Week 8 – Lecture 39

Suffices to show that:

$\exists\ \mu, \mu' \in [0,1]^K$ s.t.

$$\max\left\{R_T(\pi, \nu), R_T(\pi, \nu')\right\} \geq C\sqrt{KT}$$

Let μ be s.t.

$$\mu_i = \begin{cases} \Delta & i = 1 \\ 0 & i \neq 1 \end{cases}$$ ← For this bandit instance arm 1 is optimal

$$R_T(\pi, \nu) = \left(T - \mathbb{E}[N_1(T)]\right)\Delta$$

$\exists\ i \neq 1$ s.t. $\mathbb{E}[N_i(T)] \leq \dfrac{T - \mathbb{E}[N_1(T)]}{K-1}$

(should be at least one arm like this)

$\leq \dfrac{T}{K-1}$ (upper-bounding)

Construct $\mu' \in [0,1]^K$ as

$$\mu'_j = \begin{cases} \mu_j & j \neq i \\ 2\Delta & j = i \end{cases}$$

$$\begin{bmatrix} \mu = (\Delta, 0, 0, \overset{i}{0}, 0, 0) \\ \mu' = (\Delta, 0, 0, 2\Delta, 0, 0) \end{bmatrix}$$

The two instances differ only in the ith arm, the optimal arms are different for both cases.

Now the task is to set Δ s.t. the algorithm cannot distinguish between the two instances.

Expected Regret for 2nd bandit instance

$$R_T(\pi, \nu') = \underbrace{\mathbb{E}'[N_1(T)]\Delta}_{\text{regret due to arm 1}} + \underbrace{\sum_{j \neq i,1} \mathbb{E}'[N_j(T)]\, 2\Delta}_{\text{regret for other arms}}$$

$$R_T(\pi, \nu') \geq \Delta \mathbb{E}'[N_1(T)]$$

Let $\Delta = \sqrt{\dfrac{1}{\mathbb{E}[N_i(T)]}} \geq \sqrt{\dfrac{K-1}{T}}$

Then we expect

$$\mathbb{E}[N_1(T)] \approx \mathbb{E}'[N_1(T)]$$

Consider that :

1) $\mathbb{E}[N_1(T)] < T/2$

$$R_T(\pi, \nu) = (T - \mathbb{E}[N_1(T)])\Delta$$

$$\geq \frac{T}{2}\Delta \geq \frac{1}{2}\sqrt{T(K-1)}$$

2) $\mathbb{E}[N_1(T)] \geq T/2$

$$R_T(\pi, \nu') \geq \Delta \mathbb{E}'[N_1(T)]$$

$$\approx \Delta \mathbb{E}[N_1(T)] \geq \frac{1}{2}\sqrt{T(K-1)}$$

Some Definitions from Information Theory

$$P = (P_1, P_2, \ldots P_K)$$

$$H(P) = \sum_{i=1}^{K} P_i \log \frac{1}{P_i}$$

$\uparrow$

Entropy

Amount of information contained is inversely proportional to its probability.

Hence while encoding something :

→ code frequent events with smaller code

→ code rare events with larger code

Entropy gives the length of code for information

Let P & Q are defined on same prob. space.

Then:

$$D(P\|Q) = \sum_{i=1}^{K} P_i \log \frac{P_i}{q_i}$$

[Here P, Q defined on discreet r.v.]

$$= -H(P) + \sum_{i=1}^{K} P_i \log \frac{1}{q_i} \geq 0$$

Suppose message is generated by P_i but wrongly interpreted as from q_i, then the length assigned to the message would be $\sum_{i=1}^{K} P_i \log \frac{1}{q_i}$,

but that would be wrong, and $-H(P)$ is the best that could have been done.

K-L Divergence is not transitive, hence not a true metric, but measures divergence between 2 probability distributions.

$$D(P\|Q) = \begin{cases} \int_{\omega} \log\left(\frac{dP}{dQ}(\omega)\right) dP(\omega) & \text{if } P \ll Q \end{cases}$$

$$\begin{cases} & \\ 0 & \text{o.w.} \end{cases}$$

[defined for continous r.v.]

Let $P \sim N(\mu_1, \sigma^2)$

$Q \sim N(\mu_2, \sigma^2)$

$$D(P, Q) = \frac{(\mu_1 - \mu_2)^2}{2\sigma^2}$$

Theorem: Let P & Q be probability measures on the same measurable space $(\Omega, \mathcal{F})$.

Let $A \in \mathcal{F}$ be an arbitrary event. Then,

$$P(A) + Q(A^c) \geq \frac{1}{2} \exp(-D(P \| Q))$$

where $A^c = \Omega \setminus A$