Lecture 2

Instructor: Haipeng Luo

1 Hedge

The expert problem [Freund and Schapire, 1997] mentioned in last lecture turns out to play a fundamental role in online learning, and we will focus on this problem for a couple of lectures. The first native algorithm that comes to one's mind is probably the *follow the leader* (FTL) approach, which puts all the weights to the current best expert $i_t^\star = \mathrm{argmax}_{i\in[N]} - \sum_{\tau=1}^{t-1} \ell_s(i)$. It is not hard to see that such an approach does not work well in general, at least not in the adversarial setting.

It turns out, however, that simply replacing the "max" by some "softmax" would change to situation greatly. In fact, this leads to the the classic algorithm Hedge [Freund and Schapire, 1997] (generalizing [Littlestone and Warmuth, 1994]), also known as multiplicative weights update or simply exponential weights. Below is the pseudocode.

Algorithm 1: Hedge

Input: learning rate $\eta > 0$
Initialization: let $L_0 \in R^N$ be the all-zero vector
for $t = 1, \ldots, T$ **do**
  compute $p_t \in \Delta(N)$ such that $p_t(i) \propto \exp(-\eta L_{t-1}(i))$
  play p_t and observe loss vector $\ell_t \in [0,1]^N$
  update $L_t = L_{t-1} + \ell_t$

Recall the definition of regret for this setting:

$$\mathcal{R}_T = \sum_{t=1}^T \langle p_t, \ell_t\rangle - \min_{p\in\Delta(N)} \sum_{t=1}^T \langle p, \ell_t\rangle = \sum_{t=1}^T \langle p_t, \ell_t\rangle - \sum_{t=1}^T \ell_t(i^\star)$$

where $i^\star \in \mathrm{argmin}_i \sum_{t=1}^T \ell_t(i)$ is the best expert in hindsight. Hedge guarantees the following regret bound:

Theorem 1. *Hedge with learning rate η guarantees*

$$\mathcal{R}_T \le \frac{\ln N}{\eta} + \eta \sum_{t=1}^T \sum_{i=1}^N p_t(i)\ell_t^2(i) \tag{1}$$

$$\le \frac{\ln N}{\eta} + T\eta, \tag{2}$$

which is of order $\mathcal{O}(\sqrt{T\ln N})$ if η is optimally set to $\sqrt{(\ln N)/T}$.

There are many different proofs that lead to bound (2). Here we present a "potential-based" proof that obtains bound (1) as an intermediate step, which will turn out to be very useful later on.

Proof. Let $\Phi_t = \frac{1}{\eta}\ln\left(\sum_{i=1}^N \exp(-\eta L_t(i))\right)$. First note that

$$\Phi_t - \Phi_{t-1} = \frac{1}{\eta}\ln\left(\frac{\sum_{i=1}^N \exp(-\eta L_t(i))}{\sum_{i=1}^N \exp(-\eta L_{t-1}(i))}\right)$$

$$
\begin{aligned}
&= \frac{1}{\eta} \ln \left(\sum_{i=1}^{N} p_t(i) \exp(-\eta \ell_t(i)) \right) \\
&\le \frac{1}{\eta} \ln \left(\sum_{i=1}^{N} p_t(i) \left(1 - \eta \ell_t(i) + \eta^2 \ell_t^2(i)\right) \right) && (e^{-y} \le 1 - y + y^2 \text{ for all } y \ge 0) \\
&= \frac{1}{\eta} \ln \left(1 - \eta \langle p_t, \ell_t \rangle + \eta^2 \sum_{i=1}^{N} p_t(i) \ell_t^2(i) \right) \\
&\le - \langle p_t, \ell_t \rangle + \eta \sum_{i=1}^{N} p_t(i) \ell_t^2(i). && (\ln(1+y) \le y)
\end{aligned}
$$

Summing over t, telescoping and rearranging give

$$
\begin{aligned}
\sum_{t=1}^{T} \langle p_t, \ell_t \rangle &\le \Phi_0 - \Phi_T + \eta \sum_{t=1}^{T} \sum_{i=1}^{N} p_t(i) \ell_t^2(i) \\
&\le \frac{\ln N}{\eta} - \frac{1}{\eta} \ln \left(\exp(-\eta L_T(i^\star)) \right) + \eta \sum_{t=1}^{T} \sum_{i=1}^{N} p_t(i) \ell_t^2(i) \\
&\le \frac{\ln N}{\eta} + L_T(i^\star) + \eta \sum_{t=1}^{T} \sum_{i=1}^{N} p_t(i) \ell_t^2(i),
\end{aligned}
$$

which proves Eq. (1). Eq. (2) follows immediately by the boundedness of losses. □

Note that the regret of Hedge has only logarithmic dependence on N, which as we will see is very useful in solving many problems with huge number of experts.

2 Lower bound for the Expert Problem

Is the regret bound of Hedge good or bad? In general, how can we tell whether a regret upper bound is satisfactory or not? The notion of minimax regret can be used to answer these questions exactly. Intuitively, minimax regret is the smallest possible worst-case regret of any algorithm. For example, the minimax regret of the expert problem can be defined as

$$
\min_{\mathcal{A}} \max_{\ell_1, \ldots, \ell_T} \mathcal{R}_T
$$

where $\mathcal{A}$ is any legitimate expert algorithm. Note that $\mathcal{R}_T$ depends on both $\mathcal{A}$ and all the losses even if the dependence is not explicitly spelled out. Also note that in general $\mathcal{R}_T$ should be viewed as the expected regret if the algorithm is randomized. The existence of the Hedge algorithm already shows that

$$
\min_{\mathcal{A}} \max_{\ell_1, \ldots, \ell_T} \mathcal{R}_T \le 2\sqrt{T \ln N}.
$$

The following theorem proves that this bound is minimax optimal (up to a constant of $2\sqrt{2}$). In the proof we use an implicit and probabilistic construction of the environment, which is a very useful technique in proving lower bounds.

Theorem 2. *For any algorithm, we have*

$$
\sup_{T,N} \max_{\ell_1, \ldots, \ell_T} \frac{\mathcal{R}_T}{\sqrt{T \ln N}} \ge \frac{1}{\sqrt{2}}.
$$

Proof. Let $\mathcal{D}$ be the uniform distribution over $\{0, 1\}$. We have

$$\begin{aligned}
\max_{\ell_1,\ldots,\ell_T} \mathcal{R}_T &\geq \mathbb{E}_{\ell_1,\ldots,\ell_T \overset{iid}{\sim} \mathcal{D}^N} [\mathcal{R}_T] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\ell_1,\ldots,\ell_{t-1}} \mathbb{E}_{\ell_t} [\langle p_t, \ell_t \rangle \,|\ell_{t-1},\ldots,\ell_1] - \mathbb{E}_{\ell_1,\ldots,\ell_T} \left[\min_{i\in[N]} \sum_{t=1}^{T} \ell_t(i)\right] \\
&= \sum_{t=1}^{T} \mathbb{E}_{\ell_1,\ldots,\ell_{t-1}} \langle p_t, \mathbb{E}_{\ell_t}[\ell_t|\ell_{t-1},\ldots,\ell_1]\rangle - \mathbb{E}_{\ell_1,\ldots,\ell_T} \left[\min_{i\in[N]} \sum_{t=1}^{T} \ell_t(i)\right] \\
&= T/2 - \mathbb{E}_{\ell_1,\ldots,\ell_T} \left[\min_{i\in[N]} \sum_{t=1}^{T} \ell_t(i)\right] \\
&= \mathbb{E}_{\ell_1,\ldots,\ell_T} \left[\max_{i\in[N]} \sum_{t=1}^{T} (\tfrac{1}{2} - \ell_t(i))\right] \\
&= \frac{1}{2} \mathbb{E}_{\sigma_1,\ldots,\sigma_T} \left[\max_{i\in[N]} \sum_{t=1}^{T} \sigma_t(i)\right],
\end{aligned}$$

where $\sigma_t(i)$ for $i \in [N], t \in [T]$ are i.i.d. Rademacher random variables (i.e. -1 with probability 0.5 and 1 with probability 0.5). Using the following result from probability theory (see for example [Cesa-Bianchi and Lugosi, 2006, Chapter 3.7]) completes the proof.

$$\lim_{T\to\infty} \lim_{N\to\infty} \frac{\mathbb{E}_{\sigma_1,\ldots,\sigma_T}\left[\max_{i\in[N]} \sum_{t=1}^{T} \sigma_t(i)\right]}{\sqrt{T \ln N}} = \sqrt{2}.$$

□

3 Follow the Regularized Leader

Hedge is just one classic example of online learning. For a general OCO problem, how do we design low-regret algorithms? There are in fact several general frameworks to do this. Here we explore one of them, called *Follow the Regularized Leader* (FTRL).

To introduce FTRL, first recall the FTL algorithm for OCO: $w_t = \operatorname{argmin}_{w\in\Omega} \sum_{\tau=1}^{t-1} f_\tau(w)$. As mentioned, this is not a good algorithm generally. However, if we could cheat and play w_{t+1} at time t (which requires the knowledge of f_t), how small would the regret be? This invaild algorithm is often called *Be the Leader* (BTL) and the following lemma shows that it in fact has negative regret.

Lemma 1 (BTL lemma). *If* $w_t = \operatorname{argmin}_{w\in\Omega} \sum_{\tau=1}^{t-1} f_\tau(w)$*, then*

$$\sum_{t=1}^{T} f_t(w_{t+1}) - \min_{w\in\Omega} \sum_{t=1}^{T} f_t(w) \leq 0.$$

Proof. By definition and optimality of w_t, we have

$$\begin{aligned}
\sum_{t=1}^{T} f_t(w_{t+1}) - \min_{w\in\Omega} \sum_{t=1}^{T} f_t(w) &= \sum_{t=1}^{T} f_t(w_{t+1}) - \sum_{t=1}^{T} f_t(w_{T+1}) \\
&= \sum_{t=1}^{T-1} f_t(w_{t+1}) - \sum_{t=1}^{T-1} f_t(w_{T+1}) \\
&\leq \sum_{t=1}^{T-1} f_t(w_{t+1}) - \sum_{t=1}^{T-1} f_t(w_T) \\
&= \sum_{t=1}^{T-2} f_t(w_{t+1}) - \sum_{t=1}^{T-2} f_t(w_T) \\
&\leq \ldots \leq f_1(w_2) - f_1(w_3) \leq 0.
\end{aligned}$$

$\square$

Therefore, the regret of FTL can be bounded by:

$$\sum_{t=1}^{T} f_t(w_t) - \min_{w\in\Omega} \sum_{t=1}^{T} f_t(w) \le \sum_{t=1}^{T} \left(f_t(w_t) - f_t(w_{t+1})\right),$$

which means the regret is controlled by how close w_t and w_{t+1} are, or in other words, how stable the algorithm is. One way to see that FTL is not a low-regret algorithm is exactly by arguing that it is not stable. Therefore, to fix this issue, we should think about how to improve the stability of the algorithm.

Regularization, a widely-used technique in machine learning, turns out to be also extremely useful here in terms of stabilizing the algorithms. Specifically, FTRL plays at round t:

$$w_t = \operatorname*{argmin}_{w\in\Omega} \sum_{\tau=1}^{t-1} f_\tau(w) + \frac{1}{\eta}\psi(w) \tag{3}$$

where $\eta > 0$ is some learning rate parameter to be specified and $\psi : \Omega \to \mathbb{R}$ is the regularizer. To ensure stability, the regularizer ψ needs to be *strongly convex*, which means for any $w, u \in \Omega$, the following holds[1]

$$\psi(w) - \psi(u) \le \langle \nabla\psi(w), w - u\rangle - \tfrac{1}{2}\left\|w - u\right\|^2$$

for some norm $\|\cdot\|$. The next lemma shows that FTRL is stable and the level of stability is controlled by the parameter η.

Lemma 2 (Stability of FTRL)**.** *The FTRL strategy* (3) *ensures*

$$f_t(w_t) - f_t(w_{t+1}) \le \eta \left\|\nabla f_t(w_t)\right\|_\star^2,$$

where $\|\cdot\|_\star$ *is the dual norm of* $\|\cdot\|$.[2]

Proof. Let $F_t(w) = \sum_{\tau=1}^{t} f_\tau(w) + \frac{1}{\eta}\psi(w)$. By strong convexity of ψ, one can verify

$$F_{t-1}(w_t) - F_{t-1}(w_{t+1}) \le \langle \nabla F_{t-1}(w_t), w_t - w_{t+1}\rangle - \frac{1}{2\eta}\left\|w_t - w_{t+1}\right\|^2.$$

Since $w_t = \operatorname{argmin}_w F_{t-1}(w)$, first order optimality condition implies $\langle \nabla F_{t-1}(w_t), w_t - w_{t+1}\rangle \le 0$ and thus

$$F_{t-1}(w_t) - F_{t-1}(w_{t+1}) \le -\frac{1}{2\eta}\left\|w_t - w_{t+1}\right\|^2.$$

By the same argument, we have

$$F_t(w_{t+1}) - F_t(w_t) \le \langle \nabla F_t(w_{t+1}), w_{t+1} - w_t\rangle - \frac{1}{2\eta}\left\|w_t - w_{t+1}\right\|^2 \le -\frac{1}{2\eta}\left\|w_t - w_{t+1}\right\|^2.$$

Summing up the above two inequalities and rearranging give

$$\left\|w_t - w_{t+1}\right\|^2 \le \eta(f_t(w_t) - f_t(w_{t+1})).$$

Finally by convexity and Hölder's inequality we have

$$\begin{aligned} f_t(w_t) - f_t(w_{t+1}) &\le \langle \nabla f_t(w_t), w_t - w_{t+1}\rangle \le \left\|\nabla f_t(w_t)\right\|_\star \left\|w_t - w_{t+1}\right\| \\ &\le \left\|\nabla f_t(w_t)\right\|_\star \sqrt{\eta(f_t(w_t) - f_t(w_{t+1}))}, \end{aligned}$$

and solving for $f_t(w_t) - f_t(w_{t+1})$ finishes the proof. $\square$

With this stability lemma, we can show the following regret bound for FTRL.

[1] More precisely, this is the definition of ψ being 1-strongly convex.

[2] The definition of dual norm is $\left\|u\right\|_\star = \max_{\|w\|\le 1} \langle u, w\rangle$.

Theorem 3. *With parameter η FTRL ensures,*

$$\mathcal{R}_T \leq \frac{D}{\eta} + \eta \sum_{t=1}^{T} \|\nabla f_t(w_t)\|_\star^2 ,$$

where $D = \max_{w \in \Omega} \psi(w) - \min_{w \in \Omega} \psi(w)$. If we further have $\|\nabla f_t(w)\|_\star \leq G$ for all $w \in \Omega$, then setting $\eta = \sqrt{\frac{D}{TG^2}}$ leads to $\mathcal{R}_T = \mathcal{O}(G\sqrt{TD})$.

Proof. Define $f_0(w) = \frac{\psi(w)}{\eta}$ so that $w_t = \operatorname{argmin}_{w \in \Omega} \sum_{\tau=0}^{t-1} f_\tau(w)$. By the BTL lemma, we have for $w^\star = \operatorname{argmin}_{w \in \Omega} \sum_{t=1}^{T} f_t(w)$,

$$\sum_{t=0}^{T} f_t(w_{t+1}) - \sum_{t=0}^{T} f_t(w^\star) \leq 0.$$

Therefore, the regret of FTRL is

$$\begin{aligned}
\mathcal{R}_T &= \sum_{t=1}^{T} f_t(w_t) - \sum_{t=1}^{T} f_t(w^\star) \\
&\leq \sum_{t=1}^{T} f_t(w_t) - \sum_{t=0}^{T} f_t(w_{t+1}) + f_0(w^\star) \\
&= f_0(w^\star) - f_0(w_1) + \sum_{t=1}^{T} (f_t(w_t) - f_t(w_{t+1})) \\
&\leq \frac{D}{\eta} + \eta \sum_{t=1}^{T} \|\nabla f_t(w_t)\|_\star^2 ,
\end{aligned}$$

where the last step if by the definition of D and the stability lemma. □

References

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, August 1997.

Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108:212–261, 1994.