ColumbiaX: CSMM.101x Artificial Intelligence (AI)

Help

Week 9: Reinforcement Learning > Week 9 Discussion Questions > Week 9 DQ1

Bookmarks

Week 9 DQ1

Bookmark this page

Complete **Week 9 Optional Project** and post your findings and questions on the discussion forum.

Answer on the discussion forum or in the space below.

Week 9: DQ1

Topic: Week 9 / DQ1

Hide Discussion

Add a Post

All Posts

discussion posted 8 days ago by **RMAYORAL**

Answer to question number 1 of the Optional Project

Good afternoon everybody.

Following the class example and the policy iteration method to find optimal policy for a discount factor of **0.1**, the optimal policy is "FAST ACTION IN MOVING STATE, SLOW ELSEWHERE", as the invited Profesor kindly show to us.

Well, with a discount factor of **0.01** there is no change in the optimal policy and 2 iterations were also required to finish the iteration method.

However, with a discount factor of **0.9 or 0.99**, the optimal policy changes to the starting policy in just one iteration, that is to say, ALWAYS SLOW ACTION.

Have a nice week.

RMS

This post is visible only to Default Group.

Add a Response

2 responses

adkjeff
7 days ago

Building on what RMAYORAL wrote, I ran the policy iteration method for values of gamma from 0 to 0.99 in increments of 0.01.

```
gamma <= 0.83: the value of the fast action in the
moving state
exceeds the value of the slow action in that state, and
the
optimal policy converges to (slow, slow, fast) in 2
iterations for the
(fallen, standing, moving) states.

gamma >= 0.84: the value of the slow action overtakes
the value of the
fast action in the moving state, and the optimal policy
becomes (slow,
slow, slow).
```

Results change if a different starting policy is used. For example, the following results occur for a starting policy of (slow, fast, fast):

```
gamma <= 0.74 -- optimal policy converges to (slow, slow,
fast)
in 2 iterations

0.75 <=  gamma <= 0.83 -- optimal policy converges to
(slow, slow, fast)
in 3 iterations

gamma >= 0.84 -- optimal policy converges to (slow, fast,
fast) in
1 iteration
```

posted 7 days ago by **adkjeff**

[Week 9 Optional Project: Reinforcement Learning (not graded)]

[Week 9 Discussion Questions]

[Week 10: Logical Agents]

If we start the policy iteration process with the policy (slow, slow, slow), the "I-gamma*P" (gamma = 0.1) matrix has an eigenvector that contains a value near 0.83:

```
P matrix:
0.6    0.4    0
0.0    0.0    1
0.0    0.0    1

I-gamma*P matrix:
0.94   -0.04    0.0
0.00    1.00   -0.1
0.00    0.00    0.9

eigenvalues:
[1.00, 0.94, 0.90]

eigenvectors:
-0.5547002   1   0.5773503
0.8320503    0   0.5773503
0.0000000    0   0.5773503
```

[Staff] -- is that value of 0.8320503 in one of the eigenvectors related to where the switch in the optimal policy occurs, or is it just a coincidence?

posted 6 days ago by **adkjeff**

Add a comment

RMAYORAL

7 days ago

¡¡¡Great, precise and complete observations/remarks Adkjeff¡¡¡

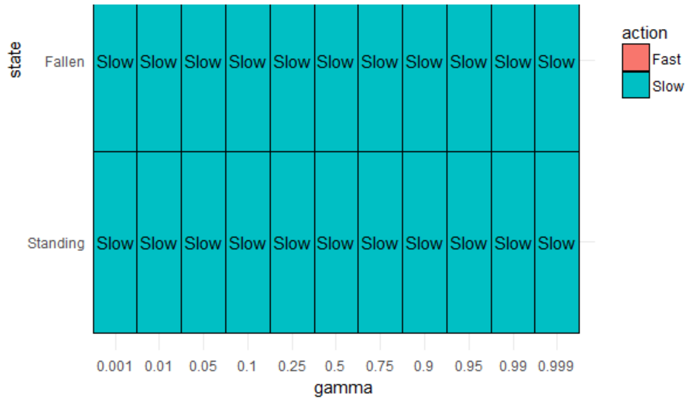
Thanks a lot,

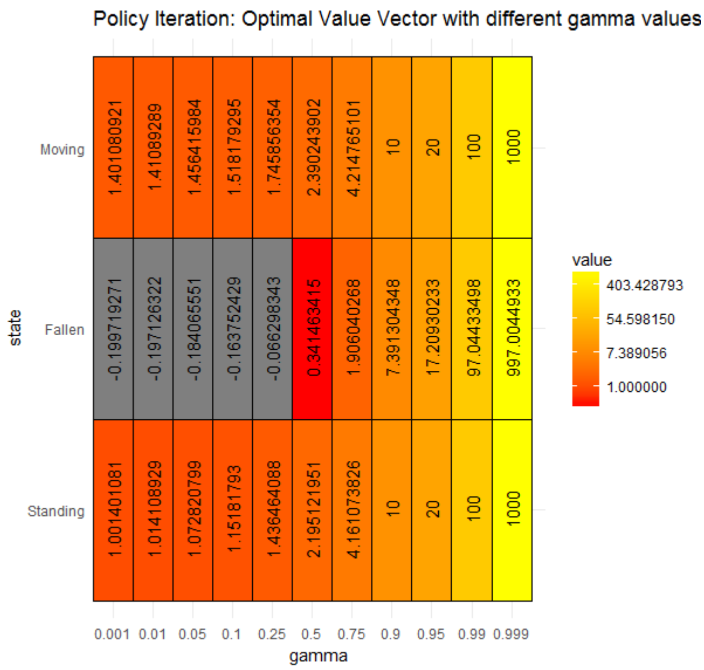
Renato Mayoral (rmayoral)

@Adkjeff @RMAYORAL do you have an intuition on the change of the optimal policy based on the value in an eigenvector? does it have to be the dominant eigenvector? does it have to be the 2nd value of the vector? is there a mathematical proof regarding this?

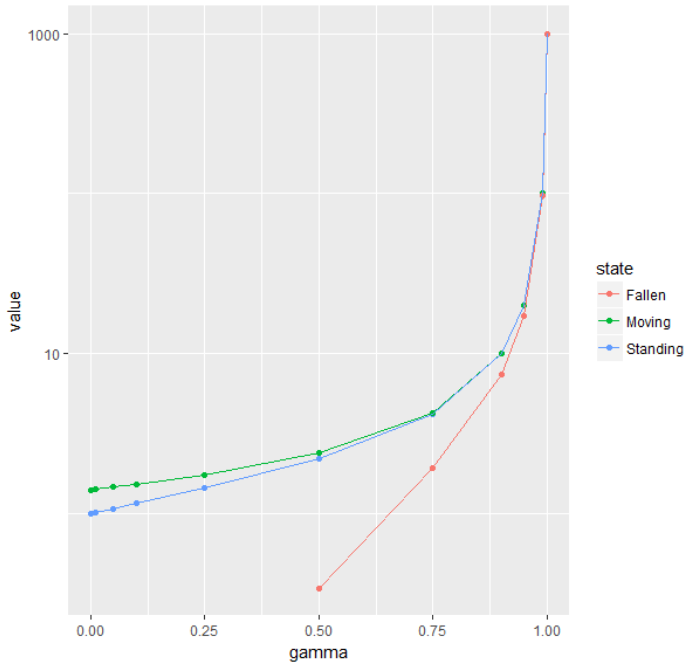
However, here are the results (visualizations) with policy iteration started at (slow, slow, slow) with different gamma values:

Policy Iteration: Optimal Policy with different gamma values





Policy Iteration changes in the value vector corresponding to the optimal policy



Can you guys verify the results i obtained with yours, so that i can be confident about the correctness of my implementation?

posted about 23 hours ago by **sandipan_dey**

@sandipan_dey: Wow -- nice way of graphically displaying the results. I had no deep intuition about any relationship between the eigenvectors and the "switching point" of the optimal policy. I simply couldn't think of anything else, and was surprised to see the 0.83 value in one of them.

posted about 22 hours ago by **adkjeff**

@sandipan_dey: Compliments about your visualizations. Do you create these professionally, or did you take some course for it? Or both?

posted about 18 hours ago by **wvdzwart**

@adkjeff @wvdzwart Thank you very much. @adkjeff A nice observation indeed. [STAFF] could you please share some intuition on this? @wvdzwart I visualized these using R, I am working on R past few years, got a chance to learn a few things by myself on work.

posted about 14 hours ago by **sandipan_dey**

Add a comment

Showing all responses

Add a response:

Preview

Submit