ROB 311 - TP2 - September 23

Authors

Simon Queyrut, Zhi Zhou

Question 1

We count two possible policies, namely

	s_0	s_1	s_2	s_3
π_1	s_1	s_0	s_0	s_0
π_2	s_2	s_0	s_0	s_0

Question 2

Knowing

$$V^*(S) = R(s) + \max_a \gamma \sum_{S\prime} T\left(S, a, S'\right) V^*\left(S'\right)$$

and

$$R(s) = \begin{cases} 10, & \text{for state } S3 \\ 1, & \text{for state } S2 \\ 0, & \text{otherwise} \end{cases}$$

we obtain

$$\begin{aligned} V^*\left(s_0\right) &= 0 + \max_a \gamma \sum_{S'} T\left(S, a, S'\right) V^*\left(S'\right) \\ V^*\left(S_1\right) &= 0 + \max_a \gamma \sum_{S'} T\left(S, a, S'\right) V^*\left(S'\right) \\ V^*\left(S_2\right) &= 1 + \max_a \gamma \sum_{S'} T\left(S, a, S'\right) V^*\left(S'\right) \\ V^*\left(S_3\right) &= 10 + \max_a \gamma \sum_{S'} T\left(S, a, S'\right) V^*\left(S'\right) \end{aligned}$$

Thus, replacing with values from $T\left(S, a0, S'\right)$, $T\left(S, a1, S'\right)$ and $T\left(S, a2, S'\right)$, we obtain

$$\begin{aligned} V^*\left(S_0\right) &= \gamma \max[V^*\left(S_1\right), V^*\left(S_2\right)] \\ V^*\left(S_1\right) &= \gamma \left[(1-x)V^*\left(S_1\right) + xV^*\left(S_3\right)\right] \\ V^*\left(S_2\right) &= \gamma \left[(1-y)V^*\left(S_0\right) + yV^*\left(S_3\right)\right] + 1 \\ V^*\left(S_3\right) &= \gamma V^*\left(S_0\right) + 10 \end{aligned}$$

Question 3

We have

$$\pi^*(s) = \arg\max_a \sum_{S'} T\left(S, a, S'\right) V^*\left(S'\right)$$

- Trivial case for $\gamma = 0$ leads to $V^*\left(S_1\right) = V^*\left(S_1\right) = 0$. Therefore

$$\pi^*(s_0) = \arg\max_a \left\{T(S_0, a_1, S_2)V^*\left(s_2\right) + T(S_0, a_1, S_3)V^*\left(s_3\right), \right.$$
$$\left. T(S_0, a_1, S_2)V^*\left(s_2\right) + T(S_0, a_1, S_3)V^*\left(s_3\right)\right\}$$

and because $T(S_0, a_1, S_2) = T(S_0, a_1, S_3) = T(S_0, a_1, S_3) = 0$ and $T(S_0, a_2, S_2) = 1$, we have

$$\begin{aligned}\pi^*(s_0) &= \arg\max_a \left\{0 + 0, V^*(s_2) + 0\right\} \\ &= \arg\max_a \{0, 1\} \\ &= a_2\end{aligned}$$

- For $\gamma > 0$, given $T(S_0, a_1, S_1) = T(S_0, a_2, S_2) = 1$ we look for $V^*(S_2) > V^*(S_1)$ for all $y \in [0, 1]$. With $x = 0$, we have $V^*(S_1) = \gamma V^*(S_1)$ so

$$V^*(S_1) = 0$$

and given $V^*(S_2) = \gamma\left[(1-y)V^*(S_0) + yV^*(S_3)\right] + 1 \geq 1$, we obtain

$$V^*(s_2) > V^*(S_1)$$

and finally $\pi^*(S_0) = a_2$ for any $\gamma \in (0, 1)$ or $y \in [0, 1]$

Question 4

We this time would like $H_1 : V^*(S_1) > V^*(S_2)$ to be true for any $\gamma \in (0, 1)$ or $x > 0$. We know that $V^*(S_2) = \gamma\left[(1-y)V^*(S_0) + yV^*(S_3)\right] + 1 \geq 1$ for any $\gamma \in (0, 1)$ or $x > 0$. We also have $V^*(S_1) = \gamma\left[(1-x)V^*(S_1) + xV^*(S_3)\right]$ which means

$$H_1 \Rightarrow V^*(S_1) = \frac{x\gamma V^*(S_3)}{1 - \gamma(1-x)} \geq 1, \forall \gamma \in (0, 1), x > 0$$

which is false since no matter $V^*(S_3)$ or x, we can always find an small enough γ inside $[0, 1]$ to make the fraction strictly smaller than 1.