

Meeting

Po Hsun Wu

August 08, 2022

Progress report

- Changing the reward function to Quadratic function.

$$r(t) = - \int_0^t x^T Q x + u^T R u d\tau$$

- Changing the range of boundary.
 - boundary: $\theta_1 = [180^\circ, -180^\circ] \rightarrow [10^\circ, -10^\circ]$
 - initial state: $\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2 = [0.05, -0.05]$ with uniform distributed
- Studing the Proximal Policy Optimization(PPO) algorithm.