

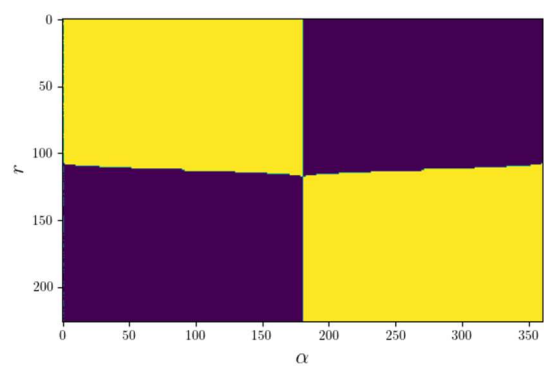
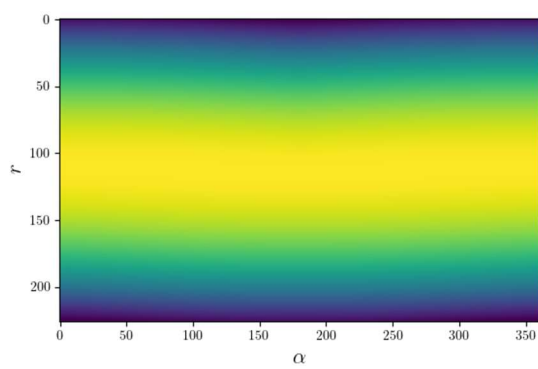
Optimal Control and Reinforcement Learning

Homework #3

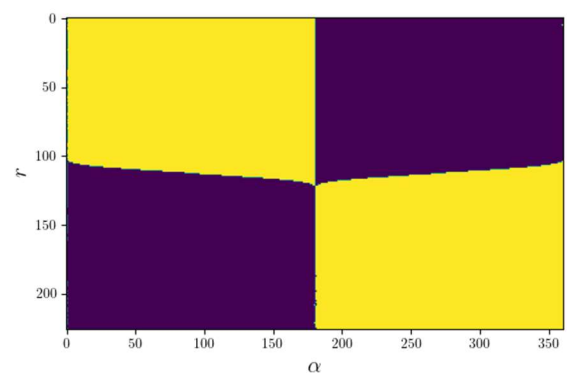
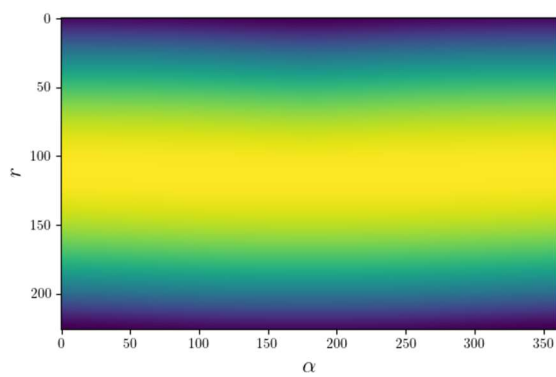
Student ID : 2016145015

Name : 이두현

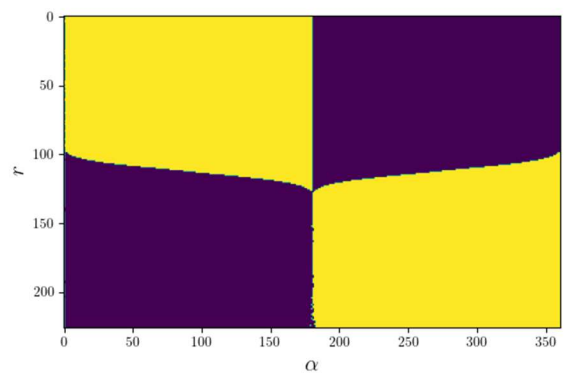
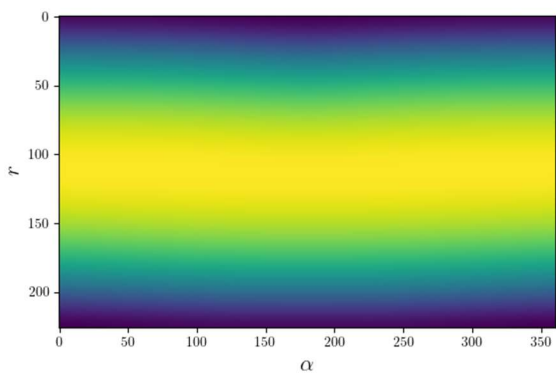
- Sigma = 0



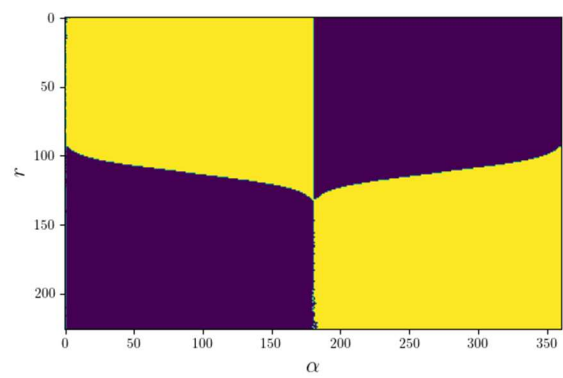
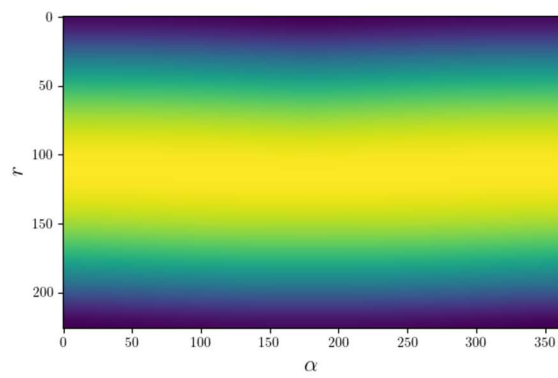
- Sigma = 2



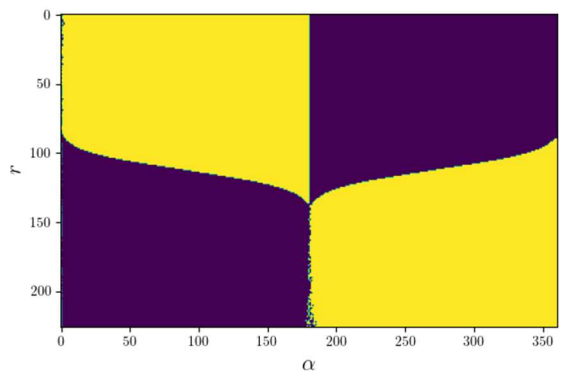
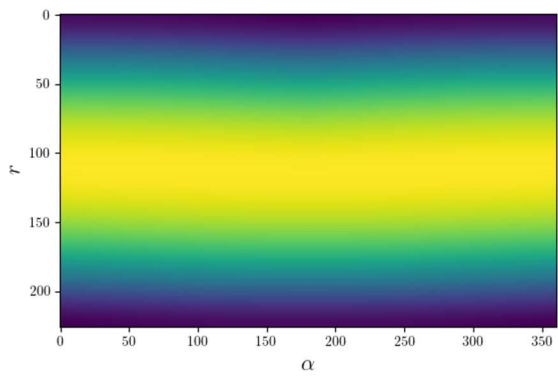
- Sigma = 4



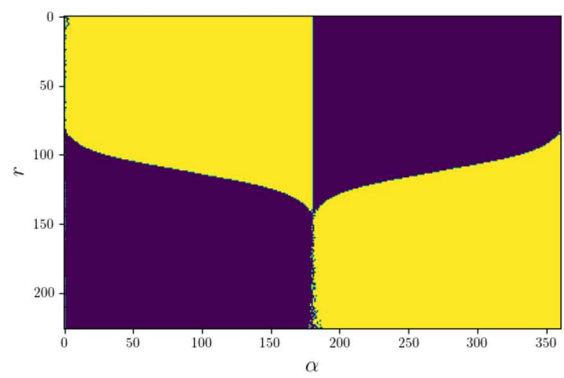
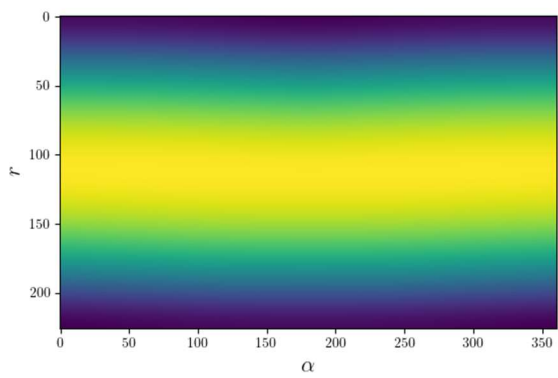
- Sigma = 6



- Sigma = 8



- Sigma = 10



1. Value map의 경우에는 $r_{\min} = 10$, $r_{\max} = 90$ 범위에서 $_rewards_tmp = (r - keep_dist) ** 2 / p_denominator$ 로 정의되어 있기 때문에 $r = 50$ (keeping distance = 50)에서 reward가 가장 높게 나오는 것을 알 수 있었습니다. 또한, sigma가 커짐에 따라서 target의 motion uncertainty가 증가하기 때문에, $r = 50$ 에서 멀 때의 reward가 더 빠르게 감소하는 것을 알 수 있고 그래서 value map 상에서 봤을 때 가운데 $r = 50$ 부분으로 좀 더 concentration 되는 것으로 보입니다..

2. Policy map의 경우에는 $\sigma = 0$, 즉 deterministic의 경우에 살펴보면 $r < d$ 인 경우에는 $\alpha < 0$ 이어야 reward가 크고, $r > d$ 인 경우에는 $\alpha > 0$ 이어야 reward가 큰 것을 알 수 있습니다. Sigma가 커지면 target의 motion uncertainty가 커지므로 $|\alpha|$ 가 큰 부분에서의 control 경계가 날카롭게 변하는 것을 알 수 있습니다. 또한, $\alpha = 0$, $r = d$ 근처 부분에서도 마찬가지로 uncertainty가 증가함에 따라서 control 경계가 날카롭게 변한 것으로 보입니다.