Lecture 17 – April 6, 2016

Prof. Ankur Moitra *Scribe: Rishad Rahman, Paolo Gentili*

1 Multiplicative Weights

1.1 Setup

We have access to n experts whose advice (e.g. financial advice, stock picks) we may choose to take on each of T days. Define f_i^t as

$$f_i^t := \text{loss of following the advice of expert } i \text{ on day } t.$$

We assume that the losses are bounded ($\|f^t\|_\infty \leq 1$) but are arbitrary.

Our goal is to perform almost as well as the best expert did in hindsight, or in more precise terms, to minimize regret where regret is defined as follows:

$$regret := \sum_{t=1}^{T} \langle p^t, f^t \rangle - \min_i \sum_{t=1}^{T} f_i^t.$$

Here, p^t represents the indicator of the expert we choose on day t or the distribution of experts that we consider on day t. The first term in the definition of regret sums over our expected loss on each day t, while the second term calculates the loss of the best performing expert.

1.2 Attempt 1: Naive Algorithm

A simple idea might be to select the best expert so far. However, this can perform very poorly if we are unlucky or if losses are chosen adversarially. For example, consider the following vectors representing the losses $f^1, f^2, \ldots$:

$$\begin{bmatrix} 0 \\ \frac{1}{n} \\ \frac{2}{n} \\ \vdots \\ \frac{n-1}{n} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots$$

The best performing expert up to a certain day t cycles through our n experts, and on each day we happen to follow the advice of the expert that does the worst that day. Our loss through T days is $T - O(1)$, while that of the best expert is $\frac{T-O(1)}{n}$.

1.3 Attempt 2: Weighted Majority Algorithm

If we are predicting binary events, we can assign weights to the experts and make our prediction according to the weighted majority. We can show that our loss is $\lesssim 2 \times$ *Best Expert*, but we did not perform the analysis of this algorithm in class. The analysis can be found here: http://courses.cs.washington.edu/courses/cse521/10wi/kale-thesis-chap2.pdf

1.4 Attempt 3: Multiplicative Weights Update

The Multiplicative Weights algorithm is as follows:

$w_i^1 \leftarrow 1 \quad \forall i$
for $t = 1 \ldots T$ **do**
  Follow expert i with probability $p_i^t := \frac{w_i^t}{\sum_j w_j^t}$
  $w_i^{t+1} \leftarrow w_i^t(1 - \epsilon f_i^t) \quad \forall i$
end for

Theorem 1. *(Multiplicative Weights Update) If $0 < \epsilon \leq \frac{1}{2}$, then using MWU gives the following bound on our regret:*

$$regret \leq \frac{\ln n}{\epsilon} + \epsilon T$$

If we know T beforehand, then we may set $\epsilon = \sqrt{\frac{\ln n}{T}}$ to get

$$regret \leq 2\sqrt{T \ln n}$$

While our total regret does increase over time, we note that our average regret per day goes to 0 as T goes to infinity.

Proof. We define the following potential function:

$$\Phi^t := \sum_{i=1}^{n} w_i^t.$$

We will show that Φ decreases as we incur loss, but Φ cannot be too small if there is a good expert. The first fact will provide an upper bound on Φ, while the second will provide a lower bound, and combining the two will yield our desired result.

We have $\Phi^{t+1} = \sum_{i=1}^n w_i^{t+1} = \sum_{i=1}^n w_i^t(1 - \epsilon f_i^t)$ and $w_i^t = p_i^t \Phi^t$, so

$$\begin{aligned}
\Phi^{t+1} &= \sum_{i=1}^{n} p_i^t \Phi^t (1 - \epsilon f_i^t) \\
&= \Phi^t \left(\sum_{i=1}^{n} p_i^t - \sum_{i=1}^{n} p_i^t \epsilon f_i^t \right) \\
&= \Phi^t \left(1 - \epsilon \langle p^t, f^t \rangle\right) \\
&\leq \Phi^t \exp(-\epsilon \langle p^t, f^t \rangle)
\end{aligned}$$

where we make use of the fact that $1 - x \leq e^{-x}$. Repeatedly applying this inequality yields

$$\Phi^{T+1} \leq \Phi^1 \exp\left(-\epsilon \sum_{t=1}^{T} \langle p^t, f^t \rangle\right) = n \exp\left(-\epsilon \sum_{t=1}^{T} \langle p^t, f^t \rangle\right),$$

providing our upper bound.

Now, we note that for all i,

$$\begin{aligned}
\Phi^{T+1} &\geq w_i^{T+1} \\
&= \prod_{t=1}^{T} (1 - \epsilon f_i^t) \\
&\geq \prod_{t=1}^{T} \exp\left(-\epsilon f_i^t - \epsilon^2 \left(f_i^t\right)^2\right) \\
&= \exp\left(-\epsilon \sum_{t=1}^{T} f_i^t - \epsilon^2 \sum_{t=1}^{T} \left(f_i^t\right)^2\right)
\end{aligned}$$

where we make use of the fact that $1 - x \geq e^{-x-x^2}$. This provides our lower bound. Combining the two, we get

$$\begin{aligned}
&\exp\left(-\epsilon \sum_{t=1}^{T} f_i^t - \epsilon^2 \sum_{t=1}^{T} \left(f_i^t\right)^2\right) \leq n \exp\left(-\epsilon \sum_{t=1}^{T} \langle p^t, f^t \rangle\right) \\
&\implies \ln\left(\exp\left(-\epsilon \sum_{t=1}^{T} f_i^t - \epsilon^2 \sum_{t=1}^{T} \left(f_i^t\right)^2\right)\right) \leq \ln\left(n \exp\left(-\epsilon \sum_{t=1}^{T} \langle p^t, f^t \rangle\right)\right) \\
&\implies -\epsilon \sum_{t=1}^{T} f_i^t - \epsilon^2 \sum_{t=1}^{T} \left(f_i^t\right)^2 \leq \ln n - \epsilon \sum_{t=1}^{T} \langle p^t, f^t \rangle \\
&\implies \epsilon \sum_{t=1}^{T} \langle p^t, f^t \rangle - \epsilon \sum_{t=1}^{T} f_i^t \leq \ln n - \epsilon^2 \sum_{t=1}^{T} \left(f_i^t\right)^2 \\
&\implies \textit{regret} \leq \frac{\ln n}{\epsilon} - \epsilon T
\end{aligned}$$

where in the final implication, we use the fact that $\sum_{t=1}^{T} \left(f_i^t\right)^2 \leq T$ because of the bounds on f^t. $\square$

We note that the only assumption we have used on the f^t is that they are bounded. Thus, we achieve this bound on regret even if an adversary knows our strategy and our probability distributions p^t.

If we have a different bound on the losses, $\|f^t\|_\infty \leq \rho$, we may set $w_i^{t+1} = w_i^t(1 - \frac{\epsilon}{\rho} f_i^t)$ to get

$$\textit{regret} \leq \frac{\rho^2 \ln n}{\epsilon} + \epsilon T$$

2 Application to Zero Sum Games

Recall that a zero sum game involves Alice and Bob picking strategies $i \in [1, m]$ and $j \in [1, n]$ respectively with Alice's payoff (and Bob's loss) being given by entry $M_{i,j}$ in the payoff matrix M. Previously, we reduced the problem of maximizing expected payoff to a linear program and showed it was equivalent to minimizing expected payoff via strong duality. We now analyze what happens when Alice and Bob both use multiplicative weights to update their mixed strategies over T days, $(p^1, q^1), (p^2, q^2) \ldots (p^T, q^T)$, to minimize their expected regret.

Our update algorithm tells us Alice's expected regret on day t with strategy q^t is

$$-\sum_{t=1}^{T} \langle p^t, Aq^t \rangle + \max_i \sum_{t=1}^{T} (Aq^t)_i \leq \frac{\ln n}{\epsilon} + \epsilon T$$

Note that we are looking at Alice's payoff instead of loss hence the sign differences and taking the maximum advice. Now Bob's expected regret is

$$\sum_{t=1}^{T} \langle p^t, Aq^t \rangle - \min_j \sum_{t=1}^{T} (A^{\mathbf{T}} p^t)_j \leq \frac{\ln n}{\epsilon} + \epsilon T$$

We set $\epsilon = \sqrt{\frac{\ln n}{T}}$ and sum up the above to obtain

$$\max_i \sum_{t=1}^{T} (Aq^t)_i - \min_j \sum_{t=1}^{T} (A^{\mathbf{T}} p^t)_j \leq \frac{2 \ln n}{\epsilon} + 2\epsilon T \leq 4\sqrt{\frac{\ln n}{T}} := \delta$$

Define the mean strategies $\overline{p} = \frac{1}{T} \sum_{t=1}^{T} p^t$ and $\overline{q} = \frac{1}{T} \sum_{t=1}^{T} p^t$ to simplify the above to

$$\max_i (A\overline{q})_i - \min_j (A^{\mathbf{T}} \overline{p})_j \leq \delta$$

Note that the two quantities being subtracted represent pure strategies Alice and Bob will take if the adversary plays with the mean strategies. Piecing this together we have

$$\min_j (A^{\mathbf{T}} \overline{p})_j \leq \overline{p}^{\mathbf{T}} A \overline{q} \leq \max_i (A\overline{q})_i$$

where the bounds are off by at most δ. For large T, δ is small so the strategies $(\overline{p}, \overline{q})$ are approximately in equilibrium with the optimum obtained by linear programming and neither player can exceed it by more than δ.