire Yahtzee: Optimal Player and Profici...	1