IE613: Online Machine Learning **Jan-Apr 2018**

Lecture 12: EXP3-IX Algorithm

Lecturer: M. K. Hanawal *Scribes: Prashant Trivedi*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

Recap

In last class we discussed about the EXP3.P algorithm and the regret bounds in term of expectation and in terms of high probability and proved that both of them have same order, thus one question that we can ask is, Is it possible to come up with any other algorithm which will give us a better regret bounds? This we are going to see in this class and will prove that there is no such algorithm which give us better regret bounds.

12.1 Introduction

In EXP3 algorithm we only focused on exploitation whereas in EXP3.P we also considered exploration along with exploitation. In last class we argued that the exploration term is essentially $\frac{\gamma}{K}$ which is given by the uniform distribution. Also we noticed that both EXP3 and EXP3.P gave us same regret bound in terms of order which is $O(\sqrt{TKlogK})$, also we argued that in EXP3 we can show the bounds in expectations whereas in EXP3.P we can show the bounds in expectation as well as in high probability. Now the question is can we get this bound in Expectation and in high probability only by doing the exploration and the answer to this question is yes! we can do this and the algorithm that we have for this EXP3-IX.

Algorithm 1 EXP3-IX

Parameters: $\eta_t > 0$, $\gamma_t > 0$.
Initialization: $w_{1,i} = 1$.
for $t = 1, 2, \ldots, T$, **repeat**

1. $p_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^{K} w_{t,j}}$.
2. Draw $I_t \sim \boldsymbol{p}_t = (p_{t,1}, \ldots, p_{t,K})$.
3. Observe loss $\ell_{\{t,I_t\}}$.
4. $\tilde{\ell}_{t,i} \leftarrow \frac{\ell_{t,i}}{p_{t,i}+\gamma_t}\mathbb{I}_{\{I_t=i\}}$ for all $i \in [K]$.
5. $w_{t+1,i} \leftarrow w_{t,i}e^{-\eta_t\tilde{\ell}_{t,i}}$ for all $i \in [K]$.

Clearly, the loss estimates are defined as

$$\tilde{\ell}_{t,i} = \frac{\ell_{t,i}}{p_{t,i} + \gamma_t}\mathbb{I}_{\{I_t=i\}}, \tag{12.1}$$

for all i and an appropriately chosen $\gamma_t > 0$. This technique of defining loss is referred to as as *Implicit eXploration*, or, in short, IX.

12.2 High-probability regret bounds via implicit exploration

Following theorem states high-probability bound on the regret of EXP3-IX.

Theorem 12.1 *Fix an arbitrary $\delta > 0$. With $\eta_t = 2\gamma_t = \sqrt{\frac{2\log K}{KT}}$ for all t, EXP3-IX guarantees*

$$\tilde{R}_T \leq 2\sqrt{2KT\log K} + \left(\sqrt{\frac{2KT}{\log K}} + 1\right)\log(2/\delta)$$

with probability at least $1-\delta$. Furthermore, setting $\eta_t = 2\gamma_t = \sqrt{\frac{\log K}{Kt}}$ for all t, the bound becomes

$$\tilde{R}_T \leq 4\sqrt{KT\log K} + \left(2\sqrt{\frac{KT}{\log K}} + 1\right)\log(2/\delta).$$

To prove this theorem we need a lemma (one of its version we proved in the previous lecture). We will use the similar lemma (stated below) for EXP3-IX also.

Lemma 12.2 *Let (γ_t) be a* fixed *non-increasing sequence with $\gamma_t \geq 0$ and let $\alpha_{t,i}$ be nonnegative $\mathcal{F}_{t-1}$-measurable random variables satisfying $\alpha_{t,i} \leq 2\gamma_t$ for all t and i. Then, with probability at least $1-\delta$,*

$$\sum_{t=1}^{T}\sum_{i=1}^{K}\alpha_{t,i}\big(\tilde{\ell}_{t,i} - \ell_{t,i}\big) \leq \log(1/\delta)$$

A particularly important special case of the above lemma is the following:

Corollary 12.3 *Let $\gamma_t = \gamma \geq 0$, then $\forall\ \delta \in (0,1)$, with probability atleast (1-δ)*

$$\sum_{t=1}^{T}\big(\tilde{\ell}_{t,i} - \ell_{t,i}\big) \leq \frac{\log(K/\delta)}{2\gamma} \tag{12.2}$$

simultaneously holds for all $i \in [K]$

Now we will proceed to prove the theorem(12.1)
Proof: Let us fix an arbitrary $\delta' \in (0,1)$. Following the standard analysis of EXP3 in the loss game and non increasing learning rates , we can obtain the bound

$$\sum_{t=1}^{T}\left(\sum_{i=1}^{K}p_{t,i}\tilde{\ell}_{t,i} - \tilde{\ell}_{t,j}\right) \leq \frac{\log K}{\eta_T} + \sum_{t=1}^{T}\frac{\eta_t}{2}\sum_{i=1}^{K}p_{t,i}\big(\tilde{\ell}_{t,i}\big)^2 \tag{12.3}$$

for any j.

Now observe that

$$\sum_{i=1}^{K}p_{t,i}\tilde{\ell}_{t,i} = \sum_{i=1}^{K}\mathbb{I}_{\{I_t=i\}}\frac{\ell_{t,i}(p_{t,i}+\gamma_t)}{p_{t,i}+\gamma_t} - \gamma_t\sum_{i=1}^{K}\mathbb{I}_{\{I_t=i\}}\frac{\ell_{t,i}}{p_{t,i}+\gamma_t} = \ell_{t,I_t} - \gamma_t\sum_{i=1}^{K}\tilde{\ell}_{t,i} \tag{12.4}$$

Substituting this back in equation(12.3) we have:

$$\sum_{t=1}^{T}\left[\left(\ell_{t,I_t}-\gamma_t\sum_{i=1}^{K}\tilde{\ell}_{t,i}\right)-\ell_{t,j}\right]\leq\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\frac{\eta_t}{2}\sum_{i=1}^{K}p_{t,i}\left(\tilde{\ell}_{t,i}\right)^2 \tag{12.5}$$

Also notice that, $\sum_{i=1}^{K}p_{t,i}\tilde{\ell}^2_{t,i}\leq\sum_{i=1}^{K}\tilde{\ell}_{t,i}$ holds by the boundedness of the losses.

$$\sum_{t=1}^{T}\left(\ell_{t,I_t}-\tilde{\ell}_{t,j}\right)\leq\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\tilde{\ell}_{t,i} \tag{12.6}$$

Adding and subtracting one $\sum_{t=1}^{T}\ell_{t,j}$ on R.H.S. we have

$$\sum_{t=1}^{T}\left(\ell_{t,I_t}-\sum_{i=1}^{K}\tilde{\ell}_{t,j}\right)+\sum_{t=1}^{T}\ell_{t,j}-\sum_{t=1}^{T}\ell_{t,j}\leq\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\tilde{\ell}_{t,i}$$

$$\Rightarrow\sum_{t=1}^{T}\left(\ell_{t,I_t}-\ell_{t,j}\right)\leq\sum_{t=1}^{T}\left(\tilde{\ell}_{t,j}-\ell_{t,j}\right)+\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\tilde{\ell}_{t,i}$$

Thus, we get that

$$\begin{aligned}\sum_{t=1}^{T}(\ell_{t,I_t}-\ell_{t,j})\leq&\sum_{t=1}^{T}\left(\tilde{\ell}_{t,j}-\ell_{t,j}\right)+\frac{\log K}{\eta_T}+\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)\sum_{i=1}^{K}\tilde{\ell}_{t,i}\\ \leq&\frac{\log(K/\delta')}{2\gamma}+\frac{\log K}{\eta}+\sum_{t=1}^{T}(\frac{\eta_t}{2}+\gamma_t)\sum_{i=1}^{K}\ell_{t,i}+\log(1/\delta')\end{aligned}$$

holds with probability at least $1-2\delta'$, where the last line follows from an application of lemma(12.2) with $\alpha_{t,i}=\eta_t/2+\gamma_t$ for all t,i and taking $j=arg\ \underset{i}{min}\ L_{T,i}$ and $\delta'=\delta/2$, and using the boundedness of the losses,

$$\tilde{R}_T\leq\frac{\log(2K/\delta)}{2\gamma_T}+\frac{\log K}{\eta_T}+K\sum_{t=1}^{T}\left(\frac{\eta_t}{2}+\gamma_t\right)+\log(1/\delta).$$

The statements of the theorem then follow immediately, noting that $\sum_{t=1}^{T}1/\sqrt{t}\leq2\sqrt{T}$. ∎