cs234.stanford.assignment3

borislav.mavrin

June 2018

2. Expected Regret Bounds

(a)

Let $e' > 0$ and $T' \geq \log_{\frac{1}{\gamma}} \frac{H}{e'}$. By definition

$$V_\pi(s) = E_\pi\big[\sum_{t=1}^{\infty} \gamma^{t-1} r_t | s_1 = s\big] = \sum_{t=1}^{\infty} \gamma^{t-1} E_\pi\big[r_t | s_1 = s\big]$$

Hence,

$$V_\pi(s) - \sum_{t=1}^{T'} \gamma^{t-1} E_\pi\big[r_t | s_1 = s\big] = \sum_{t=T'+1}^{\infty} \gamma^{t-1} E_\pi\big[r_t | s_{T'} + 1 = s\big] \tag{1}$$

Rearranging $T' \geq \log_{\frac{1}{\gamma}} \frac{H}{e'}$:

$$\frac{\gamma^T}{1-\gamma} \leq e' \tag{2}$$

From the properties of geometric series sum:

$$\sum_{t=1}^{T'} \gamma^{t-1} + \sum_{t=T'+1}^{\infty} \gamma^{t-1} = \frac{1}{1-\gamma}, \qquad \frac{\gamma^{T'}}{1-\gamma} = \sum_{t=T'+1}^{\infty} \gamma^{t-1} \tag{3}$$

Therefore, from 1, 2 and 3

$$|V_\pi(s) - \sum_{t=1}^{T'} \gamma^{t-1} E_\pi\big[r_t | s_1 = s\big]| = \tag{4}$$

(by $r_t \in 0, 1$ and 1)

$$\sum_{t=T'+1}^{\infty} \gamma^{t-1} E_\pi\big[r_t | s_{T'} + 1 = s\big] \leq \tag{5}$$

(by $r_t \in 0, 1$)

$$\sum_{t=T'+1}^{\infty} \gamma^{t-1} \leq \tag{6}$$

(by 2 and 3)

$$\leq e' \tag{7}$$

(b)

Let $\epsilon' = \frac{\epsilon}{4}$, then we have to show that: $\sqrt{T'+T} - \sqrt{T} \leq \frac{\epsilon}{2}$. Rearranging gives:

$$\left(\frac{T'}{\epsilon}\right)^2 - T + \frac{\epsilon^2}{2} - \frac{T'}{2} \leq 0 \tag{8}$$

Since $T' \geq 1$ and without loss of generality we can assume $\epsilon \leq 1$, $\frac{\epsilon^2}{2} - \frac{T'}{2} \leq 0$. Hence, to prove 8, we must have $\left(\frac{T'}{\epsilon}\right)^2 - T \leq 0$, $\left(\frac{T'}{\epsilon}\right)^2 \leq T$.

Therefore for $\forall T \geq \frac{1}{\epsilon^2}\left(\log_{\frac{1}{\gamma}} \frac{4H}{\epsilon}\right)^2$ and setting $t = T - 1$, we get the result.