IE 498: Online Learning and Decision Making, Fall 2019 09/17/2019

Lectur 07: MAB Lower Bounds

Lecturer: Yuan Zhou *Scribe: Shen Yan, Jianhao Peng*

1 Sample complexity lower bound for $(\epsilon, \delta)-$PAC learning algorithms for best-arm identification

In this lecture we consider a warm-up case for MAB lower bound. In particular, we consider a MAB with 2 arms, and the goal is to identify the best arm. The final goal is to prove the following lower bound theorem, which matches the current algorithmic upper bound,

Theorem 1. *Suppose algorithm $\mathcal{A}$ is $(\epsilon, \delta)-$PAC and uses less than T samples, then $T \gtrsim \frac{\ln(\delta^{-1})}{\epsilon^2}$.*

Recall that the definition of a $(\epsilon, \delta)-$PAC algorithm is such that

$$\Pr[\mathcal{A} \textit{ outputs an arm within } \epsilon \textit{ of the best arm}] \geq 1 - \delta$$

To further simplify the analysis, we also assume $\mathcal{A}$ is a deterministic algorithm.

Proposition 2. *Suppose a $(\epsilon, \delta)-$PAC algorithm $\mathcal{A}$ exists, then $\exists \mathcal{A}'$ which is also $(\epsilon, \delta)-$PAC, and plays arm1 T times, arm2 T times, and outputs via a function $f : [0, 1]^T \times [0, 1]^T \to \{1, 2\}$.*

Thus we can further assume that the algorithm pulls arm1 and arm2 for T times each. Finally, we only consider Bernoulli reward case, i.e. $f : \{0, 1\}^T \times \{0, 1\}^T \to \{1, 2\}$.

To begin our analysis, we consider 2 cases

- Instance 1: $\mu_1 = p$, $\mu_2 = p + \epsilon$
- Instance 2: $\mu_1 = p + \epsilon$, $\mu_2 = p$

Then by the definition of $(\epsilon, \delta)-$PAC,

(1) Instance 1 $\implies \Pr_{r \sim \mathcal{B}_p^{\otimes T} \bigotimes \mathcal{B}_{p+\epsilon}^{\otimes T}}[f(r) = 1] \leq \delta$

(2) Instance 2 $\implies \Pr_{r \sim \mathcal{B}_{p+\epsilon}^{\otimes T} \bigotimes \mathcal{B}_p^{\otimes T}}[f(r) = 2] \leq \delta$

where $\mathcal{B}_p$ represents Bernoulli distribution with mean p, and we let $\mathcal{B}_p^{\bigotimes T} \bigotimes \mathcal{B}_{p+\epsilon}^{\bigotimes T}$ be $\mathcal{D}_1$, $\mathcal{B}_{p+\epsilon}^{\bigotimes T} \bigotimes \mathcal{B}_p^{\bigotimes T}$ be $\mathcal{D}_2$. The intuition for the rest of the proof is that if T is small, then $\mathcal{D}_1$ is "close" to $\mathcal{D}_2$, which would imply that $\Pr_{\mathcal{D}_1}[f(r) = 1] \approx \Pr_{\mathcal{D}_2}[f(r) = 1]$, but this would be a contradiction.

1.1 Aside: some results on information theory

In order to measure the "closeness" of $\mathcal{D}_1$ and $\mathcal{D}_2$, we need to introduce some useful inequalities from information theory. For simplicity let's assume here all considered distributions are discrete distribution.

Definition 3. *The total variation distance between measure P, Q is*

$$\Delta(P,Q) = \frac{1}{2}\|P-Q\|_1 = \frac{1}{2}\sum_{a\in\Omega}|P(a)-Q(a)|$$

Remark 1. *The equivalent definition of total variation distance is*

$$\Delta(P,Q) = \max_{A\in\Omega}|P(A)-Q(A)|$$

Proof. Let $B = \{a \in \Omega : P(a) \geq Q(a)\}$, then for any A

$$\begin{aligned}
|P(A)-Q(A)| &= |P(A\cap B) - Q(A\cap B) + P(A\cap\bar{B}) - Q(A\cap\bar{B})| \\
&\leq \max\{P(A\cap B) - Q(A\cap B), Q(A\cap\bar{B}) - P(A\cap\bar{B})\} \\
&\leq \max\{P(B) - Q(B), Q(\bar{B}) - P(\bar{B})\} \\
&= \frac{1}{2}[P(B) - Q(B) + Q(\bar{B}) - P(\bar{B})] \\
&= \frac{1}{2}\sum_{a\in\Omega}|P(a)-Q(a)|
\end{aligned}$$

Also, if we choose $A = B$, the all the inequalities hold as equalities. □

For our current problem, we can let $A = \{f(r) = 1\}$, then by the definition of total variation, $|\mathcal{D}_1(A) - \mathcal{D}_2(A)| \leq \Delta(\mathcal{D}_1, \mathcal{D}_2)$. Also by $(\epsilon,\delta)-$PAC, $|\mathcal{D}_1(A) - \mathcal{D}_2(A)| \geq 1-2\delta$, thus our goal is to prove that if T is too small, $\Delta(\mathcal{D}_1, \mathcal{D}_2) \leq 1-2\delta$, which would give the contradiction.

1.2 Towards bounding $\Delta(\mathcal{D}_1, \mathcal{D}_2)$ with T samples

1.2.1 Approach 1: brute-force calculation

Since given T, both distributions are explicitly defined, we can calculate the total variation directly:

$$\begin{aligned}
\mathcal{D}_1(r_1, r_2, \cdots, r_T, r'_1, r'_2, \cdots, r'_T) &= p^a(1-p)^{T-a}(p+\epsilon)^b(1-p-\epsilon)^{T-b} := \mathcal{D}_1(a,b) \\
\mathcal{D}_2(r_1, r_2, \cdots, r_T, r'_1, r'_2, \cdots, r'_T) &= (p+\epsilon)^a(1-p-\epsilon)^{T-a}p^b(1-p)^{T-b} := \mathcal{D}_2(a,b)
\end{aligned}$$

where a is the number of times that 1 appears in $r_1, r_2, \cdots, r_T$, b is the number of times that 1 appears in $r'_1, r'_2, \cdots, r'_T$. Thus

$$\|\mathcal{D}_1 - \mathcal{D}_2\|_1 = \sum_{a,b}\binom{T}{a}\binom{T}{b}|\mathcal{D}_1(a,b) - \mathcal{D}_2(a,b)|$$

This could potentially produce the desired bound, but this method is problem-specific and not applicable to other problems.

1.2.2 Approach 2: leveraging the relation between KL-divergence and total variation

Definition 4. *The Hellinger distance between P and Q is:*

$$H(P,Q) \coloneqq \left[\frac{1}{2}\sum_{a\in\Omega}\left(\sqrt{P(a)}-\sqrt{Q(a)}\right)^2\right]^{\frac{1}{2}} = \frac{1}{\sqrt{2}}\left\|\sqrt{P}-\sqrt{Q}\right\|_2$$

where $\sqrt{P}$ is the vector of $\left[\sqrt{P(a)}, \forall a \in \Omega \text{ in a fixed order}\right]$

Obviously we have $0 \le H(P,Q)$, and the following fact:

Remark 2.

$$H^2(P,Q) = \frac{1}{2}\sum_{a\in\Omega}\left(P(a)+Q(a)-2\sqrt{P(a)Q(a)}\right) = 1 - \sum_{a\in\Omega}\sqrt{P(a)Q(a)}$$

With the previous definitions and remarks we have the following lemma to sandwich $\Delta(P,Q)$:

Lemma 5.

$$\begin{aligned} H^2(P,Q) &\le \Delta(P,Q) && (3)\\ &\le \sqrt{H^2(P,Q)(2-H^2(P,Q))} && (4)\\ &\le \sqrt{2}H(P,Q) \end{aligned}$$

Proof. The last line of inequalities is trivial since $2 - H^2(P,Q) \le 2$. To prove (3), we need to show:

$$\sum_{a\in\Omega}\left(\sqrt{P(a)}-\sqrt{Q(a)}\right)^2 \le \sum_{a\in\Omega}|P(a)-Q(a)|$$

which can be implied from:

$$\begin{aligned} \forall a \in \Omega,\quad P(a)+Q(a)-2\sqrt{P(a)Q(a)} &\le |P(a)-Q(a)| \\ \Leftarrow \quad \frac{P(a)+Q(a)-|P(a)-Q(a)|}{2} &\le \sqrt{P(a)Q(a)} \\ \Leftarrow \quad \min\{P(a),Q(a)\} &\le \sqrt{P(a)Q(a)} \end{aligned}$$

where the last inequality is because of: $\min\{P(a),Q(a)\} \equiv \frac{P(a)+Q(a)-|P(a)-Q(a)|}{2}$
The proof of (4) is as following:

$$\begin{aligned} \Delta^2(P,Q) &= \frac{1}{4}\left(\sum_{a\in\Omega}|P(a)-Q(a)|\right)^2 = \frac{1}{4}\left(\sum_{a\in\Omega}\left|\sqrt{P(a)}-\sqrt{Q(a)}\right|\left(\sqrt{P(a)}+\sqrt{Q(a)}\right)\right)^2 \\ \text{(C-S inequality)} &\le \frac{1}{4}\left(\sum_{a\in\Omega}\left(\sqrt{P(a)}-\sqrt{Q(a)}\right)^2\right)\left(\sum_{a\in\Omega}\left(\sqrt{P(a)}+\sqrt{P(a)}\right)^2\right) \\ &= \frac{1}{2}H^2(P,Q)\left(\sum_{a\in\Omega}\left(P(a)+Q(a)+2\sqrt{P(a)Q(a)}\right)\right) \\ &= H^2(P,Q)\left(1+\sum_{a\in\Omega}\sqrt{P(a)Q(a)}\right) = H^2(P,Q)\left(2-\left(1-\sum_{a\in\Omega}\sqrt{P(a)Q(a)}\right)\right) \\ \text{(Remark 2)} &= H^2(P,Q)\left(2-H^2(P,Q)\right) \end{aligned}$$

$\square$

Definition 6. *The Kullback–Leibler (KL) divergence between P and Q is:*

$$D_{KL}(P||Q) \coloneqq -\sum_{a\in\Omega} P(a)\ln\frac{Q(a)}{P(a)}$$

Note that KL-divergence is not a metric and asymmetric in P and Q. One important fact is:

Remark 3.

$$D_{KL}\left(P||Q\right) \iff P = Q$$

To connect the Hellinger distance with KL-divergence, we have the following bound:

Lemma 7.

$$H^2(P,Q) \leq 1 - \exp\left(-\frac{1}{2}D_{KL}\left(P||Q\right)\right)$$

Proof.

$$\begin{aligned}
1 - H^2(P,Q) &= \sum_{a\in\Omega}\sqrt{P(a)Q(a)} = \exp\left(\ln\left(\sum_{a\in\Omega}\sqrt{P(a)Q(a)}\right)\right) \\
&= \exp\left(\ln\left(\sum_{a\in\Omega}P(a)\sqrt{\frac{Q(a)}{P(a)}}\right)\right) \\
&= \exp\left(\ln\left(\mathop{\mathbb{E}}_{a\sim P}\sqrt{\frac{Q(a)}{P(a)}}\right)\right) \\
\text{(Jensen's Inequality)} &\geq \exp\left(\sum_{a\in\Omega}P(a)\ln\sqrt{\frac{Q(a)}{P(a)}}\right) \\
&= \exp\left(-\frac{1}{2}D_{KL}\left(P||Q\right)\right)
\end{aligned}$$

□

To connect the total variacne with KL-divergence, we have the following bounds:

Lemma 8. *The Pinsker's Inequality is:*

$$\Delta(P,Q) \leq \sqrt{D_{KL}\left(P||Q\right)}$$

Proof. from lemma 5 and 7 we have:

$$\begin{aligned}
\Delta(P,Q) &\leq \sqrt{2H^2(P,Q)} \\
&\leq \sqrt{2\left(1 - \exp\left(-\frac{1}{2}D_{KL}\left(P||Q\right)\right)\right)} \\
\text{(by } e^{-x} \geq 1 - x\text{)} \quad &\leq \sqrt{2\left(1 - \left(1 - \frac{1}{2}D_{KL}\left(P||Q\right)\right)\right)} \\
&= \sqrt{D_{KL}\left(P||Q\right)}
\end{aligned}$$

□

Lemma 9. *The High Probability Pinsker's Inequality is:*

$$\Delta(P,Q) \leq 1 - \frac{1}{2}\exp\left(-D_{KL}\left(P||Q\right)\right)$$

Proof. From lemma 5, we have:

$$\begin{aligned} \Delta(P,Q) &\leq \sqrt{1-\left(1-H^2(P,Q)\right)^2} \\ \text{(by } \sqrt{1-x} \leq 1-\tfrac{1}{2}x\text{)} \quad &\leq 1-\frac{1}{2}\left(1-H^2(P,Q)\right)^2 \\ \text{(by Lemma 7)} \quad &\leq 1-\frac{1}{2}\exp\left(-D_{KL}\left(P||Q\right)\right) \end{aligned}$$

□

A few comments on these two inequalities:

- when $D_{KL}\left(P||Q\right)$ is small, Lemma 8 gives a much tighter bound on $\Delta(P,Q)$ yet Lemma 9 gives loose upper bound at around 0.5.
- when $D_{KL}\left(P||Q\right)$ is large, Lemma 8 gives a useless upper bound since $\Delta(P,Q) \leq 1$. However Lemma 9 gives a more refined bound.