Question 4 revised Sat Mar 1 13:37:28 PST 2014

1. (Adapting to $T$)

We saw in lecture 9 that the Hedge algorithm with learning rate \( \eta = \sqrt{\frac{8 \ln \log K}{T}} \) has regret after \( T \) rounds bounded by \( \sqrt{\frac{T}{2 \ln K}} \). In practice, we may not know \( T \) in advance, or we may even desire an algorithm that has good guarantees for all \( T \) simultaneously, i.e. that keeps on operating forever. Consider the following idea, called the **doubling trick**, to accomplish this. We run Hedge with \( \eta \) tuned for 1 round. After that, we restart Hedge, now with \( \eta \) tuned for 2 rounds. After that, we restart Hedge again with \( \eta \) tuned for 4 rounds, and so on.

(a) Prove that the overall accumulated regret of Hedge with the doubling trick is bounded above by \((1 + \sqrt{2}) \sqrt{T \ln K}\).

(b) What about the **tripling trick** and friends? Prove a regret bound for restarting at exponentially sized segments, where the \( i \)th segment is of size \( \phi^i - 1 \) for some basis \( \phi > 1 \). Judging from your bound, what is the best choice of basis \( \phi \)?

2. (Why FTL fails)

Consider the dot-loss game (Figure 1) with \( K \) experts and \( T \) rounds.

(a) The Follow-the-Leader (FTL) strategy plays \( w_t = e_{k_t} \), where \( k_t \in \arg \min_k \sum_{s=1}^{t-1} \ell_s,k \), breaking ties arbitrarily. Construct a strategy for the adversary (that is, a recipe to choose the next loss vector given the past weights and losses, and the weights played by FTL in the current round) that ensures that FTL incurs regret \( T \frac{K-1}{K} \).

(b) We call a strategy for Learner **deterministic** if, at each round, its action \( w_t \) is concentrated on a single expert, i.e. for all \( t \) there is an expert \( k \) such that \( w_t = e_k \). Show that any deterministic strategy can be made to suffer regret \( T \frac{K-1}{K} \).

3. (An especially lucky case)

In this question we investigate the effect of having a **perfect** expert in the dot-loss game (Figure 1). We implement this as a restriction on the Adversary, that is, the Adversary has to ensure that at least one expert keeps zero cumulative loss. Construct a strategy for the Learner that, under this restriction, keeps the regret below \( R_T \leq \ln K \) for all \( T \).
4. (FPL as Hedge)

Recall that Follow the Perturbed Leader with learning rate \( \eta > 0 \) chooses the expert \( k \) that minimizes \( L_k + X_k/\eta \), where \( X_k \) are i.i.d. perturbations. Here we consider the standard Gumbel distribution, which has CDF

\[
P(X \leq x) = \exp(-\exp(-x)).
\]

Show that the probability that FPL with negative standard Gumbel perturbations selects expert \( k \) is given by

\[
P\left\{ k = \arg\min_j \left( L_j + \frac{-X_j}{\eta} \right) \right\} = \frac{e^{-\eta L_k}}{\sum_j e^{-\eta L_j}}.
\]

5. (NML, SNML, KT)

Let \( \{P_\theta | \theta \in \Theta\} \) be a parametric model. Recall that the Normalized Maximum Likelihood strategy for \( T \) rounds assigns to data \( x_1, \ldots, x_T \) probability

\[
P_{\text{NML}}(x_1, \ldots, x_T) = \frac{\sup_{\theta \in \Theta} P_\theta(x_1, \ldots, x_T)}{\sum_{x_1, \ldots, x_T} \sup_{\theta \in \Theta} P_\theta(x_1, \ldots, x_T)}.
\]

In this question we consider the model of i.i.d. Bernoulli distributions. That is, \( \mathcal{X} = \{0, 1\} \), \( \Theta = [0, 1] \) and \( P_\theta(x_1, \ldots, x_T) = \prod_{t=1}^T \theta^{x_t} (1 - \theta)^{1-x_t} \).

(a) Compute a closed-form expression for \( P_{\text{NML}}(x_1, \ldots, x_T) \). For \( 1 \leq t \leq T \), write an expression for the prediction \( P_{\text{NML}}(x_t = 1 | x_1, \ldots, x_{t-1}) \).

(b) The Sequential Normalized Maximum Likelihood (or Last Step Minimax) strategy predicts with

\[
P_{\text{SNML}}(x_t | x_1, \ldots, x_{t-1}) = \frac{\sup_{\theta \in \Theta} P_\theta(x_1, \ldots, x_t)}{\sum_{x_t} \sup_{\theta \in \Theta} P_\theta(x_1, \ldots, x_t)}
\]

Simplify the expression for the prediction \( P_{\text{SNML}}(x_t | x_1, \ldots, x_{t-1}) \) as far as possible.

(c) For \( T = 1, \ldots, 200 \), graph the worst-case regret of NML, SNML and the Krichevsky-Trofimov estimator, defined by \( P_{\text{KT}}(x_t = 1 | x_1, \ldots, x_{t-1}) = \frac{1}{t} \left( \sum_{s=1}^{t-1} x_s + \frac{1}{2} \right) \), which predicts with the \( \frac{1}{2} \)-smoothed empirical frequency.

To compute the worst-case regret of SNML and KT, use the observation that the prediction of either depends on the past data sequence only through the number of zeros and ones. Let

\[
M_P(n_0, n_1) = \max_{x_1, \ldots, x_{n_0+n_1}} \sum_{t=1}^{n_0+n_1} -\ln P(x_t | x_1, \ldots, x_{t-1})
\]

be the maximum loss of strategy \( P \) on sequences with \( n_0 \) zeros and \( n_1 \) ones. The value of \( M_P \) can be computed by making use of the recurrence

\[
M_P(n_0, n_1) = \max \{ M_P(n_0 - 1, n_1) - \ln P(0 | n_0 - 1, n_1), M_P(n_0, n_1 - 1) - \ln P(1 | n_0, n_1 - 1) \}\]

when \( n_0 > 0 \) and \( n_1 > 0 \).