

Assignment 4 – Yair Vaknin 034662072, Liran Nochumsohn 204693410

In this project we implement the actor-critic algorithm to solve the pendulum problem.

For approximation of the value function, we use RBF with 100 different possible states that are sampled from the environment. Each state represents a dot and theta dot with the expression $[\sin(\theta), \cos(\theta), \dot{\theta}]$. Each sample will be featured using RBF with $\sigma = 1$ and RBF with $\sigma = \sqrt{0.25}$.

In our algorithm we used the following parameters:

α (step size update for the critic) = 0.005

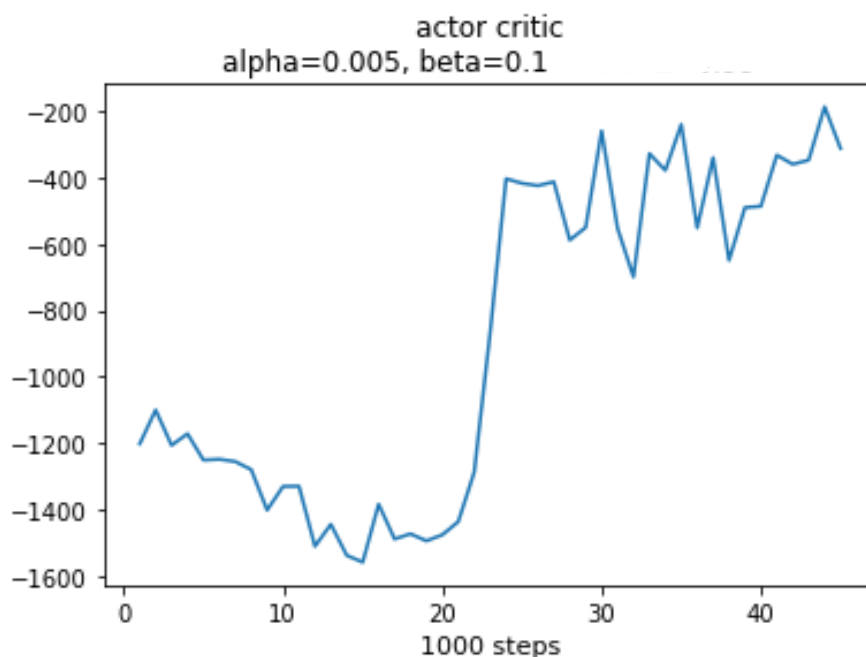
β (step size update for the actor) = 0.1

$\gamma = 0.99$ (for avoiding exploding updates)

σ (for critic) = 1.1

The algorithm finds optimality at around the -200 mark.

We must note, although γ was used for the learning process, it was not using for the evaluations is observed in the plot.



player2.py – includes the class for actor critic and all the functions.

Some of the function:

actor_critic () – the actor critic algorithm

ret_action() – return the action based on the given policy probability.

RBF() – a function for construction the gaussian feature space.

gradient_t () – a function for getting the score function for θ (critic)

main.py – is the run file with the simulation.