Assignment 3

Miao-Chin Yen

January 24, 2022

Problem 1

For a deterministic Policy, $\pi_D : \mathcal{S} \rightarrow \mathcal{A}$, i.e., $\pi_D(s) = a$, where $s \in \mathcal{S}, a \in \mathcal{A}$.
MDP(State-Value Function) Bellman Policy Equation $V^{\pi_D} : \mathcal{N} \rightarrow \mathbb{R}$:

$$V^{\pi_D}(s) = \mathcal{R}(s, \pi_D(s)) + \gamma \cdot \sum_{s' \in \mathcal{N}} \mathcal{P}\left(s, \pi_D(s), s'\right) \cdot V^{\pi_D}\left(s'\right)$$

Action-Value Function (for policy π_D) $Q^{\pi_D} : \mathcal{N} \times \mathcal{A} \rightarrow \mathbb{R}$:

$$Q^{\pi_D}(s, \pi_D(s)) = \mathcal{R}(s, \pi_D(s)) + \gamma \cdot \sum_{s' \in \mathcal{N}} \mathcal{P}\left(s, \pi_D(s), s'\right) \cdot V^{\pi_D}\left(s'\right)$$

$$V^{\pi_D}(s) = Q^{\pi_D}(s, \pi_D(s))$$

$$Q^{\pi_D}(s, \pi_D(s)) = \mathcal{R}(s, \pi_D(s)) + \gamma \cdot \sum_{s' \in \mathcal{N}} \mathcal{P}\left(s, \pi_D(s), s'\right) \cdot Q^{\pi_D}\left(s', \pi_D(s')\right)$$

Problem 2

TBD

Problem 3

TBD

Problem 4

MDP State-Value Function Bellman Optimality Equation:

$$V^*(s) = \max_{a \in \mathcal{A}} \left\{ \mathcal{R}(s, a) + \gamma \cdot \sum_{s' \in \mathcal{N}} \mathcal{P}\left(s, a, s'\right) \cdot V^*\left(s'\right) \right\}$$

Consider the myopic case ($\gamma = 0$) and $S' \sim \mathcal{N}(s,\, \sigma^2)$:

$$V^*(s) = \max_{a \in \mathcal{A}} \mathcal{R}(s, a) = \max_{a \in \mathcal{A}} \sum_{s' \in \mathbb{R}} \mathcal{P}_{S'}\left(s'\right) \cdot e^{as'} = \max_{a \in \mathcal{A}} \mathbb{E}[e^{as'}] = \max_{a \in \mathcal{A}} M_{S'}(a)$$

$$M_{S'}(a) = e^{sa + \frac{\sigma^2 a^2}{2}}$$