8803 Connections between Learning, Game Theory, and Optimization

Homework # 1 **Due: September 30th 2010**

This homework is due by the start of class on September 30th 2010. You can either submit the homework via the course page on T-Square or hand it in at the beginning of the class on September 30th 2010. Start early!

Groundrules:

- Your work will be graded on correctness, clarity, and conciseness.
- You may collaborate with others on this problem set. However, you must *write your own solutions* and *list your collaborators/sources* for each problem.

Problems:

1. **Weighted-majority.** For this problem you may use either the deterministic or randomized weighted-majority algorithm.

   a) Suppose we have some initial belief about which expert is likely to be the best one. In that case, a natural modification to the Weighted-Majority algorithm is that instead of initializing all the weights to 1, we instead initialize $w_i = p_i$, where p_i is our initial belief that expert i is going to be best ($\sum_{i=1}^{n} p_i = 1$). Show how this results in a bound where the $\ln n$ term is replaced with $\ln(1/p_i)$. For example, if you pick the randomized algorithm, you should get a bound of:

   $$M \leq \frac{1}{\epsilon}\left[m_i \ln(1/(1-\epsilon)) + \ln(1/p_i)\right],$$

   where m_i is the number of mistakes of expert i. So, this bound is better if our prior beliefs turn out to be reasonable.[1]

   b) What if we have infinitely many experts? Use your answer to part (a) to show how you can replace $\ln n$ with $O(\log i)$ in comparing our performance to that of the ith expert.

2. **Tracking a moving target.** Here is a variation on the deterministic Weighted-Majority algorithm, designed to make it more adaptive.

   (a) Each expert begins with weight 1 (as before).

   (b) We predict the result of a weighted-majority vote of the experts (as before).

   (c) If an expert makes a mistake, we penalize it by dividing its weight by 2, but *only* if its weight was at least 1/4 of the average weight of experts.

   Prove that in any contiguous block of trials (e.g., the 51st example through the 77th example), the number of mistakes made by the algorithm is at most $O(m+\log n)$, where m is the number of mistakes made by the best expert *in that block*, and n is the total number of experts.

[1] Notice that in this analysis we are *not* assuming that the best expert is actually picked from our prior. We simply are producing a bound that depends on what our beliefs were.

3. **2-Player Zero Sum Games.** In this problem, you will prove that the Nash equilibria of a 2-player zero sum game have several interesting properties. First, they all exhibit the same value. Second, they are *interchangeable*, meaning that given two Nash equilibrium points (σ_1, σ_2) and (τ_1, τ_2), the strategy pairs (τ_1, σ_2) and (σ_1, τ_2) are also Nash equilibria.

   Specifically, let G be a two person zero sum game, let A_i be the set of possible pure strategies for player i, $i = 1, 2$ and let $\pi : A_1 \times A_2 \to R$ be the function describing the payoff value for player I, or the loss value for player II. The goal of player I is to maximize π, while the goal of player II is to minimize π. Let (τ_1, τ_2) and (σ_1, σ_2) be two (mixed) Nash equilibria for G. Prove:

   (a) $\pi(\tau_1, \tau_2) = \pi(\tau_1, \sigma_2) = \pi(\sigma_1, \tau_2) = \pi(\sigma_1, \sigma_2)$

   (b) Both (σ_1, τ_2) and (τ_1, σ_2) are Nash equilibria of G.

   Conclude that the set of Nash equilibrium points of a 2-player zero sum game is the Cartesian product of the equilibrium strategies of each player.

4. **External regret vs Swap regret.** In Rock-Paper-Scissors, Rock beats Scissors (winner has loss 0, loser has loss 1), Scissors beats Paper, and Paper beats Rock; if both players play the same action, they tie (each gets loss of 1/2).[2]

   Consider playing T games of Rock-Paper-Scissors against an opponent who first plays Rock $T/3$ times, then plays Scissors $T/3$ times, then plays Paper $T/3$ times.[3]

   (a) Thinking of Rock, Paper, and Scissors as three "experts", describe in words what Randomized Weighted Majority would do against such an opponent. To be concrete, consider the version of RWM that, when expert i incurs loss ℓ, updates using $w_i \leftarrow w_i(1-\epsilon)^\ell$. Assume a learning rate $\epsilon \gg 1/T$, or if you like, you can think of $\lim_{\epsilon \to 1}$. Approximately (ignoring terms that are $o(T)$) what is the total loss of RWM and how does that compare to the loss of the best expert?

   (b) What approximately is the *swap regret* of RWM?

   (c) Since external regret is defined as the difference between the loss of the algorithm and the loss of the best expert, any two sequences of actions with the same total loss will result in the same external regret. Is this true for swap regret? In the context of this Rock-Paper-Scissors example, is there a behavior with approximately the same total loss as RWM but with much less swap regret?

Extra Credit:

1. **Minimax Theorem** Exhibit an example showing that the minimax theorem is not true in non-finite two player zero sum games (i.e., in games where players could have an infinite number of strategies.)

[2]Note the game is described as a constant sum game where the goal of each player is to minimize its own loss.

[3]Of course if we knew in advance this is going to happen we could play to win everytime, but we will instead analyze the behavior of a natural algorithm here.