

Exercise 09 [Avik Banerjee (3374885), Soumyadeep Bhattacharjee (3375428)]

Tuesday, 14. July 2020 21:47

$$1) (a) Q_{\Omega}(s, w) = \sum_a \pi_{w, \theta}(a|s) Q_U(s, w, a)$$

$$Q_U(s, w, a) = r(s, a) + \gamma \sum_{s'} P(s'|s, a) U(w, s')$$

$$\begin{aligned} \frac{\partial Q_{\Omega}(s, w)}{\partial \theta} &= \sum_a \frac{\partial}{\partial \theta} [\pi_{w, \theta}(a|s) Q_U(s, w, a)] \\ &= \sum_a \pi_{w, \theta}(a|s) \frac{\partial}{\partial \theta} Q_U(s, w, a) \\ &\quad + Q_U(s, w, a) \frac{\partial}{\partial \theta} \pi_{w, \theta}(a|s) \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial \theta} Q_U(s, w, a) &= \frac{\partial}{\partial \theta} \left[r(s, a) + \gamma \sum_{s'} P(s'|s, a) U(w, s') \right] \\ &= 0 + \gamma \sum_{s'} P(s'|s, a) \frac{\partial U(w, s')}{\partial \theta} \\ &= \gamma \sum_{s'} P(s'|s, a) \frac{\partial U(w, s')}{\partial \theta} \end{aligned}$$

$$\therefore \frac{\partial Q_{\Omega}(s, w)}{\partial \theta} = \sum_a \left[\pi_{w, \theta}(a|s) \cdot \gamma \sum_{s'} P(s'|s, a) \frac{\partial U(w, s')}{\partial \theta} + Q_U(s, w, a) \frac{\partial}{\partial \theta} \pi_{w, \theta}(a|s) \right]$$

(b)

$$\frac{\partial U(w, s')}{\partial \theta} = \frac{\partial}{\partial \theta} \left[(1 - \beta_{w, v}(s')) Q_{\Omega}(s', w) + \beta_{w, v}(s') V_{\Omega}(s') \right]$$

$$= \frac{\partial}{\partial \theta} \left[Q_{\Omega}(s', \omega) - \beta_{\omega, \nu}(s') Q_{\Omega}(s', \omega) + \beta_{\omega, \nu}(s') V_{\Omega}(s') \right]$$

$$= \frac{\partial}{\partial \theta} Q_{\Omega}(s', \omega) - \beta_{\omega, \nu}(s') \frac{\partial}{\partial \theta} Q_{\Omega}(s', \omega) + \beta_{\omega, \nu}(s') \frac{\partial V_{\Omega}(s')}{\partial \theta}$$

2)(c)

