EE6106 Lecture 10 (Date : 20th Feb 2024)

Recall :

$\mathbb{P}, \mathbb{Q} \rightarrow 2$ prob measures on $(\Omega, \mathcal{F})$

$\mathbb{P}$ absolutely continuous wrt $\mathbb{Q}$

$$\{ \mathbb{P}(A) > 0 \Rightarrow \mathbb{Q}(A) > 0 \}$$

$\exists$ measurable function $\frac{d\mathbb{P}}{d\mathbb{Q}}$ s.t.

$$\mathbb{E}_P[X] = \mathbb{E}_{\mathbb{Q}}\left[X \frac{dP}{dQ}\right]$$

If $\Omega = \{1, 2, \dots, n\}$

$$p_i = \mathbb{P}(\{i\}), \quad q_i = \mathbb{Q}(\{i\})$$

Then $\frac{d\mathbb{P}}{d\mathbb{Q}}(i) = \frac{p_i}{q_i}$

captures a likelihood ratio

We are doing all this in order to arrive at the KLDivergence

Def : For prob. measures $\mathbb{P}, \mathbb{Q}$ on $(\Omega, \mathcal{F})$, if $\mathbb{P}$ is abs. cont. wrt $\mathbb{Q}$

$$KL(\mathbb{P}, \mathbb{Q}) = D(\mathbb{P}, \mathbb{Q}) = \int \log\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right) d\mathbb{P}$$

$$= \mathbb{E}_P\left[\log\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)\right]$$

Kullback Liebler Divergence

relative entropy between $\mathbb{P}$ and $\mathbb{Q}$

If $\mathbb{P}$ is not absolutely continuous wrt $\mathbb{Q}$,

$$KL(\mathbb{P}, \mathbb{Q}) = D(\mathbb{P}, \mathbb{Q}) = \infty$$

Note : The KL divergence is as asymmetric object, hence it is not a distance. It is astounding that it still captures the closeness of the two distributions

Note : It is entirely possible that $D(\mathbb{P}, \mathbb{Q}) = \infty$ when $D(\mathbb{Q}, \mathbb{P})$ is finite

If $\Omega = \{1, 2, \ldots, n\}$, then

$$D(P, Q) = \sum_{i \in \Omega} p_i \log\left(\frac{p_i}{q_i}\right)$$

Note: The KL divergence between two random variables X and Y is understood to be the KL divergence between their laws

Eg $X \sim \text{Bernoulli}(p)$ and $Y \sim \text{Bernoulli}(q)$

$$D(X, Y) = p \log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right)$$

Eg $X \sim \mathcal{N}(\mu_1, \sigma^2)$ (variances are assumed to be the same)

$Y \sim \mathcal{N}(\mu_2, \sigma^2)$

$$\frac{dF_x}{dF_y} = \exp\left(-\frac{(x-\mu_1)^2}{2\sigma^2} + \frac{(x-\mu_2)^2}{2\sigma^2}\right) \quad \text{... special case}$$

$$KL(X, Y) = \int_{-\infty}^{\infty} f_x(x) \log\left(\frac{dF_x(x)}{dF_y}\right) dx \underset{\text{EXERCISE}}{=} \frac{(\mu_1 - \mu_2)^2}{2\sigma^2}$$

Thm If P, Q are 2 probability measures on $(\Omega, \mathcal{F})$, then for $A \in \mathcal{F}$,

$$P(A) + Q(A^c) \geq \frac{1}{2} e^{-\min(D(P, Q), D(Q, P))}$$

This is known as the Brettagnolle – Huber inequality

Note: This is true for all $A \in \mathcal{F}$.

The left and right hands are both symmetric wrt P & Q

☆ It hence suffices to show

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

What this inequality is trying to tell us is the following

① If P is very close to Q, the event on the left is sure and hence $P(A) + Q(A^c)$ is a very likely event

② $P(A) + Q(A^c)$ can be small only if A is rare under P and A^c is rare under Q. Then by this P cannot be close to Q.

Proof: STEP 1

We will show that $\mathbb{P}(A) + \mathbb{Q}(A^c) \geq \sum_{i \in \Omega} \min(p_i, q_i)$

(We assume we are working in the discrete setting)

$$\mathbb{P}(A) + \mathbb{Q}(A^c) = \sum_{i \in A} p_i + \sum_{i \in A^c} q_i \geq \sum_{i \in \Omega} \min(p_i, q_i) \quad \square$$

STEP 2

$$\sum \min(p_i, q_i) \geq \frac{1}{2}\left(\sum \sqrt{p_i q_i}\right)^2$$

$$RHS = \left(\sum \sqrt{p_i q_i}\right)^2 = \left(\sum \sqrt{\max(p_i, q_i)\min(p_i, q_i)}\right)^2$$

$$\underset{\text{Cauchy-Schwarz}}{\leq} \left[\sum \max(p_i, q_i)\right]\left[\sum \min(p_i, q_i)\right]$$

$\leq p_i + q_i \leq 2$

$$\leq 2\sum \min(p_i, q_i) \quad \square$$

STEP 3

$$\left(\sum \sqrt{p_i, q_i}\right)^2 \geq e^{-D(P, Q)}$$

$$LHS = \left(\sum \sqrt{p_i q_i}\right)^2 = e^{2\log \sum \sqrt{p_i q_i}}$$

$$= e^{2\log \sum p_i \sqrt{\frac{q_i}{p_i}}}$$

$$\geq e^{2\sum p_i \log\left(\sqrt{\frac{q_i}{p_i}}\right)}$$

$$\geq e^{-\sum p_i \log\left(\frac{p_i}{q_i}\right)} = e^{-D(P, Q)}$$

Consider a policy π, acting on bandit instance $\nu = (\nu_1, \nu_2, \ldots, \nu_k)$, we define $\mathbb{P}_\nu$, which captures the joint distribution of $(A_1, X_1, A_2, X_2, \ldots, A_n, X_n)$

Lemma (Divergence Decomposition Lemma)

Let ν, ν' be 2 MAB instances. Let $\mathbb{P}_\nu$, $\mathbb{P}_{\nu'}$ denote the corresponding measures induced by a common policy π. Then

$$D(\mathbb{P}_\nu, \mathbb{P}_{\nu'}) = \sum_{i=1}^{K} \underbrace{E_{\mathbb{P}_\nu}[T_i(n)]}_{\substack{\text{Expected times} \\ i^{\text{th}} \text{ arm is pulled under } \nu}} D(\nu_i, \nu_i')$$

Note that this is for arm i

ν_i is the distribution of the i^{th} arm in the instance ν_i

We shall use the quantity $D(\mathbb{P}_{\nu}, \mathbb{P}_{\nu'})$ by treating it with the Brettagnole-Huber inequality and choosing the event A strategically.

Note
There are no restrictions on the two instances expect such that the two supports are conveniently taken to be the same.

Proof

$$D(\mathbb{P}_{\nu}, \mathbb{P}_{\nu'}) = \mathbb{E}_{\nu}\left[\log \frac{d\mathbb{P}_{\nu}}{d\mathbb{P}_{\nu'}}\right]$$

$$\frac{d\mathbb{P}_{\nu}}{d\mathbb{P}_{\nu'}}(a_1, x_1, \dots, a_n, x_n) = \frac{\prod_{t=1}^{n} \underbrace{\pi_t(a_t \mid a_1, x_1, \dots, a_{t-1}, x_{t-1})}_{\mathbb{P}(\text{taking action based on history})} \underbrace{P_{a_t}(x_t)}_{\mathbb{P}(\text{outcome given action})}}{\prod_{t=1}^{n} \pi_t(a_t \mid a_1, x_1, \dots, a_{t-1}, x_{t-1}) P'_{a_t}(x_t)}$$

$$= \prod_{t=1}^{n} \frac{P_{a_t}(x_t)}{P'_{a_t}(x_t)}$$ ... since the policies are identical

$$\mathbb{E}_{\nu}\left[\log \frac{d\mathbb{P}_{\nu}}{d\mathbb{P}_{\nu'}}\right] = \mathbb{E}_{\nu}\left[\log\left(\frac{\prod_{1}^{n} P_{a_t}(x_t)}{\prod_{1}^{n} P'_{a_t}(x_t)}\right)\right]$$

$$= \sum_{t=1}^{n} \mathbb{E}_{\nu}\left[\log \frac{P_{A_t}(x_t)}{P'_{A_t}(x_t)}\right]$$

$$= \sum_{t=1}^{n} \mathbb{E}_{\nu}\left[\mathbb{E}_{\nu}\left[\log \frac{P_{A_t}(X_t)}{P'_{A_t}(X_t)} \,\middle|\, A_t\right]\right]$$

$$= \sum_{t=1}^{n} \mathbb{E}_{\nu}\left[D(\nu_{A_t}, \nu'_{A_t})\right]$$

$$= \sum_{t=1}^{n} \sum_{i=1}^{k} \mathbb{E}_{\nu}\left[D(\nu_{A_t}, \nu'_{A_t}) \mathbb{1}_{\{A_t = i\}}\right]$$

$$= \sum_{i=1}^{k} \sum_{t=1}^{n} D(v_i, v_i') \mathbb{P}_v (A_t = i)$$

$$= \sum_{i=1}^{k} \mathbb{E}_v [T_i(n)] D(v_i, v_i') \qquad \square$$

Syllabus : Everything till now

HW2 : Stochastic Bandits and Preliminaries for IT Bounds (Submission post Midsem)

It is open notes !! (Handwritten notes can be referred)