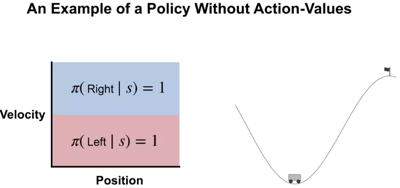
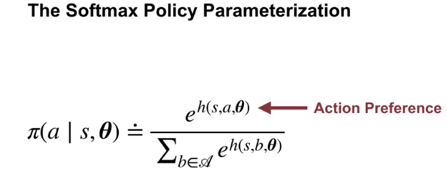
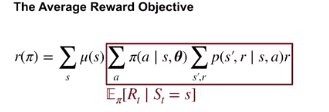
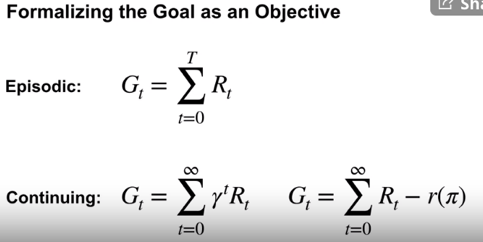
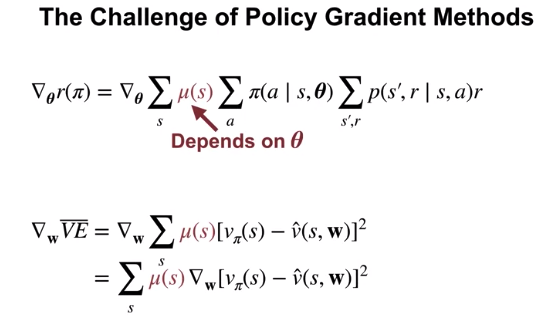
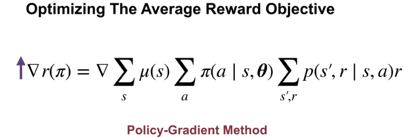
# Week4 Notes

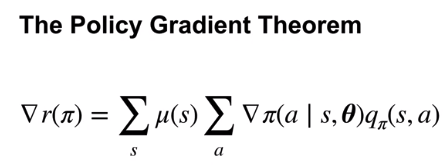
**Lesson 1: Learning Parameterized Policies**

* Understand how to define policies as parameterized functions
  + 
* Define one class of parameterized policies based on the softmax function
* Understand the advantages of using parameterized policies over action-value based methods
  + Flexibility of stochastic policies
  + Stochastic policies might be better than deterministic under fn approximation

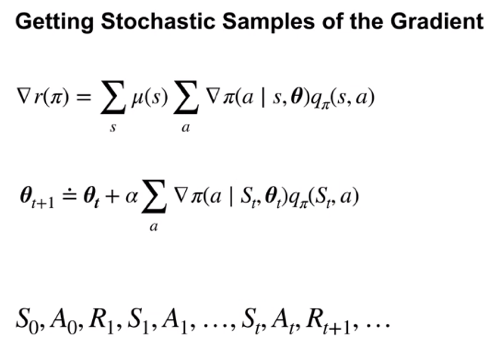
**Lesson 2: Policy Gradient for Continuing Tasks**

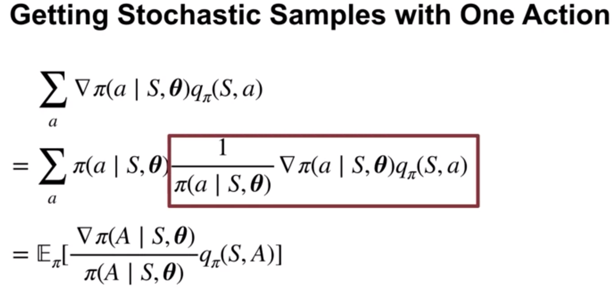
* Describe the objective for policy gradient algorithms
* 

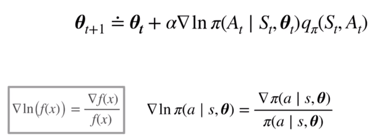
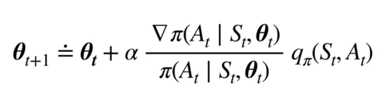


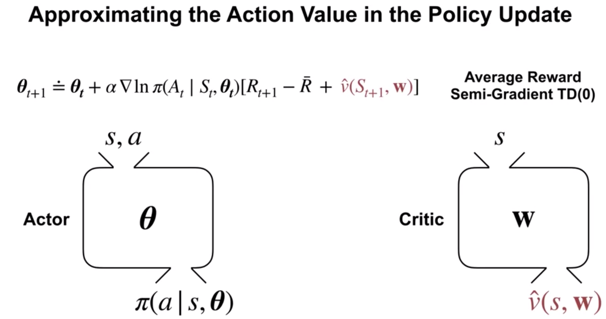
* Describe the results of the policy gradient theorem
  + 
* Understand the importance of the policy gradient theorem

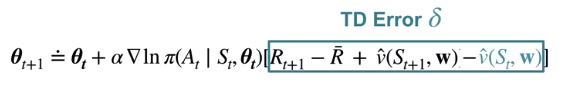
**Lesson 3: Actor-Critic for Continuing Tasks**

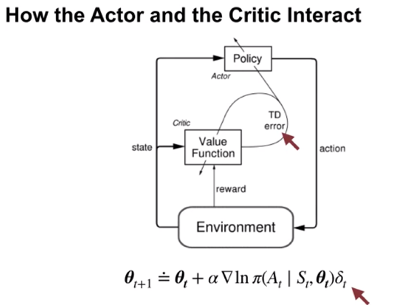
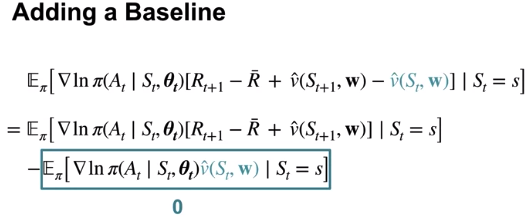
* Derive a sample-based estimate for the gradient of the average reward objective
* 

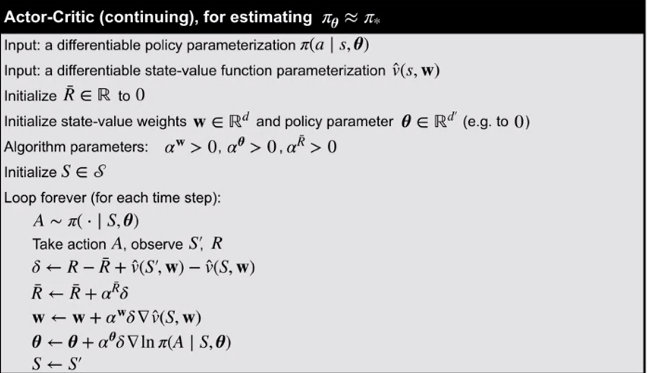




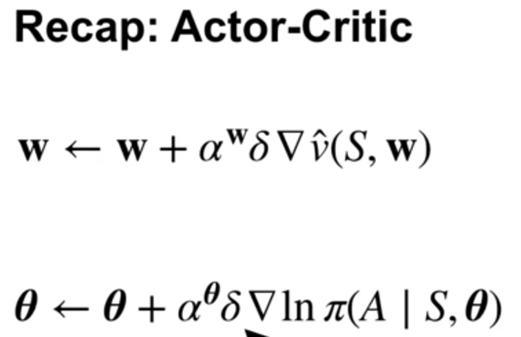
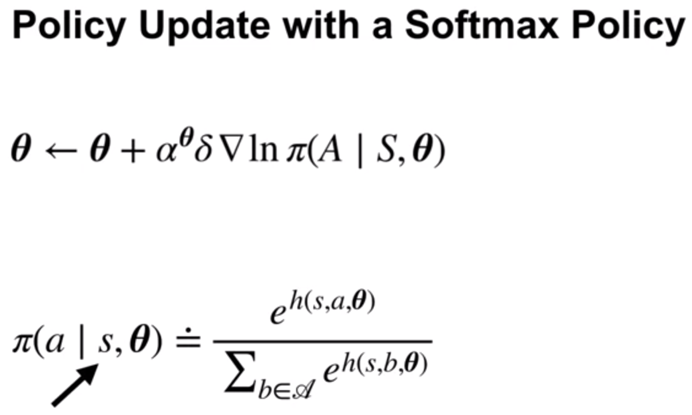
* Describe the actor-critic algorithm for control with function approximation, for continuing tasks
* 

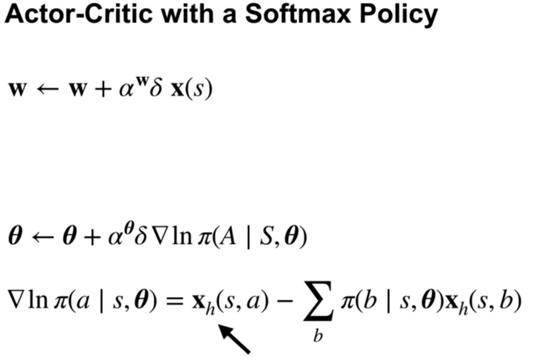


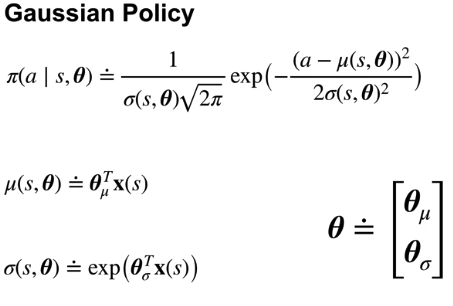
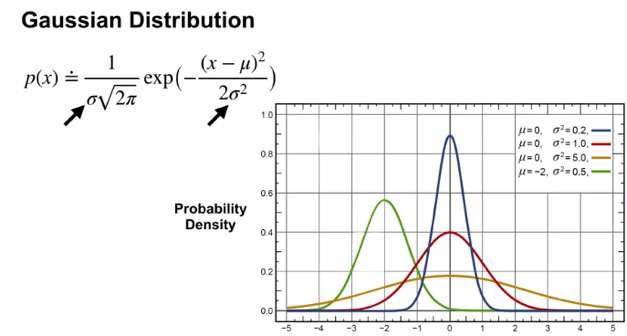
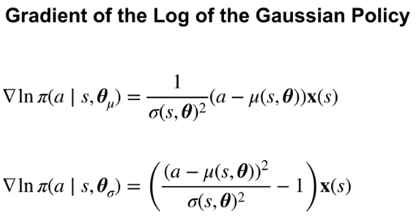




**Lesson 4: Policy Parameterizations**

* Derive the actor-critic update for a softmax policy with linear action preferences
* 
* 



* Implement this algorithm
* Design concrete function approximators for an average reward actor-critic algorithm
* Analyze the performance of an average reward agent
* Derive the actor-critic update for a gaussian policy
* 
* Apply average reward actor-critic with a gaussian policy to a particular task with continuous actions
* 
* 