Markov Decisions, Powerful RL Tools

Shoaib Akthar
@theshoaibakthar

Follow for more Data Analytics, Data Science and AI insights

Understanding Markov Property

- A state's future depends only on the present state.
- No memory of past states affects decisions.
- Simplifies problem-solving and learning.
- Makes learning more efficient.
- Useful in many real-world scenarios.

Markov Decision Processes Defined

- A process with a state, action, reward, and transition probability.
- State is the current situation.
- Action is the choice made.
- Reward is the feedback received.
- Transition probability defines how we move from one state to another.

State Transitions

- The probability of moving to a new state given the current state and action is the transition probability.
- This probability is often denoted as P(next state | current state, action).
- States can be discrete (distinct) or continuous.
- Continuous states require discretization for modeling.
- This process is often represented by a transition matrix.

Value Functions

- Value function estimates the expected cumulative reward from a given state following a specific policy.
- Value of a state V(s) indicates how good it is to be in that state.
- V(s) = E[R_t | S_t = s] where R_t is reward at time t.
- Different types of value functions include state-value and action-value functions.

Policy Functions

- A policy function defines the strategy for selecting actions in a given state.
- A policy π(a | s) gives the probability of taking action a in state s.
- Optimal policy maximizes cumulative reward.
- Finding the optimal policy is a key goal in reinforcement learning.
- Examples include greedy policies and epsilon-greedy policies.

Policy Iteration

- A method for finding the optimal policy iteratively.
- Value iteration is a predecessor to policy iteration.
- It involves alternating between policy evaluation and policy improvement.
- Policy Evaluation finds the optimal value function.
- Policy Improvement finds a better policy based on the value function.

Value Iteration

- An iterative algorithm for finding the optimal state-value function.
- Uses the Bellman equation to update value estimates.
- Ensures convergence to the optimal value function.
- Guaranteed to find the optimal value function if the Markov property holds.

Applications & Future

- Many applications in robotics, game playing, and resource management.
- Can be used for planning, decision-making in complex environments.
- Continues to be a very active area of research.
- Improves AI's ability to learn optimal strategies.
- A fundamental concept in modern Reinforcement Learning.

Follow for More

Like this post? Repost!

Share your thoughts in the comments!

I share little hacks and tips I actually use in Data & AI — follow if you wanna see more!