Online learning

Note Title 2/12/2013

There are n "experts"

Repeat
- → Algorithm picks an expert
- Each expert recieves a payoff.

Payoff(expert i, round t) $= V_{i,t} \in [0,1]$

Maximize total payoff.

Differences from online matching:

- Pick before $V_{i,t}$'s are revealed
- Only local constraint.

What is OPT?

- Say optimal choice on hindsight. What does OPT do?

Answer! OPT picks best expert in each round

- What is the worst scenario for ALG?
- Deterministic & 2 experts. ALG → 1.
  - Answer!- $V_1 = 0$, $V_2 = 1$.
- Randomized: ALG → 1 w/ p, 2 w/ q. say $p \geq q$
  - Ans:- again $V_1 = 0$ $V_2 = 1$, $E[ALG] = q \leq \frac{1}{2}$

- Randomized & n experts. ALG → $p_1, p_2, \ldots p_n$!

Ans:- $i = \arg\min_j \{p_j\}$ $V_i = 1$, $V_j = 0$ $\forall j \neq i$.

$E[ALG] \leq \frac{1}{n}$. $OPT = 1$. Can repeat this.

OPT is too powerful!

Redefine. $OPT = \max_i \left\{ \sum_{t=1}^{T} v_{i,t} \right\}$

optimal single expert in hindsight.

OPT can't change its choice for each round

Let $w_i = \sum_{t=1}^{T} v_{i,t}$. $OPT = \max_i \{w_i\}$

Smooth approx. to max

Proof based on properties of ϕ

$\phi(w_1, \dots w_n) := \frac{1}{\lambda} \log \sum_i e^{\lambda w_i}$

for some $\lambda > 0$. $\vec{w} = (w_1, \dots w_n)$

Lemma 1:- $\phi(\vec{w}) \geq OPT$

Proof. $\phi \geq \frac{1}{\lambda} \log \max_i e^{\lambda w_i} = \frac{1}{\lambda} \log e^{\lambda OPT}$

$= \frac{1}{\cancel{\lambda}} \cancel{\lambda} \cdot OPT = OPT$.

Lem 2:- $\phi \leq OPT + \frac{1}{\lambda} \log n$.

$\sum_i e^{\lambda w_i} \leq \sum_i e^{\lambda OPT} = n e^{\lambda OPT}$.

$\therefore \log\left(\sum_i e^{\lambda w_i}\right) \leq \log n + \lambda OPT$

$\therefore \phi \leq \frac{1}{\lambda} \log n + OPT$

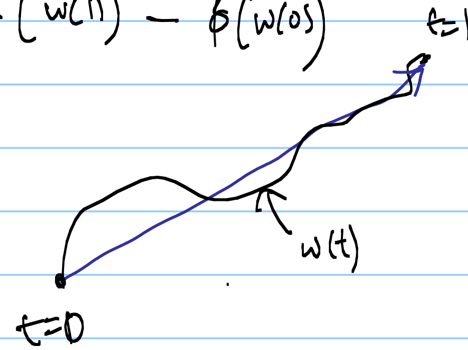
Lem 3:- $\frac{\partial \phi}{\partial w_i} = \frac{1}{\cancel{\lambda}} \frac{1}{\sum_i e^{\lambda w_i}} \cdot \cancel{\lambda} e^{\lambda w_i}$

Note:- $\sum_i \frac{\partial \phi}{\partial w_i} = \frac{\sum_i e^{\lambda w_i}}{\sum_i e^{\lambda w_i}} = 1$.

Def: $\vec{\nabla}\phi = \left(\frac{\partial \phi}{\partial w_1}, \frac{\partial \phi}{\partial w_2}, \dots, \frac{\partial \phi}{\partial w_n}\right)$.

Thm 1: $\forall$ "path" in $\mathbb{R}^n$ given by $w:[0,1]$

$$\int_{t=0}^{1} \left\langle \vec{\nabla}\phi(w(t)), \frac{d\vec{w}}{dt} \right\rangle dt = \phi(w(1)) - \phi(w(0))$$



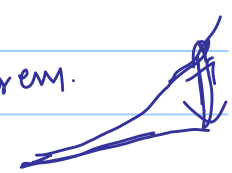
Proof: By chain rule, $\frac{d\phi}{dt} = \left\langle \vec{\nabla}\phi(w(t)), \frac{d\vec{w}}{dt} \right\rangle$

$$\int_{t=0}^{1} \frac{d\phi}{dt}\, dt = \phi(w(1)) - \phi(w(0))$$

Theorem 2: $\langle \vec{\nabla}\phi(w_0), w_1 - w_0 \rangle \geq \phi(w_1) - \phi(w_0) - \lambda \|w_1 - w_0\|_\infty$

If $w_1 \geq w_0$, co-ordinate-wise.

Discrete Analog of previous theorem.



Assume Theorem 2, define Algorithm.

$\vec{w}_0 = (0, 0, \dots, 0)$

In round $t = 1..T$

- pick expert i w.p. $\nabla_i \phi(\vec{w}_{t-1})$
- $\vec{w}_t = \vec{w}_{t-1} + \vec{v}_t$

$$E[ALG] = \sum_{t=1}^{T} \langle \nabla\phi(\vec{w}_{t-1}), \vec{v}_t \rangle$$

Theorem 3:- $E[ALG] \geq OPT - \frac{1}{\lambda}\log n - \lambda T$

$$\text{Regret} = OPT - E[ALG] \leq \frac{1}{\lambda}\log n + \lambda T$$

Proof:- $E[ALG] = \sum_{t=1}^{T} \langle \nabla\phi(\vec{w}_{t-1}), \vec{v}_t \rangle$

From Thm 2, $\geq \sum_{t=1}^{T} \phi(\vec{w}_t) - \phi(\vec{w}_{t-1}) - \lambda \|\vec{v}_t\|_\infty$

$$\geq \phi(\vec{w}) - \phi(\vec{0}) - \lambda T$$

$$\geq OPT - \frac{1}{\lambda}\log n - \lambda T$$

$\because \phi(\vec{0}) = \frac{1}{\lambda}\log \sum_i e^0 = \frac{1}{\lambda}\log n$

Pick λ to minimize regret.

$$\text{Regret} = \frac{1}{\lambda}\log n + \lambda T$$

Set $\frac{1}{\lambda}\log n = \lambda T$ i.e $\lambda^2 = \frac{\log n}{T}$

$$\lambda = \sqrt{\frac{\log n}{T}}$$

$$\text{Regret} = 2\sqrt{T\log n}$$

Proof of Theorem 2:- Say $\|w_1 - w_0\|_\infty \leq 1$.

wnt $\phi(w_1) - \phi(w_0) - \langle \nabla\phi(w_0), w_1 - w_0 \rangle \leq \lambda$

$$\Leftrightarrow \frac{1}{\lambda}\log \sum_i e^{\lambda w_{1i}} - \frac{1}{\lambda}\log \sum_i e^{\lambda w_{0i}}$$

$$\leq \lambda + \frac{\sum_i e^{\lambda w_{0i}} \cdot \Delta w_i}{\sum_i e^{\lambda w_{0i}}}$$

$$\Leftarrow \quad \log \frac{\sum_i e^{\lambda \omega_i}}{\sum_i e^{\lambda \omega_{0i}}} \leq \lambda^2 + \frac{\sum_i e^{\lambda \omega_{0i}} \cdot \lambda \Delta \omega_i}{\sum_i e^{\lambda \omega_{0i}}}$$

Let $\dfrac{e^{\lambda \omega_{0i}}}{\sum_i e^{\lambda \omega_{0i}}} =: \mu_i$, $\lambda \Delta \omega_i =: \lambda_i$

$$\Leftrightarrow \quad \log\left(\sum_i \mu_i e^{\lambda_i}\right) \leq \lambda^2 + \sum_i \mu_i \lambda_i$$

$$\Leftrightarrow \quad \sum_i \mu_i e^{\lambda_i} \leq e^{\lambda^2} \cdot e^{\sum_i \mu_i \lambda_i}$$

$\left(\sum_i \mu_i = 1\right)$

$\lambda_i \leq \lambda$

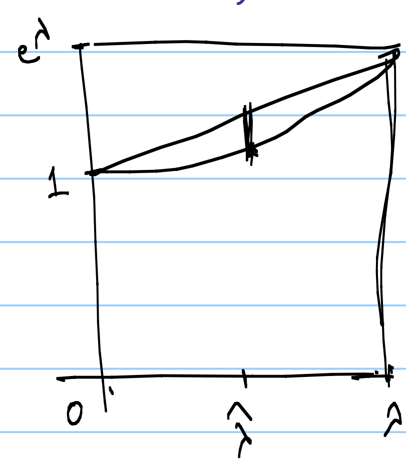
Convexity of e^x $\Rightarrow$

$$\sum_i \mu_i e^{\lambda_i} \geq e^{\sum_i \mu_i \lambda_i}$$

Convexity of $e^x \leq e^{\lambda^2}$ in $[0, \lambda]$

Claim: $\forall \; \hat{\lambda} \in [0, \lambda]$,

$$1 + \frac{e^\lambda - 1}{\lambda} \cdot \hat{\lambda} \leq e^{\lambda^2} \cdot e^{\hat{\lambda}}$$



Continue proof.

By convexity of e^x,

$$e^{\lambda_i} \leq 1 + \frac{e^\lambda - 1}{\lambda} \cdot \lambda_i$$

$$\therefore \sum_i \mu_i e^{\lambda_i} \leq \sum_i \mu_i \left[1 + \frac{e^\lambda - 1}{\lambda} \cdot \lambda_i\right]$$

$$= 1 + \frac{e^\lambda - 1}{\lambda} \hat{\lambda} \qquad \left[\hat{\lambda} = \sum_i \mu_i \lambda_i\right]$$

$$\leq e^{\lambda^2} \cdot e^{\hat{\lambda}}$$

∎

Proof of Claim: If $\lambda \geqslant 1$, $e^{\hat\lambda^2} \geqslant e^{\hat\lambda} \geqslant \frac{e^{\hat\lambda}-1}{\hat\lambda}$

Slope of $e^{\hat\lambda^2}, e^{\hat\lambda}$ is bigger.

Also at $\hat\lambda = 0$, is bigger.

If $\lambda < 1$, $e^{\lambda} < 1+\lambda+\lambda^2$.

$$\frac{e^{\lambda}-1}{\lambda} < 1+\lambda$$

$$\therefore 1+\hat\lambda\left(\frac{e^{\lambda}-1}{\lambda}\right) < 1+\hat\lambda+\lambda\hat\lambda < 1+\hat\lambda+\lambda^2$$

$$\leq e^{\lambda^2+\hat\lambda}$$

$\square$

Applications:-

- Boosting in ML.

  Learn a function $f : X \to \{-1,1\}$
  "Is this a cat?"

  Use a hypothesis class $\mathcal{H}$ of fn-s.

  $\mathcal{D}$ is a probability distribution on X.

Strong learning algorithm:- $\forall\, \mathcal{D}$, find $h \in \mathcal{H}$ s.t.

$$\Pr_{x\sim D}\left[h(x) = f(x)\right] \geqslant 1-\varepsilon$$

Weak —"— —"— find $h\in\mathcal{H}$ s.t.

—"— $\geqslant \frac{1}{2}+\delta$

Boosting : Weak $\longrightarrow$ Strong.

fn X = "training set". $\forall x \in X$, know $f(x)$

D = uniform

Ada-Boost :-

- X = Experts.
- Run MWU algo, Let D_t be the distr^n over X used by MWU
- Use WL to find h_t s.t $\Pr_{x \sim D_t}[h_t(x) = f(x)] \geq \frac{1}{2} + \delta$
- $\forall x \in X \quad V_{x,t} = \mathbb{1}\{h_t(x) \neq f(x)\}$

x managed to fool $h_t \Rightarrow$ bigger weight on x.

Repeat T times.

Strong-Learner :- $h(x) = \text{sign}\left\{ \sum_{t=1}^{T} h_t(x) \right\}$

Majority. of h_t's.

$$OPT = \max_{x \in X} \left\{ \sum_t V_{x,t} \right\} = \max_{x \in X} \left\{ \sum_t \mathbb{1}\{h_t(x) \neq f(x)\} \right\}$$

$$E[ALG] = \sum_{t=1}^{T} \Pr_{x \in D_t} [V_{x,t}] = \sum_{t=1}^{T} \Pr_{x \sim D_t} [h_t(x) \neq f(x)]$$

$$\leq \sum_{t=1}^{T} \left(\frac{1}{2} - \delta\right) = \left(\frac{1}{2} - \delta\right) T$$

$$\Phi_T - ALh \leq \lambda T + \frac{1}{\lambda} \log n$$

$$\Phi_T \leq \left(\frac{1}{2} - \delta\right) T + \lambda T + \frac{1}{\lambda} \log n$$

Suppose $h(x) \neq f(x)$. for εn x's.

Then for $\geq \frac{T}{2}$ t's, $h_t(x) \neq f(x)$

i.e $\#\{x : w_x \geq T/2\} = \varepsilon n$

$$\therefore \underset{T}{\Phi} = \frac{1}{\lambda} \log \left(\Sigma_x e^{\lambda w_x}\right) \geq \frac{1}{\lambda} \log\left[\varepsilon n \cdot e^{\frac{\lambda T}{2}}\right] \quad \frac{nT}{2}$$

$$= \frac{1}{\lambda} \log n + \frac{\log \varepsilon}{\lambda} + \frac{T}{2}$$

$$\therefore \log \varepsilon \leq (\lambda - \delta)\lambda T$$

$$\text{or } \log \frac{1}{\varepsilon} \geq (\delta - \lambda)\lambda T \qquad \lambda = \frac{\delta}{2}$$

$$= \frac{\delta^2}{4} T$$

$$T = 4 \log {}^1/_\varepsilon / \delta^2$$

]

<u>Application</u>: Playing a zero-sum game.

Game between 2 players, Max & Min.

Max has n strategies, $i = 1..n$.

Min has m strategies $j = 1..m$.

If Max plays i, & min plays j, then Min has to pay Max a_{ij} dollars.

Example :- Rock, paper, scissors.

$$\begin{array}{cc|c|c|c|} & & \multicolumn{3}{c}{\text{Min}} \\ & & R & P & S \\ \hline \text{Max} & R & 0 & -1 & +1 \\ & P & 1 & 0 & -1 \\ & S & -1 & 1 & 0 \\ \hline \end{array} = A \in \mathbb{R}^{n \times m}$$

Randomized Strategies :- Max plays i w.p p_i

Min plays j w.p q_j

$$E[\text{payment}] = \sum_{i,j} p_i q_j a_{ij} = p^T A q$$

Max's problem: $\underset{\vec{p}}{\text{Max}} \ \underset{\vec{q}}{\text{Min}} \left\{ p^T A q \right\} = \Lambda^*_{max}$

$$\overset{\shortparallel}{\underset{\vec{p}}{\text{Max}} \ \underset{j}{\text{Min}}} \left\{ \sum_i p_i a_{ij} \right\}$$

"Min knows max's strategy & optimizes accordingly."

Min's problem: $\underset{\vec{q}}{\text{Min}} \ \underset{\vec{p}}{\text{Max}} \left\{ p^T A q \right\} = \Lambda^*_{min}$

$$\shortparallel$$

$$\underset{\vec{q}}{\text{Min}} \ \underset{i}{\text{Max}} \left\{ \sum_j q_j a_{ij} \right\}.$$

Max knows min's strategy & optimizes.

Lemma :- $\Lambda^*_{min} \geq \Lambda^*_{max}$

Use expert's algo. to find optimal strategy:

- Experts = Strategies
- Repeat $t=1..T$
  - Max picks a distribution p_t over strategies
  - Min picks $j_t = \arg\min_{j'} \left\{ \sum_i p_{it} a_{ij'} \right\}$
  - Strategy / Expert i gets $v_{i,t} = a_{ij_t}$

$$ALG = \sum_{t=1}^{T} \sum_i p_{it} a_{ij_t}$$

Since $\max_p \min_j \left\{ \sum_i p_i a_{ij_t} \right\} = \lambda^*_{max}$

$\forall t \quad \sum_i p_{it} a_{ij_t} \le \lambda^* \quad \therefore ALG \le T \cdot \lambda^*_{max}$

$$OPT = \max_i \left\{ \sum_{t=1}^{T} a_{ij_t} \right\} = T \max_i \left\{ \sum_{t=1}^{T} \frac{1}{T} a_{ij_t} \right\}$$

Prob. distribution over $j = 1..m$

Since $\min_q \max_i \left\{ \sum_{j=1}^{m} q_j a_{ij} \right\} = \lambda^*_{min}$,

$OPT \ge T\lambda^*_{min}$

$OPT - ALG \le \sqrt{T \log n}$

$\therefore T\lambda^*_{min} - T\lambda^*_{max} \le \sqrt{T \log n}$

$\therefore \lambda^*_{min} - \lambda^*_{max} \le \sqrt{\frac{\log n}{T}} \to 0$ as $T \to \infty$

.D