AI-3001: Advanced Topics in Machine Learning **Aug-Nov 2022**

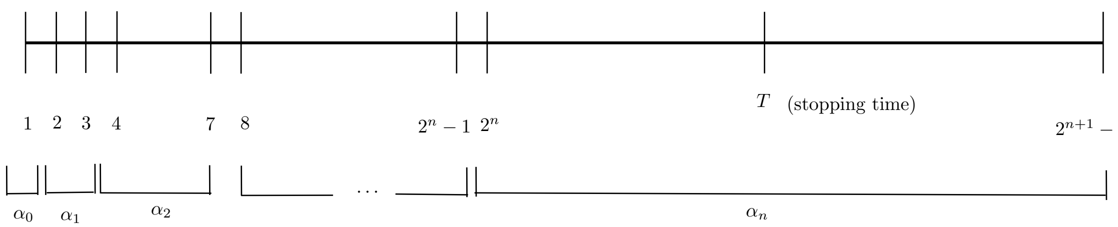
Lecture 3: Doubling Trick and Follow the Best Expert

Lecturer: Ganesh Ghalme *Scribes: Ganesh Ghalme*

We observed in the last lecture that one can achieve sub-linear regret guarantee when the algorithm know the value of stopping time T in advance by setting the parameter value α accordingly. In many practical applications the value of T is not known in advance or considered to be set adversarially. Doubling trick, presented next, helps in guaranteeing *any-time* guarantee without prior knowledge of T.

3.1 Doubling Trick and Follow the Best Expert

3.1.1 Doubling Trick



The algorithm updates its parameters in every epoch. Each epoch is of twice the size of the previous one. That is, the first epoch consists of only one time period. Second epoch is 2 time periods long and ith epoch is 2^{i-1} time periods long. The algorithm is restarted in epoch $i+1$ with the parameter value set to α_i with the stopping time $T_i = 2^i$.

Claim 3.1. *Let an algorithm* $\textsc{Alg}$ *satisfies* $\mathcal{R}_T(\textsc{Alg}) = c\sqrt{T}$ *with a-priori knowledge of stopping time* T*. Then the regret upper bound with doubling trick is given by* $\frac{\sqrt{2}}{\sqrt{2}-1}c\sqrt{T}$*.*

Proof. Let $T_i = 2^i$ for every i and let the stopping time T be such that

$$2^n \leq T \leq 2^{n+1} - 1 \text{ for some } n \in \mathbb{N}. \tag{3.1}$$

We have

$$\begin{aligned}
\mathcal{R}_T(\textsc{Alg}) &\leq \sum_{i=0}^{n} c\sqrt{T_i} && \text{(why ?)} \\
&= c \cdot \sum_{i=0}^{n} 2^{i/2} = c \cdot \frac{2^{\frac{n+1}{2}} - 1}{\sqrt{2} - 1} \leq c \cdot \frac{\sqrt{2}\sqrt{T} - 1}{\sqrt{2} - 1} < c\frac{\sqrt{2}}{\sqrt{2} - 1}\sqrt{T}
\end{aligned}$$

□

3.1.2 Follow the Best Expert

We saw that randomization helps for the extension of the WEIGHTEDMAJORITY algorithm. We will now extend our NAIVE algorithm for this setting. Since we cannot completely eliminate the expert after she makes a single mistake in the best expert setting, in our extension, we will follow the advice of the expert which made least mistakes so far. We will analyze the regret of follow the best expert strategy given by

$$p_t = f_{i_t,t} \quad \text{where} \quad i_t \in \arg\min_{i=1,2,\cdots N} \sum_{s=1}^{t-1} \ell(f_{i,s}, y_s) \tag{3.2}$$

First define a *Lookahead* forecaster who has a knowledge of tth outcome before making optimal forecast. This forecaster is not real as it needs access to future outcome, it will be used in the analysis. The lookahead forecaster is given by

$$p_t^\star = f_{i_t^*,t} \quad \text{where} \quad i_t^* \in \arg\min_{i=1,2,\cdots N} \sum_{s=1}^{t} \ell(f_{i,s}, y_s) \tag{3.3}$$

We begin with the following useful lemma.

Lemma 3.2. *For any $y_1, y_2, \cdots y_T$ we have*

$$\sum_{t=1}^{T} \ell(p_t^*, y_t) \leq \sum_{t=1}^{T} \ell(p_T^*, y_t)$$

Proof. We will prove it by induction on stopping time T. For $T = 1$, the result is trivially true. Say that it is true for $T - 1$, that is

$$\sum_{t=1}^{T-1} \ell(p_t^*, y_t) \leq \sum_{t=1}^{T-1} \ell(p_{T-1}^*, y_t) \tag{3.4}$$

We have

$$\begin{aligned}
\sum_{t=1}^{T} \ell(p_T^\star, y_t) &= \sum_{t=1}^{T-1} \ell(p_T^*, y_t) + \ell(p_T^*, y_T) \\
&\geq \sum_{t=1}^{T-1} \ell(p_{T-1}^*, y_t) + \ell(p_T^*, y_T) && \text{(Since } p_{T-1}^* \text{ in the infimum)} \\
&\geq \sum_{t=1}^{T-1} \ell(p_t^*, y_t) + \ell(p_T^*, y_T) && \text{(Equation 3.4)} \\
&= \sum_{t=1}^{T} \ell(p_t^*, y_t)
\end{aligned}$$

□

Above lemma gives us

$$\begin{aligned}
\mathcal{R}_T &\leq \sum_{t=1}^{T} (\ell(p_t, y_t) - \ell(p_t^*, y_t)) \\
&\leq \sum_{t=1}^{T} \varepsilon_t && (\ \ell(p_t, y_t) - \ell(p_t^*, y_t) \leq \varepsilon_t \text{ for all } t)
\end{aligned}$$

If $\varepsilon_t \leq \frac{1}{t}$ then we have that the regret grows as $O(\log(T))$ (why?). Note that this bound is true irrespective of the number of experts (N) and irrespective of the size of outcome space $\mathcal{Y}$.

3.1.3 Constant Experts with Squared loss

Let us assume that $\mathcal{D} = \mathcal{Y} = \{x \in \mathbb{R}^d : ||x|| \leq 1\}$ and $\ell(p, y) = \sum_{i=1}^{d}(p_i - y_i)^2$. Furthermore, assume for exposition that each point in the unit ball is a "constant" expert; i.e. the one which always gives the same recommendation. Hence the set of experts is $\mathcal{D}$ (uncountably infinite constant experts). It is easy to see that the loss minimizing prediction policy is the one which predicts the average i.e.

$$p_t = \frac{1}{t-1}\sum_{s=1}^{t-1} y_s \text{ and } p_t^* = \frac{1}{t}\sum_{s=1}^{t} y_s.$$

We have

$$\begin{aligned}
\ell(p_t, y) - \ell(p_t^\star, y) &= ||p_t - y||^2 - ||p_t^* - y||^2 \\
&= (p_t - p_t^*) \cdot (p_t + p_t^* - 2y) \\
&\leq 4||p_t - p_t^*||
\end{aligned}$$

(Cauchy-Shwartz inequality (given in below section) and the fact that $\mathcal{D}$ is a unit ball)

Next we bound the norm $||p_t - p_t^*||$ as follows

$$\begin{aligned}
||p_t - p_t^*|| &= ||\Big(\frac{1}{t-1} - \frac{1}{t}\Big)\sum_{s=1}^{t-1} y_s - \frac{y_t}{t}|| \\
&= \frac{1}{t}||\Big(\frac{\sum_{s=1}^{t-1} y_s}{t-1} - \frac{y_t}{t}\Big)|| \\
&\leq \frac{1}{t}\Big(||\frac{\sum_{s=1}^{t-1} y_s}{t-1}|| + ||\frac{y_t}{t}||\Big) && \text{(Triangle inequality)} \\
&\leq \frac{2}{t} && \text{(since } \mathcal{Y} \text{ is a unit ball)}
\end{aligned}$$

We obtained that $\ell(p_t, y) - \ell(p_t^\star, y) \leq \frac{8}{t}$. This gives us logarithmic regret. To obtain logarithmic regret we used following two powerful assumptions 1) we restricted the experts' advice to a constant advice and 2) we considered convex loss functions (in fact, squared loss satisfied much stronger condition which we will explore more in detail in later classes). We will study this actions setting next. Before that, we will quickly review some preliminaries.