# Move 37
# Policy Gradients - Study Guide

**Policy gradient methods...**

- Use gradient ascent to adjust toward a policy with greater reward
- Are model-free
- Are 'on policy'
- Are a form of policy search

**Main types of policy gradient methods:** (source)

- Finite difference methods
- Likelihood ratio methods (REINFORCE)
- Natural policy gradients

**How do policy gradients compare to other methods?**

- They are prefered to DQN (source)
- They are commonly used as an actor for actor-critic methods (source)

Policy gradients compared to value based methods: (source)

**Advantages**:

- Better convergence properties (guarantees local convergence)
- Effective in high dimensional/ continuous action spaces
- Can learn stochastic policies

**Disadvantages**:

- Tends to converge on local rather than global optimum
- Evaluating the policy can be very inefficient

**Extra facts:**

- *AlphaGo* uses policy gradients in combination with monte carlo tree search (source)
- REINFORCE was the first policy gradient method introduced in 1992 (source)
- REINFORCE is sometimes called monte-carlo policy gradient (source)
- Policy gradients can be used as an actor in actor-critic (source)
- The original DQN authors prefer policy gradients (source)

For an extensive list of methods based on policy gradients, see this blogpost.