



EE932 Assignment-2 Solution

eMasters in Communication Systems, IITK

EE932: Introduction to Reinforcement Learning

Instructor: Prof. Subrahmanya Swamy Peruru

Student Name: Venkateswar Reddy Melachervu

Roll No: 23156022

Question 9:

In the Bellman expectation equation, we related V^π in terms of V^π . Write down an equation that expresses Q^π in terms of Q^π .

Hint: Refer to the Week 3 (Part 2) slide with the title “Relating Q^π and V^π ”

Solution:

$$\begin{aligned}
Q_\pi(s, a) &= \mathbb{E}_\pi\{G_t \mid S_t = s, A_t = a\} \\
&= \mathbb{E}_\pi\{R_{t+1} + \gamma G_{t+1} \mid S_t = s, A_t = a\} \\
&= \mathbb{E}\{R_{t+1} \mid S_t = s, A_t = a\} + \gamma \mathbb{E}_\pi\{G_{t+1} \mid S_t = s, A_t = a\} \\
&= R_s^a + \gamma \sum_{s'} P_{ss'}^a \mathbb{E}_\pi\{G_{t+1} \mid S_{t+1} = s', A_{t+1:\infty} \sim \pi\} \\
&= R_s^a + \gamma \sum_{s'} P_{ss'}^a V_\pi(s')
\end{aligned}$$

Substituting $V_\pi(s) = \sum_a \pi(a \mid s) Q_\pi(s, a)$ into above Q_π equation,

$$Q_\pi(s, a) = R_s^a + \gamma \sum_{s'} P_{ss'}^a \sum_{a'} \pi(a' \mid s') Q_\pi(s', a')$$

Substituting $\sum_{a'} P_{ss'}^{a'} \pi(a' \mid s') = P_{ss'}^\pi$,

$$Q_\pi(s, a) = R_s^a + \gamma \sum_{s', a'} P_{ss'}^\pi Q_\pi(s', a')$$

----- End of the Document -----

