Aug, 3, 2020

Exp3-IX Algorithm

- let $y_{ti} = 1 - x_{ti}$ (recall $x_{ti} \in [0,1]$)
- recall that ① $E_{t-1}[\hat{X}_{ti}] = x_{ti}$

$$V_{t-1}[\hat{X}_{ti}] = \frac{x_{ti}^2 (1-P_{ti})}{P_{ti}}$$

- problematic when P_{ti} small, $x_{ti} \geq c > 0$
- fix: can choose a "modified estimator"

$$\hat{X}_{ti} = 1 - \frac{\mathbb{1}\{A_t = i\}}{P_{ti}}(1 - X_t)$$

$$E_{t-1}[\underbrace{\hat{Y}_{ti}}_{\text{loss}}] = y_{ti}, \quad V_t[\hat{Y}_{ti}] = y_{ti}^2 \frac{(1-P_{ti})}{P_{ti}}$$

- proof of Exp 3 considers following def^n's

$$\hat{S}_{ni} = \sum_{s=1}^{n} \hat{X}_{si} \qquad \left(\hat{X}_{si} = \frac{A_{si}}{P_{ti}} x_{ti}\right)$$

$$\hat{S}_n = \sum_s \sum_i P_{si} \hat{X}_{si}$$

and also,

$$\hat{S}_{ni} - \hat{S}_n \leq \frac{\log k}{\eta} + \eta \sum_{t=1}^{n} \sum_{j=1}^{k} P_{tj} \hat{X}_{tj}^2$$

observe that $E[\hat{S}_{ni} - \hat{S}_n] = R_{ni} = \sum_t x_{ti} - E[\sum_t X_t]$

- instead, if we define estimators in terms of losses, we have

$$\hat{L}_n - \hat{L}_{ni} \leq \frac{\log k}{\eta} + \frac{\eta}{2} \sum_{j=1}^{k} \hat{L}_{nj}$$

where, $\hat{L}_{ni} = \sum_{t=1}^{n} \hat{Y}_{ti}$, $\hat{L}_n = \sum_{t=1}^{n} \sum_{j=1}^{k} P_{tj} \hat{Y}_{tj}$

- next ingredient, losses for fixed action,

$$\tilde{L}_n = \sum_{t=1}^{n} y_{tA_t}, \quad L_{ni} = \sum_{t=1}^{n} y_{ti}$$

let the random regret, $\hat{R}_{ni} = \sum_{t=1}^{n} x_{ti} - \sum_{t=1}^{n} X_t$

$$= \tilde{L}_n - L_{ni}$$

thus,

$$\hat{R}_{ni} = \tilde{L}_n - L_{ni} = (\tilde{L}_n - \hat{L}_n) + (\hat{L}_n - \hat{L}_{ni}) + (\hat{L}_{ni} - L_{ni})$$

$$\leq \frac{\log k}{\eta} + \frac{\eta}{2} \sum_{j=1}^{k} \hat{L}_{nj} + \underbrace{(\tilde{L}_n - \hat{L}_n)} + \underbrace{(\hat{L}_n - L_{ni})}$$

need to bound these

- pick biased estimator,

$$\hat{Y}_{ti} = \frac{\mathbb{1}\{A_t = i\} Y_t}{P_{ti} + \gamma}, \quad \gamma > 0$$

a. 'bias-variance' trade off

b. need to pick optimal γ

c. γ leads to implicit exploration (IX)

Algo : Exp3-IX

- Input : n, k, η, γ
- set $\hat{L}_{0i} = 0 \quad \forall i$
- for $t = 1, \dots, n$

  - $P_{ti} = \dfrac{\exp(\eta \hat{L}_{t-1,i})}{\sum_j \exp(\eta \hat{L}_{t-1,j})}$

  - sample $A_t \sim P_t$, observe X_t

  - $\hat{L}_{t,i} = \hat{L}_{t-1,i} + \dfrac{\mathbb{1}\{A_t = i\}(1 - X_t)}{P_{ti} + \gamma}$

Thm: let $\delta \in (0,1)$. define

$$\eta_1 = \sqrt{\frac{2\log(k+1)}{nk}}, \qquad \eta_2 = \sqrt{\frac{\log k + \log\left(\frac{k+1}{\delta}\right)}{nk}}$$

1. if $\eta = \eta_1$ and $\gamma = \eta/2$, then Exp3-IX ~~satisfies~~ satisfies following with probability at most δ,

$$\hat{R}_n \geq \sqrt{8nk\log(k+1)} + \log(1/\delta)\sqrt{\frac{nk}{2\log(k+1)}} + \log\left(\frac{k+1}{\delta}\right)$$

2. if $\eta = \eta_2$ and $\gamma = \eta/2$, then with prob $\leq \delta$,

$$\hat{R}_n \geq 2\sqrt{(2\log(k+1) + \log(1/\delta))nk} + \log\left(\frac{k+1}{\delta}\right)$$

Scanned with CamScanner