Machine Learning Theory (CS 6783)

Lecture 19: Stochastic Multi-armed Bandit Lower Bound

1 Lower Bound Tools

For this section just to ease notation, assume losses on each arm take on a finite set of values (for what we care you can just assume that losses are 0 or 1). To deal with more general case you will need to deal with Radon Nykodym derivative and differential entropy etc. But the basic ideas remain.

Key Idea: In these lower bounds, we will consider two loss distribution that are "close" and the idea is that in one of them the optimal arm is different from the other one by the desired sub-optimality level. To be able to do well, for instanced drawn from distribution one we need to pick arm (after looking at samples) that is different from the one picked when we get instances from distribution 2. However, if the two distributions are close, it is likely to get similar runs of instances from both the distributions. Specifically, we will provide lower bounds via KL divergence used to measure differences between the two distributions. We will demonstrate both instance independent lower bound of $\sqrt{Kn}$ and instance dependent lower bound that almost matches the instance dependent upper bound of UCB (LCB).

First, let us start with a bound on KL divergence over distributions of runs under two different product distribution over losses of arms.

Lemma 1. *Let $\mathcal{D}$ and $\mathcal{D}'$ be two distributions on the losses with $(\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_K)$ and $(\mathcal{D}'_1, \mathcal{D}'_2, \ldots, \mathcal{D}'_K)$ being the corresponding marginal distributions for the K arms. Let $\mathbf{A}$ be any stochastic bandit algorithm (that is, the algorithm picks the next action to play (possibly randomly) based on past actions and losses observed). In this case,*

$$\mathrm{KL}(P_{\mathbf{A},\mathcal{D},n} | P_{\mathbf{A},\mathcal{D}',n}) = \sum_{k=1}^{K} \mathbb{E}_{R \sim P_{\mathbf{A},\mathcal{D},n}} [n_{k,n}] \, \mathrm{KL}(\mathcal{D}_k | \mathcal{D}'_k)$$

Proof. Consider any sequence R of action losses pairs we encounter, say

$$R = (I_1, \ell_1[I_1], \ldots, I_1, \ell_n[I_n])$$

Now

$$P_{\mathbf{A},\mathcal{D},n}(R) = \prod_{t=1}^{n} P_{\mathcal{A}}(I_t | I_1, \ell_1[I_1], \ldots, I_{t-1}, \ell_{t-1}[I_1]) \cdot P_{\mathcal{D}_{I_t}}(\ell_t[I_t]) \qquad \text{and similarly,}$$

$$P_{\mathbf{A},\mathcal{D}',n}(R) = \prod_{t=1}^{n} P_{\mathcal{A}}(I_t | I_1, \ell_1[I_1], \ldots, I_{t-1}, \ell_{t-1}[I_1]) \cdot P_{\mathcal{D}'_{I_t}}(\ell_t[I_t])$$

(I am abusing notation here to make life simpler. Basically we want to look at probability of the run under the two distributions). Hence we can conclude that:

$$\mathrm{KL}(P_{\mathbf{A},\mathcal{D},n}|P_{\mathbf{A},\mathcal{D}',n}) = \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}\left[\log\left(\frac{\prod_{t=1}^{n} P_{\mathcal{A}}(I_t|I_1,\ell_1[I_1],\ldots,I_{t-1},\ell_{t-1}[I_1])\cdot P_{L\sim\mathcal{D}}(L[I_t]=\ell_t[I_t])}{\prod_{t=1}^{n} P_{\mathcal{A}}(I_t|I_1,\ell_1[I_1],\ldots,I_{t-1},\ell_{t-1}[I_1])\cdot P_{L\sim\mathcal{D}'}(L[I_t]=\ell_t[I_t])}\right)\right]$$

$$= \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}\left[\log\left(\frac{\prod_{t=1}^{n} P_{\mathcal{A}}(I_t|I_1,\ell_1[I_1],\ldots,I_{t-1},\ell_{t-1}[I_1])\cdot P_{\mathcal{D}_{I_t}}(\ell_t[I_t])}{\prod_{t=1}^{n} P_{\mathcal{A}}(I_t|I_1,\ell_1[I_1],\ldots,I_{t-1},\ell_{t-1}[I_1])\cdot P_{\mathcal{D}'_{I_t}}(\ell_t[I_t])}\right)\right]$$
$$= \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}\left[\sum_{t=1}^{n}\log\left(\frac{P_{\mathcal{D}_{I_t}}(\ell_t[I_t])}{P_{\mathcal{D}'_{I_t}}(\ell_t[I_t])}\right)\right]$$
$$= \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}\left[\sum_{t=1}^{n}\mathbb{E}_{\ell_t[I_t]\sim D_{I_t}}\left[\log\left(\frac{P_{\mathcal{D}_{I_t}}(\ell_t[I_t])}{P_{\mathcal{D}'_{I_t}}(\ell_t[I_t])}\right)\right]\right]$$

$$= \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}\left[\sum_{t=1}^{n}\mathrm{KL}(P_{\mathcal{D}_{I_t}}|P_{\mathcal{D}'_{I_t}})\right]$$
$$= \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}\left[\sum_{k=1}^{K} n_{k,n}\mathrm{KL}(P_{\mathcal{D}_k}|P_{\mathcal{D}'_k})\right]$$
$$= \sum_{k=1}^{K}\mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}\left[n_{k,n}\right]\mathrm{KL}(P_{\mathcal{D}_k}|P_{\mathcal{D}'_k})$$

□

Another useful result that I will write down here (look up online for proof) is the relationship between KL divergence and total variation distance. Total variation distance between two distribution P and P' is written as

$$\|P-P'\|_{TV} = \sup_{A\in\mathcal{F}}\left|P(A)-P'(A)\right|$$

where $\mathcal{F}$ above is the sigma algebra. It is easy to check that for countable sets, $\|P-P'\|_{TV} = \frac{1}{2}\sum_{\omega\in\Omega}|P(\omega)-P'(\omega)|$.

The following two inequalities are useful:

1. Pinsker's Inequality

$$\|P-P'\|_{TV} \le \sqrt{\frac{1}{2}\mathrm{KL}(P|P')}$$

2. An exponential bound:

$$\|P-P'\|_{TV} \le \sqrt{1-\exp\left(-\mathrm{KL}(P|P')\right)}$$

2 Instance Independent Lower Bound

Theorem 2. *For all $K \geq 2$ and any $n > K$, there exists a distribution over losses of arms such that:*

$$\mathbb{E}\left[\mathrm{Reg}_n\right] \geq \frac{1}{8}\sqrt{\frac{K-1}{8n}}$$

Proof. Let $D_1 = B_{1/2-\Delta}$ and for all other $k \in [2, \ldots, K]$, $D_i = B_{1/2}$ where B_θ is the Bernoulli distribution with probability θ of a 1 and $1-\theta$ of producing 0. Now the way we produce D' is as follows. Pick $j^* = \operatorname*{argmin}_{k \in [K]: k \neq 1} \mathbb{E}_{R \sim P_{\mathbf{A},\mathcal{D},n}}\left[n_{k,n}\right]$. Now set D' to be such that for all $k \neq j^*$, $D'_k = D_k$ and $D'_{j^*} = D_{1/2-2\Delta}$ Note that for D' optimal arm is j^* while for D it is 1. Define event A to be

$$A = \{\omega \in \Omega : n_{1,n} \leq n/2\}$$

That is the event that arm 1 is played less than 1/2 the times. Now note two things. First,

$$\mathbb{E}_{\mathcal{D}}\left[\mathrm{Reg}_n | A\right] \geq \frac{1}{2}\Delta$$

and

$$\mathbb{E}_{\mathcal{D}'}\left[\mathrm{Reg}_n | A^c\right] \geq \frac{1}{2}\Delta$$

The first is true because we play optimal arm less than 1/2 the number of times and so we are Δ suboptimal more than 1/2 the time. The second is also true because when we are in complement of A, and distribution is D', then we are playing again a Δ sub-optimal arm more than 1/2 the times. Hence we conclude that:

$$\begin{aligned}
\mathbb{E}_{\mathcal{D}}\left[\mathrm{Reg}_n\right] + \mathbb{E}_{\mathcal{D}'}\left[\mathrm{Reg}_n\right] &\geq \frac{\Delta}{2}\left(P_{\mathbf{A},\mathcal{D}',n}(A^c) + P_{\mathbf{A},\mathcal{D},n}(A)\right) \\
&= \frac{\Delta}{2}\left(1 - P_{\mathbf{A},\mathcal{D}',n}(A) + P_{\mathbf{A},\mathcal{D},n}(A)\right) \\
&\geq \frac{\Delta}{2}\left(1 - \left|P_{\mathbf{A},\mathcal{D}',n}(A) - P_{\mathbf{A},\mathcal{D},n}(A)\right|\right) \\
&\geq \frac{\Delta}{2}\left(1 - \sqrt{\frac{1}{2}\mathrm{KL}(P_{\mathbf{A},\mathcal{D},n} | P_{\mathbf{A},\mathcal{D}',n})}\right)
\end{aligned}$$

Now using lemma from section 1 we have for this case that

$$\mathrm{KL}(P_{\mathbf{A},\mathcal{D},n} | P_{\mathbf{A},\mathcal{D}',n}) = \mathbb{E}_{R \sim P_{\mathbf{A},\mathcal{D},n}}\left[n_{j^*,n}\right] \mathrm{KL}(B_{1/2-\Delta} | 1/2 - 2\Delta) \leq \frac{2n}{k-1}\Delta^2$$

Hence we conclude that:

$$\mathbb{E}_{\mathcal{D}}\left[\mathrm{Reg}_n\right] + \mathbb{E}_{\mathcal{D}'}\left[\mathrm{Reg}_n\right] \geq \frac{\Delta}{2}\left(1 - \sqrt{\frac{2n}{k-1}\Delta^2}\right)$$

Hence we can conclude that:

$$\max\{\mathbb{E}_{\mathcal{D}}\left[\mathrm{Reg}_n\right], \mathbb{E}_{\mathcal{D}'}\left[\mathrm{Reg}_n\right]\} \geq \frac{\Delta}{4}\left(1 - \sqrt{\frac{2n}{k-1}\Delta^2}\right)$$

Setting $\Delta = \sqrt{\frac{K-1}{8n}}$ yields the result that:

$$\max\{\mathbb{E}_{\mathcal{D}}\left[\text{Reg}_n\right], \mathbb{E}_{\mathcal{D}'}\left[\text{Reg}_n\right]\} \geq \frac{\Delta}{8}$$

and hence lower bound.

□

3 Instance Dependent Lower Bound

We use similar style proof technique now to provide an instance specific lower bound. This one needs a more careful statement.

Theorem 3. *For any Stochastic bandit algorithm (with binary losses) that guarantees a regret bound of:* $\mathbb{E}\left[\text{Reg}_n\right] \leq Cn^{-\beta}$*, we have that for any distribution* $\mathcal{D}$*,*

$$\mathbb{E}_{\mathcal{D}}\left[\text{Reg}_n\right] \geq \frac{1}{n} \sum_{k:\Delta_k \neq 0} \frac{2\log\left(\frac{\Delta_k n^{\beta}}{2C}\right)}{\Delta_k} \approx \frac{\beta}{n} \sum_{k:\Delta_k \neq 0} \frac{\log\left(\frac{n}{2C}\right)}{\Delta_k}$$

Proof. We already showed that

$$\mathbb{E}_{\mathcal{D}}\left[\text{Reg}_n\right] = \frac{1}{n} \sum_{i:i \neq 0} \Delta_i \mathbb{E}\left[n_{i,n}\right]$$

Hence, for each k, we need to prove a lower bound on $\mathbb{E}\left[n_{k,n}\right]$ for any sub-optimal k. Now we prove a lower bound for a given k.

To this end, given a distribution $\mathcal{D}$ with mean losses of arms $\mu = (\mu_1, \ldots, \mu_K)$ where $\mu_k = \mathbb{E}_{\ell \sim D}\left[\ell[k]\right]$. Note that since the losses are binary, the marginal distribution of loss of each arm k is a Bernoulli distribution B_{μ_k}. Given an arm k, consider the product distribution $\mathcal{D}' = (B_{\mu_1}, \ldots, B_{\mu_{k-1}}, B_{\mu_k - 2\Delta_k}, B_{\mu_{k+1}}, \ldots, B_{\mu_K})$. That is, for this new distribution $\mathcal{D}'$, arm k is now the optimal arm with margin Δ_k. Now similar to the previous section, define event:

$$A = \{\omega \in \Omega : n_{k,n} > n/2\}$$

Note that conditioned on this event, clearly, regret under distribution $\mathcal{D}$ is larger than $\Delta_k/2$. Hence, overall,

$$\mathbb{E}_{\mathcal{D}}\left[\text{Reg}_n\right] \geq \frac{\Delta_k}{2} P_{\mathbf{A},\mathcal{D},n}[A]$$

On the other hand, conditioned on the complement of A, regret under distribution $\mathcal{D}'$ is lower bounded by $\Delta_k/2$ and so

$$\mathbb{E}_{\mathcal{D}'}\left[\text{Reg}_n\right] \geq \frac{\Delta_k}{2} P_{\mathbf{A},\mathcal{D}',n}[A^c]$$

Hence, we have that

$$\max\{\mathbb{E}_{\mathcal{D}'}\left[\text{Reg}_n\right], \mathbb{E}_{\mathcal{D}}\left[\text{Reg}_n\right]\} \geq \frac{\Delta_k}{2}\left(P_{\mathbf{A},\mathcal{D}',n}[A^c] + P_{\mathbf{A},\mathcal{D},n}[A]\right)$$

However, since $\max\{\mathbb{E}_{\mathcal{D}'}[\text{Reg}_n], \mathbb{E}_{\mathcal{D}}[\text{Reg}_n]\} \leq Cn^{-\beta}$ (as the algorithm has a regret guarantee), we have that:

$$\begin{aligned}
Cn^{-\beta} &\geq \frac{\Delta_k}{2}\left(P_{\mathbf{A},\mathcal{D}',n}[A^c] + P_{\mathbf{A},\mathcal{D},n}[A]\right) \\
&\geq \frac{\Delta_k}{2}\left(1 + P_{\mathbf{A},\mathcal{D},n}[A] - P_{\mathbf{A},\mathcal{D}',n}[A]\right) \\
&\geq \frac{\Delta_k}{2}\left(1 - \|P_{\mathbf{A},\mathcal{D},n} - P_{\mathbf{A},\mathcal{D}',n}\|_{TV}\right) \\
&\geq \frac{\Delta_k}{2}\left(1 - \sqrt{1 - \exp\left(-\text{KL}\left(P_{\mathbf{A},\mathcal{D},n} | P_{\mathbf{A},\mathcal{D}',n}\right)\right)}\right) \\
&\geq \frac{\Delta_k}{2}\exp\left(-\frac{1}{2}\text{KL}\left(P_{\mathbf{A},\mathcal{D},n} | P_{\mathbf{A},\mathcal{D}',n}\right)\right)
\end{aligned}$$

Hence we conclude that:

$$\text{KL}\left(P_{\mathbf{A},\mathcal{D},n} | P_{\mathbf{A},\mathcal{D}',n}\right) \geq 2\log\left(\frac{\Delta_k}{2Cn^{-\beta}}\right)$$

On the other hand, note that by Lemma 1, we have that

$$\begin{aligned}
\text{KL}\left(P_{\mathbf{A},\mathcal{D},n} | P_{\mathbf{A},\mathcal{D}',n}\right) &= \sum_{i=1}^{K} \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}[n_{i,n}]\,\text{KL}(\mathcal{D}_i|\mathcal{D}'_i) \\
&= \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}[n_{k,n}]\,\text{KL}(\mathcal{D}_k|\mathcal{D}'_k) \\
&= \mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}[n_{k,n}]\,\text{KL}(B_{\mu_k}|B_{\mu_k - 2\Delta_k})
\end{aligned}$$

Now as long as $\mu_k \in (0.1, 0.9)$ we can conclude that

$$\text{KL}(B_{\mu_k}|B_{\mu_k-2\Delta_k}) \leq \Delta_k^2$$

Hence putting it all together we can conclude that:

$$\mathbb{E}_{R\sim P_{\mathbf{A},\mathcal{D},n}}[n_{k,n}] \geq \frac{2\log\left(\frac{\Delta_k}{2Cn^{-\beta}}\right)}{\Delta_k^2}$$

Using this with the fact that $\mathbb{E}_{\mathcal{D}}[\text{Reg}_n] = \frac{1}{n}\sum_{i:\Delta_i\neq 0}\Delta_i \mathbb{E}[n_{i,n}]$ we can conclude that:

$$\mathbb{E}_{\mathcal{D}}[\text{Reg}_n] \geq \frac{1}{n}\sum_{k:\Delta_k\neq 0}\frac{2\log\left(\frac{\Delta_k n^{\beta}}{2C}\right)}{\Delta_k}$$

□