Assignment 3 Writeup

Evan Borras

February 19, 2021

Policy Gradient Methods: Writeup

Best Arm Identification in Multiarmed Bandit

(a)

Prove: $\Pr\left(\exists a \in \mathcal{A} \;\; s.t. \; |\hat{r}_a - \bar{r}_a| > \sqrt{\frac{\log(2/\delta)}{2n_{des}}}\right) < A\delta$

Define: $A_a = |\hat{r}_a - \bar{r}_a| > \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$

$\Pr\left(\exists a \in \mathcal{A} \;\; s.t. \; |\hat{r}_a - \bar{r}_a| > \sqrt{\frac{\log(2/\delta)}{2n_{des}}}\right) = \Pr\left(\exists a \in \mathcal{A} \;\; s.t. \, A_a\right) = \Pr\left(\bigcup_{a \in \mathcal{A}} A_a\right)$

$\Pr\left(\bigcup_{a \in \mathcal{A}} A_a\right) = \sum_{a \in \mathcal{A}} \Pr\left(A_a\right) = \sum_{a \in \mathcal{A}} \Pr\left(|\hat{r}_a - \bar{r}_a| > \sqrt{\frac{\log(2/\delta)}{2n_{des}}}\right)$

Using Hoeffding's inequality: $\Pr\left(|\hat{x} - \bar{x}| > \sqrt{\frac{\log(2/\delta)}{2n}}\right) < \delta$

$\sum_{a \in \mathcal{A}} \Pr\left(|\hat{r}_a - \bar{r}_a| > \sqrt{\frac{\log(2/\delta)}{2n_{des}}}\right) < \sum_{a \in \mathcal{A}} \delta = A\delta$

Therefore: $\Pr\left(\exists a \in \mathcal{A} \;\; s.t. \; |\hat{r}_a - \bar{r}_a| > \sqrt{\frac{\log(2/\delta)}{2n_{des}}}\right) < A\delta$

QED

(b)

Given a margin of error ϵ and a probability of failure δ'. What value of n_{des} assures: $\Pr\left(\bar{r}_{a^\dagger} \geq \bar{r}_{a^\star} - \epsilon\right) \geq 1 - \delta'$ (What value of n_{des} assures ϵ **is** a margin of error, and δ' **is** a probability for our margin of error). Where $a^\dagger = argmax_a(\hat{r}_a)$ and $a^\star = argmax_a(\bar{r}_a)$

Rewriting Prompt:

Given ϵ and δ'. What value of n_{des} assures: $\Pr\left(\bar{r}_{a^\star} - \bar{r}_{a^\dagger} \leq \epsilon\right) \geq 1 - \delta'$.

Rewriting Prompt Again:

Define: Event $O' = \left\{ \hat{r}_a | \bar{r}_{a^\star} - \bar{r}_{a^\dagger} \leq \epsilon \right\}$

Given ϵ and δ'. What value of n_{des} assures: $\Pr(O') \geq 1 - \delta'$

Section 1

From part (a), bound: $\Pr\left(\exists a \in \mathcal{A} \;\; s.t. \; |\hat{r}_a - \bar{r}_a| > \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \right) < A\delta$ is true.

taking the complement

bound: $\Pr\left(\forall a \in \mathcal{A} \left(|\hat{r}_a - \bar{r}_a| \leq \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \right) \right) \geq 1 - A\delta$ is true as well.

Define: Event $O = \left\{ \hat{r}_a | \forall a \in \mathcal{A} \left(|\hat{r}_a - \bar{r}_a| \leq \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \right) \right\}$.

Therefore:

From part (a), bound: $\Pr(O) \geq 1 - A\delta$ is true for our scenario.

Section 2

Assume Event O has taken place, therefore $\forall a \in \mathcal{A} \left(|\hat{r}_a - \bar{r}_a| \leq \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \right)$

Choosing $a = a^\dagger = argmax_a(\hat{r}_a)$, and $a = a^\star = argmax_a(\bar{r}_a)$ implies:

$$1) \;\; |\hat{r}_{a^\dagger} - \bar{r}_{a^\dagger}| \leq \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ and } 2) \;\; |\hat{r}_{a^\star} - \bar{r}_{a^\star}| \leq \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$$

From the above we get:

$$1) \;\; \hat{r}_{a^\dagger} - \bar{r}_{a^\dagger} \leq \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ and } \hat{r}_{a^\dagger} - \bar{r}_{a^\dagger} \geq -\sqrt{\frac{\log(2/\delta)}{2n_{des}}}$$

$$2) \;\; \hat{r}_{a^\star} - \bar{r}_{a^\star} \leq \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ and } \hat{r}_{a^\star} - \bar{r}_{a^\star} \geq -\sqrt{\frac{\log(2/\delta)}{2n_{des}}}$$

Rewritting results:

$$1) \;\; \bar{r}_{a^\dagger} \geq \hat{r}_{a^\dagger} - \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ and } \bar{r}_{a^\dagger} \leq \hat{r}_{a^\dagger} + \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$$

$$2) \;\; \bar{r}_{a^\star} \geq \hat{r}_{a^\star} - \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ and } \bar{r}_{a^\star} \leq \hat{r}_{a^\star} + \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$$

By def: $\bar{r}_{a^\star} \geq \bar{r}_{a^\dagger}$ since $a^\star = argmax_a(\bar{r}_a) \implies \bar{r}_{a^\star} \geq \bar{r}_{a^\dagger} \geq \hat{r}_{a^\dagger} - \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$

$$\implies 1) \;\; \bar{r}_{a^\star} \geq \hat{r}_{a^\dagger} - \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ and } \bar{r}_{a^\dagger} \leq \hat{r}_{a^\dagger} + \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$$

Subtracting the second inequality from the first yields:

$$\text{a) } \bar{r}_{a^\star} - \bar{r}_{a^\dagger} \geq -2\sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ under Event } O$$

Also

By def: $\hat{r}_{a^\dagger} \geq \hat{r}_{a^\star}$ since $a^\dagger = argmax_a(\hat{r}_a) \implies \hat{r}_{a^\star} \leq \hat{r}_{a^\dagger} \leq \bar{r}_{a^\dagger} + \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$

$$\implies \text{1) } \bar{r}_{a^\dagger} \geq \hat{r}_{a^\star} - \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ and } \bar{r}_{a^\dagger} \leq \hat{r}_{a^\dagger} + \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$$

$$\text{2) } \bar{r}_{a^\star} \geq \hat{r}_{a^\star} - \sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ and } \bar{r}_{a^\star} \leq \hat{r}_{a^\star} + \sqrt{\frac{\log(2/\delta)}{2n_{des}}}$$

Subtracting the first inequality of 1) from the second inequality of 2) yields:

$$\text{b) } \hat{r}_{a^\star} - \hat{r}_{a^\dagger} \leq 2\sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ under Event } O$$

Compressing a) and b) yields:

$$-2\sqrt{\frac{\log(2/\delta)}{2n_{des}}} \leq \hat{r}_{a^\star} - \hat{r}_{a^\dagger} \leq 2\sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ under Event } O$$

or

$$|\hat{r}_{a^\star} - \hat{r}_{a^\dagger}| \leq 2\sqrt{\frac{\log(2/\delta)}{2n_{des}}} \text{ under Event } O$$

Section 3

If we choose to lower bound ϵ by $2\sqrt{\frac{\log(2/\delta)}{2n_{des}}}$ then $2\sqrt{\frac{\log(2/\delta)}{2n_{des}}} \leq \epsilon$

$\implies \hat{r}_{a^\star} - \hat{r}_{a^\dagger} \leq 2\sqrt{\frac{\log(2/\delta)}{2n_{des}}} \leq \epsilon$ under Event O

$\implies \hat{r}_{a^\star} - \hat{r}_{a^\dagger} \leq \epsilon$ under Event O

$\implies$ Event O' under Event O

Asserting this constraint on ϵ implies Event O' occurs only when Event O occurs.

$\implies \Pr(O') = \Pr(O) \geq 1 - A\delta$

$\implies \Pr(O') \geq 1 - A\delta$

If we choose to constrain δ' to $\delta' = A\delta$ then

$$\Pr(O') \geq 1 - \delta' \text{ when } \epsilon \geq 2\sqrt{\frac{\log(2/\delta)}{2n_{des}}}, \text{ and } \delta' = A\delta$$

Section 4

$$\delta = \frac{\delta'}{A} \implies \epsilon \geq 2\sqrt{\frac{\log(2A/\delta')}{2n_{des}}}$$

$$\epsilon^2 \geq \frac{2\log(2A/\delta')}{n_{des}}$$

$$n_{des} \geq \frac{2\log(2A/\delta')}{\epsilon^2}$$

Therefore:

Given a margin of error ϵ and a probability of failure δ'. $n_{des} \geq \frac{2\log(2A/\delta')}{\epsilon^2}$ assures: $\Pr(\bar{r}_{a^\dagger} \geq \bar{r}_{a^\star} - \epsilon) \geq 1 - \delta'$. For $a^\dagger = argmax_a(\hat{r}_a)$, $a^\star = argmax_a(\bar{r}_a)$, and $A = |\mathcal{A}|$

QED