Lecture ~~16~~ 17: Online Decision Making II

Difficulties not (just) computational

but lack of information

Decision Making in the face of Uncertainty

Approx + Lack of Info

find Alg st $\left(\max_{I} \frac{Alg(I)}{OPT(I)} \right) \leq$ small

cost

competitive ratio

Today

Online Algos
- Sleator + Tarjan
- competitiveness

Next

Online Learning
- Blackwell & Hannan
- regret

Mistake Bound Model / Experts Problem

$\forall$ day $t = 1, 2 \ldots T \ldots$

a_t $\begin{cases} \circ \text{ listen to } n \text{ "experts" predictions} \\ \circ \text{ make a prediction} \end{cases}$ (Y/N) "is the weather good" "is the market going up"

o_t $\circ$ after that, hear the actual outcome

$$\text{Mistake}_t = \text{"} a_t \neq o_t \text{"}$$

$$\boxed{\text{Minimize \#mistakes}}$$

① Sps a perfect expert.

$\Rightarrow \exists$ algo that makes $\leq \lg_2 n$ mistakes

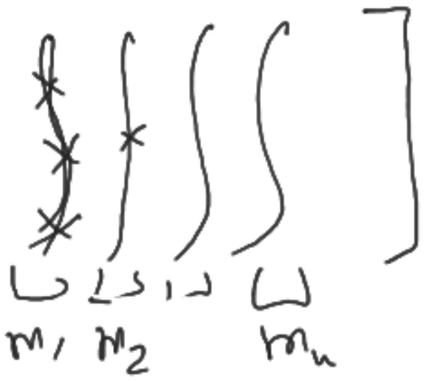
Predict the majority opinion, discard all incorrect experts.

every mistake $\Rightarrow$ throw away majority of existing experts

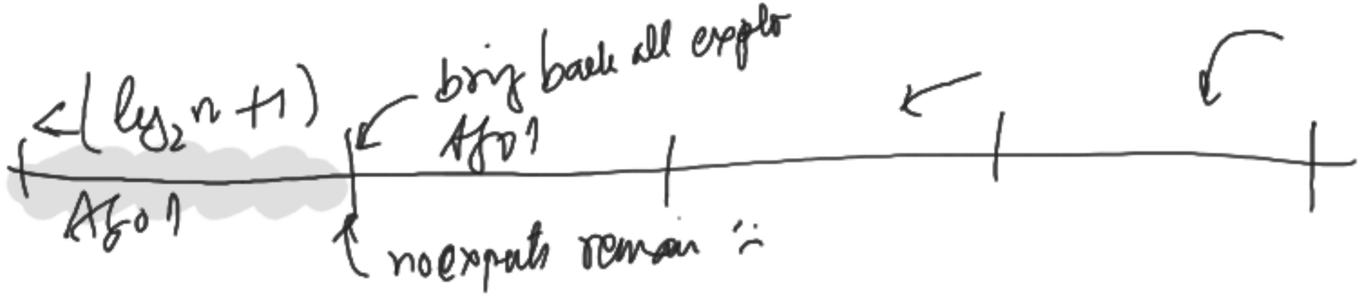
$\Rightarrow \leq \lg_2 n$ mistakes

② Sps. the best expert makes $\underline{\underline{m^*}}$ mistakes

$\Rightarrow \exists$ algo that makes $\leq (m^*+1)(\lg_2 n + 1)$ mistakes.



$m^* = \min_i m_i$



$\text{Algo} \leq (\#\text{rounds})(\lg_2 n + 1)$

$m^* \geq (\#\text{rounds} - 1)$ mistakes

$\forall$ expert i, $\quad$ Algo's mistakes $\leq \overset{\text{mult}}{\underset{??}{(\lg_2 n + 1)}}$ (Expert i is mistakes) $+ \overset{\text{add.}}{\underset{\text{necc.}}{(\lg_2 n)}}$

Margin Hence

Thm $\quad \forall$ expert i, $\quad$ Algo's mistakes $\leq (1+\varepsilon)$ (Expert i is mistakes) $+ \left(\frac{\lg_2 n}{\varepsilon}\right)$

$w_i^{(1)} = 1 \quad \forall$ expert i

for $t = 1, 2, 3 \ldots T \ldots$

- Predict the weighted majority

$$\begin{cases} \text{if } \sum_{i:\, \text{predict}(i,t) = Y} w_i^t \geq \sum_{i:\, \text{predict}(i,t) = N} w_i^t \\ \text{predict } Y, \text{ else } N. \end{cases}$$

- $$w_i^{t+1} = \begin{cases} w_i^t & i \text{ correct} \\ w_i^t \cdot (1/2) \end{cases}$$

Weighted Majority (WM)
Multiplicative Weights (MW)

Thm: $\forall i$ Mistakes of WM $\leq \underbrace{2.4}$ (Mistakes of i) $+ \underbrace{O(\log n)}$

Pf: "potential" $\Phi_t = \sum_i w_i^{(t)}$ $\qquad \Phi_0 = n$

sps 9 make no mistake in t $\quad \Phi^{t+1} \leq \Phi^t$

9 make mistakes $\quad \Phi^{t+1} \leq \left(\frac{3}{4}\right) \Phi^t$

$$\underline{\underline{\left(\frac{1}{2}\right)^{m_i}}} \leq \underline{\underline{\Phi^T}} \leq \left(\frac{3}{4}\right)^{\text{my mistakes}} \cdot \Phi^0 \leq \underline{\underline{n \cdot \left(\frac{3}{4}\right)^{\text{most}}}}$$

$$\left(\frac{1}{2}\right)^{m_i} \leq n \left(\frac{3}{4}\right)^{Ag} \Rightarrow \left(\frac{4}{3}\right)^{Ag} \leq n \cdot 2^{m_i}.$$

take logs $\Rightarrow Ag\left(\log_2 \frac{4}{3}\right) \leq \log n + m_i$ ☺

① Basic MW/WM. with halving gives $\forall i$ $A_j \le \underline{\underline{2.4}}(m_i) + O(\log n)$

② $w_i^{t+1} \leftarrow w_i^t (1-\varepsilon)$ $\le (2+\varepsilon)(m_i) + O\left(\frac{\log n}{\varepsilon}\right)$

<u>Bad news</u>: no deterministic algo can do better than $2 m_i$

<u>Pf</u> 2 experts Always YES / Always NO.

$\text{outcome}_t \neq \text{alg}_t$ $\Rightarrow$ Algo makes T mistakes.

better of these experts $\le T/2$ mistakes.

Model (Refined)

$\forall$ day t,

simultaneously
- algo predicts a_t [randomize OK]
- outcome is revealed O_t

mistake = "$a_t \neq O_t$"

"learning rate"

Thm $\forall i$ $\mathbb{E}[\text{Mistakes RWM}] \leq (1+\epsilon)\, m_i + \frac{\lg n}{\epsilon}$

Rand WM : [Littlestone/Warmuth]

- start with weights $w_i^1 = 1$
- at time t,

$$\begin{cases} \text{Predict } Y \text{ w.p. } \dfrac{\sum_{i:\, \text{predict } Y} w_i^t}{\sum_i w_i^t} \\ N \text{ otherwise} \end{cases}$$

- $$w_i^{t+1} = \begin{cases} w_i^t \cdot (1-\epsilon) & \text{if } i \text{ made mistake} \\ w_i^t & \text{otherwise} \end{cases}$$

exactly same as before

Extension 1: Change of Perspective

$$\frac{w_i^t}{\sum_j w_j^t} = p_i^t = \text{"prob of listening to expert } i \text{ @ time } t\text{"}$$

at each timestep t,

Simultaneously

- Algorithm produces a vector $p^t \in \Delta^n$ ← prob. simplex in $\mathbb{R}^n$ $= \{x \in [0,1]^n \mid \sum_i x_i = 1\}$
- Nature produces a "loss" vector $\ell^t \in [0,1]^n$

My loss cost @ time t $= \langle \ell^t, p^t \rangle \overset{\text{def}}{=} \sum_i \ell_i^t p_i^t$

Want: My loss over time $\sum_t \langle \ell^t, p^t \rangle$ $\leq$ loss of best vector $q^* \in \Delta^n$

$$\leq (1+\varepsilon) \min_{q^* \in \Delta^n} \sum_t \langle \ell^t, q^* \rangle + \frac{\log n}{\varepsilon}$$

Ex: Check that linear function model (dot product model) captures experts model.

Extension 2: "Bandits" instead of "experts" / partial info model.

- In previous models, see entire loss vector.
- now suppose I only see the loss value $\langle \ell^t, p^t \rangle$
  - or in RWM model, picked a random expert, get see its loss.
- Good news: Experts algos extend to bandits

$$\forall q^* \quad E[\text{loss of algo in bandit setting}] \le (1+\varepsilon) \sum_t \langle \ell^t, q^* \rangle + \underbrace{O(n/\varepsilon)}_{\text{bad news}}$$

explore vs exploit