





MITx 6.86x

Machine Learning with Python-From Linear Models to Deep Learning[Course](#)[Progress](#)[Dates](#)[Discussion](#)[Resources](#)[Course](#) / [Unit 5. Reinforcement Learning \(2 we...](#) / [Lecture 17. Reinforcement L](#)[< Previous](#)

7. Value Iteration

Bookmark this page

Exercises due May 3, 2023 08:59 -03 Past due

Value Iteration**Video**
 [Download video file](#)
Transcripts
 [Download SubRip \(.srt\) file](#)
 [Download Text \(.txt\) file](#)

Recall from lecture the **value iteration update rule** :

$$V_{k+1}^*(s) = \max_a \left[\sum_{s'} T(s, a, s') (R(s, a, s') + \gamma V_k^*(s')) \right],$$

where $V_k^*(s)$ is the expected reward from state s after acting optimally for k steps.

Recall the example discussed in the lecture.

| | | | | |
|------------------------------|--|--|--|----|
| Agent's starting state | | | | +1 |
|------------------------------|--|--|--|----|

Then, using the value iteration update rule, we get

Note: Note that as soon as the agent takes the first action to reach cell , it stays for 0 and does not take any more action, so we set for all .

Value Function Update

1.0/1 point (graded)

Run the iteration of the value iteration algorithm to get and answer the following

Enter the value of as an array .

(For example, type `[0, 2, 0, 3, 4]` for the array .)



Submit

You have used 3 of 3 attempts

Number of Steps to Convergence

1/1 point (graded)

Enter below the number of steps it takes starting from for the value function update to reach the optimal value function :



Submit

You have used 2 of 2 attempts

Complexity of Value Iteration

Another Example of Value Iteration (Software Implementation)

3 points possible (graded)

Consider the same one-dimensional grid with reward values as in the first few problems. However, consider the following change to the transition probabilities: At any given grid location, the agent chooses to either stay at the location or move to an adjacent grid location. If the agent chooses to move at a given location, such an action is successful with probability $\frac{1}{2}$ and

- if the agent is at the leftmost or rightmost grid location it ends up at its neighboring location with probability $\frac{1}{2}$,
- if the agent is at any of the inner grid locations it has a probability $\frac{1}{4}$ each of ending up at either of its two neighboring locations.

If the agent chooses to move (either left or right) at any of the inner grid locations, such an action is successful with probability $\frac{1}{2}$ and with probability $\frac{1}{2}$ it fails to move, and

- if the agent chooses to move left at the leftmost grid location, then the action ends up at the leftmost location as choosing to stay, i.e., staying at the leftmost grid location with probability $\frac{1}{2}$, and at the neighboring grid location with probability $\frac{1}{2}$,
- if the agent chooses to move right at the rightmost grid location, then the action ends up at the rightmost location as choosing to stay, i.e., staying at the rightmost grid location with probability $\frac{1}{2}$, and at the neighboring grid location with probability $\frac{1}{2}$.

Note in this setting, we assume that the game does not halt after reaching the rightmost location.

Let V be the value function.

Run the value iteration algorithm for 100 iterations. Use any computational software of your choice.

Enter the value of V at each location as an array

(For example, type `[0, 2, 0, 3, 4]` for the array $[0, 2, 0, 3, 4]$. Type at least 4 decimal places.)

Are the values different if we iterate 200 times? Consider only the first three decimal digits of the values.

☐ Yes

☐ No

Discussion

Topic: Unit 5. Reinforcement Learning (2 weeks) :Lecture 17. Reinforcement Learning
1 / 7. Value Iteration

[< Previous](#)[Next >](#)

Show all posts ▾



edX

[About](#)[Affiliates](#)[edX for Business](#)[Open edX](#)[Careers](#)[News](#)

Legal

[Terms of Service & Honor Code](#)[Privacy Policy](#)[Accessibility Policy](#)[Trademark Policy](#)[Sitemap](#)[Cookie Policy](#)[Do Not Sell My Personal Information](#)

Connect

[Blog](#)[Contact Us](#)[Help Center](#)[Security](#)[Media Kit](#)