

State:

Everything working Q1-Q7. Didn't manage to do Q8. For Q6 I am not sure if answered it correctly. I put it is not possible to get desired effect with only controlling epsilon and learning rate.