Title of this document

Firstname Lastname

January 6, 2015

Contents

0.1 Abstract

The problem we worked on can be describe like this : We consider a multi-armed bandit where the arms are divided up in several classes. Every time the learner chooses an arm, not only he observes the loss of this arm but can also have informations about losses of non-chosen arms. In our case, we consider that when the learner choose a arm of the class i, the other arms of the class i have the probability r_{ii} to reveal their loss and the other arms of a class j have the probability r_{ij}. So the problem is characterized by the matrix R which represents the probabilies of communication between two arms of differents classes.

0.2 Case with one class

We first consider the case where we have one class to explain the algorithms we adapted to fit our model.
This can be identified as a social graph with Facebook-like settings where we consider that for a person, his group of friends is homogenous enough.

Then the problem is reduced to the situation where all non-chosen arms reveal their loss with and unknown probability $r_{11} = r$. This can be modeled by Erd?s-Rényi random graphs with parameter r; i.e. at each step t we are in the situation as if from N vertices, we constructed an edge between i and j with the probability r. Then after chosing and arm I_t, all the information about losses recieved additionnaly at this step t correspond to the arms (or vertices in the graph) that are connected to I_t with an edge.

0.2.1 Problem definition

We consider a sequential set on iteractions with the multi-armed bandit we assume to have N arms, for each step t=1,..,T these are the actions performed by the environment :

1. The environment chooses losses for every arm noted $l_{t,i}$ for the arm i at the step t.
2. Following the algorithm we hope would minize as much as possible the regret the player draws an arm I_t.
3. The player recieves the losse l_{t,I_t}.
4. $O_{t,I_t} = 1$ and $\forall i,\ i \neq I_t$, Ot, i is drawn from a Bernouilli distribution with mean r. These other $O_{t,i}$ corresponds to the value of the logic expression $i\ is\ neighboor\ of\ I_t$ in the Erd?s-Rényi random graph of the step t.
5. $\forall i$ such that $O_{t,i} = 1$ the player can observe the loss $l_{t,i}$.

We write $p_{t,i} = \mathbb{P}[I_t = i|\mathcal{F}_{t-1}]$ where $\mathcal{F}_{t-1}$ corresponds to all the actions and observations the player had until the step t. Then intuitively the probability of observing the loss of the arm i at the step t would be $q_{t,i} = p_{t,i} + (1 - p_{t,i})r$ and to use the Exp3 the the loss estimate :

$$\hat{l}_{t,i} = \frac{O_{t,i} l_{t,i}}{q_{t,i}}.$$

But the main problem resides in the fact that r is unknown so the algorithms presented use tricks to obtain loss estimates such that we keep the property :

$$\mathbb{E}[\hat{l}_{t,i}|\mathcal{F}_{t-1}] = l_{t,i}$$

The principal idea is to have access to two **independant** geometrically distributed random variables M_t^* and $K_{t,i}$ with respective parameters r and $p_{t,i}$, then the variable $G_{t,i}^* = min\ \{K_{t,i}, M_t^*\}$ is also geometrically distributed with the parameter $q_{t,i}$ previously defined. Then if we have $G_{t,i}^*$ **independant** of $O_{t,i}$ we can replace in the definition of $\hat{l}_{t,i}$, $\frac{1}{q_{t,i}}$ by $G_{t,i}^*$.

0.2.2 DuplExp3 for large values of r

We assume $r \geq \frac{log(T)}{2N}$, which implies that the probability of having no additionnal observations in round t is bounded by $\frac{1}{\sqrt{T}}$.
This algorithm needs two Exp3 sub-algorithms, one for the round when t is even and the other one for the rest so that we can construct independant M_t^* and $K_{t,i}$ and indenpendant $G_{t,i}^*$ and $O_{t,i}$. Then the algorithm compute :

$$p_{t+2,i} \propto w_{t+2,i} = \frac{1}{N} exp\left(-\eta_{t+2}\hat{L}_{t,i}\right)$$

Where $\hat{L}_{t,i} = \sum_{k=0}^{t/2} \hat{l}_{t-2k,i}$ the cumulative sum of the loss estimates for the arm i for one of the Exp3 sub-algorithms. Then M_t^* is constructed like this :

- We define $O_{t,i}^{'}\ i = 1, ..., N-1$
- $\forall i < I_t\ O_{t,i}^{'} = O_{t,i}$
- $\forall N \geq i > I_t\ O_{t,i-1}^{'} = O_{t,i}$

$$M_t^* = min\ \{1 \leq i < N : O_{t-1,i}^{'} = 1\} \cup \{N\}$$

Then M_t follows a truncated geometric law. We also define $K_{t,i}$ as a geometric random variable with parameter $p_{t,i}$ computed a the step $t-2$. The since M_t depends of O_{t-1} and $p_{t,i}$ of O_{t-2} they are obviously independant. That's why we can consider :

$$G_{t,i} = min\ \{K_{t,i}, M_t\}$$

And then as expected $G_{t,i}$ is independant of $O_{t,i}$. Finally, the loss estimate becomes :

$$\hat{l_{t,i}} = G_{t,i} O_{t,i} l_{t,i} q_{t,i}.$$

With the duplex algorithm. In this case with the right η_t we find the upperbound for the regret :

$$R_T \leq 4\sqrt{\left(\frac{T}{r} + N^2\right) logN} + sqrtT$$

0.2.3 Estimate r

However, a problem remains in this algorithm ; since we don't know a priori what is the value of r, we can't ensure that $r \geq \frac{log(T)}{2N}$ as we did the assumption previously. So we need to find a lower bound on r to know in which case we probably are. The algorithm *Estimating* $\underline{r}$

List of Figures

List of Tables