CSE 736

Online Learning!

Atri Rudra

Problem 1': n experts

Algo's task: $t \geq 1, \ldots$ pick $x_t \in [n]$

$\rightarrow$ incur a cost $c_t(x) \in [0,1]$ for $x \in [n]$

Goal: $\forall T$, minimize $\sum_{t=1}^{T} c_t(x_t)$

Hedge(ε):

$$\mathbb{E}\left[\sum_{t=1}^{T} c_t(x_t)\right] \leq \frac{1}{1-\varepsilon} \cdot \mathbb{E}\left[\min_{x \in [n]} \sum_{t=1}^{T} c_t(x)\right] + \frac{1}{\varepsilon} \cdot \ln n$$

$\frac{1}{1-\varepsilon} < 1 + 2\varepsilon \quad (\varepsilon < \tfrac{1}{2})$

$v_t \in [1, h]$

Problem 2: Selling a digital good

For $t \geq 1 \ldots$

$\rightarrow$ A buyer bids b_t

unique

$\rightarrow$ Propose a price p_t

$\rightarrow$ If $p_t > b_t$, $r_t = 0$

o/w $r_t = p_t$

Buyers are **rational**

Max utility

$u_t = \max(v_t - p_t, 0)$

Goals: (i) Max $\sum_{t=1}^{T} r_t$

(ii) Truthful algo $(b_t = v_t)$

correct value

TODAY:

$c_t(x) = 1 - g_t(x) \equiv g_t(x) \in [0,1]$

$\max \sum_{t=1}^{T} g_t(x_t)$

(i) Max Hedge (ε) $\leftarrow$ max version

(ii) Max Hedge(ε) solve problem 2.

Read Bobby's remarks in W2 notes

(iii) Notion of regret.

(iv) Formal definition of a game, Nash Equilibria

(v) 2 player zero-sum games, von Neumann's theorem

(Pf via Hedge algo)

MaxHedge (ε)

(i) $w_x \leftarrow 1, \quad x \in [n]$

(ii) for $t = 1, 2, \ldots$

(a) $p_t(x) = w_x / \sum_{y=1}^{n} w_y \quad \forall x \in [n]$

(b) Pick $x_t \leftarrow x$ w.p. $p_t(x)$

(c) Observe g_t

(d) $w_x \leftarrow w_x \cdot (1+\varepsilon)^{g_t(x)}$

$$\mathbb{E}\left[\sum_{t=1}^{T} g_t(x_t)\right] \geq (1-\varepsilon) \cdot \mathbb{E}\left[\max_{x \in [n]} \sum_{t=1}^{T} g_t(x)\right] - \frac{1}{\varepsilon} \cdot \ln n$$

Pf: "Same as before"

MaxHedge to solve Problem 2.

Assume $v_t \in [h]$

→ $n = h$.

→ Expert $x \in [h]$ propose price of x

→ $g_t(x) = \frac{1}{h} \cdot \begin{cases} x & \text{if } x \le b_t \\ 0 \end{cases}$

Q (i) Revenue guarantee?

(ii) Why truthful?

Q (ii) Truthfulness

Main Obs: Price p_t determined by Hedge does not depend on b_t.

Utility $u_t = \max(v_t - p_t, 0)$ ← independent of b_t.

(as every buyer is unique)

$\Rightarrow$ Truthful!

Q (i) Revenue

$$\mathbb{E}\left[\sum_{t=1}^{T} r_t\right] = \mathbb{E}\left[h \cdot \sum_{t=1}^{T} g_t(x_t)\right]$$

$$\sum_{t=1}^{T} r_t = h \cdot g_t(x_t) > (1-\varepsilon)\cdot \underbrace{\mathbb{E}\left[\max_{x\in[h]} h \cdot \sum_{t=1}^{T} g_t(x)\right]}_{\text{best fixed price revenue}} - \frac{h}{\varepsilon}\ln(h)$$

Optimal?

$\Omega(h)$ lower bound on the additive factor

[Blum, Kumar, R., Wu 03]

b, hence the algo

Hedge

$1 \to \frac{1}{h}$ $\quad v_1 = h$

$2 \quad \cdot$

$\vdots$

$h \to \frac{1}{h}$

Improving $O\left(\frac{h}{\varepsilon}\ln h\right) \to O\left(\frac{h}{\varepsilon^2}\ln\ln(h)\right)$

$\to O(h)$ [Blum, Hartline 04]

Currenty: set of ~~experts~~ $\{1, 2, \ldots, h\}$

Let $0 < \alpha < 1$.

$S = \{1, (1+\alpha), (1+\alpha)^2, \ldots\ldots, (1+\alpha)^{\lfloor \log_{1+\alpha} h \rfloor}\}$

$n = \left\lfloor \frac{\ln h}{\ln 1+\alpha} \right\rfloor + 1 \leq O\left(\frac{\ln h}{\alpha}\right)$ for small δ $h(1+\delta)$ $\sim \delta$

$$\mathbb{E}[\text{Hedge}(\varepsilon)\text{ revenue}] > (1-\varepsilon)\,\mathbb{E}\Big[\max_{x\in S}\sum_{t=1}^{T} g_t(x)\Big] - \underbrace{\frac{h}{\varepsilon}\cdot\ln\Big(\frac{\ln h}{\alpha}\Big)}_{O\left(\frac{h\ln\ln h}{\varepsilon^2}\right)}$$

$$\mathbb{E}\Big[\max_{x\in S}\sum_{t=1}^{T} g_t(x)\Big] \geq \frac{1}{1+\alpha}\cdot\mathbb{E}\Big[\max_{x\in[h]}\sum_{t=1}^{T} g_t(x)\Big]$$

$$\alpha = \Theta(\varepsilon)$$

$$\Rightarrow \forall\ \varepsilon > 0$$

$$\mathbb{E}[\text{Revenue}] > (1-\varepsilon)\cdot\mathbb{E}[\text{best fixed price rev.}] - O\Big(\frac{h\ln\ln h}{\varepsilon^2}\Big)$$

$$\underbrace{\mathbb{E}\left[\max_{x\in S}(\ldots)\right]}_{L} \geq \frac{1}{1+\alpha}\cdot \mathbb{E}\left[\max_{x\in[n]}(\ldots)\right]$$

Pf: Main idea: Let x^* be the optimal price

Let $\bar{x}$ be the largest price in S

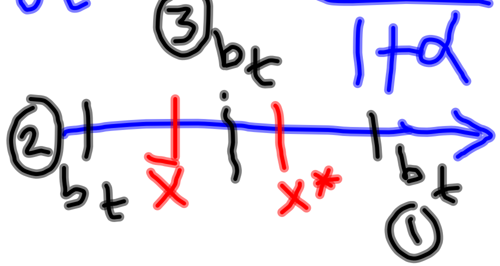
$$g_t(x) = \frac{1}{h}\cdot\begin{cases} x & x\le b_t\\ 0 & o/w\end{cases}$$

$$\frac{x^*}{1+\alpha}\le \bar{x}\le x^* \leftarrow$$

Pf: $\bar{x} < \frac{x^*}{1+\alpha}$

$(1+\alpha)\cdot\bar{x} < x^*$

$$g_t(\bar{x}) \geq \frac{g_t(x^*)}{1+\alpha}$$



$$L \geq \mathbb{E}\left[\sum_{t=1}^{T} g_t(\bar{x})\right]$$

$$\geq \frac{1}{1+\alpha}\,\mathbb{E}\left[\sum_{t=1}^{T} g_t(x^*)\right]$$

$\frac{1}{1-\varepsilon} < 1+2\varepsilon$ for $\varepsilon < \frac{1}{2}$

(Back to Bobby's notes)

REGRET: (Informally) Regret of an online Algo A is $\mathbb{E}[\text{A's value} - \text{OPT value}]$

$(0<\varepsilon<\frac{1}{2})$

$$\mathbb{E}[\text{Cost}(\text{Hedge}(\varepsilon))] < (1+2\varepsilon)\,\mathbb{E}[\text{Cost}(\text{OPT})] + \frac{1}{\varepsilon}\cdot \ln n$$

$$\Rightarrow \text{Regret}(\text{Hedge}) < (2\varepsilon)\cdot \underbrace{\mathbb{E}[\text{Cost}(\text{OPT})]}_{\text{Upper bound?}} + \frac{1}{\varepsilon}\cdot \ln n$$

Regret(Hedge) $\leq o(T)$

$\to$ Sequences up to time T.

$\forall \varepsilon < \frac{1}{2}$

~~$\Rightarrow \text{Cost}(OPT) \leq T.$~~ **Next week: Doubling!**

$$\Rightarrow \text{Regret}(\text{Hedge}^T(\varepsilon)) \leq 2\varepsilon \cdot T + \frac{1}{\varepsilon} \cdot \ln n$$

What if T is not known?

$$\left(\varepsilon = \sqrt{\frac{\ln n}{2T}}\right)$$

T large enough

$$\leq 2\sqrt{2T \cdot \ln(n)}$$

$ax + \frac{b}{x}$ $\quad 2\sqrt{ba}$

Min if x s.t.

$ax = \frac{b}{x} \Leftarrow x = \sqrt{\frac{b}{a}}$

average regret $= \lim_{T\to\infty} \frac{\text{Regret}}{T} \leq 2\sqrt{\frac{2\ln n}{T}} \to 0$