On-policy Prediction with Approximation

Practice Assignment • 30 min

← Back

English

Your grade: 98.33%

Your latest: **98.33%** • Your highest: **98.33%**

To pass you need at least 80%. We keep your highest score.

Next item →

1. Which of the following statements is true about function approximation in reinforcement learning? (**Select all that apply**) 1 / 1 point

- [ ] We only use function approximation because we have to for large or continuous state spaces. We would use tabular methods if we could, and learn an independent value per state.
- [x] It allows faster training by generalizing between states.

  Correct
  Correct! Function approximation allows the agent to generalize to unseen but similar states, and can learn the value function more quickly. Furthermore, in continuous state/action spaces the agent may never see the same state twice and we need such generalization to accurately estimate the values.

- [x] It can be more memory efficient.

  Correct
  Correct! We cannot enumerate and store all states in a table for large or continuous state spaces. By using function approximation, we can use fewer parameters to represent the value function.

- [x] It can help the agent achieve good generalization with good discrimination, so that it learns faster and represent the values quite accurately.

  Correct
  Correct. Recall the 2D plot of generalization and discrimination. Tabular methods discriminate between different states perfectly but with no generalization. Alternatively, one could treat all states as the same, with each update generalizing to all states but with no discrimination. Ideal function approximation methods achieves both good generalization and good discrimination.

2. We learned how value function estimation can be framed as supervised learning. But not all supervised learning methods are suitable. What are some characteristics of reinforcement learning that can make it harder to apply standard supervised learning methods? 1 / 1 point

- [x] When using bootstrapping methods like TD, the target labels change.

  Correct
  Correct. Targets depend on our own estimates, and these estimates change as learning progresses.

- [ ] Data is available as a fixed batch.
- [x] Data is temporally correlated in reinforcement learning.

  Correct
  Correct!

3. Value Prediction (or Policy Evaluation) with Function Approximation can be viewed as supervised learning mainly because _____. [choose the most appropriate completion of the proceeding statement] 1 / 1 point