| | |
|---|---|
| **Advanced Topics in ML and AGT** | **Fall Semester, 2017/12** |

# Lecture 8: December 11

*Lecturer: Yishay Mansour* *Scribe: Amit Zohar, Amir Barda, Olya Sirkin*

## 8.1 Lecture Overview

In this lecture we discuss sample size lower bounds. For this we need some background in Information Theory.

## 8.2 Distance between Distributions

**Definition** The distance between two distributions P and Q is defined as:

$$\|P - Q\|_1 = \sum_x |P(x) - Q(x)|$$

**Lemma 8.1** *For every function $f : x \to [-1, +1]$ :*

$$E_{x\sim Q}[f(x)] - E_{x\sim P}[f(x)] \leq \|P - Q\|_1$$

*and there exists $f$ such that:*

$$E_{x\sim Q}[f(x)] - E_{x\sim P}[f(x)] = \|P - Q\|_1$$

We can think of $f$ as a probability of $x$ coming from $Q$. The larger $\|P - Q\|_1$ is, the better $f$ can distinguish between $P$ and $Q$.

**Proof:** Let $f$ be defined as $f(x) = sign(P(x) - Q(x))$

$$\begin{aligned} E_p[f] - E_q[f] &= |\sum_x |f(x)(P(x) - Q(x)| \\ &= |\sum_x |(P(x) - Q(x)| \\ &= \|P - Q\|_1 \end{aligned}$$

And for any $f$:

$$|E_P[f] - E_Q[f]| = |\sum_x f(x)(P(x) - Q(x))| \leq \sum_x |(P(x) - Q(x))| \leq \|P - Q\|_1$$

□

We can think about $f$ giving a "correct" answer as $Pr_Q[f(x) = Q]$ and of $f$ giving a "wrong" answer as $Pr_P[f(x) = Q]$. We would like the difference between the two to be large so that $f$ could distinguish between $P$ and $Q$.

**Definition** $P^m$, $Q^m$

$$P^m = P_1 \times P_2 \ldots P_m$$
$$Q^m = Q_1 \times Q_2 \times \ldots Q_m$$

**Lemma 8.2**

$$\|P^m - Q^m\|_1 \leq \sum_{i=1}^{m} \|P_i - Q_i\|_1$$

**Proof:** We are going to prove the lemma by induction on $m$. For $m = 1$:

$$\|P^m - Q^m\|_1 = \|P - Q\|_1$$

This follows trivially from definition of $P^m$ and $Q^m$.
For $m > 1$:

$$\|P^m - Q^m\|_1 = \sum_{x_1} \sum_{x_2} \cdots \sum_{x_m} |\prod_i P_i(x_i) - \prod_i Q_i(x_i)|$$

Let us define $\alpha(x) = \prod_{x=2}^{m} P_i(x_i)$ and $\beta(x) = \prod_{x=2}^{m} Q_i(x_i)$ and also $\overrightarrow{x} = x_2 \ldots x_m$.

$$\|P^m - Q^m\|_1 \leq \sum_{x_1} \sum_{\overrightarrow{x}} \alpha(x)|P_1(x_1) - Q_1(x_1)| + \sum_{\overrightarrow{x}} \sum_{x_1} Q_1(x_1)|\alpha(x) - \beta(x)|$$

Note that $\sum_{\overrightarrow{x}} \alpha(x) = 1$ and $\sum_{x_1} Q_1(x_1) = 1$ because $\alpha$ and $Q_1$ are distributions. Therefore:

$$\begin{aligned}
\|P^m - Q^m\|_1 &\leq \sum_{x_1} |P_1(x_1) - Q_1(x_1)| + \sum_{\overrightarrow{x}} |\alpha(x) - \beta(x)| \\
&\leq \|P_1 - Q_1\|_1 + \sum_{i=2}^{m} \|P_i - Q_i\|_1 \\
&= \sum_{i=1}^{m} \|P_i - Q_i\|_1
\end{aligned}$$

□

We'll use the above lemma to derive a lower bound on $m$.
Suppose $P_i \sim Br(\frac{1}{2})$, $Q_i \sim Br(\frac{1+\epsilon}{2})$. Where $Br(p)$ is a Bernoulli random variable of probability $p$.

$$\|P_i - Q_i\|_1 = \epsilon$$

$$1 - 2\delta \leq \|P^m - Q^m\|_1 \leq m \times \epsilon$$
$$m \geq \frac{1-2\delta}{\epsilon} = \Omega(\frac{1}{\epsilon})$$

This bound is not tight, since we know that the right rate is $\frac{1}{\epsilon^2}$.
Next we are going to derive tighter bound using KL-Divergence.

## 8.3 KL-Divergence

The Kullback-Leibler (KL) divergence is a measure of the difference between two distributions $P$ and $Q$.

**Definition** The KL Divergence between two distributions $P$ and $Q$ is defined as:

$$KL(P||Q) = E_{x\sim P}[\log \frac{P(x)}{Q(x)}] = \sum_x P(x) \log(\frac{P(x)}{Q(x)}$$

Note that KL divergence is not symmetric; i.e., $KL(P||Q) \neq KL(Q||P)$.

### 8.3.1 Properties of KL divergence.

1. Positivity: $KL(P||Q) \geq 0$, $KL(P||Q) = 0 \leftrightarrow KL = 0$.

   **Proof:** Suppose $f(y) = yln(y)$. It is easy to show that $f$ is convex, therefore:

$$KL(P||Q) = \sum_x P(x) log(\frac{P(x)}{Q(x)}) = \sum_x Q(x) f(\frac{P(x)}{Q(x)})$$

$$\geq f(\sum_x Q(x) \frac{P(x)}{Q(x)}) = f(\sum_x P(x)) = f(1) = 0$$

□

2. Chain Rule for KL-divergence. Under independence assumption it holds that:

$$KL(P^m||Q^m) = \sum_{i=1}^{m} KL(P_i||Q_i)$$

   **Proof:** Let us define $h_i(x_i) = ln(\frac{P_i(x_i)}{Q_i(x_i)})$ and $x = (x_1 \dots x_m)$, $x_i$ $i.i.d$

$$\begin{aligned}
KL(P||Q) &= \sum_x P(x) ln(\frac{P(x)}{Q(x)}) \\
&= \sum_{i=1}^{m} \sum_x P(x) h_i(x_i) \\
&= \sum_{i=1}^{m} \sum_{x_i} P_i(x_i) h_i(x_i) \sum_{x_j \neq x_i} \prod_{j \neq i} P_j(x_j) \\
&= \sum_{i=1}^{m} \sum_{x_i} P_i(x_i) h_i(x_i) \\
&= \sum_{i=1}^{m} \sum_{x_i} P_i(x_i) ln \left[ \frac{P_i(x_i)}{Q_i(x_i)} \right] \\
&= \sum_{i=1}^{m} KL(P_i||Q_i)
\end{aligned}$$

□

3. The Pinsker Inequality:

$$\forall A \in \Omega : 2(P(A) - Q(A))^2 \leq KL(P||Q)$$

Before proving the inequality, we first begin with a helpful lemma:

**Lemma 8.3** $\sum_{x \in A} P(x) \log \frac{P(x)}{Q(x)} \geq P(A) \log \frac{P(A)}{Q(A)}$

**Proof:** Denote: $\forall x \in A : P_A(x) = \frac{P(x)}{P(A)}, Q_A(x) = \frac{Q(x)}{Q(A)}$. Therefore:

$$\sum_{x \in A} P(x) \log \frac{P(x)}{Q(x)} = P(A) \cdot \sum_{x \in A} P_A(x) \log \frac{P(A) \cdot P_A(x)}{Q(A) \cdot Q_A(x)} =$$

$$= P(A) \cdot \sum_{x \in A} P_A(x) \log \frac{P_A(x)}{Q_A(x)} + P(A) \log \frac{P(A)}{Q(A)} \cdot \sum_{x \in A} P_A(x)$$

Since $\sum_{x \in A} P_A(x) = 1$ and $P_A(x) \log \frac{P_A(x)}{Q_A(x)} = KL(P_A||Q_A) \geq 0$:

$$\sum_{x \in A} P(x) \log \frac{P(x)}{Q(x)} \geq P(A) \log \frac{P(A)}{Q(A)}$$

□

We now prove the Pinsker Inequality:

**Proof:** We apply the previous lemma twice and get:

$$\sum_{x \in A} P(x) \log \frac{P(x)}{Q(x)} \geq P(A) \log \frac{P(A)}{Q(A)}$$

$$\sum_{x \notin A} P(x) \log \frac{P(x)}{Q(x)} \geq (1 - P(A)) \log \frac{1 - P(A)}{1 - Q(A)}$$

Denote $a = P(A), b = P(B)$. By summing the two inequalities above we get:

$$KL(P||Q) \geq a \log \frac{a}{b} + (1-a) \log \frac{1-a}{1-b} = \int_a^b -\frac{a}{x} + \frac{1-a}{1-x} dx = \int_a^b \frac{x-a}{x(1-x)} dx$$

Since $x(1-x) \geq \frac{1}{4}$:

$$KL(P||Q) \geq \int_a^b 4(x-a)dx = 2(b-a)^2$$

□

## 8.4 A Better Lower Bound

Now we can get a much better bound than before.

Assume that $P = Br(\frac{1+\epsilon}{2}), Q = Br(\frac{1}{2})$. Like before, $||P - Q||_1 = \epsilon$. Furthermore:

$$KL(P||Q) = \frac{1+\epsilon}{2} \log(1+\epsilon) + \frac{1-\epsilon}{2} \log(1-\epsilon) = \frac{1}{2} \log(1-\epsilon^2) + \frac{\epsilon}{2} \log(\frac{1+\epsilon}{1-\epsilon}) \leq 2\epsilon^2$$

By using the 2nd and 3rd properties of the KL divergence we get:

$$2(P^m(A) - Q^m(A))^2 \leq KL(P^m||Q^m) = m \cdot KL(P||Q) \leq 2m\epsilon^2$$

By setting the lower bound to $1 - 2\delta$ we get:

$1 - 2\delta \leq |P(A) - Q(A)| \leq \epsilon\sqrt{m}$

Therefore we derive the bound: $m \geq \frac{(1-2\delta)^2}{\epsilon^2} = \Omega(\frac{1}{\epsilon^2})$

## 8.5 A Lower Bound for Best Arm Identification

We begin with a short reminder about the Best-Arm-Identification problem (for simplicity we assume the distribution is Bernouli):

- The algorithm uses actions $T$ times and then outputs a guess $y_T \in A$ for the best action

- We focus only on the quality of the guess - the probability for success: $Pr[y_T = a^*]$
- Each action $a \in A$ has a reward $r(a) \in [0, 1]$
- For every action $a \in A$ denote $\mu(a) = E[r(a)]$.
- An actions profile is defined as $I = \{\mu(a) : a \in A\}$
- The criterion for success (we ignore $\delta$ for simplicity): $Pr[y_T \; correct|I] \geq 0.99$

We'll look at a group of profiles:

$$I_j = \begin{cases} \mu(i) = \frac{1}{2} & i \neq j \\ \mu(i) = \frac{1+\epsilon}{2} & i = j \end{cases}$$

We'll want: $Pr[y_T = i|I_i] \geq 0.99$

**Lemma 8.4** *Suppose $T < \frac{ck}{\epsilon^2}$ for a small enough c. Then there exist at least $\lceil \frac{k}{3} \rceil$ actions j for which $Pr[y_T = j|I_j] < \frac{3}{4}$*

Before the proof, we define some notations.
Given a profile I, we sample each action $T$ times. Therefore the probability space is $\Omega = \{0, 1\}^{KxT}$ and each $w \in \Omega$ is a realization.

For every $A \subseteq \Omega$ denote: $P_j(A) = Pr[A|I_j]$
Denote the probability for action $a \in A$ at time $t \in [T]$ by $P_j^{a,t}$
Therefore: $P_j = \Pi_{a \in A} \Pi_{t \in [T]} P_j^{a,t}$

We prove for two cases:

1. For $k = 2$ we prove the bound $T = \Omega(\frac{1}{\epsilon^2})$
2. For $k \geq 24$ we prove the bound $T = \Omega(\frac{k}{\epsilon^2})$

## 8.6 Proof for $K = 2$

**Proof:** There are two actions $\{1, 2\}$. Denote by A the realizations where the algorithm predicts 1: $A = \{w \in \Omega : y_T = 1\}$. For correctness, we demand $P_1(A) \geq 0.99$ and $P_2(A) \leq 0.01$.

Then:

$$2(P_1(A) - P_2(A))^2 \leq KL(P_1||P_2) = \sum_{a \in \{1,2\}} \sum_{t \in [T]} KL(P_1^{a,t}||P_2^{a,t})$$

$$\leq 2T \cdot 2\epsilon^2 = 4T\epsilon^2$$

By requiring $|P_1(A) - P_2(A)| \geq 0.98$ we get:

$0.98 \leq |P_1(A) - P_2(A)| \leq \epsilon\sqrt{2T}$

Therefore: $T = \Omega(\frac{1}{\epsilon^2})$ □

## 8.7 Proof for $K \geq 24$

**Proof:** We define another profile: $I_0 = \{\mu(a) = \frac{1}{2}\}$. Intuitively, this is a "fair" profile, which assigns the same expected value to all arms, in contrast to the "unfair" profiles $I_j$, which give the $j^{th}$ arm's expected value a positive bias.

We make the following claims:

**Claim 8.5** $\exists K_1 \subseteq K : |K_1| \geq \frac{2}{3}|K| \wedge E[T_j|I_0] = E_0[T_j] \leq \frac{3T}{K}$

**Proof:** By contradiction, assume that there is a group $K' \subseteq K : |K'| > |K/3|$ arms with $\forall j \in K', E_0[T_j] > 3T/K$. This implies that the arms in $K'$ are played strictly more than $T$ times. □

**Claim 8.6** $\exists K_2 \subseteq K : |K_2| \geq \frac{2}{3}|K| \wedge \forall j \in K_2 : Pr[y_T = j|I_0] = Pr_0[y_T = j] \leq \frac{3}{K}$

**Proof:** By contradiction, assume that there is a group $K' \subseteq K : |K'| > |K/3|$ arms with $\forall j \in K', Pr_0[y_T = j] > \frac{3}{K}$. This implies that the combined probability that the arms in $K'$ are selected to be $y_T$ is strictly greater than 1. □

Now, we use Markov's inequality and $E_0[T_j] \leq \frac{3}{K}T$ from claim 8.5 to deduce that $Pr_0[T_j \geq \frac{24}{K}T] \leq \frac{1}{8}$. This is equivalent to $Pr_0[T_j \leq \frac{24}{K}T] \geq \frac{7}{8}$

We define $K_3 = |K_1 \cap K_2|$. Due to the pigeon hole principle we have $|K_1 \cap K_2| \geq \frac{1}{3}K$.

It follows from the definition of $K_3$ and claim 8.5 and 8.6 that:

$$\forall j \in K_3 : Pr_0[T_j \leq \frac{24}{K}T] \geq \frac{7}{8} \wedge Pr_0[y_T = j] \leq \frac{3}{K}$$

We choose a $j \in K_3$ and define a "reduced" sample space $\Omega^* = \Omega_j^m \times \prod_{a \neq j} \Omega_a^T$, where arm $j$ is played only $m = \frac{24T}{K}$ times. The reason for defining this sample space will become apparent later on.
We define for every action and for each event $A \subseteq \Omega^*$:

$$P_l^*(A) = Pr[A|I_l]$$

Using pinsker's rule:

$$2(P_0^*(A) - P_j^*(A))^2 \leq KL(P_0^*||P_j^*) = \sum_{a \neq j} \sum_t KL(P_0^{*a,t}||P_j^{*a,t}) + \sum_{t=1}^{m} KL(P_0^{*j,t}||P_j^{*j,t}) \leq 2m\epsilon^2$$

Rearranging terms, we get (for $c < \frac{1}{1600}$):

$$\forall A : |P_0^*(A) - P_j^*(A)| \leq \epsilon\sqrt{m} = \epsilon\sqrt{\frac{24T}{K}} = \epsilon\sqrt{\frac{24cK}{\epsilon^2 K}} = \sqrt{24c} \leq \frac{1}{8} \tag{8.1}$$

Notice the role of $m$ in the simplification.

We note the bound in (8.1) holds only for events $A \subseteq \Omega^*$ . Therfore, we can not use it the bound the event $A = \{y_T = j\}$ directly. To overcome this, we now define the following two events:

$$A = \{y_T = j \wedge T_j \leq m\}, A' = \{T_j > m\}$$

both of these events are in $\Omega^*$ ($A' \subseteq \Omega^*$ because whether the algorithm samples arm $j$ more than $m$ times is completely determined by the first $m$ coin tosses)

Using $A$ and $A'$, we bound $P_j^*(A)$ :

$$P_j^*(A) \leq \frac{1}{8} + P_0^*(A) \leq \frac{1}{8} + P_0^*(y_T = j) \leq \frac{1}{8} + \frac{3}{K} \leq \frac{1}{4}$$

$$P_j^*(A') \leq \frac{1}{8} + P_0^*(A') \leq \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

$$P_j(y_T = j) \leq P_j^*(y_T = j \wedge T_j \leq m) + P_j^*(T_j > m) \leq \frac{1}{2}$$

□

## 8.8 Proving a Lower Bound on Multi Armed Bandits Algorithms

**Theorem 8.7 (MAB Lower Bound)** *For every MAB algorithm:*

$$E[regret] = \Omega(\sqrt{TK})$$

**Proof:** Assume $T \leq \frac{cK}{\epsilon^2}$

For every round $t$, we define a set $S_t$ which contains all the "good" arms which the MAB algorithm will not pick with the required high probability (define high probability as say, 99%), as defined in lemma **??**.

$$S_t = \{j \in K : Pr[a_t = j|I_j\} \leq \frac{3}{4}$$

Due to lemma 8.4, we have the lower bound of $S_t$:

$$|S_t| \geq \frac{K}{3}$$

Now, we run a profile $I_j$. For every $a \neq j$ we have $\Delta = \epsilon/2$ (difference in bias between fair coin $RC_0$ and unfair coin $RC_\epsilon$)

We proceed to derive the required lower bound:

$$E[\mu(a_t)|a^* \in S_t] \leq \frac{1}{2}Pr[a_t \neq a^*|a^* \in S_t] + \frac{1+\epsilon}{2}Pr[a_t = a^*|a^* \in S_t] =$$

$$\frac{1}{2} + \frac{\epsilon}{2}Pr[a_t = a^*|a^* \in S_t] \leq \frac{1}{2} + \frac{\epsilon}{2} \cdot \frac{3}{4} = \frac{1}{2} + \frac{3\epsilon}{8} = \mu^* - \frac{\epsilon}{8}$$

$$E[\mu(a_t)] \leq Pr[a^* \in S_t](\mu^* - \frac{\epsilon}{8}) + Pr[a^* \in S_t]\mu^* = \mu^* - \frac{\epsilon}{8}Pr[a^* \in S_t] \leq \mu^* - \frac{\epsilon}{24}$$

$$E[R_{on}] = \sum_t E[\mu(a_t)] \leq \mu^* \cdot T - \frac{\epsilon}{24} \cdot T$$

$$E[regret] \geq \frac{\epsilon}{24} \cdot T \geq c \cdot \sqrt{KT}$$

Where we used $\epsilon = \sqrt{\frac{k}{T}}$ □