6. Bellman Equations

Bellman Equations

Start of transcript. Skip to the end.

0:00 / 0:00 1.0x

So now we will start introducing a tiny bit more notation.

Just three.

I promise to discuss Bellman equations.

So the first annotation that I will introduce,

as I already alluded, is called V. V star.

This is the value of state.

What does it mean?

It tells you the value of the expected reward

if you're starting at state s and act optimally.

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:

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

$$Q^*(s, a_4) = 11.$$

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

11 ✔ **Answer:** 11

Solution:

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

$$\begin{aligned} V^*(s) &= \max_a Q^*(s, a) \\ V^*(s) &= \max(10, -1, 0, 11) = 11. \end{aligned}$$

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Bellman Equation for Q function

1/1 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:

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

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

$$\begin{aligned} T(s, a_1, s') &= 1 \\ R(s, a_1, s') &= 5 \\ \gamma &= 0.5. \end{aligned}$$

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

10 ✔ **Answer:** 10

Solution:

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

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

. Also,

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

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

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

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

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

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 5 Reinforcement Learning (2 weeks) :Lecture 17. Reinforcement Learning 1 / 6. Bellman Equations