Week 3

Lecture 14

Adversarial MAB

Pseudo regret: $(\{L_t\} = \text{loss})$

$$\bar{R}(n,\pi) = \mathbb{E}\left[\sum_{t=1}^{n} L_{tI_t}\right] - \min_i \mathbb{E}\left[\sum_{t=1}^{n} L_{ti}\right] \quad (3)$$

Exp 3 (Exponential Weights for exploration & exploitation) Algorith

Input: $(\eta_t)_{t\in\mathbb{N}}$, number of arms K

Initialize: P_1 is uniform over $[K] = \{1, 2, \dots K\}$

For rounds $t = 1, 2, \dots n$

(1) Draw an arm I_t from distribution P_t and observe loss l_{tI_t}

(2) $\forall\, i \in [K]$, estimate $\hat{l}_{it} = \dfrac{l_{it}}{P_{it}} \mathbb{1}\{I_t = i\}$

and update cumulative loss

$$\hat{L}_{ti} = \hat{L}_{(t-1)i} + \hat{l}_{it}$$

(3) Update probability

$$\forall i \quad P_{(t+1)i} = \frac{\exp\left(-\eta_t \hat{L}_{ii}\right)}{\sum_{k=1}^{K} \exp\left(-\eta_t \hat{L}_{kt}\right)}$$

Theorem :— (1) If Exp3 is run with

$$\eta := \eta_t = \sqrt{\frac{2\log K}{nK}}, \quad \text{then}$$

[If we know how many total number of rounds $= n$
η_t is a const. here

$$\bar{R}(n, \text{Exp3}) \leq \sqrt{2nk\log K}$$

(2) if $\eta_t = \sqrt{\frac{\log K}{tK}}$, then [We don't know horizon

$$\bar{R}(n, \text{Exp3}) \leq 2\sqrt{nK\log K} \longrightarrow \text{Anytime algorithm}$$

To prove the theorem, we will show that :

$$\bar{R}_n \leq \frac{K}{2}\sum_{t=1}^{n} \eta_t + \frac{\ln K}{\eta_n}$$

if (1) holds then $\bar{R}_n \leq \sqrt{2n\log K}$ (by inputting the value of η_t and simplifying)

if (2) holds ; $\bar{R}_n \leq \frac{K}{2}\sum_{t=1}^{n}\sqrt{\frac{\log K}{tK}} + \frac{\ln K}{\sqrt{\frac{\log K}{nK}}}$ ← as η_t for $t=n$ will be this

$$\left[\sum_{t=1}^{n}\frac{1}{\sqrt{t}} = \int \frac{1}{\sqrt{t}}\,dt = 2\sqrt{n}\right]$$

$$\therefore \bar{R}_n \leq \frac{K}{\cancel{2}}\sqrt{\frac{\log K}{K}} \times \cancel{2}\sqrt{n} + \sqrt{nK\log K}$$

$$= 2\sqrt{nK\log K}$$