

ECEN 689: RL: Fall 2018

Midterm Programming Question

This programming assignment is to learn a control algorithm for the Lunar Lander! Details regarding the environment is [here](#). The environment consists of a lander with 4 discrete actions and a continuous state space. Play with the environment to get an idea of the reward function and the task.

Solve this using two approaches.

1. Q-Learning with Linear Function Approximation
2. Q-Learning with Neural Networks

You can use standard library (such as tensorflow, keras, torch or others) functions. Feel free to modify the reward signal if you think it is necessary to converge and indicate your reasoning if you've chosen to modify.

Submission Details: Generate an HTML report that contains

1. Video of the resultant policy execution.
2. Plots showing the convergence (error, weights, . . .).
3. Mathematical representation of your linear function approximation and gradient.
4. The performance of the linear function approximation method depends on the features used. It is possible that the policy might not converge for some feature vectors. If so, indicate why it didn't converge and what modifications might make the policy converge.
5. Details of the NN model used.
6. Comment which method worked better and why? Comment on the advantages and disadvantages for each method.

Please be patient with the training. It can take a while!