RL class - quiz on chapter 1

Total des points 3/7

Student last name *

IGBIDA

Student first name *

Rayanne

Question 1 *0/1

How many stationary deterministic memory-less policies are there in an MDP with 5 states and 3 actions?

- ○ 8
- ● 5 ✗
- ○ 15
- ○ 20

Bonne réponse

- ● 15

Question 2 1/1

Consider the family of finite horizon MDPs of horizon H equipped with a discounted criterion $\mathbb{E}\left(\sum_{t=0}^{H}\gamma^t R_t\right)$ on trajectories. Is there always an optimal stationary policy for such MDPs?

- ○ Yes
- ◉ No ✓

Question 3 0/1

An MDP equipped with a policy $\pi = (\pi_t)_{t\in\mathbb{N}}$ defines a sequence on state random variables $(S_t)_{t\in\mathbb{N}}$ that is a Markov chain.

- ○ True
- ◉ False ✗

Bonne réponse

- ◉ True

Question 4 0/1

Take a finite state space MDP, a stationary, memoryless, stochastic policy π and an initial state distribution $\mathbb{P}(S_0 = s) = \mu(s)$. What is the size of the transition matrix p^π whose p^π_{ij} element is $\mathbb{P}(s' = s_j | s = s_i)$? What is the dependence of p^π_{ij} on π, on the transition model $p(s'|s,a)$, on the initial state distribution μ?

$|S| \times |S|$ and $p_{ij} = \int_A p(s'|s,a)\pi(a|s)da$

Option 1

$|S| \times |A|$ and $p_{ij} = \int_A p(s'|s,a)\pi(a|s)da$

Option 2

$|S||A| \times |S|$ and $p_{ij} = \int_A p(s'|s,a)\pi(a|s)\mu(s)dsda$

Option 3

$|S| \times |S|$ and $p_{ij} = \int_A p(s'|s,a)\pi(a|s)\mu(s)dsda$

- (●) Option 4 ✗

Bonne réponse

- (●) Option 1

✓ **Question 5** * 1/1

Take a finite state and action space MDP and a fixed policy. Suppose the corresponding Markov chain on states is irreducible and aperiodic. What does $(p^{\pi})^k$ tend to when $k \to \infty$?

- ( ) A diagonal matrix
- (●) A matrix whose lines are all equal to the stationary distribution ✓
- ( ) A matrix whose columns are all equal to the stationary distribution

✓ **Question 6** * 1/1

Take a finite state and action space MDP, a fixed policy, and a given probability distribution $\mu(s)$ on starting states. What is the probability distribution on states after k transitions?

$$\mu\left((p^{\pi})^k\right)^{-1}$$

Option 1

$$\mu\left(p^{\pi}\right)^{k}$$

Option 2 ✓

$$\mu$$

Option 3

$$\left(p^{\pi}\right)^{k}\mu$$

Option 4

Question 7 *

0/1

Take a finite state and action space MDP, a fixed policy, and a given probability distribution $\mu(s)$ on starting states. Pick the true statement(s) about the state occupancy measure ρ_μ^π.

$$\rho_\mu^\pi(s) = \lim_{t\to\infty} \mu\,(p^\pi)^t$$

- [ ] Option 1

$$\rho_\mu^\pi(s) = \sum_{t=0}^{\infty} \gamma^t \mu\,(p^\pi)^t$$

- [ ] Option 2

$\rho_\mu^\pi(s)$ sums to 1

- [ ] Option 3

$\rho_\mu^\pi(s)$ sums to $1-\gamma$

- [ ] Option 4

$\rho_\mu^\pi(s)$ sums to $\frac{1}{1-\gamma}$

- [x] Option 5 ✓

$\mathbb{E}\left(\sum_{t=0}^\infty \gamma^t R_t \middle| S_0 \sim \mu\right) = \langle \rho_\mu^\pi, r^\pi \rangle$

- [ ] Option 6

Bonne réponse

- [x] Option 2
- [x] Option 5
- [x] Option 6