

## 1 Context

Deep Deterministic Gradient Descent or DDPG is a type of off policy, model free Deep RL method. It takes it's basis from the actor-critic method but applied in the setting of continuous control.

## 2 Algo description

As mentionned DDPG uses an Actor-Critic construction. The policy will be defined as the actor network. In this case, since it is deterministic, the output of the network will be a valid action that our agent can take in the environment:

$$\mathcal{N}_\pi : \mathcal{X} \rightarrow \mathcal{A} \tag{1}$$

the parameters defining network  $\mathcal{N}_\pi$  will be addressed as  $\theta^\pi$ . The critic role will be played by a function approximator, whose goal is to estimate the state action value function:

$$\mathcal{Q} : \mathcal{X} \times \mathcal{A} \rightarrow \tag{2}$$

The reason we call this method an off policy method is because the weights of the policy network are updated outside of the current environment run through.

The weights are initialized randomly for both the actor and the critic networks. We let the agent discover the environment. This is accomplished by applying an action