

Exercise set #9

Solution should be submitted in teams of two if possible. Due to the current COVID-19 pandemic please submit your solution online using the sciebo file-drop folder. The link will be available in ILIAS. Please submit a single zip file with the following naming scheme: `username1-username2.zip` (e.g. `jadoe101-jodoe108.zip`). Allowed file extensions (of files within the zip file) are: `.pdf`, `.txt`, `.py` and `.ipynb`. Make sure the total file size does not exceed 10 MB.

Exercises are adapted from Richard Sutton's CMPUT609 [1].

1. REINFORCE eligibility vector (Exercise 13.3 from Sutton and Barto [2])

Let $\mathbf{x}(s, a) \in \mathbb{R}^n$ be a feature vector. For

$$h(a|s, \theta) = \theta^T \mathbf{x}(s, a), \text{ where } \theta \in \mathbb{R}^n$$

and

$$\pi(a|s, \theta) = \frac{e^{h(a|s, \theta)}}{\sum_b e^{h(b|s, \theta)}}$$

prove that

$$\nabla_{\theta} \ln \pi(a|s, \theta) = \mathbf{x}(s, a) - \sum_b \pi(b|s, \theta) \mathbf{x}(s, b).$$

Hint: for $f : \mathbb{R}^n \rightarrow \mathbb{R}$ it holds that $\nabla_{\theta} \ln f(\theta) = \frac{\nabla_{\theta} f(\theta)}{f(\theta)}$ and $\nabla_{\theta} e^{f(\theta)} = e^{f(\theta)} \nabla_{\theta} f(\theta)$.

40 points

2. Short-Corridor Gridworld (programming task)

Implement the Short-Corridor Gridworld described in Example 13.1 from Sutton and Barto [2] as a Gym¹ environment. Follow the instructions in `exercises09.ipynb`.

20 points

3. REINFORCE (programming task)

Implement REINFORCE for the Short-Corridor Gridworld from the previous exercise. Use the policy parametrization from Exercise 1 and the feature vector $\mathbf{x}(s, 0) = [1, 0]^T$ and $\mathbf{x}(s, 1) = [0, 1]^T$ for all states s . Follow the instructions in `exercises09.ipynb`.

40 points

References

- [1] CMPUT 609: reinforcement learning for artificial intelligence. <http://incompleteideas.net/rlai.cs.ualberta.ca/RLAI/RLAICourse/2009.html>.
- [2] Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT press, 2018.

¹<https://gym.openai.com>