The Dan & Bobby Lecture Series

Day 2

The EXP3 Algorithm

Multi-Armed Bandit: Review.

n experts/slot machines, T time steps

at time t, choose x_t & discover $c_t(x_t)$

but do not know $c_t(x)$, $x \neq x_t$.

PSim (Phased Simulation)

p phases, of Length $L = \frac{T}{p}$.

randomly choose n of L steps for EXPLORE

$\longrightarrow$ build $\hat{c}_j(x)$

For remaining $L-n$ steps, sample x_t from BEX using $\hat{c}_1, \ldots, \hat{c}_{t-1}$

$$\hat{R}(\mathrm{PSim}, T) = O\left(\sqrt[3]{\frac{n \log n}{T}}\right)$$

The EXP3 algorithm (Exponential Exploration & Exploitation).

We want to maximize the exploit steps from PSim, so we want $L = 1$.

Let's define $\Pr[x_t] = \hat{p}_t(x)$

$$x_t = \begin{cases} \text{sample uniformly w/ Pr } \gamma \\ \text{from BEX w/ Pr } 1-\gamma. \end{cases}$$

So, $\hat{p}_t(x) = \gamma \cdot \frac{1}{n} + (1-\gamma) p_t(x)$.

(Note, every step is a phase)

Let EXPLORE be $\{t \mid \text{choose uniform sampling}\}$

$$\text{EXPLOIT} = [T] \setminus \text{EXPLORE}.$$

Let $$\hat{c}_t(x) = \begin{cases} c_t(x)/\hat{p}_t(x), & x = x_t \\ 0 \end{cases}$$

$c_t(x) < 1, \hat{p}_t(x) > \frac{\gamma}{n}$
$\Rightarrow \hat{c}_t(x) < \frac{n}{\gamma}$

__Lemma 6__ $\mathbb{E}\left[\hat{c}_t \mid \mathcal{F}_{<t}\right] = c_t$.

__Proof:__ $= \sum_x \hat{c}_t(x)\,\hat{p}_t(x) \stackrel{\text{def}}{=} c_t(x)$.

__Lemma 7.__ $\mathbb{E}\left[\hat{c}_t(x_t) \mid \mathcal{F}_{<t}\right]$

$$= \mathbb{E}\left[c_t(x_t) \mid \mathcal{F}_{<t}\right].$$

__Proof.__ $= \sum_x \mathbb{E}\left[\hat{c}_t(x) \mid \mathcal{F}_{<t}\right] p_t(x) =$

$$= \sum_x c_t(x)\,p_t(x) = \mathbb{E}\left[c_t(x_t) \mid \mathcal{F}_{<t}\right]$$

Lemma 8. For all x in $[n]$,

$$\frac{1}{T}\mathbb{E}\left[\sum_{t\in EXPLOIT}\left(c_t(x_t)-c_t(x)\right)\right] < \frac{\varepsilon n}{\delta} + \frac{n\log n}{\varepsilon\delta T}.$$

By Lemmas 6 & 7,

$$= \frac{1}{T}\mathbb{E}\left(\sum_{t\in EXPLOIT}\hat{c}_t(x_t)-\hat{c}_t(x)\right)$$

& multiply previous proof by factor of $\frac{n}{\delta}$.

Lemma ?: $\frac{1}{T}\mathbb{E}\left[\sum_{t \in EXPLORE}\left(c_t(v_t) - c_t(v)\right)\right] < \gamma.$

$$\leq \frac{1}{T} \cdot \mathbb{E}\left[\,|EXPLORE|\,\right]$$

$$= \frac{1}{T} \cdot \gamma T = \gamma.$$

__Theorem 10.__ Let $\gamma = 4\sqrt{\frac{n \log n}{T}}$ and $\xi = \gamma^2$.

Then $\hat{R}(\mathrm{EXP3}, T) = O(\gamma)$.

Pretty good, but we can do better!

Lemma 11. If $0<\alpha<1$ and $\hat{c}_t : [n] \to \mathbb{R}_+$.

$$\sum_{t=1}^{T} \sum_{x\in[n]} p_t(x)\hat{c}_t(x) \leq$$

$$\max_{x\in[n]} \sum_{t=1}^{T} \hat{c}_t(x) - \log(1-\alpha) \sum_{t=1}^{T} \sum_{x\in[n]} p_t(x)\hat{c}_t(x)^2$$

$$+ \frac{\log(n)}{\alpha}.$$

Proof.

First, we want to bound $(1-\alpha)^x$

(1) $$\underbrace{(1-\alpha)^x}_{\substack{\| \\ f(x)}} \leq \underbrace{1 + \log(1-\alpha)x + \log^2(1-\alpha)x^2}_{g(x)}.$$

$$f(0) = g(0) = 1$$

$$f'(0) = g'(0) = \log(1-\alpha)$$

$$f''(0) = \log^2(1-\alpha) < 2\log^2(1-\alpha) = g''(0)$$

$$f''(x) < f''(0) < g''(0) = g''(x).$$

Define $h(x) = g(x) - f(x)$.

$h(0) = h'(0) = 0$

$h''(x) > 0 \implies f(x) \leq g(x)$ for all x, Hedge says:

$$w_t(x) = (1-\alpha)^{\sum_{s=1}^{t} \hat{c}_s(x)}$$

&

$$W_t = \sum_x w_t(x)$$

$$P_t(x) = \frac{w_{t-1}(x)}{W_{t-1}}.$$

$$\frac{W_t}{W_{t-1}} = \sum_x p_t(x)\,(1-\alpha)^{\hat{\ell}_t(x)}$$

$$= \sum_x \frac{w_{t-1}(x)}{W_{t-1}}\,(1-\alpha)^{\hat{\ell}_t(x)}$$

$$= \frac{1}{W_{t-1}} \sum_x \left((1-\alpha)^{\sum_{s=1}^{t} \hat{\ell}_s(x)} \cdot \cancel{(1-\alpha)^{\hat{\ell}_t(x)}} \right)$$

$$= \frac{W_t}{W_{t-1}}$$

$\frac{w_t}{w_{t-1}} = \sum_x p_t(x)(1-\alpha)^{\hat{c}_t(x)}$

$\log(1+b) \leq b$ if $b > -1$

$\leq 1 + \log(1-\alpha)\sum_x p_t(x)\hat{c}_t(x) + \log^2(1-\alpha)\sum_x p_t(x)\hat{c}_t(x)^2$

$|\log(1-\alpha)\hat{c}_t(x)| < 1$ or $|\log(1-\alpha)\hat{c}_t(x)| \geq 1$

$$\log(\overline{W}_t) - \log(\overline{W}_{t-1}) \leq \log(1-\alpha)\sum_x p_t(x)\hat{c}_t(x) + \log^2(1-\alpha)\sum_x p_t(x)\hat{c}_t^2(x)$$

$$\log(W_T) - \log(\underset{\underset{n}{\parallel}}{W_0}) \le \log(1-\alpha) \sum_{t,x} p_t(x)\hat{\ell}_t(x)$$

$$+ \log^2(1-\alpha) \sum_{t,x} p_t(x)\hat{\ell}_t^2(x)$$

ALSO

$$\log(W_T) \ge \log(1-\alpha) \sum_{t=1}^{T} \hat{\ell}_t(x)$$

$$W_T \ge (1-\alpha)^{\sum_{t=1}^{T} \hat{\ell}_t}$$

$$0 < \alpha < 1$$

$$\frac{-1}{\log(1-\alpha)} \leq \frac{1}{\alpha}$$

$$-1 \geq \frac{\log(1-\alpha)}{\alpha}$$

Lemma 12

If $0 < \gamma \leq 1/2$

AND Hedge(γ/n) is used

$\Rightarrow \alpha = \gamma/n$

$$\frac{1}{T}\mathbb{E}\left[\sum_{t \in I_{\text{Exploit}}} \ell_t(x_t) - \ell_t(x)\right] \leq (4\ln 2)\gamma T + \frac{n \log n}{\gamma}$$

$$\frac{1}{T}\mathbb{E}\left[\sum_{\text{Exploit}} \ell_t(x_t) - \ell_t(x)\right] \le$$

$$-\log\left(1-\frac{\gamma}{n}\right)\mathbb{E}\left[\sum_{t,x} p_t(x)\,\hat{\ell}_t^{\,2}(x)\right] + \frac{n\log n}{\gamma}$$

$$\alpha\ \frac{\gamma}{n} \le \frac{1}{2} \Rightarrow -\log\left(1-\gamma/n\right) \le \left[2\ln(2)\right]\frac{\gamma}{n}$$

$$p_t(x) = \frac{1}{1-\gamma}\left(\hat{p}_t(x) - \frac{\gamma}{n}\right) < 2\hat{p}_t(x)$$

So

$$p_t(x)\,\hat{\ell}_t(x) \le 2\hat{p}_t(x)\,\hat{\ell}_t(x) \le 2$$

$$-\log\left(1-\frac{\gamma}{n}\right)\mathbb{E}\left[\sum_{x,t} P_t(x)\,\hat{c}_t^2(x)\right]$$

$$< (4\ln 2)\frac{\gamma}{n}\sum_{x,t}\mathbb{E}\left[\hat{c}_t(x)\right]$$

$$= (4\ln 2)\frac{\gamma}{n}\sum_{x,t} c_t(x)$$

$$\le 4\ln(2)\frac{\gamma}{n}(nT)$$

$$= (4\ln 2)\gamma T$$

THM 13

Regret of Exp3 is $O(\sqrt{T n \log n})$

lets pick $\gamma = \sqrt{\frac{n \log n}{T}}$

then we see thm 13 true

Example / Demo

or Something

(for those of you who cared enough to show up)