E1 245: Online Prediction and Learning	Fall 2015
Lecture 6 — August 20	
Lecturer: Aditya Gopalan	*Scribe: Sindhu P R*

6.1 Recap

In the last lecture, we focussed on expert advice based online learning algorithms wherein the loss function was convex. We specifically looked at the Exponential Weights algorithm for convex decision space and loss functions. The algorithm enjoys a regret bound of $\frac{\eta T}{8} + \frac{\log N}{\eta}$, which was established by employing Jensen's inequality and Hoeffding's Lemma. Today, we will look at the scenario when the decision space and the loss function need not have the nice convex structure.

6.2 Regret for a Deterministic Prediction Algorithm

Let us consider $\mathscr{D} = \mathscr{Y} = \{0,1\}$ and the loss function to be $l(p,y) = \mathbb{1}\{p \neq y\}$. It can be seen that the decision space as well as the range of the loss function is a discrete space (and non-convex). Suppose we consider two experts - one who always recommends 0 and the other who always recommends 1 to the agent. Thus, $|\mathscr{E}| = 2$.

Theorem 6.1. *For $\mathscr{D}$, $\mathscr{Y}$, l and $\mathscr{E}$ as before, suppose $\mathscr{A}$ is a deterministic prediction algorithm. If the algorithm predicts the outcome for T time instants, then, $R_T(\mathscr{A}) \geq \frac{T}{2}$.*

The general idea of the proof is to construct an outcome sequence $(y_1, y_2, \ldots, y_T)$ for which the algorithm suffers a large loss and hence large regret.

Proof: The prediction of the algorithm at time t is denoted as p_t. The prediction of Expert j at time t is given by f_{jt}. Note that $p_t, f_{jt} \in \mathscr{D}, i \in \{1,2\}$ and $\forall t \geq 1$. Consider the outcome sequence $(y_1, y_2, \ldots, y_T)$ defined as follows:

$$y_1 = 1 - \hat{p}_1(f_{11}, f_{21})$$

and

$$\forall t \geq 1 \qquad y_t = 1 - \hat{p}_t(f_{11}, f_{21}, y_1, f_{12}, f_{22}, y_2, \ldots, y_{t-1}, f_{1t}, f_{2t}).$$

Every outcome is inverse of whatever the algorithm predicts. In such a scenario, $\mathscr{A}$ makes T number of mistakes. Hence $L_T = T$. From the definition of the experts, we also know that

$$\min_{i \in \{1,2\}} \sum_{t=1}^{T} \mathbb{1}\{f_{it} \neq y_t\} \leq \frac{T}{2}$$

$$\Rightarrow L_T - \min_i L_{i,T} \geq T - \frac{T}{2} = \frac{T}{2}.$$

□

6.3 Randomized Exponential Weights Algorithm

Algorithm 1 REXP-Wts(η)

Input: Set of expert indices: $\mathscr{E}$
Convex decision space: $\mathscr{D}$
Outcome space: $\mathscr{Y}$
Loss function $l : \mathscr{D} \times \mathscr{Y} \to \mathbb{R}^+$, l is convex on $\mathscr{D}$

Parameters: $\eta \geq 0$

Initialise: $W_{i,1} = 1 \; \forall i \in \{1, 2, \ldots, N\}$

For each round $t = 1, 2, \ldots$

- Get expert advice $F_{i,t}$, $i \in \{1, 2, \ldots, N\}$
- Draw $I_t \in \{1, 2, \ldots, N\}$ according to $P(I_t = i) = \frac{W_{i,t}}{\sum_{j=1}^{N} W_{j,t}}$
- Predicted outcome $\hat{p}_t = f_{I_t,t}$
- See y_t. Algorithm suffers loss $l(f_{I_t}, y_t)$
- Update weights:

$$W_{i,t+1} = W_{i,t} \exp(-\eta \, l(f_{i,t}, y_t))$$

End

Theorem 6.2. *Let $\mathscr{D}, \mathscr{Y}, l$ be arbitrary and $|\mathscr{E}| = N$. Denote the expected regret of the Randomized Exponential Weights algorithm as $R_T(REXP\text{-}Wts(\eta))$. Then,*

$$R_T(REXP\text{-}Wts(\eta)) = \sup_{\{y\},\{f\}} \mathbb{E}\left[L_T - \min_{i \in \mathscr{E}} L_{i,T}\right] \leq \frac{\log N}{\eta} + \frac{\eta T}{8}$$

where $L_T = \sum_{t=1}^{T} l(f_{i,t}, y_t)$.

We will connect this problem to a new learning problem (in the spirit of reductions in computer science) and transform $(\mathscr{D}, \mathscr{Y}, l, \mathscr{E})$ to $(\mathscr{D}', \mathscr{Y}', l', \mathscr{E}')$ so that

- $\mathscr{D}^{'}, \mathscr{Y}^{'}$ are convex
- Expected Regret of REXP-Wts(η) running on $(\mathscr{D}, \mathscr{Y}, l, \mathscr{E})$ = Deterministic Regret of EXP-Wts(η) running on $(\mathscr{D}^{'}, \mathscr{Y}^{'}, l^{'}, \mathscr{E}^{'})$

Proof: Step I: Construction

- $\mathscr{D}^{'} = \{\pi \in \mathbb{R}^N : \sum_{i=1}^{N}, \pi_i \geq 0 \, \forall i\}$, i.e $\mathscr{D}^{'}$ is a N-dimensional simplex and hence $\mathscr{D}^{'}$ is convex in $\mathbb{R}^N$
- $\mathscr{Y}^{'} = \mathscr{Y} \times \mathscr{D}^N$
- $l^{'}(\pi, (y, f_1, f_2, \ldots, f_N)) = \sum_{i=1}^{N} \pi_i l(f_i, y)$. So, $l^{'}$ is convex in $\mathscr{D}^{'}$

Given $\{y_t\} \in y$, $\{f_{i,t} \in \mathscr{D}\}$, $1 \leq t \leq T$, $1 \leq i \leq N$ we define a corresponding sequence:

$$\begin{aligned} f_{i,t}^{'} &\in \mathscr{D}^{'} \\ f_{i,t}^{'} &= e_i = (0, 0, \ldots, 1, \ldots, 0) \in \mathscr{D}^{'} \\ y_t^{'} &\in \mathscr{Y}^{'}, y_t^{'} = (y_t, f_{1t}, f_{2t}, \ldots, f_{Nt}) \end{aligned}$$

Step II: Transformation

Let us now run EXP-Wts(η) on $(\mathscr{D}^{'}, \mathscr{Y}^{'}, l^{'}, \mathscr{E}^{'})$ with $\{y_t^{'}\}_{1 \leq t \leq T}$.

<u>Observation 1:</u> Upon transformation the per-round losses of the experts remain the same.

$$\begin{aligned} l^{'}(f_{i,t}^{'}, y_t^{'}) &= l^{'}(e_i, y_t^{'}) \\ &= l^{'}(e_i, (y_t, f_{1,t}, \ldots, f_{N,t})) \\ &= l(f_{i,T}, y_t) \\ &\Rightarrow L_{i,T} = L_{i,T}^{'}. \end{aligned}$$

So,

$$\min_i L_{i,T} = \min_i L_{i,T}^{'}$$

$\Rightarrow$ Weights used by both algorithms are the same. $\left[\because W_{i,t} = \exp\left[-\eta \sum_{s=1}^{t-1} l(f_{i,s}, y_s) \right] \right]$

<u>Observation 2:</u>

$$\begin{aligned} \mathbb{E}\left[l(f_{I_t,t}, y_t)\right] &= \sum_{i=1}^{N} \mathbb{E}\left[l(f_{I_t,t}, y_t)\right] P(I_t = i) \\ &= \sum_{i=1}^{N} \frac{W_{i,t}}{\sum_j W_{j,t}} l(f_{i,t}, y_t) \\ &= l^{'}((\pi_{1t}, \pi_{2t}, \ldots, \pi_{Nt}), (y_t, f_{1t}, f_{2t}, \ldots, f_{Nt})) \\ &= l^{'}(\hat{p}_t^{'}, y_t^{'}). \end{aligned}$$

By EXP-Wts result for convex losses we get,

$$\left[\sum_{t=1}^{T} l^{'}(\hat{p}_t^{'}, y_t^{'}) - \min_i \sum_{t=1}^{T} l^{'}(f_{it}^{'}, y_t^{'})\right] \leq \frac{\log N}{\eta} + \frac{\eta T}{8}$$

$$\Rightarrow \left[\sum_{t=1}^{T} \mathbb{E}[l(f_{I_t,t}, y_t)] - \min_i \sum_{t=1}^{T} l(f_{it}, y_t)\right] \leq \frac{\log N}{\eta} + \frac{\eta T}{8}$$

□

6.3.1 Bounding the Tail of the Regret

The previous result helped bound the (worst case) expected regret of the REXPWTS algorithm. One could ask - can the variance or tail of the regret be nontrivially large? It turns out that this is not the case - we can control the deviation of the regret from its mean as follows.

Theorem 6.3. *Let $(\mathscr{D}, \mathscr{Y}, l, \mathscr{E})$ be arbitrary and $|\mathscr{E}| = N$. Then, for any $0 < \delta < 1$, the loss of REXP-Wts(η) satisfies the following:*

$$\forall\{y_t\}, \forall\{f_{i,t}\}\, P\left[\hat{L}_T - \mathbb{E}[\hat{L}_T] \geq \sqrt{\frac{T}{2}\log(\frac{1}{\delta})}\right] \leq \delta$$

and with probability $\geq (1-\delta)$:

$$\hat{L}_T - \min_i L_{i,T} \leq \sqrt{\frac{T\log N}{2}} + \sqrt{\frac{T\log(\frac{1}{\delta})}{2}}$$

References

[1] Gabor Bartok, David Pal, Csaba Szepesvari, and Istvan Szita. *Online learning - CMPUT 654 Course Notes*. 2011.

[2] Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning and Games*. Cambridge University Press, 2006.