E1 245: Online Prediction & Learning **Fall 2014**

Lecture 3 — August 12

Lecturer: Aditya Gopalan *Scribe: Aadirupa Saha*

3.1 Recap

In the last lecture, we focused on expert advice based online learning algorithms, more specifically Exponential Weights algorithm for convex decision spaces and loss functions. In this lecture, we will study the regret guarantee of this algorithm, and also analyze the limitation of expert advice based prediction models in a more general setting, in particular, for non-convex decision spaces and loss functions.

3.2 Quick summary on convexity

1. A set $\mathscr{X} \subseteq \mathbb{R}^d$ is said to be convex if $\forall x, y \in \mathscr{X}$, and $\lambda \in [0,1]$,
$$\lambda x + (1-\lambda)y \in \mathscr{X}.$$
2. Let $\mathscr{X}$ be a convex set, a function $f : \mathscr{X} \mapsto \mathbb{R}$ is said to be convex if $\forall x, y \in \mathscr{X}$, and $\lambda \in [0,1]$,
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$
3. Let $\mathscr{X}$ be a convex set, a function $g : \mathscr{X} \mapsto \mathbb{R}$ is said to be concave if $(-g)$ is convex.
4. Let $\mathscr{X}$ be a convex set and $\sigma > 0$ be any constant. A function $f : \mathscr{X} \mapsto \mathbb{R}$ is said to be σ-Exp concave if $x \longmapsto \exp(-\sigma f(x))$ is a concave function, $\forall x \in \mathscr{X}$.
5. Jensen's Inequality (finite version): Let $\mathscr{X}$ be a convex set and $f : \mathscr{X} \mapsto \mathbb{R}$ be a convex function. Consider $\lambda_1, \lambda_2, \ldots, \lambda_N \in \mathbb{R}$ s.t. $\lambda_i \geq 0 \ \forall i \in \{1, 2, \ldots, N\}$, and $\sum_{i=1}^N \lambda_i = 1$. Then for $x_1, x_2, \ldots, x_N \in \mathscr{X}$,
$$f(\sum_{i=1}^N \lambda_i x_i) \leq \sum_{i=1}^N \lambda_i f(x_i).$$
Jensen's Inequality (general version): Let $\mathbf{Z}$ be a random variable taking values in $\mathscr{X}$, then
$$f(\mathbb{E}[\mathbf{Z}]) \leq \mathbb{E}[f(\mathbf{Z})].$$

3.3 Exponential Weights Algorithm

Algorithm **EXP-Wts**

Inputs:

Set of expert indices: $\mathscr{E}$

Convex decision space: $\mathscr{D}$

Outcome space: $\mathscr{Y}$

Loss function $l : \mathscr{D} \times \mathscr{Y} \mapsto \mathbb{R}^+$, l is convex on $\mathscr{D}$

Parameter:

Learning rate $\eta > 0$

Initialize:

Weight on each expert $i \in \mathscr{E}$, $w_{i,0} := 1$

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

– Environment chooses $y_t \in \mathscr{Y}$

– Experts' decide $(f_{i,t})_{i=1}^{|\mathscr{E}|}, f_{i,t} \in \mathscr{D}$

– Predict $\hat{p}_t := \frac{\sum_{i \in \mathscr{E}} w_{i,t-1} f_{i,t}}{\sum_{j \in \mathscr{E}} w_{j,t-1}} \in \mathscr{D}$

– Receive y_t

– Incur loss $l(\hat{p}_t, y_t)$

– Re-weight each expert $i \in \mathscr{E}$, $w_{i,t} := \exp\left(-\eta \sum_{s=1}^{t} l(f_{i,s}, y_s)\right)$

End

The regret of **EXP-Wts** algorithm w.r.t. the best expert in $\mathscr{E}$, at the end of round T is

$$R_T(\textbf{EXP-Wts}) := \sum_{t=1}^{T} l(\hat{p}_t, y_t) - \min_{i \in \mathscr{E}} \sum_{t=1}^{T} l(f_{i,t}, y_t)$$

Theorem 3.1. *Let $\mathscr{D}$ be a convex decision space, $\mathscr{Y}$ be the outcome space, and $l : \mathscr{D} \times \mathscr{Y} \mapsto [0,1]$ be a convex loss function w.r.t. its first argument, then for* ***EXP-Wts*** *algorithm,*

$$R_T(\textbf{\textit{EXP-Wts}}) \leq \frac{\log|\mathscr{E}|}{\eta} + \frac{\eta T}{8}.$$

In particular, setting $\eta = \sqrt{\frac{8 \log|\mathscr{E}|}{T}}$,

$$R_T(\textbf{\textit{EXP-Wts}}) \leq \sqrt{\frac{T}{2} \log|\mathscr{E}|}.$$

Remarks:

1. If the time horizon T is known in advance, with $\eta = \sqrt{\frac{8 \log|\mathscr{E}|}{T}}$, the average per trial regret of **EXP-Wts** algorithm $\frac{R_T(\textbf{EXP-Wts})}{T} = O(\frac{1}{\sqrt{T}}) \to 0$ as $T \to \infty$.

2. In general T is not known beforehand, however one can still get an $O(\sqrt{T})$ regret bound with time varying learning rate $\eta_t = \frac{1}{\sqrt{t}}$, $\forall t = 1,2,\dots$ which is independent of T. Another approach to make η horizon free (i.e. independent of T) is the so called "doubling trick".

3. Convex losses are commonly used in the context of online learning problems as the corresponding regret guarantees can be made sub-linear in T.

Proof of Theorem 3.1. We will use a potential function based argument to proof the result. Let us consider the potential function $W_t = \sum_{i\in\mathscr{E}} w_{i,t}$, $\forall t \in \{1,2,\dots,T\}$. Clearly, $W_0 = |\mathscr{E}|$. Thus

$$\begin{aligned}\log\left(\frac{W_t}{W_{t-1}}\right) &= \log\left(\frac{\sum_{i\in\mathscr{E}} w_{i,t-1}\exp\left(-\eta\, l(f_{i,t},y_t)\right)}{\sum_{i\in\mathscr{E}} w_{i,t-1}}\right)\\ &= \log\mathbb{E}[\exp(-\eta\mathbf{X})], \end{aligned} \tag{3.1}$$

where $\mathbf{X}$ is a random variable s.t. $\mathbb{P}(\mathbf{X} = l(f_{i,t},y_t)) = \frac{w_{i,t-1}}{\sum_{j\in\mathscr{E}} w_{j,t-1}}$, $\forall i \in \mathscr{E}$. Note that, since, $l : \mathscr{D}\times\mathscr{Y} \mapsto [0,1]$, $\mathbf{X} \in [0,1]$.

Lemma 3.2. *Let $\mathbf{X}$ be a random variable s.t. $\mathbf{X} \in [a,b]$. Then $\forall z \in \mathbb{R}$,*

$$\log\mathbb{E}[\exp(z\mathbf{X})] \le z\mathbb{E}[\mathbf{X}] + \frac{z^2(b-a)^2}{8}.$$

Now applying *Lemma 3.2* in (3.1) with $a = 0$, $b = 1$, and $z = -\eta$ we get,

$$\begin{aligned}\log\left(\frac{W_t}{W_{t-1}}\right) &\le -\eta\mathbb{E}[\mathbf{X}] + \frac{\eta^2}{8}\\ &= -\eta\sum_{i\in\mathscr{E}}\left(\frac{w_{i,t-1}}{\sum_{j\in\mathscr{E}} w_{j,t-1}}\right) l(f_{i,t},y_t) + \frac{\eta^2}{8}\\ &= -\eta\sum_{i\in\mathscr{E}}\left(\frac{w_{i,t-1}}{W_{t-1}}\right) l(f_{i,t},y_t) + \frac{\eta^2}{8}\\ &\le -\eta\, l\left(\sum_{i\in\mathscr{E}}\left(\frac{w_{i,t-1}}{W_{t-1}}\right) f_{i,t}, y_t\right) + \frac{\eta^2}{8} \quad \text{[Since } l \text{ is convex in its first argument.]}\\ &= -\eta\, l\left(\left(\frac{\sum_{i\in\mathscr{E}} w_{i,t-1} f_{i,t}}{\sum_{j\in\mathscr{E}} w_{j,t-1}}\right), y_t\right) + \frac{\eta^2}{8}\\ &= -\eta\, l\left(\hat{p}_t, y_t\right) + \frac{\eta^2}{8}.\end{aligned}$$

Summing over $t = 1,2,\dots,T$ then gives

$$\log\left(\frac{W_T}{W_0}\right) \le -\eta\sum_{t=1}^{T} l\left(\hat{p}_t, y_t\right) + \frac{\eta^2 T}{8}. \tag{3.2}$$

Again,

$$
\begin{aligned}
\log\left(\frac{W_T}{W_0}\right) &= \log\left(\frac{\sum_{i\in\mathscr{E}}\exp\big(-\eta\sum_{t=1}^{T} l(f_{i,t},y_t)\big)}{|\mathscr{E}|}\right)\\
&= \log\left(\sum_{i\in\mathscr{E}}\exp\big(-\eta\sum_{t=1}^{T} l(f_{i,t},y_t)\big)\right)-\log|\mathscr{E}|\\
&\geq \log\left(\max_{i\in\mathscr{E}}\left(\exp\big(-\eta\sum_{t=1}^{T} l(f_{i,t},y_t)\big)\right)\right)-\log|\mathscr{E}|\\
&= \max_{i\in\mathscr{E}}\left(\log\left(\exp\big(-\eta\sum_{t=1}^{T} l(f_{i,t},y_t)\big)\right)\right)-\log|\mathscr{E}|\\
&= \max_{i\in\mathscr{E}}\left(-\eta\sum_{t=1}^{T} l(f_{i,t},y_t)\right)-\log|\mathscr{E}|\\
&= -\min_{i\in\mathscr{E}}\left(\eta\sum_{t=1}^{T} l(f_{i,t},y_t)\right)-\log|\mathscr{E}|. \qquad (3.3)
\end{aligned}
$$

Combining (3.2) and (3.3) we get

$$-\min_{i\in\mathscr{E}}\left(\eta\sum_{t=1}^{T} l(f_{i,t},y_t)\right)-\log|\mathscr{E}| \leq -\eta\sum_{t=1}^{T} l\left(\hat{p}_t,y_t\right)+\frac{\eta^2 T}{8}.$$

$$\text{Thus,}\quad R_T(\textbf{EXP-Wts}) = \sum_{t=1}^{T} l(\hat{p}_t,y_t)-\min_{i\in\mathscr{E}}\sum_{t=1}^{T} l(f_{i,t},y_t) \leq \frac{\log|\mathscr{E}|}{\eta}+\frac{\eta T}{8}.$$

□

3.4 Online learning with non-convex losses/decisions

Consider the following setting of 1 bit prediction problem with expert advice:

- Decision space $\mathscr{D} = \{0,1\}$
- Outcome space $\mathscr{Y} = \{0,1\}$
- Set of two experts $\mathscr{E} = \{0,1\}$, where $f_{0,t} = 0$ and $f_{1,t} = 1$, $\forall t = 1,2,\ldots$
- 0-1 loss function $l : \mathscr{D}\times\mathscr{Y} \mapsto [0,1]$ s.t. $l(\hat{p},y) = \mathbf{1}(\hat{p} \neq q)$

Now, let us assume that we have a deterministic forecasting algorithm $\mathscr{A}$ which predicts $\hat{p}_t \in \{0,1\}$, and the environment sets $y_t = 1 - p_t$ in each round $t = 1, 2, \ldots, T$. Clearly, $\sum_{t=1}^{T} l(\hat{p}_t, y_t) = T$. Moreover in each round t, only one of the expert suffers 1 unit of loss as $y_t \in \{0,1\}$. Hence, $\sum_{t=1}^{T} l(f_{0,t}, y_t) + l(f_{1,t}, y_t) = T$, and $\min_{i \in \{0,1\}} \sum_{t=1}^{T} l(f_{i,t}, y_t) \leq \frac{T}{2}$. Thus regret of algorithm $\mathscr{A}$ becomes

$$R_T(\mathscr{A}) = \sum_{t=1}^{T} l(\hat{p}_t, y_t) - \min_{i \in \mathscr{E}} \sum_{t=1}^{T} l(f_{i,t}, y_t) \geq \frac{T}{2}.$$

Clearly $R_T(\mathscr{A}) \to \infty$ as $T \to \infty$, and this leads to the conclusion that in the current setting, no online learning algorithm can achieve a non-trivial regret guarantee which is sub-linear in T.

3.5 Next Lecture

In the next lecture, we will focus on randomized online prediction problems where in each round the forecasting algorithm plays a distribution over decision space $\mathscr{D}$, instead of making a single deterministic prediction, and also analyze the regret guarantees of such algorithms with arbitrary loss functions.

References

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

2. The Multiplicative Weights Update Method: A Meta-Algorithm and its Applications. Sanjeev Arora, Elad Hazan, Satyen Kale. *Theory of Computing*, 2012.