EE6106 Online Learning and Optimization (Date : 19/01/2024)

Generalization of prev setting

$Y_{i,t}, A_t \in D \longrightarrow$ convex (D can be arbitrary as long as it is convex)

D: last class : assumed it was $[0,1]$

$X_t \in E$ (environment) E can be arbitrary

$\ell : E \times D \longrightarrow [0,1]$

$\ell(x, \cdot) \sim$ convex wrt second argument

Protocol:

1. At time $t \geq 1$,
   - $Y_{i,t}$; $1 \leq i \leq k$ seen
   - Action A_t played
   - X_t revealed
   - Algo loss : $\ell(X_t, A_t)$
   - Expert i's loss is $\ell(X_t, Y_{i,t})$

Also note that randomization does not require the loss to be convex

In the earlier setting, convexity was used in 2 places

1) the Jensen's inequality
2) for Hoeffding's lemma ($\in [0,1]$)

Under EWMA, with

$$\beta = \sqrt{\frac{8 \log K}{T}}, \quad R_T \leq \sqrt{\frac{T \log K}{2}}$$

Also, the convexity of D is being used in the working of the algorithm itself (action chosen is a weighted average of the experts' predictions)

Now, we shall drop 2 assumptions :
- D convex (D is still known to us)
- $\ell(x, \cdot)$ convex (it is still bounded between 0 and 1)

We shall show that we are screwed without randomization

Basic Idea : Choose expert randomly using the distribution as provided by the weights

REWMA [Randomized EWMA]
- Initialize weights $w_{i,1} = 1 \ \forall i$
- At $t \geq 1$
  - Choose expert I_t randomly
    $$P(I_t = i) = \frac{w_{i,t}}{\sum w_{i,t}}$$
  - Play $Y_{i,t} = A_t$
  - $w_{i,t+1} = w_{i,t} \, e^{-\beta \ell(X_t, Y_{i,t})}$

REWMA :

EWMA : $\frac{(\cdot)}{\Sigma(\cdot)}$

everything else is the same

$$\underbrace{\text{Define } D' = \{ p \in [0,1]^{k} : \textstyle\sum p_i = 1 \}}$$

k: # of experts

space of probability distributions on experts

$E' = E \times D^{k}$

E: arbitrary, D^k: arbitrary

Note :
$Y_{i,t}, A_t \in D$
$X_t \in E$

$l' : E' \times D' \longrightarrow [0,1]$

Example : $l'((x, \underbrace{y_1, \dots y_k}_{D^k}), p) = \sum_{i=1}^{k} p_i l(x, y_i)$

x: E; $(x, y_1, \dots, y_k)$: E'; p: probability dist. space

$Y_{i,t}' = (0, \dots, 1, 0, 0) = e_i \qquad (Y_{i,t}' \in D')$

i^{th} entry

constant vector

this is NOT an initializat[ion] use the consta[nt] vector $\forall t$

$X_t' = (X_t, Y_{1,t}, Y_{2,t} \dots, Y_{k,t}) \in E'$

X_t: environment signal

$Y_{1,t}, \dots, Y_{k,t}$: signal given by experts

Note that l' is convex (trivially so, it is linear) wrt D'

Using EWMA in this setting,

$$A_t' = \frac{\sum w_{i,t} Y_{i,t}'}{\sum w_{i,t}} = \left(\frac{w_{1,t}}{\sum w_{i,t}}, \frac{w_{2,t}}{\sum w_{i,t}}, \dots, \frac{w_{k,t}}{\sum w_{i,t}} \right)$$

$l'(X_t', Y_{i,t}') = l(X_t, Y_{i,t})$

Thus the loss of the experts does not change , hence the evolution of weights does not change and hence the action taken by the algorithm does not change.

$$\ell'(X'_t, A'_t) = \sum \frac{w_{i,t}}{\sum w_{j,t}} \ell(X_t, Y_{i,t})$$

When EWMA applies to this setting, the deterministic loss of EWMA is the same as the expected loss of REWMA in the setting

$$= \mathbb{E}[\ell(X_t, Y_{I_t,t})]$$

Theorem: Under REWMA, with $\beta = \sqrt{\frac{8\log K}{T}}$

$$\mathbb{E}[R_T] \leq \sqrt{\frac{T\log K}{2}}$$

Note: $R_T = L_T - \left(\min_i \sum_1^T \ell(X_t, Y_{i,t})\right)$

Also note that in REWMA, the losses of the experts are still deterministic

The losses at every instant are mutually independent (not identically distributed) because of the randomization step

$\ell_{1,1}$	$\ell_{1,2}$		
$\ell_{2,1}$	$\ell_{2,2}$		
		$\cdots$	
$\ell_{k,1}$	$\ell_{k,2}$		

$t \longrightarrow$

Idea: The table is filled apriori by the adversary, the only way in which the randomized step comes into picture is through the probability distribution

HOEFFDING'S INEQUALITY: If $(X_1, \ldots, X_n)$ are independent random variables over $[0, 1]$, then for $\varepsilon > 0$,

$$\mathbb{P}\left(\sum_i^n X_i - \sum_i^n \mathbb{E}[X_i] \geq \varepsilon\right) \leq e^{-2\varepsilon^2/n}$$

sum of the r.v.s

Expectation of the sum = Sum of the expectations

Say $\delta = e^{-2\varepsilon^2/n}$

$$\Rightarrow \varepsilon = \sqrt{\frac{n\log(1/\delta)}{2}}$$

Under REWMA, w.p. $\geq 1-\delta$,

$$L_T - \mathbb{E}[L_T] < \sqrt{\frac{n}{2}\log\left(\frac{1}{\delta}\right)} \quad \text{where } n = T$$

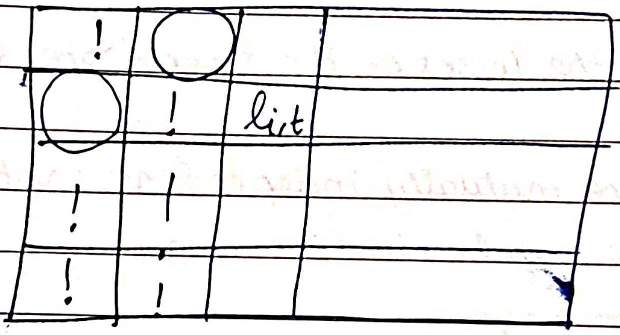
$$\therefore L_T - \mathbb{E}[L_T] < \sqrt{\frac{T}{2}\log\left(\frac{1}{\delta}\right)}$$

$$\boxed{\therefore R_T \leq \sqrt{\frac{T\log K}{2}} + \sqrt{\frac{T\log}{2}\left(\frac{1}{\delta}\right)}}$$

This intuition is similar to what one would use during invoking the Central Limit Theorem

with high probability, in REWMA, the regret is order of $\sqrt{T}$

i

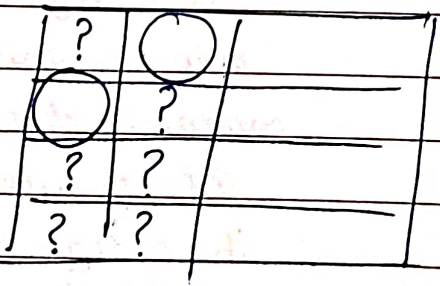


t

You come to know the performance of ALL the experts

MULTIARMED BANDITS

i



t

ONLY the CHOSEN entry is revealed

The loss benchmark is still not changed even in this case
CANNOT avoid exploration even in this case
We won't be using the idea of a "LOSS FUNCTION" from now on

Protocol :

At time $t \geq 1$

- Algo picks expert/arm A_t (has to do it randomly)

  $A_t \in \{1, 2, \dots, K\}$

- Loss $Y_{A_t, t}$ is revealed (the loss for all experts is still well-defined)

Goal minimize $E[L_T] = E\left[\sum_{1}^{T} Y_{A_t, t}\right]$ or equivalently

$$E[R_T] = E\left[\sum_{1}^{T} Y_{A_t, t}\right] - \left(\min_i \sum Y_{i,t}\right)$$

$$\underbrace{I_{t-1}}_{\text{history (this is also a random variable)}} = (A_1, Y_1, A_2, \dots, A_{t-1}, Y_{t-1})$$

$$P_{jt} = \mathbb{P}(A_t = j \mid I_{t-1})$$

Note : $Y_{j,t}$ is not known unless we choose $Y_{j,t}$'s arm

IDEA

Design a conditionally unbiased estimator

Claim : $\hat{Y}_{j,t} = \dfrac{\mathbb{1}\{A_t = j\}\, Y_t}{P_{jt}}$ is an unbiased estimator of Y_t ? (conditioned on the history)

Y_t: loss seen by algo at time t

$$= \begin{cases} \dfrac{Y_t}{P_{jt}} & , \text{if } A_t = j \\ 0 & , \text{if } A_t \neq j \end{cases}$$

$$E[\hat{Y}_{j,t}, I_{t-1}] = Y_{j,t}$$