IE613: Online Machine Learning **Jan-Apr 2016**

Lecture 10: EXP3.P Algorithm

Lecturer: M. K. Hanawal *Scribes: Vishal Dawar*

Recapitulation

- We define Regret as follows:

$$\tilde{R}_T = \sum_{t=1}^{T} l_{I_t,t} - \underset{k}{min} \sum_{t=1}^{T} l_{k,t} \tag{10.1}$$

  As this is a random quantity, we consider the Expected value of Regret and we denote it as $R_T = E\big[\tilde{R}_T\big]$. The problem with this quantity is $R_T = E\left[\sum_{t=1}^{T} l_{I_t,t}\right] - E\left[\underset{k}{min} \sum_{t=1}^{T} l_{k,t}\right]$. We need to take the ***min*** out of the expectation. For that we can write R_T as follows:

$$R_T \geq E\left[\sum_{t=1}^{T} l_{I_t,t}\right] - \underset{k}{min}\ E\left[\sum_{t=1}^{T} l_{k,t}\right]$$

- and hence Pseudo Regret is defined as follows:

$$\bar{R}_T = E\left[\sum_{t=1}^{T} l_{I_t,t}\right] - \underset{k}{min}\ E\left[\sum_{t=1}^{T} l_{k,t}\right] \tag{10.2}$$

$$\Rightarrow R_T \geq \bar{R}_T$$

  T = number of rounds we are running the algorithm for,
  $l_{k,t}$ = loss incurred by the player for choosing k^{th} arm in t^{th} round.

- We will say that the adversary is *oblivious* if the adversary doesn't know $I_1, I_2, ..., I_{t-1}$(arms picked by the player).
- Also, if the Adversary chooses the loss l_t knowing $I_1, I_2, ..., I_{t-1}$, then we call him *non-oblivious.*

10.1 Dealing directly with the Random quantity $\widetilde{R}_T$

We have been dealing with the Expected Regret values all along in the *EXP3* algorithm, but what if someone prefers to not use expectations and rather use the random quantity $\tilde{R}_T$ instead?

So, now we would try to find an upper bound on the random quantity $\widetilde{R_T}$. But because $\tilde{R}_T$ being a random quantity, we can only talk about an upper bound on it with an associated probability.

10.1.1 High Probability Bounds

In the previous lecture, we defined an unbiased estimator for the loss $l_{i,t}$ corresponding to i^{th} arm and t^{th} round as follows:

$$\tilde{l}_{i,t} = \frac{l_{i,t}}{P_{i,t}}\mathbb{1}_{\{I_t=i\}}$$

As mentioned that $\tilde{l}_{i,t}$ is an unbiased estimator of $l_{i,t}$, we are sure that the expected value of this unbiased estimator is exactly equal to $l_{i,t}$, but that might give us very fluctuating upper bounds corresponding to each realization as variance can be really high. To show this, we consider the second moment:

$$E_{I_t \sim P_t}\left[\tilde{l}_{i,t}^2\right] = \sum_{I_t=1}^{K}\left(\frac{l_{i,t}}{P_{i,t}}\mathbb{1}_{\{I_t=i\}}\right)^2 P_{i,t}$$

$$= \frac{l_{i,t}^2}{P_{i,t}}$$

As per the expression, if $P_{i,t}$ is arbitrarily close to zero, then the second moment can blow up and tend to infinity, which is a quite an issue because the variance then will also tend to infinity. To tackle this, we can try to devise an estimator with less variance.

This leads to the motivation of a modified algorithm which will give us a proper upper bound on the value of the Regret along with the probability it holds with.

10.1.2 A Modified Algorithm: EXP3.P

In this section, we introduce a modified algorithm which is based on EXP3, but gives us a better result in terms of upper bounds on the Regret $\tilde{R}_T$.

Algorithm 1 EXP3.p

1: **Input:** Parameters = η , γ , $\beta \in$ [0,1].
2: Let P_1 be the uniform distribution over $\{1, 2, ...K\}$.
3: **for** t **do**=1,2,...,T
4: Draw an arm I_t from distribution P_t
5: Compute the estimated loss

$$\tilde{l}_{i,t} = \frac{l_{i,t}\mathbb{1}_{\{I_t=t\}}+\beta}{P_{i,t}} \;\forall\; i = 1, 2, ..., K$$

6: Update the estimated cumulative loss

$$\tilde{L}_{i,t} = \tilde{L}_{i,t-1} + \tilde{l}_{i,t} \;\forall\; i = 1, 2, ..., K$$

7: Update probability distribution over arms $P_{t+1} = (P_{1,t+1}, P_{2,t+1}, ..., P_{K,t+1})$

$$P_{i,t+1} = (1-\gamma)\frac{exp(-\eta\tilde{L}_{i,t})}{\sum_k exp(-\eta\tilde{L}_{k,t})} + \gamma\frac{1}{K} \;\forall\; i = 1, 2, ..., K$$

8: **Output:** Estimated loss vectors $\tilde{l}_t$; $\forall t = 1, 2, ...T$

By adding the additional $\frac{\gamma}{K}$ term, we are making sure that $P_{i,t}$ never becomes arbitrarily close to zero, which leads to the second moment of $\tilde{l}_{i,t}$ never blow up.

Now we can move towards finding an upper bound on the Regret $\tilde{R}_T$ with some probability associated with it. To do that, we first state and prove a lemma.

Lemma 10.1 *For $\beta \leq 1$, let*

$$P_{i,t+1} = (1-\gamma)\frac{exp(\eta \tilde{L}_{i,t})}{\sum_k exp(\eta \tilde{L}_{k,t})} + \frac{\gamma}{K}$$

Then $\forall \delta \in (0,1)$, with probability at least 1-δ

$$\sum_{t=1}^{T} l_{i,t} - \sum_{t=1}^{T} \tilde{l}_{i,t} \leq \frac{\log \frac{1}{\delta}}{\beta} \qquad \forall i = 1,2,...,K \tag{10.3}$$

Proof: Let E_t be the expectation conditional on $I_1, I_2, ..., I_{t-1}$, then to prove the lemma, we need a result. So we claim the following:

$$E_t\Big[exp\Big(\beta\ l_{i,t} - \beta\ \tilde{l}_{i,t}\Big)\Big] \leq 1$$

We will be requiring the following two inequalities in the proof:

$$e^x \leq 1 + x + x^2 \qquad \forall x \leq 1$$

$$1 + x \leq e^x \qquad \forall x \in \mathbb{R}$$

Consider the following:

$$\begin{aligned}
E_t\Big[exp\Big(\beta l_{i,t} - \beta\tilde{l}_{i,t}\Big)\Big] &= E_t\Big[exp\Big(\beta l_{i,t} - \beta\Big(\frac{l_{i,t}\mathbb{1}_{\{I_t=i\}} + \beta}{P_{i,t}}\Big)\Big)\Big] \\
&= E_t\Big[exp\Big(\beta l_{i,t} - \beta\Big(\frac{l_{i,t}\mathbb{1}_{\{I_t=i\}}}{P_{i,t}}\Big)\Big)exp\Big(\frac{-\beta^2}{P_{i,t}}\Big)\Big] \\
&= E_t\Big[exp\Big(\beta l_{i,t} - \beta\Big(\frac{l_{i,t}\mathbb{1}_{\{I_t=i\}}}{P_{i,t}}\Big)\Big)\Big]E_t\Big[exp\Big(\frac{-\beta^2}{P_{i,t}}\Big)\Big]
\end{aligned}$$

$\triangleright$ Using $e^x \leq 1 + x + x^2$

$$\leq E_t\Big[1 + \Big(\beta l_{i,t} - \beta\Big(\frac{l_{i,t}\mathbb{1}_{\{I_t=i\}}}{P_{i,t}}\Big)\Big) + \Big(\beta l_{i,t} - \beta\Big(\frac{l_{i,t}\mathbb{1}_{\{I_t=i\}}}{P_{i,t}}\Big)\Big)^2\Big]exp\Big(\frac{-\beta^2}{P_{i,t}}\Big)$$

$$= \Big[1 + E_t\Big(\beta l_{i,t} - \beta\Big(\frac{l_{i,t}\mathbb{1}_{\{I_t=i\}}}{P_{i,t}}\Big)\Big) + E_t\Big(\beta l_{i,t} - \beta\Big(\frac{l_{i,t}\mathbb{1}_{\{I_t=i\}}}{P_{i,t}}\Big)\Big)^2\Big]exp\Big(\frac{-\beta^2}{P_{i,t}}\Big)$$

$$= \Big[1 + \beta E_t(l_{i,t}) - \beta E_t\Big(\frac{l_{i,t}\mathbb{1}_{\{I_t=i\}}}{P_{i,t}}\Big) + E_t(\beta^2 l_{i,t}^2) + E_t\Big(\beta^2\frac{l_{i,t}^2\mathbb{1}_{\{I_t=i\}}^2}{P_{i,t}^2}\Big) - 2E_t\Big(\beta^2\frac{l_{i,t}^2\mathbb{1}_{\{I_t=i\}}}{P_{i,t}}\Big)\Big]exp\Big(\frac{-\beta^2}{P_{i,t}}\Big)$$

$$= \Big[1 + \beta l_{i,t} - \beta l_{i,t} + \beta^2 l_{i,t}^2 + \beta^2\frac{l_{i,t}^2}{P_{i,t}} - 2\beta^2 l_{i,t}^2\Big]exp\Big(\frac{-\beta^2}{P_{i,t}}\Big)$$

$\triangleright$ Step (A)

$$= \Big[1 - \beta^2 l_{i,t}^2 + \beta^2\frac{l_{i,t}^2}{P_{i,t}}\Big]exp\Big(\frac{-\beta^2}{P_{i,t}}\Big)$$

$$\leq \left(1 + \frac{\beta^2}{P_{i,t}}\right) exp\left(\frac{-\beta^2}{P_{i,t}}\right)$$

$\triangleright$ Eliminating the -ive terms by an upper bound and also $l_{i,t} \leq 1$

$$\leq exp\left(\frac{\beta^2}{P_{i,t}}\right) exp\left(\frac{-\beta^2}{P_{i,t}}\right) = 1$$

$\triangleright$ Using $1+x \leq e^x$

$$\Rightarrow E_t\left[exp\left(\beta l_{i,t} - \beta\tilde{l}_{i,t}\right)\right] \leq 1$$

We also have used the results $E(\tilde{l}_{i,t}) = l_{i,t}$ and $E(\tilde{l}^2_{i,t}) = \frac{l^2_{i,t}}{P_{i,t}}$ in Step (A) above.

Now we will prove the lemma.

Consider the following:

$$E\left[exp\left(\sum_{t=1}^{T} \beta l_{i,t} - \sum_{t=1}^{T} \beta\tilde{l}_{i,t}\right)\right] = E_1\left[exp\left(\beta l_{i,1} - \beta\tilde{l}_{i,1}\right)\right] \times ..... \times E_T\left[exp\left(\beta l_{i,T} - \beta\tilde{l}_{i,T}\right)\right] \tag{10.4}$$

$$\tag{10.5}$$

$$\leq 1 \times 1 \times ..... \times 1 = 1$$

$$\Rightarrow E\left[exp\left(\sum_{t=1}^{T} \beta l_{i,t} - \sum_{t=1}^{T} \beta\tilde{l}_{i,t}\right)\right] \leq 1$$

We know that the if X is a non-negative random variable and ϵ is a positive real number, then by Markov's inequality,

$$P(X \geq \epsilon) \leq \frac{E[X]}{\epsilon}$$

$$P\left(X \geq \log\frac{1}{\delta}\right) \leq \frac{E[X]}{log\frac{1}{\delta}} \tag{10.6}$$

$$P\left(exp(X) \geq \frac{1}{\delta}\right) \leq \delta E[exp(X)]$$

Therefore, using this inequality, we can write,

$$P\left(exp\left(\sum_{t=1}^{T} \beta l_{i,t} - \sum_{t=1}^{T} \beta\tilde{l}_{i,t}\right) \geq \frac{1}{\delta}\right) \leq \delta E\left[exp\left(\sum_{t=1}^{T} \beta l_{i,t} - \sum_{t=1}^{T} \beta\tilde{l}_{i,t}\right)\right] \tag{10.7}$$

$$\leq \delta 1 = \delta$$

$$\Rightarrow P\left(\sum_{t=1}^{T} \beta l_{i,t} - \sum_{t=1}^{T} \beta\tilde{l}_{i,t} \geq \log\frac{1}{\delta}\right) \leq \delta$$

$$\Rightarrow P\left(\sum_{t=1}^{T} \beta l_{i,t} - \sum_{t=1}^{T} \beta\tilde{l}_{i,t} \leq \log\frac{1}{\delta}\right) \geq 1 - \delta$$

So, the quantity $\left(\sum_{t=1}^{T} \beta l_{i,t} - \sum_{t=1}^{T} \beta\tilde{l}_{i,t}\right) \leq \log\frac{1}{\delta}$ with at least probability $1 - \delta$.
Hence Proved. ■

Through the proof, we can conclude that the difference in the total actual loss and total estimated loss when multiplied by β is always upper bounded by $\log \frac{1}{\delta}$ with probability at least $1 - \delta$.

Through some computations and analysis of the values of parameters, the following result can be stated.

Theorem 10.2 *For any $\delta \in (0, 1)$ if* EXP3.p *is run with*

$$\beta = \sqrt{\frac{\log(K\delta^{-1})}{TK}}, \qquad \eta = 0.95\sqrt{\frac{\log K}{TK}}, \qquad \gamma = 1.05\sqrt{\frac{K \log K}{T}}$$

Then with probability at least $1 - \delta$

$$\tilde{R}_T \leq 5.15\sqrt{TK \log(K\delta^{-1})}$$

and if $\beta = \sqrt{\frac{\log K}{TK}}$ and η , γ are as above, then with probability at least 1-δ

$$\tilde{R}_T \leq \sqrt{\frac{TK}{\log K}} \log \frac{1}{\delta} + 5.15\sqrt{TK \log K}$$