Week 3 — Lecture 17

In each round

learner picks p_t

Adversary picks l_t

Types of Adversary :

- Oblivious : Before game begins Adversary chooses the losses for all rounds
  $\{ l_1, l_2, \dots l_n \}$

- Non-Oblivious : losses can be assigned to each action based on the past observations

Exp3 holds for both kinds of adversary.
Two kinds of $\mathbb{E}$ were taken, one for learner & one for adversary, holds for non-oblivious case, which is the worse of the two.

true loss incurred

$$R(n,\pi) = \sum_{t=1}^{n} l_{tI_t} - \min_i \sum_{t=1}^{n} l_{ti} \quad \text{——} \ (1)$$

$$\mathbb{E}\left[R(n,\pi)\right] = \mathbb{E}\left[\sum_{t=1}^{n} l_{tI_t}\right] - \min_i \mathbb{E}\left[\sum_{t=1}^{n} l_{ti}\right] \quad \text{——} \ (2)$$

The bound on (2) (already derived) will be very different, may be higher, than that for (1)

We want to bound $R(n,\pi)$. For a good algorithm we would want to know

$$P_r\left\{R(n,\pi) \leq \mathbb{E}\left[R(n,\pi)\right]\right\}$$ ← should be high

While finding the bound for (1) we will only consider the Oblivious adversary.

For the bound on a quantity, single value, to be low we would want the estimates to have <u>low variance</u>.

<u>Estimates of Exp3</u>

$$\hat{l}_{ti} = \frac{l_{ti}}{P_{ti}} 1\{I_t = i\}$$

$$\mathbb{E}\left[\hat{l}_{ti}\right] = l_{ti}$$

$$Var\left(\hat{l}_{ti}\right) = \frac{1}{P_{ti}}$$

If P_{ti} is small then variance will be very high.

<u>Exp3.P Algorithm</u>

<u>Input</u>: $\{\eta_t\}, K, \gamma \quad [\gamma \in (0,1)], \beta, \delta$

<u>Initialize</u>: $\hat{L}_{0i} = 0$

for $t = 1, 2, \ldots$

$$\forall i \quad P_{ti} = \frac{(1-\gamma)\exp(-\eta_t \hat{L}_{ti})}{\sum_j \exp(-\eta_t \hat{L}_{tj})} + \frac{\gamma}{K}$$

for $\gamma = 1$ only uniform distr. applies & $\gamma = 0$ is back to Exp3

Draw $I_t \sim P_t$

$$\forall i \quad \hat{l}_{ti} = \frac{l_{ti} \mathbb{1}\{I_t = i\} + \beta}{P_{ti}}$$

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

Also ensures that P_{ti} is not arbitrarily close to zero, it is bound by γ/K, thus ensuring low variance

High probability bound for Exp3.P

For any $\delta \in (0,1)$ if Exp3P is run with

$$\beta = \sqrt{\frac{\log K/\delta}{nK}}, \quad \eta = 0.95\sqrt{\log K}, \quad \gamma = \sqrt{\frac{K \log K}{n}}$$

Then,

$$R(n, \text{Exp3.P}) \le 5.15\sqrt{nK\log(K/\delta)}$$ with probability $(1-\delta)$

because we are considering individual regret & not $\mathbb{E}$

Further if $\beta = \sqrt{\frac{\log K}{nK}}$ *

$$R(n, \text{Exp3.P}) \le \sqrt{\frac{nK}{\log K}} \log(1/\delta) + 5.15\sqrt{nK\log K}$$ with probability $(1-\delta)$

(stop after n rounds)

* This situation is when we don't specify δ a-priori

High Probability bounds are more desired that expected bounds, as it bounds the actual regret we face at every round. But these bounds are difficult to maintain because they are guarenteed by $\Pr(1-\delta$

High Probability bounds can be converted to expected bound by considering :

$$\mathbb{E}[X] \leq \int_0^1 \frac{1}{\delta} \Pr\left\{x \geq \ln \frac{1}{\delta}\right\} d\delta$$

$$\left[R(n, \text{Exp3P}) - 5.15\sqrt{nK\log K}\right]\sqrt{\frac{\log K}{nK}} > \log(1/\delta$$

holds ~~w.p.~~ w.p. δ

Let $\left[R(n, \text{Exp3P}) - 5.15\sqrt{nK\log K}\right]\sqrt{\frac{\log K}{nK}} = X$

Then

$$\mathbb{E}\left[R(n, \text{Exp3P}) - 5.15\sqrt{nK\log K}\right]\sqrt{\frac{\log K}{nK}} \leq 1$$

$$\mathbb{E}\left[R(n, \text{Exp3P})\right] \leq \sqrt{\frac{nK}{\log K}} + 5.15\sqrt{nK\log K}$$

Exp3IX Algorithm

Forces implicit exploration

Input: $\{\eta_t\}$, K, γ, β

Initialize: $\hat{L}_{0i} = 0$

for $t = 1, 2, \ldots$

$$\forall i \quad P_{ti} = \frac{\exp(-\eta_t \hat{L}_{ti})}{\sum_j \exp(-\eta_t \hat{L}_{tj})}$$

Draw $I_t \sim P_t$, incur loss l_{ti}

$$\forall i \quad \hat{l}_{ti} = \frac{l_{ti} \mathbb{1}\{I_t = i\}}{P_{ti} + \gamma}$$

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