Lecture 120: Online Learning/Prediction Part 2: Randomization & Generalizations

Recall: Game: Each day $t = 1, 2, \ldots, T$
- experts $1 \ldots N$ predict $e^t_1, \ldots, e^t_N \in \{0,1\}$
- you take action $a^t \in \{0,1\}$
- outcome $o^t$ occurs. Mistake: $a^t \neq o^t$ (you)
  $e^t \neq o^t$ (expert)

"Weighted Majority Alg":
- Initialize weights $w^t_1, w^t_2, \ldots, w^t_N = 1$.
- $a^t = \text{weighted \ major \ of } e^t_1, \ldots, e^t_N$ \text{ with } \text{ } w^t_i \text{ weights}
- $w^{t+1}_i = \begin{cases} w^t_i & \text{if } e^t_i = o^t \\ (1-\epsilon)w^t_i & \text{if } e^t_i \neq o^t. \end{cases}$ (old eq. was $\epsilon = 3$)

Analysis: #Your Mistakes $\leq (2 + \frac{5}{\epsilon})(\text{#Best experts mistakes}) + \frac{2}{\epsilon} \ln N$

Fact: For any deterministic strat, can't beat factor 2.
Because if $\begin{array}{c|c|c} e^t_1 & e^t_2 \\ \hline 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ \vdots \end{array}$, you do some fixed $a$, and $o$ might be $\begin{array}{c|c|c|c|c|c} a^0 & a^1 & a^2 & a^3 & a^4 \\ \hline a^0 & -a^1 & -a^2 & -a^3 & -a^4 \end{array}$

Rem.: If you do W.M. alg, your $a^t$'s alternate, $o^t$'s alternate, experts are "equally bad", $w^t \propto w^t \forall t$.

Idea: Randomize! \[ \text{If it's a "close call" on what to predict, flip a coin!} \]
Random model: As usual, if the "outcome-revealing adversary" can see what action you play, it's just like the deterministic case.

"Adversary" picking outcomes can:
- know your randomized strategy
- see experts predictions
- can't: see algorithm's "coin flips"
  - see algorithm's actions.

Actually: it'll be okay if Adv. can see $a_1^1, \ldots, a_{t-1}$ before deciding on $o^t$; but never mind for now.

So can:
- fix randomized alg
- then look at worst possible table
- then compute $E[\# \text{your mistakes}]$.

What's most natural randomized alg? Usually,

Randomized Weighted Majority

At time $t$, if total weight is $\Phi^t = w_1^t + w_2^t + \cdots + w_N^t$,

say total weight of experts predicting 1 is $q \cdot \Phi^t$.

Then choose $a^t = 1$ with prob $q$, // old rule was $1$ if $g^t \geq t$

$0$ if $g^t < t$.

Still penalize wrong experts weight by $1$-$e$ factor.

Analysis: Fix $r \text{ expert predictions} & \text{ outcomes For all time & worst case.}$

Play Randomized Weighted Maj, define random variables

$M^t = \begin{cases} 1 & \text{if } a^t \neq o^t \text{ [mistake by you at time } t] \\ 0 & \text{if } a^t = o^t \end{cases}$

$t = 1, \ldots, T.$
\[ E[\text{your mistakes}] = E[M_1 + M_2 + \ldots + M_T] \]
\[ = E[M_1] + \ldots + E[M_T] \]  \[\text{[linearity of expectation]}\]
\[ = f_1 + \ldots + f_T, \quad f_t = \Pr[\text{you mistake at time } t] \]

But \( f_t \) has another interpretation.

By def. of the alg., \( f_t = \frac{\text{fraction of } \Phi^t \text{ coming from experts making mistake on day } t.}{\Phi^t \text{ goes down by } e \cdot f_t \cdot \Phi^t} \)

\[ \Phi^{t+1} \leq (1 - e \cdot f_t) \Phi^t \]

\[ 1 - e f_t \approx e^{-e f_t} \]
\[ 1 + x \approx e^x \]
In fact, \( 1 - x \approx e^{-x} \)

\[ \Rightarrow \Phi^{t+1} \leq e^{-e f_t} \Phi^t \]

\[ \Phi^{t+1} \leq \frac{\Phi^{t+1}}{\Phi^t} \leq \ldots \leq \frac{\Phi^1}{\Phi^t} = N e^{-e f_1 - e f_2 - \ldots - e f_T} \]

Still have \( \Phi^{t+1} \leq W_{t+1} \)

\[ \Phi^{t+1} \leq N e^{-e f_1 - e f_2 - \ldots - e f_T} \]

\[ \leq N e^{-e (f_1 + \ldots + f_T)} \]

\[ \leq N e^{-e} \cdot E[\text{your mistakes}] \]

\[ N e^{-e} E[\text{YM}] \geq (1 - e)^{\#BM} \]

\[ \ln N - e E[\text{YM}] \geq \ln (1 - e) \cdot \#BM \]
\[ \leq -(\#BM)(\varepsilon + \frac{1}{2} \varepsilon^2) \]
\[ E[\text{YM}] \leq (1 + \frac{e}{2}) \cdot \#BM + \ln N \]
\[ E[\text{YM}] \leq (1 + \frac{3}{2}) \cdot \#BM + \ln N \]

\[ \varepsilon E[\text{YM}] \leq (1 + \frac{3}{2}) \cdot \#BM + \frac{\ln N}{\varepsilon} \]

\[ \text{Ratio is } 1 + \frac{3}{2} \]
Now want to generalize the setup somewhat. Once we get a good alg. in generalized setup, we'll see it can be used to... solve flows, 2-player zero-sum games, LPs...

R.W.M.'s strategy is equivalent to...
- at time t, let $p_i^t = \frac{w_i}{\Phi^t}$ for $i = 1, \ldots, N$ [these form a "prob. distribution" if they sum to 1]
- play expert is prediction with prob. $p_i^t$

I could imagine Adversary's strat. is... pick an expert $i$, do their prediction or its opposite.

A weird way to look at things. Granted, I could also let adversary

New game: 
- Expert $i$ is like a slot machine
  - Alg. chooses $p_i^t, \ldots, p_i^t$ at time $t$; interpretation is "play slot machine $i$ with prob. $p_i^t$"
  - "Adversary" sets $n$ loss/reward $l_i^t, \ldots, l_i^t$ between $-1,1$
  - "Loss" at time $t$ is $p_i^t l_i^t = p_i^t l_i^t + \cdots + p_i^t l_i^t$

Goal: small loss compared with best (least) value of $\sum_{i=1}^N l_i^t$ over $1 \leq i \leq N$

"Hedge" Alg., aka "Multiplicative Weights":
- Same as R.W.M., but every weight $w_i^t$ changed each round:
  
  $w_i^{t+1} = (1 - \epsilon_i l_i^t) w_i^t$

  $1 - \epsilon$ when $i$th "loss" is $+1$, but $1 + \epsilon$ when "loss" is $-1$.

What you'd suffer if you just did 100% on expert/slot you're i* every time.
E.g., can choose $\varepsilon = \sqrt{\frac{\ln N}{T}}$

$\Rightarrow$ your avg. loss $\leq$ avg. loss of best fixed $i^*$

$$+ \frac{2\ln N}{\sqrt{T}}$$

decays to 0 as $T \to \infty$

"learning to play perfectly over time"