1 Solution for HW1 / CSE254 / 2020

1.1 Setup

We observe a bit sequence $Y_1, Y_2, \ldots \in \{0, 1\}$. We generate a sequence of predictions $p_1, p_2, \ldots \in [0, 1]$ and would like to minimize the cumulative loss $L_T = \sum_{t=1}^{T} |p_t - Y_t|$.

We consider two scenarios:

1. We assume that the sequence is generated by IID coin flips where $P(Y_t = 1) = q$. q is unknown.

2. We don't assume anything about how the sequence is generated. We want the cumulative loss to not be much larger than the cumulative loss of the best (in hindsight) fixed prediction $h \in [0, 1]$

1.2 Analysis for IID scenario

Suppose that we are in the IID scenario, and that we know the bias q. It is not hard to convince yourself that if $q > 1/2$ the best prediction is $p = 1$ and if $q < 1/2$ the best prediction is $p = 0$. If $q = 1/2$ the expected loss is $1/2$ independent of the value of p

As we don't know the value of q, we need to estimate it, the minimal variance estimator is

$$\hat{q}_t = \frac{\sum_{i=1}^{t} Y_i}{t}$$

The best prediction, given this estimate, is

$$p_t = \begin{cases} 0 & \text{if } \hat{q}_t \leq 1/2 \\ 1 & \text{if } \hat{q}_t > 1/2 \end{cases}$$

How would this predictor behave? Note that the estimate $\hat{q}_t$ gets closer and closer to q as t increases.

Consider the case $q < 1/2$ (the case $q > 1/2$ is symmetric). As t increases, the probability that $\hat{q}_t > 1/2$ converges to zero.

How does this convergence effect the regret? On step t, the expected loss of the optimal prediction (which is 0), is equal to q, the expected loss of the other prediction (which is 1) is equal to $1 - q$. The regret if $\hat{q}_t > 1/2$ is $(1 - q) - q = 1 - 2q$. The expected regret on iteration t is therefor $P(\hat{q}_t > 1/2)(1 - 2q)$. Defining $\gamma = |1/2 - q|$ we can rewrite this as $P(\hat{q}_t > 1/2)2\gamma$. The last expression is correct also for $q > 1/2$.

Using Hoeffding bound we get

$$P\left(|\hat{q}_t - q| \geq \gamma\right) \leq e^{-2t\gamma^2}$$

Summing the per-iteration regrets over all

$$Regret \leq 2\gamma \sum_{t=0}^{\infty} e^{-2t\gamma^2} = \frac{2\gamma}{1 - e^{-2\gamma^2}} \approx \frac{1}{\gamma}$$

We found a bound on the cumulative regret that depends only on γ. In particular, it does not depend on the length of the sequence.

The hardest sequences are therefor the ones where q is close to $1/2$. However note that the regret for $q = 1/2$ is zero, because in that case all predictions suffer an expected loss of $1/2$ per iteration.

1.3 Analysis for worst case scenario

Note first that, just as in the IID scenario, the only predictions that matter are 0 and 1. If we look at the sequence $Y_1, \ldots, Y_T$ in hindsight and ask what would have been the best constant prediction. The answer is that the best prediction is determined by wheter $\frac{1}{T} \sum_{t=1}^{T} Y_t$ is larger or smaller than $1/2$. And again, if it is equal to $1/2$ it does not matter which prediction we choose.

So we care about the regret relative to two experts: one that always predicts 1 and one that always predicts zero. As the loss is bounded in $[0, 1]$ we can use Hedge to combine these two experts.

Assuming we know the length of the sequence we can set the learning rate appropriately and get the regret bound:

$$L_A - \min_i L_i \leq \sqrt{2T \ln 2} + \ln 2$$

We get a bound that does not depend on the gap γ but increases with the length of the sequence like $\sqrt{T}$.

This is the price we pay for using an algorithm that works for all sequences, not just for the IID case.

An interesting question is whether the problem is with the bound, or with the algorithm itself. My hunch is that the problem is with the algorithm itself - it keeps trying the suboptimal action whereas the algorithm for the IID case will make it's mind whether $q > 1/2$ or $q < 1/2$ around $t = 1/\gamma^2$ and will stick to its choice. Have fun trying to prove that!