Short Assignment on Bandits

Ajinkya Ambatwar
EE16B104
Dept. Of Electrical Engineering

February 9, 2019

Abstract

Here I am attempting to prove the functionality and estimate the ARM SAMPLE COMPLEXITY of Modified Median Elimination Algorithm for (ϵ, δ) - PAC. At every round instead of half of the worst estimated arms, we are going to eliminate only one-fourth of the worst estimated arms.

Algorithm:

Algorithm 1 Modified MEA(ϵ, δ)

1: Set $S = A$.
2: $\epsilon_1 = \epsilon/4$, $\delta_1 = \delta/2$, $l = 1$.
3: Sample every arm $a \in S$ for $\frac{2}{\epsilon_l^2} ln(\frac{7}{3\delta_l})$ times, let $Q_l(a)$ denote its estimated empirical value
4: Find the median of $Q_{l2}(a)$(the lower half of the $Q_l(a)$ arranged ascending as per their values), denote it by m_l.
5: $S_{l+1} = S_l \backslash \{a : Q_l(a) < m_l\}$
6: If $|S_l| = 1$, Then output S_l,
Else $\epsilon_{l+1} = \frac{3}{4}\epsilon_l$; $\delta_{l+1} = \delta_l/2$; $l = l + 1$; $n_{l+1} = \frac{3}{4} n_l$

Theorem:

The modified MEA(ϵ, δ) is an $(\epsilon, \delta)-PAC$ algorithm with arm sample complexity $O(\frac{n}{\epsilon^2} ln(\frac{1}{\delta})$

Lemma 1:

For the MEA(ϵ, δ) we have that

$$Pr[\max_{j \in S_l} q_*(j) \leq \max_{i \in S_{l+1}} q_*(i) + \epsilon_l] \geq (1 - \delta_l)$$

Proof:

Without Loss Of Generality consider $l = 1$

$E_1 = \{Q(a_l^*) < q_*(a^*) - \epsilon/2\}$

Here we are considering the case of underestimating the true value of the optimal arm in the l_{th} round

So now

$$\begin{aligned}
Pr[E_1] &= Pr[Q(a_l^*) < q_*(a^*) - \epsilon_l/2] && (1)\\
&\leq exp(-2\frac{\epsilon_l^2}{4}n_l) && (2)\\
&= exp(-2\frac{\epsilon_l^2}{4} * \frac{2}{\epsilon_l^2} ln(\frac{7}{3\delta_l})) && (3)\\
&= \frac{3\delta_l}{7} && (4)
\end{aligned}$$

Hence $Pr[E_1] \leq \frac{3\delta_l}{7}$

Now consider the case where we are not underestimating the true value of the optimal arm but still commiting error by estimating the true value of some other arm greater than the true value of the optimal arm. This brings into a CHANCE of the true optimal arm at the l_{th} round getting eliminated.

So the required probability is $Pr[Q_l(j) > Q_l(a_l^*)|E_1^c]$. Let's try to find it out -

$$P[Q_l(j) \geq Q_l(a_l^*)|E_1^c] = P[Q_l(j) \geq q_*(j) + \epsilon_l/2 \qquad (5)$$

So the event of concern will happen only in either of two cases:

1. We are either underestimating the true value of the optimal arm at l_{th} round
2. We are overestimating the true value of the suboptimal arm j by a margin $\epsilon_l/2$

Given that the scenario 1 does not happen, the equation (5) clearly imply scenario 2. So now the equation (5) by Chernoff's inequality turns out to be -

$$\begin{aligned}
P[Q_l(j) \geq q_*(j) + \epsilon_l/2] &\leq exp(-2\frac{\epsilon^2}{4}n_l)\\
&\leq \frac{3\delta_l}{7}
\end{aligned}$$

Now to possibly eliminate the best arm at the l_{th} round, there has to be atleast $3S_l/4$ (S_l is the number of all the non-eliminated arms in the l_{th}round) such j arms. Let us represent the number of such bad arms by $\#bad$. Hence the expected value of such bad arms given the event E_1 does not hold is

$$E[\#bad|E_1^c] \leq (|S_l| - 1)\frac{3\delta_l}{7} \leq |S_l|\frac{3\delta_l}{7}$$

Now probability of the optimal arm getting eliminated is the probability that $\#bad \geq 3|S_l|/4$

$$P[A_{bad}] = P[\#bad \geq \frac{3|S_l|}{4}|E_1^c] \leq \frac{E[\#bad|E_1^c]}{3|S_l|/4} \tag{6}$$

$$\leq |S_l|\frac{3\delta_l}{7} * \frac{4}{3|S_l|} \tag{7}$$

$$\leq \frac{4\delta_l}{7} \tag{8}$$

Hence for each round probability of commiting error by eliminating the optimal arm is

$$\begin{aligned} P[A_{bad} \bigcup E_1] &\leq P[A_{bad}] + P[E_1] \\ &\leq \frac{4\delta_1}{7} + \frac{3\delta_l}{7} \\ &\leq \delta_l \end{aligned}$$

Hence the probility of not commiting an error to eliminate the optimal arm in the l_{th} round is

$$Pr[\max_{j \in S_l} q_*(j) \leq \max_{i \in S_{l+1}} q_*(i) + \epsilon_l] \tag{9}$$

which means that the optimal arm chosen in $(l+1)_{th}$round is still $\epsilon-optimal$ to the arm choosen in the l_{th}round and the best arm is not eliminated in the l_{th} round.

The equation (9) is the probability of the event where the optimal arm is not eliminated in the l_{th} round. Hence the probability is

$$Pr[\max_{j \in S_l} q_*(j) \leq \max_{i \in S_{l+1}} q_*(i) + \epsilon_l] \geq (1 - \delta_l)$$

Hence at every l_{th} round the MEA shows that it is an $(\epsilon, \delta)-PAC$ algorithm. Now at every round we are losing atmost ϵ_l from the value of the best optimal arm with the probility of maximum of δ_l. So summing up over all the rounds($log_{4/3}(n)$, where n is the total number of arms at the start), ϵ_l accumulates to a suitable ϵ and the probability will accumulate to δ and the last remaining arm at the end of all the rounds will be the $\epsilon - optimal$ arm for the true best arm such that -

$$Pr[q_*(a) \geq q_*(a^*) - \epsilon] \geq (1 - \delta)$$

Hence the lemma 1 is proved. Now trying to estimate the arm sample complexity of the MEA. Noting that each arm $a \in S$ is sampled $\frac{2}{\epsilon_l^2} ln(\frac{7}{3\delta_l})$ times. Hence the number of arm samples in the l_{th} round is $\frac{2n_l}{\epsilon_l^2} ln(\frac{7}{3\delta_l})$

Now with each round

$$n_1 = n; n_l = \frac{3}{4}n_{l-1} = (\frac{3}{4})^{l-1}n$$

$$\epsilon_1 = \frac{\epsilon}{4}; \epsilon_l = \frac{3}{4}\epsilon_{l-1} = (\frac{3}{4})^{l-1}\frac{\epsilon}{4}$$

$$\delta_1 = \frac{\delta}{2}; \delta_l = \frac{\delta_{l-1}}{2} = \frac{\delta}{2^l}$$

Now noting that the rounds end when $n_l = 1$ which gives $\#l = log_{4/3}(n)$
Therefore we have

$$\begin{aligned}
\sum_{l=1}^{log_{4/3}(n)} \frac{2n_l}{\epsilon_l^2} ln(\frac{7}{3\delta_l}) &= 2 \sum_{l=1}^{log_{4/3}(n)} \frac{n_l}{\epsilon_l^2} ln(\frac{7}{3\delta_l}) \\
&= 32 \sum_{l=1}^{log_{4/3}(n)} (\frac{4}{3})^{l-1} n[\frac{ln(7/3)}{\epsilon^2} + l * \frac{ln(2)}{\epsilon^2} + \frac{ln(\frac{1}{\delta})}{\epsilon^2}] \\
&\leq \frac{32n}{\epsilon^2} ln(\frac{1}{\delta}) \sum_{l=1}^{\infty} (\frac{4}{3})^{l-1} [1 + \frac{ln(7/3)}{ln(1/\delta)} + l * \frac{ln(2)}{ln(1/\delta)}] \\
&\leq O(\frac{n}{\epsilon^2} ln(\frac{1}{\delta}))
\end{aligned}$$

Hence the arm sample complexity turns out to be of the order $O(\frac{n}{\epsilon^2} ln(\frac{1}{\delta}))$