CS 170 Final Review: Multiplicative Weights

1 Tell Me Weather It Will Rain

There are 1000 mysterious email addresses which have each been sending you an email every morning at 6 am. (Each morning, you get 1000 emails, one from each address). Each email contains only the single word ”rain” or ”sun.” Spoooooky! You realize that these are each predictions for the day's weather, and that one of the email addresses (you don't know which) is the Infallible Weather Channel, which is always correct. The other email addresses are trolls.

Every say, you must decide whether or not to take an umbrella on your commute; afterwards, you learn whether it actually rains. You have made a mistake that day if you bring an umbrella and it does not rain, or you do not bring an umbrella and it does. This process goes on forever.

Find an algorithm, which, given the history of predictions up until the current day, outputs whether or not to take an umbrella that day. Your algorithm should guarantee that, no matter how long the process goes on, you will make at most 10 mistakes, ever, despite the fact that the trolls have hacked into your computer and will know your algorithm.

(a) Clearly describe your algorithm and justify its correctness.

(b) Argue that the algorithm you gave for part (a) is the same as the Multiplicative Weights algorithm for a particular choice of ϵ. Also give the correct formula for the weight of every email address (expert) i at time t.

Solution:

(a) We maintain a list of experts (email addresses) we will consider listening to, initially containing all the email addresses. Each day, we will pick whichever option (rain or sun) has the most votes from experts on our list, breaking ties arbitrarily. After observing whether it rains or shines that day, we will eliminate every expert who was wrong from our list. Each time we make a mistake, we eliminate at least half the experts from our list. But there exists a perfect expert, so our list can never be empty. This means that we can make at most $\log n$ mistakes. Since $\log 1000 < 10$, you will make at most 10 mistakes, at which point only the Infallible Weather Channel will remain.

(b) The weight of expert i at time t is equal to $w_i^{(t+1)} = w_i^{(t)} \cdot (0)^{l_i^{(t)}}$, which is equivalent to $w_i^{(t+1)} = w_i^{(t)} \cdot (1-\epsilon)^{l_i^{(t)}}$ with $\epsilon = 1$. For every expert, if they gave the correct prediction, then their loss $l_i^{(t)}$ for that day is 0, and if they were wrong, $l_i^{(t)}$ is 1.

2 Zero Sum Games

In this problem we will prove the Minimax Theorem for two-player zero-sum games. Recall that the minimax theorem states:

$$\text{max}_x \text{min}_y \Sigma_{i,j} G_i j x_i y_j = \text{min}_y \text{max}_x \Sigma_{i,j} G_i j x_i y_j$$

Duality equalizes the two equations. By assuming weak duality, we can construct a lower bound. In this question, we will prove the upper bound.

Suppose that A is the payoff matrix for the first player in a two-player zero-sum game. Let the two players play repeated games against each other while both running the multiplicative weight algorithm, using the win/losses of the game as their respective loss vector. In other words, at each time step t, player 1 comes up with a probability vector $x^{(t)}$ and player 2 comes up with a probability vector $y^{(t)}$.

(a) What is each player's loss vector on day t? What is each player's loss on day t?

(b) The total regret of Multiplicative Weights run for T steps with parameter $0 < \epsilon \leq \frac{1}{2}$ is

$$R_T \leq (b - a) * (\epsilon T + \frac{\ln n}{\epsilon})$$

where b is the biggest value the loss can be and a is the smallest.

Given this knowledge, what is the tightest regret bound for player 1? What is the tightest regret bound for player 2? For simplicity write the bound in big-O notation.

Hint: What is the epsilon that makes the regret bound the tightest?

(c) Using the previous part, what is the tightest bound you can give on the average regret per round for each player?

(d) What is the formula that defines average regret? Plug in the loss functions for each player from part (a), and rewrite player 1's average regret so that it uses a maximum instead of a minimum.

(e) Now suppose we want to look at the sum of the average losses of each player. Using the previous parts, what is the formula for this quantity? (Hint: some terms should cancel) What big-O bound can we give on it?

(f) Rewrite your formula from the previous part using $\overline{x}$ and $\overline{y}$, the average strategies of each player, in order to eliminate the summations.

(g) We have that

$$\min_{y\in\Delta}\max_{x\in\Delta} x^T Ay - \max_{x\in\Delta}\min_{y\in\Delta} x^T Ay \geq 0$$

Use the previous parts to give an upper bound on this quantity.

(h) Prove that the quantity from the previous part must, in fact, be zero.

Solution:

(a) Player 1's loss vector is $-Ay^{(t)}$. Player 2's loss vector is $A^T x^{(t)}$.
Player 1's loss is $-(x^{(t)})^T Ay^{(t)}$. Player 2's loss is $(x^{(t)})^T Ay^{(t)}$.

(b) To get the tightest bound, you want to minimize the bound $\epsilon T + \frac{\ln n}{\epsilon}$. To do this take the derivative with respect to ϵ, set it equal to zero.

$$\frac{d}{d\epsilon} = (b-a)(T - \frac{\ln n}{\epsilon^2}) = 0$$

Solving this you get that ϵ should be $\sqrt{\frac{\ln n}{T}}$. For this choice of ϵ, we get the regret bound $2(b-a)\sqrt{T \ln n}$. Ignoring the constants, we get a total regret bound of $O(\sqrt{T \ln n})$ for each player.

(c) Since there are T days, we simply divide our answers to the previous part by T. Thus, we get a regret bound of $O\left(\sqrt{\frac{\ln n}{T}}\right)$ for each player.

(d) The definition of regret is the loss we incurred minus the minimum loss of any fixed strategy. Formally, if our strategy at time t is s_t and our loss function at step t is L_t, we have our average regret as

$$\frac{1}{T}\left(\sum_{t=1}^{T} L_t(s_t) - \min_{s \in \Delta} \sum_{t=1}^{T} L_t(s)\right)$$

Plugging in our loss function for player 2, this becomes

$$\frac{1}{T}\left(\sum_{t=1}^{T} (x^{(t)})^T A y^{(t)} - \min_{y \in \Delta} \sum_{t=1}^{T} (x^{(t)})^T A y\right)$$

Similarly for player 1, our average regret becomes

$$\frac{1}{T}\left(\sum_{t=1}^{T} -(x^{(t)})^T A y^{(t)} - \min_{x \in \Delta} \sum_{t=1}^{T} -x^T A y^{(t)}\right)$$

We note that $\min_x f(x) = -\max_x -f(x)$, so we can rewrite this as

$$\frac{1}{T}\left(\sum_{t=1}^{T} -(x^{(t)})^T A y^{(t)} + \max_{x \in \Delta} \sum_{t=1}^{T} x^T A y^{(t)}\right)$$

(e) Adding together the two formulae from the previous part, and noting that the first summations in each are negations of each other, we get that the sum of the two average regrets is

$$\frac{1}{T}\left(\max_{x \in \Delta} \sum_{t=1}^{T} x^T A y^{(t)} - \min_{y \in \Delta} \sum_{t=1}^{T} (x^{(t)})^T A y\right)$$

In terms of a big-O bound, we know that each player's average loss is bounded by $O\left(\sqrt{\frac{\ln n}{T}}\right)$, so their sum will also be bounded by this.

(f) We note that x, A, and y do not depend on t, so we can factor them out of the relevant summations. This gives us the formula

$$\frac{1}{T}\left(\max_{x \in \Delta} x^T A \left(\sum_{t=1}^{T} y^{(t)}\right) - \min_{y \in \Delta} \left(\sum_{t=1}^{T} (x^{(t)})^T\right) A y\right)$$

We can then distribute the $\frac{1}{T}$ out, and noting that $\overline{x} = \frac{1}{T}\sum_t x^{(t)}$ and $\overline{y} = \frac{1}{T}\sum_t y^{(t)}$, this becomes

$$\max_{x \in \Delta} x^T A \overline{y} - \min_{y \in \Delta} \overline{x}^T A y$$

(g) The previous parts tell us that

$$\max_{x\in\Delta} x^T A\overline{y} - \min_{y\in\Delta} \overline{x}^T Ay \leq O\left(\sqrt{\frac{\ln n}{T}}\right)$$

If we replace $\overline{y}$ with the y that minimizes $\max_x x^T Ay$, we'll certainly shrink the first term. Similarly, if we replace $\overline{x}$ with the x that maximizes $\min_y x^T Ay$, we'll grow the second term; since that second term is subtracted, this will shrink the overall expression. Thus, we get

$$\min_{y\in\Delta}\max_{x\in\Delta} x^T Ay - \max_{x\in\Delta}\min_{y\in\Delta} x^T Ay \leq \max_{x\in\Delta} x^T A\overline{y} - \min_{y\in\Delta} \overline{x}^T Ay \leq O\left(\sqrt{\frac{\ln n}{T}}\right)$$

(h) The previous part tells us that the difference we're interested in is bounded by $c\sqrt{\frac{\ln n}{T}}$ for all T, where c is some universal constant. But we note that the difference does not depend on T, so we can take T to be arbitrarily large, and hence make our upper bound arbitrarily small. The only way for a non-negative number to have arbitrarily small upper bounds is if it is zero, so we in fact have

$$\min_{y\in\Delta}\max_{x\in\Delta} x^T Ay - \max_{x\in\Delta}\min_{y\in\Delta} x^T Ay = 0$$