CS269: Machine Learning Theory
Lecture 12: Finishing Up FTRL and Alternate Goals for Online Learning
November 3, 2010

Lecturer: Jennifer Wortman Vaughan
Scribe: David Nicholaeff and Daniel Sebastian Pan

1 Randomized Weighted Majority Regret Bound, alternate proof

Last time we began an alternate proof of the Randomized Weighted Majority regret bound using the general regret bound that we derived for the family of *Follow the Regularized Leader* algorithms (parameterized by η and the regularizer). The regret bound is as follows. For all $\vec{w} \in \mathcal{K}$,

$$\sum_{t=1}^{T} l_t(\vec{w}_t) - \sum_{t=1}^{T} l_t(\vec{w}) \leq \sum_{t=1}^{T} \big(l_t(\vec{w}_t) - l_t(\vec{w}_{t+1})\big) + \frac{1}{\eta}\big(R(\vec{w}) - R(\vec{w}_1)\big)$$

where the weight vector at time step t is defined as

$$\vec{w}_t = \arg\min_{\vec{w} \in \mathcal{K}} \left(\eta \sum_{s=1}^{t-1} l_s(\vec{w}) + R(\vec{w}) \right)$$

The first term in the $\arg\min$ represents the empirical loss based on the data encountered up to time step t, while the second term provides a means of stabilizing the data. Note that the regret bound above holds in a more generalized optimization setting as well. The first term in the bound, the sum over t of $l_t(\vec{w}_t) - l_t(\vec{w}_{t+1})$, captures the change in the weights from one time step to the next, while the second term measures the difference between the regularizer's output for $\vec{w}$ and the minimum value of the regularizer. We proceed to bound both of these terms, starting with the regularizer.

1.1 Negative Entropy Regularizer

In prior lectures, we discussed the expert setting, which refers to the online convex optimization problem when we restrict $\mathcal{K}$ to be the probability simplex and the loss function to be linear (so we can write the algorithm's loss as a dot product). In other words, it is the setting we first described when we started talking about adversarial online learning, where the algorithm must choose a weighting over n "experts" at each time step. There are lots of different algorithms that you could choose to run in this setting. Randomized Weighted Majority is a particular algorithm that achieves a low regret. The weights for the Randomized Weighted Majority algorithm are given by

$$w_{i,t} = \frac{e^{-\eta L_{i,t-1}}}{\sum_{j=1}^{n} e^{-\eta L_{j,t-1}}}$$

It was shown in the last lecture that the Randomized Weighted Majority is equivalent to Follow the Regularized Leader when the regularizer is defined as negative entropy:

$$R(\vec{w}) = -\sum_{i=1}^{n} w_i \log \frac{1}{w_i}$$

This provided an intuitive notion of the regularizer as a form of equalizer akin to physical entropy. In particular, weight vectors that spread out their weight are rewarded, while weights that cause clumping on a single vector are punished. Denoted as H, it was also shown last lecture that the negative entropy obeys the following bound ($\forall \vec{w}$):

$$0 \leq H(\vec{w}) \leq \log n$$

which implies

$$\frac{1}{\eta}\big(R(\vec{w}) - R(\vec{w}_1)\big) \leq \frac{1}{\eta}\log n$$

1.2 How the weights change from one time step to the next

Lemma 1. *For the Randomized Weighted Majority, for all time steps t,*

$$\vec{l}_t \cdot \vec{w}_t - \vec{l}_t \cdot \vec{w}_{t+1} \leq 2\eta$$

Proof: First we note that this bound holds trivially for all $\eta \geq \frac{1}{2}$. Namely, $\vec{0} \leq \vec{w}_t \leq \vec{1}$ and $\vec{0} \leq \vec{l}_t \leq \vec{1}$ for all t, by definition of the weights and losses. Hence, the left side of the inequality has a maximum value of 1, corresponding to $\eta = \frac{1}{2}$. We therefore assume $\eta \leq \frac{1}{2}$ for the remainder of the proof.

Next, we examine the cumulative loss of expert i at time step t:

$$e^{-\eta L_{i,t}} = e^{-\eta(L_{i,t-1}+l_{i,t})} = e^{-\eta L_{i,t-1}} \cdot e^{-\eta l_{i,t}} \leq e^{-\eta L_{i,t-1}} \tag{1}$$

since $e^{-\eta l_{i,t}} \leq 1$ with equality holding when the loss is 0. Now we also note that this term is smallest when the loss is 1, giving us the lower bound

$$e^{-\eta L_{i,t-1}} \cdot e^{-\eta} \leq e^{-\eta L_{i,t}} \tag{2}$$

Combining these two bounds, we can find an upper bound on the weight of expert i at time step t. We do this by maximizing the numerator with the lower bound found in (2), and by minimizing each term in the sum of the denominator with the upper bound found in (1):

$$w_{i,t} = \frac{e^{-\eta L_{i,t-1}}}{\sum_{j=1}^{n} e^{-\eta L_{j,t-1}}} \leq \frac{e^{\eta} \cdot e^{-\eta L_{i,t}}}{\sum_{j=1}^{n} e^{-\eta L_{j,t}}} = e^{\eta} \cdot w_{i,t+1} \tag{3}$$

Finally, we evaluate the left side of the inequality we wish to prove using the inequality of (3), since it holds for all experts i:

$$\begin{aligned}
\vec{l}_t \cdot \vec{w}_t - \vec{l}_t \cdot \vec{w}_{t+1} &\leq \vec{l}_t \cdot \big(e^{\eta} \cdot \vec{w}_{t+1}\big) - \vec{l}_t \cdot \vec{w}_{t+1} \\
&= \vec{l}_t \cdot \vec{w}_{t+1} \cdot \big(e^{\eta} - 1\big) \\
&\leq (e^{\eta} - 1) \\
&\leq (e - 1) \cdot \eta \\
&\leq 2\eta
\end{aligned}$$

where the second inequality holds because we are only considering the case in which $\eta \leq 1/2$.

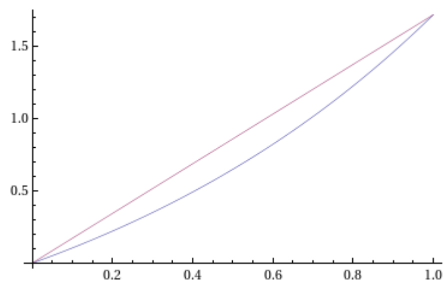

Figure 1: $e^\eta - 1 \leq (e-1)\cdot\eta$ for $\eta \leq 1$

One final note: we are being very loose with the constants here, because our goal is to use this bound to show that we can get the same basic dependence on T and n using the regularization based proof technique. $\square$

1.3 The Regret Bound of The Randomized Weighted Majority

Combining the two results of the previous sections we can now bound the regret:

$$
\begin{aligned}
\sum_{t=1}^{T} l_t(\vec{w}_t) - \sum_{t=1}^{T} l_t(\vec{w}) &\leq \sum_{t=1}^{T}\big(l_t(\vec{w}_t) - l_t(\vec{w}_{t-1})\big) + \frac{1}{\eta}\big(R(\vec{w}) - R(\vec{w}_1)\big) \\
&\leq (2\eta)\cdot T + \frac{1}{\eta}\log n \\
\text{Regret RWM} &\leq 2\sqrt{2T\log n} \qquad \left(\text{for } \eta^2 = \frac{\log n}{2T} \rightarrow \eta = \sqrt{\frac{\log n}{2T}}\right)
\end{aligned}
$$

This bound is sublinear in T and agrees with the bound for the Randomized Weighted Majority found in previous lectures. As well, it gives a nice dependence on the number of experts $(\log n)$. The significance here is that a completely different technique was used to derive the result. Namely, we appealed to the framework of regularization, and we used negative entropy as the regularizer.

As long as the regularizer is strongly convex, the upper bound on the regret will always be sublinear in T. Further, depending on the choice of regularizer, we can derive equivalences to other online learning algorithms. For example, the negative entropy regularizer developed an equivalence with the bound for the Randomized Weighted Majority algorithm. If we were to take $R(\vec{w}) = ||\vec{w}||_2$, we would find an equivalence in the bound on regret to that for the Online Gradient Descent algorithm. This demands significant consideration of the choice of regularizer to say the least.

2 Considering Dual Goals

Now let's recall the example that motivated our introduction of the regularizer in the first place. We considered the case where the best expert is changing from one time step to the next. For simplicity, we took the case of two experts oscillating in their losses. Suppose we were to run Randomized Weighted Majority on this sequence. Our weights at each time step would be defined by:

$$\begin{aligned} w_{1,t} = & \quad \frac{e^{-\eta L_{1,t-1}}}{e^{-\eta L_{1,t-1}} + e^{-\eta L_{2,t-1}}} & = \frac{e^{-\eta\left(L_{1,t-1} - L_{2,t-1}\right)}}{e^{-\eta\left(L_{1,t-1} - L_{2,t-1}\right)} + 1} \\ w_{2,t} = & \quad 1 - w_{1,t} & = \frac{1}{e^{-\eta\left(L_{1,t-1} - L_{2,t-1}\right)} + 1} \end{aligned}$$

Notice that we can write these weights in terms of the different between the two experts' losses. The following table summarizes the loss of Randomized Weighted Majority at each round.

t	1	2	3	4	5	...
Losses	(1,0)	(0,1)	(1,0)	(0,1)	(1,0)	...
$L_{1,t-1} - L_{2,t-1}$	0	1	0	1	0	...
Alg Loss	$\frac{1}{2}$	$(e^{-\eta} + 1)^{-1}$	$\frac{1}{2}$	$(e^{-\eta} + 1)^{-1}$	$\frac{1}{2}$	...

By inspection we see that the cumulative loss will be

$$\frac{T}{2} \cdot \frac{1}{2} + \frac{T}{2} \cdot \left(e^{-\eta} + 1\right)^{-1} \approx \frac{T}{2} + c\sqrt{T} \qquad (\text{for } \eta \ll 1)$$

since $\left(e^{-\eta} + 1\right)^{-1} \approx \frac{1}{2} + c\eta$ for a constant c and $\eta \sim T^{-1/2}$. We have therefore achieved a regret of $\mathcal{O}(\sqrt{T})$ to the best expert. But in this simple example, if we had chosen one of the experts rather than apply the RWM algorithm, our losses would have been lower, thus begging the question can we do better?

Well in the example above, if we chose *any* fixed distribution over the two experts and used that distribution on every time step, we would have a loss of only $T/2$, which is much better than what the Weighted Majority gets. This motivates the goal of trying to keep our guarantee of $\sqrt{T}$ regret to the best expert while simultaneously maintaining a constant regret to a fixed distribution over experts of our choosing (such as the average of all experts). This way, if there is one expert that does really well, we will still have reasonably good performance compared to this expert. On the other hand, if no experts do well, i.e., they all have comparable performance as in this example, then we will have only constant regret to them.

2.1 Dual Goal

Still working in the expert setting, we maintain the regret bound to the best expert at each time step, but now we establish a constant regret with respect to some distribution of experts. That is, we want to simultaneously satisfy the following conditions:

- <u>constant</u> regret to $\vec{w}_D$
- (close to) $\sqrt{T}$ regret to the best expert

where we define the regret to $\vec{w}_D$ in the standard way:

Definition 1.

$$\textit{Regret } \vec{w}_D = \sum_{t=1}^{T} \vec{l}_t \cdot \vec{w}_t - \sum_{t=1}^{T} \vec{l}_t \cdot \vec{w}_D$$

From this definition, we see that if we didn't care about the best expert, we would just pick the distribution, and then our regret would be equivalently 0. To implement this new strategy, we proceed by adding an additional "fake" expert chosen via the distribution. This is the D-Prod algorithm, a modification of the Prod algorithm[1]:

1. Create expert 0 as a weighted sum of the other, real experts. Then the loss of expert 0 at time t is $l_{0,t} = \vec{l}_t \cdot \vec{w}_D$.

2. Start with a non-uniform weighting ($\forall i \in \{0, 1, \ldots, n\}$):

$$w_{i,1} = \mu_i \quad \left(\text{where } \sum_{i=0}^{n} \mu_i = 1\right)$$

3. Maintain unnormalized weights for each of the experts, $\widetilde{w}_{i,t}$, where at each round:

$$\begin{aligned} \widetilde{w}_{i,t} &= \widetilde{w}_{i,t-1} \cdot \big(1 + \eta(l_{0,t-1} - l_{i,t-1})\big) \\ w_{i,t} &= \frac{\widetilde{w}_{i,t}}{\sum_{j=0}^{n} \widetilde{w}_{j,t}} \end{aligned}$$

Notice that for a higher loss relative to the artificial expert ($l_{i,t-1} > l_{0,t-1}$), we update the unnormalized weight by multiplying by a value less than 1, thus giving less stength to the i^{th} expert's contribution. Likewise, if there is a lower loss ($l_{i,t-1} < l_{0,t-1}$), we multiply by a value greater than 1, thereby increasing the contribution of the i^{th} expert to the artificial expert. Also note that the unnormalized weight for expert 0 will always remain unchanged, since the difference in losses for expert 0 will always be 0.

We now state the following lemma (not proved here).

Lemma 2. *For D-Prod* $(\vec{\mu}, \eta)$ *if* $\eta < \frac{1}{2}$, *then for any* $i \in \{0, 1, \ldots, n\}$

$$\sum_{t=1}^{T} \vec{l}_t \cdot \vec{w}_t - \sum_{t=1}^{T} l_{i,t} \leq \eta \sum_{t=1}^{T} (l_{0,t} - l_{i,t})^2 - \frac{\log \mu_i}{\eta}$$

By choosing μ and η in the appropriate manner, we can can obtain our goals above. Set

$$\eta = \sqrt{\frac{\log n}{T}}, \quad \mu_0 = 1 - \eta, \quad \mu_i = \frac{\eta}{n} \ \ \big(\forall i \in \{1, \ldots, n\}\big)$$

[1]See `http://www.springerlink.com/content/k670m78m75652854/fulltext.pdf` for more information.

We then find by applying the lemma:

$$
\begin{aligned}
\text{Regret to expert } i \neq 0 &\leq \eta T + \frac{1}{\eta}\log\left(\frac{1}{\mu_i}\right) \\
&= \sqrt{T\log n} + \sqrt{\frac{T}{\log n}} \cdot \log\left(\frac{n}{\eta}\right) \\
&= 2\sqrt{T\log n} + \sqrt{\frac{T}{\log n}} \cdot \log\sqrt{\frac{T}{\log n}} \\
&= \sqrt{T} \cdot \left[2\sqrt{\log n} + \sqrt{\frac{1}{\log n}}\log\sqrt{\frac{T}{\log n}}\right] \\
\text{Regret to } \vec{w}_D &\leq 0 - \frac{\log\mu_0}{\eta} \\
&= \frac{1}{\eta}\log\left(\frac{1}{1-\eta}\right) \\
&\leq \frac{1}{\eta}\left(\frac{1}{1-\eta} - 1\right) \qquad \text{(from } \log(1+x) \leq x) \\
&\leq 2(2-1) = 2 \qquad \text{(from } \eta \leq \frac{1}{2})
\end{aligned}
$$

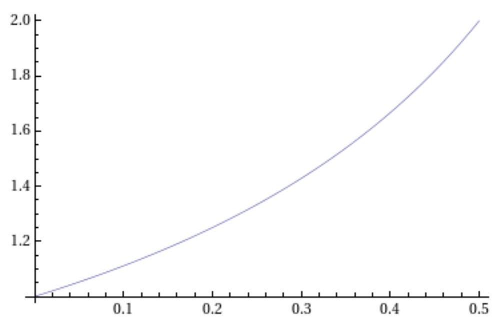

Figure 2: Plot of $\frac{1}{\eta}\left(\frac{1}{1-\eta} - 1\right)$ for $\eta \leq \frac{1}{2}$.

The purpose of this exercise in algebra with the lemma is that we obtain a result close to what we found previously, namely a $\sqrt{T}$ regret to the best expert with only an additional logarithmic factor added on, and now we have constant regret (at most two!) to a fixed distribution of experts.