CSE599s, Spring 2014, Online Learning Lecture 19 - 06/03/2014

Information Theoretic Lower Bounds

Lecturer: Ofer Dekel *Scribe: Ajay Menon*

1 Introduction

Today we are going to prove the information theoretic lower bounds on regret for the experts setting. Proving an information theoretic lower bound on regret is equivalent to saying that there does not exist any algorithm that can get a better regret.

In the experts setting (full information, k = 2 experts) with an oblivious adversary, we are going to show that Regret $= \Omega(\sqrt{T})$.

In other words:

$$\forall A, \exists \ell = \ell_1, \ldots, \ell_T \text{ with } \ell_t \in [0,1]^k \text{ s.t. } \mathrm{Regret}(A, \ell) = \Omega(\sqrt{T}).$$

We can come up with a sequence of an arbitrary number ($\leq T$) of loss functions, and let the algorithm know the length of the sequence in advance. During the game, the algorithm will incur a variable amount of loss and suffer a variable amount of regret based on the number of loss functions chosen.

We use the notation $\mathrm{Regret}(A(T), \ell^{(T)}) = \Omega(\sqrt{T})$ to describe the regret of the algorithm A which gets to know T in advance and uses the sequence of T loss functions $\ell^{(T)}$. Finding such a sequence $\ell^{(T)}$ which incurs $\Omega(\sqrt{T})$ Regret is computationally difficult, but evaluating such a sequence is computationally easy.

First, we are going to use the Minimax Theorem to bound the regret over all algorithms and loss functions in this setting. Then we will discuss a stochastic loss function which guarantees Regret $= \Omega(\sqrt{T})$ and prove that the bound holds against any algorithm.

As an aside, since the bandits setting can only be harder than the experts setting, we can generalize these results to the bandit setting (k = 2 arms) with an oblivious adversary to get the same lower bound on Regret $= \Omega(\sqrt{T})$.

2 Using the Minimax Theorem (Inequality Version)

First some notation:

- A is a distribution over deterministic algorithms (equivalent to a randomized algorithm).
- a is some deterministic online learning algorithm.
- $\mathcal{L}$ is some stochastic loss sequence.
- ℓ is the particular series of loss functions to be played.
- $G(\mathcal{L}) = \min_a \mathbb{E}_{\ell \sim \mathcal{L}}[\mathrm{Regret}(a, \ell)]$.

Theorem 1 (Minimax Theorem)**.** .

$$\min_A \max_\ell \ \mathbb{E}_{a \sim A}[\mathrm{Regret}(a, \ell)] \geq \max_{\mathcal{L}} \min_a \ \mathbb{E}_{\ell \sim \mathcal{L}}[\mathrm{Regret}(a, \ell)]$$

Proof.

$$\forall A \ \min_a \ \mathbb{E}_{\ell\sim\mathcal{L}^*}[\text{Regret}(a,\ell)] \leq \mathbb{E}_{a\sim A}[\mathbb{E}_{\ell\sim\mathcal{L}}[\text{Regret}(a,\ell)|a]].$$

Then, the order of the nested expectations can be reversed as follows:

$$\mathbb{E}_{a\sim A}[\mathbb{E}_{\ell\sim\mathcal{L}}[\text{Regret}(a,\ell)|a]] = \mathbb{E}_{\ell\sim\mathcal{L}^*}[\mathbb{E}_{a\sim A}[\text{Regret}(a,\ell)|\ell]].$$

This expected regret over loss functions and algorithms will be no greater than the regret suffered by the maximum-loss sequence of loss functions:

$$\mathbb{E}_{\ell\sim\mathcal{L}^*}[\mathbb{E}_{a\sim A}[\text{Regret}(a,\ell)|\ell]] \leq \max_\ell \ \mathbb{E}_{a\sim A}[\text{Regret}(a,\ell)].$$

The intuition behind this is that taking the average over a bunch of algorithms will produce a larger expected regret than the single best algorithm.

Specifically, this holds for $A^* = \text{argmin}_a \max_\ell \mathbb{E}_{a\sim A}[\text{Regret}(a,\ell)]$, which gives us the left hand side of the minimax theorem. □

3 Stochastic Loss Sequence Proofs

To prove the $\Omega(\sqrt{T})$ regret, it suffices to show for any T, there exists a stochastic loss sequence that guarantees $\Omega(\sqrt{T})$ expected regret against any deterministic algorithm.

Theorem 2. *For the $k = 2$ experts setting, a stochastic loss sequence which satisfies this would be:*

- *Draw $Y \in -1, +1$ with equal probability. Pr(Y = +1) = Pr(Y = -1) = $\frac{1}{2}$.*
- *Set expert 1's losses iid from $\mathcal{N}(\frac{1}{2}(1+Y\epsilon), \frac{\sigma^2}{2})$.*
- *Set expert 2's losses iid from $\mathcal{N}(\frac{1}{2}(1-Y\epsilon), \frac{\sigma^2}{2})$.*

To prove the information theoretic lower bound, we will need to show that there is not enough information to determine the value of Y until Regret $= \Omega(\sqrt{T})$.

3.1 Proof of Regret Bound

On each round, the difference between the losses of each expert is the difference between two Gaussian distributions

$$\ell_{t2} - \ell_{t1} \simeq \mathcal{N}(Y\epsilon, \sigma^2).$$

Depending on the value of Y, this equation will equal either $\mathcal{N}(\epsilon, \sigma^2)$ or $\mathcal{N}(-\epsilon, \sigma^2)$.

In order to determine the value of Y, we will need $\Theta(\frac{1}{\epsilon^2})$ samples. If we set $\epsilon = \frac{1}{\sqrt{T}}$, then $\frac{1}{\epsilon^2} = \Theta(T)$ samples will be required before we can confidently (with high probability) estimate the value of Y.

On each round, with probability $\frac{1}{2}$ we incur no loss, and with probability $\frac{1}{2}$ we incur ϵ loss (the gap between the Gaussians). Thus the expected loss we incur on each round is $\Theta(\epsilon)$. On those first $\Theta(T)$ rounds, we suffer $\Theta(\epsilon T) = \Theta(\frac{T}{\sqrt{T}}) = \Theta(\sqrt{T})$ regret.

Now we need to show is that for our chosen value of ϵ, we will need $\Theta(T)$ rounds until we can accurately determine Y.

3.2 Proof of Required Number of Samples

After k rounds, we have observed k i.i.d samples from either $P = \mathcal{N}^k(\epsilon, \sigma^2)$ or $Q = \mathcal{N}^k(-\epsilon, \sigma^2)$.

The Total Variational Distance (equivalent: Statistical Distance) between two distributions is

$$TV(P||Q) = sup_{e \in events}|P(e) - Q(e)|. \tag{1}$$

We will show that for $k << \frac{1}{\epsilon^2}$, $TV(P||Q) \simeq 0$. This is equivalent to saying that for small enough k, the statistical distance between the two hypotheses is essentially 0 - all measurable events almost equally probable for the 2-arm setting.

Pinsker's Inequality states that

$$TV(P||Q) \leq \sqrt{\frac{1}{2}\,\mathrm{KL}(P||Q)}. \tag{2}$$

Here, $\mathrm{KL}(P||Q)$ is the Kullback-Leibler divergence (also known as information gain or relative entropy) between P and Q. If we define p and q to be the pdf of P and Q, respectively, then

$$\mathrm{KL}(P||Q) = \int p(x) log(\frac{p(x)}{q(x)})dx. \tag{3}$$

$$= \mathbb{E}_{x \sim p}[log(p(x)) - log(q(x))]$$

$$= \mathbb{E}_{x \sim p}[log\Pi_{i=1}^{k} p(x_i) - log\Pi_{i=1}^{k} q(x_i)]$$

$$= k * \mathbb{E}_{x \sim \mathcal{N}(\epsilon, \sigma^2)}[log(p(x)) - log(q(x))].$$

Now, we can simplify the inner term of the expectation for any p and q from a normal distribution. The pdf of the normal distribution $\mathcal{N}(\mu, \sigma)$ is

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

For any $p = \mathcal{N}(\mu_1, \sigma_1)$ and $q = \mathcal{N}(\mu_2, \sigma_2)$

$$log(p(x)) - log(q(x)) = (-\frac{1}{2}log(2\pi) - log(\sigma_1) - \frac{1}{2\sigma_1^2}(x-\mu_1)^2) + (\frac{1}{2}log(2\pi) - log(\sigma_2) + \frac{1}{2\sigma_2^2}(x-\mu_2)^2)$$

$$= -log(\frac{\sigma_1}{\sigma_2}) - \frac{1}{2\sigma_1^2}(x-\mu_1)^2 + \frac{1}{2\sigma_2^2}(x-\mu_2)^2.$$

Plugging this into the inner term of the expectation in $\mathrm{KL}(P||Q)$, we get

$$\mathbb{E}_p[log(p(x)) - log(q(x))] = -log(\frac{\sigma_1}{\sigma_2}) - \frac{1}{2\sigma_1^2}\mathbb{E}_p[(x-\mu_1)^2] + \frac{1}{2\sigma_2^2}\mathbb{E}_p[(x-\mu_2)^2]$$

$$= -log(\frac{\sigma_1}{\sigma_2}) - \frac{\sigma_1^2}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}\mathbb{E}_p[((x-\mu_1) - (\mu_2 - \mu_1))^2]$$

$$= -log(\frac{\sigma_1}{\sigma_2}) - \frac{\sigma_1^2}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}\mathbb{E}_p[((x-\mu_1) + (\mu_1 - \mu_2))^2]$$

$$= -log(\frac{\sigma_1}{\sigma_2}) - \frac{1}{2} + \frac{1}{2\sigma_2^2}\mathbb{E}_p[(x-\mu_1)^2 + (\mu_1-\mu_2)^2 + 2(x-\mu_1)(\mu_1-\mu_2)]$$

$$-log(\frac{\sigma_1}{\sigma_2}) - \frac{1}{2} + \frac{1}{2\sigma_2^2}(\sigma_1^2 + (\mu_1 - \mu_2)^2 + 0)$$

$$= -log(\frac{\sigma_1}{\sigma_2}) - \frac{1}{2} + \frac{\sigma_1^2}{2\sigma_2^2} + \frac{(\mu_1 - \mu_2)^2}{2\sigma_2^2}.$$

If $\sigma_1 = \sigma_2 = \sigma$, then

$$\text{KL}(N(\mu_1, \sigma)||N(\mu_2, \sigma)) = \frac{(\mu_1 - \mu_2)^2}{2\sigma^2}.$$

In our specific case we know that P and Q have the same value for σ, so we can now simplify the KL divergence to be

$$TV(P^k||Q^k) \leq \sqrt{\frac{1}{2}k\frac{4\epsilon^2}{2\sigma^2}}.$$

With our chosen value of $\epsilon = \frac{1}{\sqrt{T}}$ and some constant c, we get

$$TV(P^k||Q^k) \leq \sqrt{c\frac{k}{T}}.$$

If $k < T$, then

$$\frac{k}{T} \to 0.$$

This implies that

$$TV(P^k||Q^k) \to 0.$$

Thus, with a sub-linear number of rounds, the total variational distance between the two hypotheses is approximately zero.

4 Conclusion

As a result of our proofs regarding our chosen stochastic loss function, we know that the only way to confidently estimate the value of Y is to play $\Theta(T)$ rounds, which results in the algorithm suffering an expected regret of $\Theta(\sqrt{T})$.

Since this applies for any deterministic algorithm (and consequently, any stochastic algorithm), we have proved an information theoretic lower bound on regret of $\Omega(\sqrt{T})$ for the experts setting with two experts.