E1 245 - Online Prediction and Learning, Aug-Dec 2015
Homework #2

(10 marks per problem)

1. *The* doubling trick *for obtaining "anytime" learning algorithms*
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.

2. *The Exponentially Weighted Forecaster with non-uniform initial weights*
Consider running the Exponentially Weighted Forecaster (EXPWTS) across N actions with losses bounded in $[0,1]$ (note: randomization is implicit as usual). Let the initial weight distribution over actions be arbitrary (and possibly non-uniform), denoted by $w_{i,1} \geq 0$, $i \in \{1, 2, \ldots, N\}$, $\sum_{i=1}^N w_{i,1} = 1$, and let the learning rate be $\eta > 0$. If $p_{t,i}$ and $l(i, y_t)$ denote the probability that EXPWTS plays action action i and the loss of action i, respectively, in round $t \geq 1$, then show that the cumulative (expected) loss suffered by EXPWTS after $T \geq 1$ rounds satisfies
$$\sum_{t=1}^T \sum_{i=1}^N p_{t,i} \cdot l(i, y_t) \leq \frac{1}{\eta} \log \left[\frac{1}{\sum_{i=1}^N w_{i,1} \exp\left(-\eta \sum_{s=1}^T l(i, y_s)\right)} \right] + \frac{\eta T}{8}.$$
(note: we used this result to derive a bound for the Fixed Share forecaster)

3. *Exp-Concavity and common loss functions*
   (a) Show that if for a $y \in \mathscr{Y}$ and $\eta > 0$ the function $F(z) := e^{-\eta l(z,y)}$ is concave, then $l(z,y)$ is a convex function of z.
   (b) Show that the relative entropy loss $l(x,y) := y \log \frac{y}{x} + (1-y) \log \frac{1-y}{1-x}$, $x, y \in [0,1]$, is 1-exp-concave for all valid values[1] of y.
   (c) Show that the squared loss $l(x,y) := (x-y)^2$, $x, y \in [0,1]$, is $\frac{1}{2}$-exp-concave for all valid values of y.
   (d) Show that the absolute value loss $l(x,y) := |x-y|$, $x, y \in [0,1]$, *cannot* be η-exp-concave for any $\eta > 0$.

4. *Improved regret with exp-concave losses*
Show that if the Exponentially Weighted Forecaster (EXPWTS) is run in the prediction-with-expert-advice setting with a σ-exp-concave loss function $l : \mathscr{D} \times \mathscr{Y} \to [0,1]$ (over $\mathscr{D}$)

[1] By convention, we take $\frac{0}{0} := 0$ & $0 \cdot \log 0 := 0$.

and the learning rate $\eta = \sigma > 0$ over N experts, then the algorithm enjoys the regret bound

$$\sum_{t=1}^{T} l(p_t, y_t) - \min_{i \in [N]} \sum_{t=1}^{T} l(f_{i,t}, y_t) \leq \frac{\log N}{\sigma}.$$

(note: regret does not grow with time T!)

5. *Lower bound for tracking regret*
Show that in the actions game with $N > 1$ actions, there is no algorithm whose worst-case expected regret is sublinear when competing against *all* switching experts. More precisely, show that there exists a constant $c > 0$ such that for $\mathscr{D} := [0,1]$, $\mathscr{Y} := \{0,1\}$, and $l(p,y) := |p-y|$, for any $N > 1$, for any algorithm, there exists a set of base experts of size N and a time horizon $T \geq 1$ such that the regret with respect to all switching experts is at least cT.

6. Establish the following properties used to prove a regret bound for Cover's Universal Portfolio algorithm.

   (a) Let $b^* \in \Delta_m$ represent a Constantly Rebalancing Portfolio (CRP) on the (non-negative) unit simplex in $\mathbb{R}_+^m$. Let $\text{Ball}_\varepsilon(b^*) := \{(1-\varepsilon)b^* + \varepsilon b : b \in \Delta_m\}$ for $\varepsilon \in [0,1]$. If $\text{Vol}(A)$ denotes the $(m-1)$-dimensional volume[2] of a set $A \subseteq \Delta_m$, then show that $\text{Vol}(\text{Ball}_\varepsilon(b^*)) = \varepsilon^{m-1}\text{Vol}(\Delta_m)$.

   (b) Show that the CRP strategy $b \in \text{Ball}_\varepsilon(b^*)$ achieves wealth $S_T(b, x^T) \geq S_T(b^*, x^T)(1-\varepsilon)^T$ in T investment periods.

[2] $\text{Vol}(A)$ can be taken to be the $(m-1)$-dimensional "surface area" of the surface defined by $x_m = f(x_1, \ldots, x_{m-1}) := 1 - \sum_{1=1}^{m-1} x_i$, for $x_1, \ldots, x_{m-1} \geq 0$, $\sum_{1=1}^{m-1} x_i \leq 1$. Alternatively, $\text{Vol}(A)$ can be defined to be the probability of a point lying in the set A when it is drawn from the uniform probability distribution over Δ_m.