CS446 / ECE 449: Machine Learning, Fall 2020, Homework 5

Name: Saud Alrasheed (Sauda2)

Worked individually

Problem (3):

Problem 3.a:

Since $0 \leq r_t \leq R_{max}$, for all $t > 0$, we we have that,

$$\sum_{t=1}^{\infty} \gamma^{t-1} r_t \ \leq \sum_{t=1}^{\infty} \gamma^{t-1} R_{max} = R_{max} \sum_{t=1}^{\infty} \gamma^{t-1} = \frac{R_{max}}{1-\gamma}$$

Also, we have that since $r_t \geq 0$,

$$\sum_{t=1}^{\infty} \gamma^{t-1} r_t \geq 0$$

Thus,

$$0 \leq \sum_{t=1}^{\infty} \gamma^{t-1} r_t \leq \frac{R_{max}}{1-\gamma}$$

Problem 3.b:

We will start by the definition of $V^{\pi}(s)$,

$$V^{\pi}(s) = \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^{t-1} r_t | \pi, s_1 = s\right] = \mathbb{E}\left[r_1 + \gamma \sum_{t=2}^{\infty} \gamma^{t-2} r_t | \pi, s_1 = s\right]$$

$$\implies V^{\pi}(s) = r_1 + \gamma \mathbb{E}\left[\sum_{t=2}^{\infty} \gamma^{t-2} r_t | \pi, s_1 = s\right]$$

Again,s by the definition of expected value operator, we have that,

$$\mathbb{E}\left[\sum_{t=2}^{\infty} \gamma^{t-2} r_t | \pi, s_1 = s\right] = \sum_{s'} P(s'|s, \pi(s))\ V(s')$$

Thus, combining both results we obtain,

$$V^{\pi}(s) = r_1 + \gamma \sum_{s'} P(s'|s, \pi(s))\ V(s') = R(s, \pi(s)) + \gamma \sum_{s'} P(s'|s, \pi(s))\ V(s')$$

Thus, writing this result in the matrix notation, we obtain the following,

$$V^\pi = R^\pi + \gamma P^\pi V^\pi$$

Problem 3.c:

We will denote $I_{|S|}$ by I, to make the solution look cleaner!

We will show that for any nonzero $x \in R^{|S|}$, we have that $(I - \gamma P^\pi)x \neq 0$. Consider $||(I - \gamma P^\pi)x||_p$ for all p.

$$\begin{aligned}
||(I - \gamma P^\pi)x|| &= ||Ix - \gamma P^\pi x|| \\
&= ||x - \gamma P^\pi x|| \\
&\geq ||x|| - ||\gamma P^\pi x|| && \text{(Triangular inequality)} \\
&= ||x|| - \gamma ||P^\pi x|| \\
&\geq ||x|| - \gamma ||x|| && \text{(Since each element of } P^\pi x \text{ is convex average of } x \implies ||P^\pi x|| \leq ||x||) \\
&= (1 - \gamma)||x|| \\
&\geq 0 && \text{(Since } 1 - \gamma \geq 0)
\end{aligned}$$

Thus $(I - \gamma P^\pi)$ is inverible.

Problem 3.d:

We will denote $I_{|S|}$ by I, to make the solution look cleaner!

The solution here is a direct consequence of problem 3.b and 3.c. We have that,

$$\begin{aligned}
V^\pi &= R^\pi + \gamma P^\pi V^\pi \\
V^\pi - \gamma P^\pi V^\pi &= R^\pi \\
(I - \gamma P^\pi)V^\pi &= R^\pi
\end{aligned}$$

And by the invertability of $(I - \gamma P^\pi)$, we have that,

$$V^\pi = (I - \gamma P^\pi)^{-1} R^\pi$$