E1 245: Online Prediction & Learning	Fall 2014
Lecture 4 — August 14	
Lecturer: Aditya Gopalan	*Scribe: Geethu Joseph*

4.1 Recap

In the last lecture, we showed that regret bound for ExpWeights for convex decision spaces and convex losses (over decisions), is $\mathcal{O}(\sqrt{T})$, where T is the number of rounds. Towards the end of the last lecture, we saw that no deterministic algorithm for a binary prediction problem and 0-1 loss, can achieve sublinear regret. This motivated us to introduce randomization in forecasting, and this lecture introduces randomized prediction, where forecaster chooses each outcome according to some probability distribution. The lecture presents randomized version of ExpWeights algorithm and expected regret of the algorithm. We use Azuma-Hoeffding inequality to derive regret bounds with high probability, and further, better regret bounds are obtained for the special case of exp-concave losses.

4.2 Actions model

The system model under consideration consists of the following:

- a set of available actions $\mathcal{A} = \{1, 2, \ldots, N\}$
- an outcome space $\mathcal{Y}$
- a loss function $l : \mathcal{A} \times \mathcal{Y} \to \mathbb{R}^+$

At each round, $t = 1, 2, \ldots$, the player or the algorithm plays an action $I_t \in \mathcal{A}$, and the environment plays $Y_t \in \mathcal{Y}$. At each round t, the player observes y_t, and suffers a loss of $l(I_t, y_t)$. The regret function of an algorithm after T rounds of prediction is defined in terms of forecaster's best constant prediction,

$$R_T = \sum_{t=1}^{T} l(I_t, Y_t) - \min_{i \in \mathcal{A}} \sum_{t=1}^{T} l(i, Y_t).$$

The player can possibly include an internal randomization to predict the outcomes, and the environment does not see this randomization. Randomized forecasters have been considered in various different setups; see, for example, Feder, Merhav, and Gutman [1], Foster and Vohra [2], Cesa-Bianchi and Lugosi [3], Merhav and Feder [4], and Vovk [5]. The constraint on the environment can be formally expressed as $Y_t \in \mathcal{F}_{t-1}$, where $\mathcal{F}_{t-1} = \sigma(\mathcal{H}_{t-1})$ is the sigma algebra generated by the history[1], $\mathcal{H}_{t-1}$. History is the collection of all past predictions and environment states, $\mathcal{H}_{t-1} = \{(I_1, Y_1), (I_2, Y_2), \ldots, (I_{t-1}, Y_{t-1})\}$. We note that the environment states Y_t can themselves be random. Since the player's strategy is random, we look at expected regret of the algorithm,

$$\begin{aligned}\mathbb{E}\left\{l\left(I_t, Y_t\right)|\mathcal{F}_{t-1}\right\} &= \mathbb{E}\left\{\sum_{i=1}^{N} \mathbb{I}\left\{I_t = i\right\} l\left(i, Y_t\right)|\mathcal{F}_{t-1}\right\} \\ &= \sum_{i=1}^{N} l\left(i, Y_t\right) \mathbb{E}\left\{\mathbb{I}\left\{I_t = i\right\}|\mathcal{F}_{t-1}\right\} \\ &= \sum_{i=1}^{N} l\left(i, Y_t\right) p_{i,t},\end{aligned}$$

where $p_{i,t}$ is the probability with which i^{th} action is chosen by the algorithm, and $\sum_{i=1}^{N} p_{i,t} = 1$. Thus, last step follows because, $\mathbb{E}\left\{\mathbb{I}\left\{I_t = i\right\}|\mathcal{F}_{t-1}\right\} = p_{i,t}$ Now, we write,

$$\mathbb{E}\left\{l\left(I_t, Y_t\right)|\mathcal{F}_{t-1}\right\} = \boldsymbol{l}_t^{\mathrm{T}} \boldsymbol{p}_t. \tag{4.1}$$

where $\boldsymbol{l}_t$ and $\boldsymbol{p}_t$ are N dimensional vectors, with $\boldsymbol{p}_t = [p_{i,t}]_{i=1}^{N}$ and $\boldsymbol{l}_t = [l(i, Y_t)]_{i=1}^{N}$.

4.2.1 Expected regret

Consider the randomized version of ExpWeights running on the following setting. Let $\bar{\mathcal{Y}} = \mathcal{Y}$, and $\mathcal{D} = \Delta_N$ be N-dimensional simplex.

$$\mathcal{D} = \left\{\pi_1, \pi_2, \ldots, \pi_N : \forall i \pi_i \geq 0, \sum_{i=1}^{N} \pi_i = 1\right\}.$$

The loss function, $\bar{l} : \bar{\mathcal{D}} \times \bar{\mathcal{Y}} \to \mathbb{R}$ is defined as

$$\bar{l}(\pi, y) = \sum_{i=1}^{N} \pi_i l(i, y).$$

The exponential weights come from Δ_N to pick an action I_t at random. We note that the problem of prediction with expert advice can be reduced to the actions model, if we consider a set of experts $\{1, 2, \ldots, N\}$ recommending one constant action, $f_{i,t} = i$. Thus, $\bar{\mathcal{E}} = \{u_1, u_2, \ldots, u_N\}$, where u_i is the standard unit vector with 1 at i^{th} location, and 0 elsewhere. Using standard regret bound of ExpWeights algorithm, which we derived in the last lecture, for any outcome sequence $(y_1, y_2, \ldots, y_T)$,

$$\sum_{t=1}^{T} \bar{l}(\hat{p}_t, y_t) - \min_{v \in \mathcal{E}} \sum_{t=1}^{T} \bar{l}(v, y_t) \leq \sqrt{\frac{T}{2} \log N}.$$

[1] Sigma algebra generated by a random vector X, denoted by $\sigma(X)$, is the smallest σ-field on which X is measurable

Using the definition of loss function, $\bar{l}$,

$$\sum_{t=1}^{T}\sum_{i=1}^{N}\hat{p}_t(i)l(i,y_t) - \min_{i=1,2,\ldots,N}\sum_{t=1}^{T}\bar{l}(i,y_t) \le \sqrt{\frac{T}{2}\log N}.$$

Switching back to the action model,

$$\sum_{t=1}^{T}\sum_{i=1}^{N}\hat{p}_t(i)l(i,y_t) - \min_{i\in\mathcal{A}}\sum_{t=1}^{T}\bar{l}(i,y_t) \le \sqrt{\frac{T}{2}\log N}.$$

4.2.2 High probability regret analysis

Next, we obtain a high probability regret bound for the randomized ExpWeights algorithm. To derive the result, we use Azuma-Hoeffding inequality which is a generalization of Chernoff bound. Before we state the theorem, we need to define martingale difference sequence.

Let $X_1, X_2, \ldots$ be a sequence of random variables. Let $X_1, X_2, \ldots$ be a sequence of random variables. Another sequence of random variables, $v_1, v_2, \ldots$ is called a *martingale difference sequence* with respect to $X_1, X_2, \ldots$, if $\forall i \ge 1$

1. $v_i \in \sigma(X_1, X_2, \ldots, X_i)$,
2. $\mathbb{E}\{V_i | X_1, X_2, \ldots, X_i\} = 0$.

For example, let $\{X_i\}$ be a sequence of *iid* zero mean random variables. Then, the sequence of random variables $\{V_i\}$, with $V_i = X_i$, $\forall i$ forms a martingale difference sequence. Consider another example, where $X_1, X_2, \ldots$ is any sequence of random variables. The sequence of random variable $\{V_i\}$ defined as $V_{i+1} = W_{i+1} - \mathbb{E}\{W_{i+1}|X_1, X_2, \ldots, X_i\}$ forms a martingale difference sequence, where $W_1, W_2, \ldots$ any sequence of random variables such that $W_i \in \sigma(X_1, X_2, \ldots X_i)$

Theorem 4.1 (Azuma-Hoeffding inequality). *Let $V_1, V_2, \ldots$ be a martingale difference sequence such that $V_i \in [a_i, b_i]$ almost surely, $\forall i \ge 1$. Then, $\forall z > 0$,*

$$\mathbb{P}\left\{\sum_{i=1}^{m} V_i > z\right\} \le \exp\left\{\frac{-2z^2}{\sum_{i=1}^{m}(b_i - a_i)^2}\right\},$$
$$\mathbb{P}\left\{\sum_{i=1}^{m} V_i < -z\right\} \le \exp\left\{\frac{-2z^2}{\sum_{i=1}^{m}(b_i - a_i)^2}\right\}.$$

Theorem 4.2. *With probability greater than $1-\delta$,*

$$R_T(\textit{R-EXP}) \le l_{max}\sqrt{\frac{T}{2}\log N} + l_{max}\sqrt{2T\log\frac{1}{\delta}}.$$

Proof: We note that, the sequence $\{l(I_t, Y_t) - \hat{\boldsymbol{p}}_t^{\mathrm{T}} \boldsymbol{l}_t\}$ is a martingale difference sequence, which follows directly from (4.1). Applying Lemma 4.1, we have $\forall z > 0$

$$\mathbb{P}\left\{\sum_{t=1}^{T} l(I_t, Y_t) - \hat{\boldsymbol{p}}_t^{\mathrm{T}} \boldsymbol{l}_t > z\right\} \leq \exp\left\{\frac{-2z^2}{4Tl_{\max}^2}\right\},$$

where $l_{\max} = \max_{(a,y)} l(a, y)$. Substituting $z = l_{\max}\sqrt{2T \log\left(\frac{1}{\delta}\right)}$,

$$\mathbb{P}\left\{\sum_{t=1}^{T} l(I_t, Y_t) - \hat{\boldsymbol{p}}_t^{\mathrm{T}} \boldsymbol{l}_t > l_{\max}\sqrt{2T \log\left(\frac{1}{\delta}\right)}\right\} \leq \delta.$$

Thus, the result follows. □

4.2.3 Regret bounds for exp-concave loss functions

We derive better regret bounds for randomized ExpWeights algorithm, if we further restrict the loss function to be a exp-concave function. This constraint is much stronger than convexity. The following are some of the examples for exp-concave loss functions for which $\mathcal{D} = \mathcal{Y} = [0, 1]$.

1. Relative entropy or logarithmic loss,

$$l(\hat{p}, y) = y \log\left(\frac{y}{\hat{p}}\right) + (1 - y) \log\left(\frac{1 - y}{1 - \hat{p}}\right)$$

   is 1-exp-concave and unbounded.

2. Square loss function, $l(\hat{p}, y) = (\hat{p} - y)^2$, is 1/2-exp-concave.

We also note that, absolute loss, $l(\hat{p}, y) = |\hat{p} - y|$ is not a σ-exp-concave for any $\sigma > 0$.

Theorem 4.3. *Run the exp-wts algorithm with a σ-exp-concave loss, $\eta = \sigma$, and*

$$\hat{p}_t = \frac{\sum_{i \in \mathcal{E}} \exp\left\{-\sigma \sum_{s=1}^{t-1} l(f_{i,s}, y_s)\right\} f_{i,t}}{\sum_{i \in \mathcal{E}} \exp\left\{-\sigma \sum_{s=1}^{t-1} l(f_{i,s}, y_s)\right\}},$$

then

$$R_T\left(\textit{exp-wts}(\sigma)\right) \leq \frac{\log |\mathcal{E}|}{\sigma}.$$

Proof: Let $w_{i,t} = \exp\left\{-\sigma \sum_{s=1}^{t-1} l(f_{i,s}, y_s)\right\}$. Then,

$$\hat{p}_t = \frac{\sum_{i \in \mathcal{E}} w_{i,t} f_{i,t}}{\sum_{i \in \mathcal{E}} w_{i,t}}. \tag{4.2}$$

Consider the potential function,

$$\begin{aligned} \phi_t &= \frac{\sum_{i\in\mathcal{E}} w_{i,t}}{\sum_{i\in\mathcal{E}} w_{i,t-1}} \\ &= \frac{\sum_{i\in\mathcal{E}} w_{i,t-1} e^{-\sigma l(f_{i,t-1},y_t)}}{\sum_{i\in\mathcal{E}} w_{i,t-1}}. \end{aligned} \tag{4.3}$$

Using the exp-concavity of loss function, we can upper bound the right hand side of the equation as,

$$\phi_t \le \exp\left\{-\sigma \sum_{i\in\mathcal{E}} \frac{w_{i,t-1}}{\sum_{i\in\mathcal{E}} w_{i,t-1}} l(f_{i,t-1}, y_{t-1})\right\}.$$

Further, we can bound using the convexity property of loss function(exp-concave functions are convex)

$$\begin{aligned} \phi_t &\le \exp\left\{-\sigma l\left(\sum_{i\in\mathcal{E}} \frac{w_{i,t-1}}{\sum_{i\in\mathcal{E}} w_{i,t-1}} f_{i,t-1}, y_{t-1}\right)\right\} \\ &= \exp\left\{-\sigma l\left(\hat{p}_{t-1}, y_{t-1}\right)\right\}. \end{aligned} \tag{4.4}$$

Last steps follows from (4.2). Consider the sum of logarithm of potential function, and using (4.3),

$$\begin{aligned} \sum_{t=1}^{T} \log \phi_t &= \sum_{t=1}^{T} \left\{\log \sum_{i\in\mathcal{E}} w_{i,t} - \log \sum_{i\in\mathcal{E}} w_{i,t-1}\right\} \\ &= \log \sum_{i\in\mathcal{E}} w_{i,T} - \log N \\ &\ge \log \max_{i\in\mathcal{E}} w_{i,T} - \log N \\ &= -\sigma \min_{i\in\mathcal{E}} \sum_{s=1}^{t-1} l(f_{i,s}, y_s) - \log N. \end{aligned}$$

Last step follows from the definition of $w_{i,t}$. Substituting from (4.4) we get,

$$-\sigma \min_{i\in\mathcal{E}} \sum_{s=1}^{t-1} l(f_{i,s}, y_s) - \log N \le \sum_{t=1}^{T} -\sigma l\left(\hat{p}_{t-1}, y_{t-1}\right).$$

On rearranging,

$$\begin{aligned} R_T\left(\text{exp-wts}(\sigma)\right) &= \sum_{t=1}^{T} l\left(\hat{p}_{t-1}, y_{t-1}\right) - \min_{i\in\mathcal{E}} \sum_{s=1}^{t-1} l(f_{i,s}, y_s) \\ &\le \frac{\log N}{\sigma}. \end{aligned}$$

Thus, the proof is complete. □

We note that the regret bound is independent of number of rounds, T, which implies strong learnability of the algorithm.

Bibliography

[1] M. Feder, N. Merhav, and M. Gutman,“ Universal prediction of individual sequences”, *IEEE Transactions on Information Theory*, 38:1258–1270, 1992.

[2] D. Foster and R. Vohra, “Regret in the on-line decision problem”, *Games and Economic Behavior*, 29:7–36, 1999.

[3] N. Cesa-Bianchi and G. Lugosi,“ On prediction of individual sequences”, *Annals of Statistics*, 27:1865–1895, 1999.

[4] N. Merhav and M. Feder,“ Universal prediction”, *IEEE Transactions on Information Theory*, 44:2124-147, 1998.

[5] V. Vovk, “Competitive on-line statistics”, *International Statistical Review*, 69:213–248, 2001.

[6] Chapter-4, Nicolo Cesa-Bianchi and Gabor Lugosi,“Prediction, Learning and Games”, Cambridge University Press, 2006