Topics in Reinforcement learning
(Quiz 1, 60 mins)

Each Question is for 10 marks

1. Describe limiting distribution and stationary distributions of a Markov chain ? Do both need to exist ? Why does the stationary distribution exists always for a Markov chain ? Give an example of a chain where the limiting distribution does not exist but the stationary distribution does.

2. State and prove the finite horizon MDP Bellman optimality equations.

3. Give an algorithmic procedure to simulate successive states of an arbitrary Markov chain with M states and having a transition probability matrix P and initial distribution of μ. Justify why this process results in valid sample paths of the Markov chain.

4. Obtain the stationary distribution for a Markov chain with state space $S = \{1, 2\}$ with the following transition probability matrix $P = \begin{bmatrix} 0.1 & 0.9 \\ 0.1 & 0.9 \end{bmatrix}$.

   Now suppose you get a reward of 10 everytime you visit state 1 and reward of 0 otherwise. Assuming a discount factor $\beta = 0.5$, obtain the value function ($V(s)$ for $s = 1, 2$) associated with this Markov reward process.