E1 245 - Online Prediction and Learning, Aug-Dec 2014
Homework #2

1. *Exp-Concavity of common loss functions (16 points)*
   (a) (4 points) 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) (4 points) 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 values[1] of y.
   (c) (4 points) Show that the square loss $l(x,y) := (x-y)^2$, $x,y \in [0,1]$, is $\frac{1}{2}$-exp-concave for all values of y.
   (d) (4 points) Show that the absolute value loss $l(x,y) := |x-y|$, $x,y \in [0,1]$, cannot be η-exp-concave for any $\eta > 0$.

2. *Competing with switching sequences (12 points)*
   Consider learning to play one out of N actions at each time $t = 1,2,\ldots,T$ in the full-information setting, and where each action incurs a loss bounded in $[0,1]$.
   (a) (6 points) Suppose an ambitious algorithm designer wants to design a strategy that can perform as well as the best *offline solution*, i.e., the best possible sequence of actions that can be chosen with *advance information* about all actions' losses across all time[2]. Let $\mathscr{I} = \{1,2,\ldots,N\}^T$ represent the set of *all* possible sequences of actions at times $t = 1,2,\ldots,T$. Show that, when T and N are large, any algorithm that the designer produces *cannot* achieve sublinear regret (in expectation over the algorithm's randomness) with respect to the best-performing sequence in the class $\mathscr{I}$. (You may assume that the minimax lower bound we proved in class for the absolute loss also applies to linear losses as in our case.)
   (b) (6 points) Consider the following compromise. For a sequence of actions $(i_1,\ldots,i_T) \in \mathscr{I}$, define its *complexity* to be $c(i_1,\ldots,i_T) := \sum_{t=2}^{T} \mathbb{1}\{i_t \neq i_{t-1}\}$, i.e., the number of times the sequence switches actions, and let $\mathscr{I}_m := \{x \in \mathscr{I} : c(x) \leq m\}$, where $m \leq T-1$. Give an online algorithm[3] that achieves (expected) regret $O\left(\sqrt{\frac{T}{2}\left((m+1)\log N + m \log\frac{T}{m}\right)}\right)$ with respect to the best sequence of actions[4] in $\mathscr{I}_m$.

3. (4 points) Show that the entropic regularizer $R(w) := \sum_{i=1}^{N} w_i \log w_i$ is $\frac{1}{B}$-strongly convex over $\{w \in \mathbb{R}_+^N : ||w||_1 \leq B\}$ with respect to the $||\cdot||_1$ norm, for $B > 0$.

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

[2] Denoting the loss of action i at time t by $l(i,t)$, the best offline solution, with information in advance, plays arm $\arg\min_i l(i,t)$ at each time t. This is the least loss that can ever be suffered assuming only one arm can be played at any instant.

[3] Disregard issues of computational efficiency.

[4] Feel free to use the bound $\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$, $1 \leq k \leq n$.

4. *Sequential probability estimation (5 points)*
Consider estimating the probability of a string of symbols from an alphabet $\mathscr{Y}$, $|\mathscr{Y}| = m$, with respect to the log loss. Let $\mathscr{F}$ be the class of all i.i.d. experts, i.e., all experts f that predict conditional probability $f_t(j|y^{t-1}) = f(j)$ (with $f(j) > 0$, $\sum_{j=1}^{m} f(j) = 1$) independently of t and y^{t-1}. For a particular sequence $y^T \in \mathscr{Y}^T$, determine the best expert (i.e., the one having the smallest cumulative loss) in $\mathscr{F}$ and its cumulative loss.

5. (8 points) Establish the following lemmas used in the proof of the regret bound for the Universal Portfolio algorithm.

   (a) (5 points) 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[5] of a set $A \subseteq \Delta_m$, then show that $\text{Vol}(\text{Ball}_\varepsilon(b^*)) = \varepsilon^{m-1}\text{Vol}(\Delta_m)$.

   (b) (3 points) Show that any 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.

6. *Online Gradient Descent for portfolio selection (6 points)*
Consider running (projected) Online Gradient Descent for the universal portfolio problem with m stocks. Under the condition that the market vectors $x_t \equiv (x_{1,t}, \ldots, x_{m,t})$ satisfy $1 \geq x_{i,t} \geq \varepsilon > 0$ for each $1 \leq i \leq m$ and $t \geq 1$, show that (projected) OGD, with an appropriately tuned learning rate, gets regret at most $\frac{\sqrt{2mT}}{\varepsilon}$ at the end of T rounds[6].

[5] Alternatively, $\text{Vol}(A)$ can be defined to be the probability of a point lying in the set A when it is drawn uniformly from Δ_m.

[6] Note that OGD is an efficient algorithm as compared to Cover's Universal Portfolio algorithm, but its regret performance is worse (UP enjoys $O(m \log T)$ regret).