E1 245 - Online Prediction and Learning, Aug-Dec 2014
Homework #1

---

1. *Generalizing the MAJORITY algorithm (5 points)*
We showed that the MAJORITY algorithm for binary prediction makes at most $\log_2 N$ mistakes using the advice of N experts whenever some expert is always predicting correctly. Show that a straightforward modification of MAJORITY makes at most $O((m+1)\log_2 N)$ mistakes[1] when some expert makes $m \geq 0$ mistakes.

2. *The Doubling Trick for obtaining "anytime" learning algorithms (5 points)*
Suppose an online learning algorithm with a parameter $\eta > 0$ enjoys a regret bound of $\frac{\beta}{\eta} + \gamma\eta T$ for a total of T rounds, where β and γ are some positive constants (think of the Exponential Weights forecaster for instance). If the time horizon T is known in advance, then setting $\eta := \sqrt{\frac{\beta}{\gamma T}}$ minimizes the bound. Consider the following tweak to obtain an algorithm (and bound) that does NOT require knowing the horizon T beforehand (i.e., an "anytime" algorithm). Time is divided into periods: the m-th period is formed by rounds $2^m, 2^m+1, \ldots, 2^{m+1}-1$, where $m = 0, 1, 2, \ldots$ In every m-th period, starting at round 2^m, the original algorithm is re-initialized and run with a parameter $\eta_m := \sqrt{\frac{\beta}{\gamma 2^m}}$. Prove that for *any* T, this modified algorithm enjoys a regret bound which is at most $\frac{\sqrt{2}}{\sqrt{2}-1}$ times the original optimal regret bound.

3. *A Bayesian interpretation of the Exponential Weights algorithm (5 points)*
Consider the Exponential Weights forecaster in the 1-bit setting with $\mathscr{D} = \mathscr{Y} = \{0,1\}$, the $0-1$ loss $l(d,y) := \mathbb{1}\{d \neq y\}$, and experts $\{1,2,\ldots,N\}$. Assume the following probabilistic model for the outcome sequence $y_1, y_2, \ldots$, where we view each outcome y_t as a realization of a Bernoulli random variable Y_t. First, an expert I is drawn uniformly at random from the set of experts. Then, for each $t \in \{1,2,3,\ldots\}$, the random variable Y_t is set to $f_{I,t}$ with probability p and to $1 - f_{I,t}$ with probability $1-p$, where $p := 1/(1+e^{-\eta})$ and $f_{i,t}$ denotes the advice of expert i at time t. Show that the normalized weight $w_{i,t}/\sum_j w_{j,t}$ used by the algorithm is in fact equal to the posterior probability $\mathbb{P}[I = i \mid Y_1 = y_1, Y_2 = y_2, \ldots, Y_{t-1} = y_{t-1}]$ that expert i was drawn given the observed outcomes so far.

4. *Exponential inequality (3 points)*
Prove (we used this to show a regret bound for the Randomized Weighted Majority algorithm): $-\log(1-x) \leq x + x^2$, $x \in [0, 1/2]$.

5. *A smarter ExpWeights algorithm when the best expert's loss is known beforehand (17 points)*
Consider prediction with expert advice with a convex loss (in the first argument) bounded in $[0,1]$. Suppose you know in advance what the best expert's total loss is going to be at time T (for instance, some quantity much less than T) . Can you utilize this information to tune ExpWeights better and get an improved regret bound?

   (a) (5 points) First, prove that $\log \mathbb{E}[e^{sX}] \leq (e^s - 1)\mathbb{E}[X]$ for any random variable $X \in [0,1]$.

   (b) (2 points) Let the experts be indexed by $\{1,2,\ldots,N\}$. Use the previous result instead of (the weaker) Hoeffding's inequality to show that $L_T(\text{ExpWts}) \leq (\eta L_T^* + \log N)/(1 - e^{-\eta})$ for ExpWeights run with parameter $\eta > 0$. Here, $L_T(\text{ExpWts})$ is the cumulative

[1] Big-Oh notation: We say that $f(m) = O(g(m))$ if there exist constants α, m_0 such that $f(m) \leq \alpha g(m)$ $\forall m \geq m_0$.

loss of the algorithm and $L_T^* := \min_{i=1,\dots,N} L_{i,T}$ is the cumulative loss of the best expert, over T rounds.

(c) (5 points) Use the elementary inequality $\eta \le (e^{\eta} - e^{-\eta})/2$ in the above bound to obtain a further bound. Then, assuming that the value of L_T^* is known beforehand, show that setting the ExpWeight learning rate to $\eta := \log(1 + \sqrt{(2\log N)/L_T^*})$ gives regret at most $\sqrt{2L_T^* \log N} + \log N$, which can be significantly small when the best expert's cumulative loss is small.

(d) (5 points) What if the best expert's loss L_t^* is not known beforehand but available only at time t for each t? Taking a cue from Problem 2, can you design an algorithm that does not require advance knowledge of the cumulative loss of the best expert, and show that its regret bound is only worse by a constant factor compared to the one in part (c) above?