10-716: Advanced Machine Learning	**Spring 2019**

Lecture 22: April 18

Lecturer: Pradeep Ravikumar — *Scribes: Derun Gu, Yufeng Shen*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

We have studied weight update in convex decision space during the last lecture. In general, we first calculate the weight $w_{i,t} = \bigtriangledown_i \Phi(R_{t-1})$, and then generate the decision as a convex combination of experts' advice $\hat{p}_t = \frac{\sum_{i=1}^N w_{i,t} f_{i,t}}{\sum_{i=1}^N w_{i,t}}$. However, when the decision space is not convex, $\hat{p}_t$ may not be a valid decision.

In this class, we will solve this problem in non-convex decision space.

22.1 Follow the Best Expert

Definition 22.1 *Let the best expert decision is* $\hat{p} = f_{E_t,t}$, *where* $E_t \in \arg\inf_{E\in\mathbf{\Sigma}} \sum_{s=1}^n l(f_{E,s}, y_s)$.

Proposition 22.2 *Let* $p_t^* = f_{E_t^*,t}$, *where* $E_t^* \in \arg\inf_{E\in\mathbf{\Sigma}} \sum_{s=1}^n R_{E,s}$. *Then*

$$\sum_{t=1}^n l(p_t^*, y_t) \leq \inf_{E\in\mathbf{\Sigma}} L_{E,n}$$

Let $\epsilon_t = l(p_t, y_t) - l(p_t^*, y_t)$. A direct result of the above proposition is that

$$L_n - \inf_{E\in\mathbf{\Sigma}} L_{E,n} \leq \sum_{t=1}^n l(p_t, y_t) - l(p_t^*, y_t) = \sum_{t=1}^n \epsilon_t$$

Therefore for a specific problem, we only need to show that ϵ_t decays w.r.t. t.

Empirically, if $\epsilon_t \asymp \frac{1}{t}$, then $L_n - \inf_{E\in\mathbf{\Sigma}} L_{E,n} \asymp \ln n$. The per round regret in this case is in the order of $\frac{\ln n}{n}$, which is a very good result.

Example Suppose $l(p, y) = \|p - y\|_2^2$, $D = \{z \mid \|z\|_2 \leq 1\}$, $\mathcal{Y} = \{y \mid \|y\|_2 \leq 1\}$.

$$p_t^* \in \arg\inf_{p\in D} \sum_{s=1}^t \|p - y_s\|_2^2$$

By taking derivative of the loss w.r.t. p, we have

$$p_t = \frac{1}{t}\sum_{s=1}^t y_s$$

Similarly,

$$\hat{p}_t \in \arg\inf_{p\in D}\sum_{s=1}^{t}\|p-y_s\|_2^2 = \frac{1}{t-1}\sum_{s=1}^{t-1}y_s$$

Therefore,

$$\|p_t^* - \hat{p}_t\|_2 = \|(\frac{1}{t}-\frac{1}{t-1})\sum_{s=1}^{t-1}y_s + \frac{y_t}{t}\|_2 \leq \left|(\frac{1}{t}-\frac{1}{t-1})(t-1)\right| + \frac{1}{t} = \frac{2}{t}$$

Notice that $l(p,y)$ is 4-Lipschitz w.r.t. p, we have

$$\epsilon_t = l(\hat{p}_t, y_t) - l(p_t^*, y_t) \leq 4\|\hat{p}_t - p_t^*\|_2 \leq \frac{8}{t}$$

Eventually, we get

$$L_n - \inf_{E\in\mathbf{\Sigma}} L_{E,n} \leq \sum_{t=1}^{n}\epsilon_t \leq 8(1+\ln n)$$

Notice that the regret does not depends on the number of experts, compared with exponential weighted average forecaster that we discussed in the last class. In fact, since there are infinite many experts in this example, it is also impossible for the result to be dependent on the number of experts.

22.2 Greedy Forecaster

In "following the best expert" strategy, we only take previous expert performance into account. However, in "greedy forecaster", we also take the current round performance into account assuming the nature plays the worst case for experts. Formally, we have the following discussions.

The first try would be

$$\hat{p}_t = \arg\inf_{p\in D}\sup_{y_t\in\mathcal{Y}}\max_{i=1}^{N}\{R_{i,t-1} + l(p,y_t) - l(f_{i,t},y_t)\}$$

This is a slight modification based on the "following the best expert" strategy. However, in practice this does not work well. A correct "greedy forecaster" is defined as below.

Definition 22.3

$$\hat{p}_t = \arg\inf_{p\in D}\sup_{y_t\in\mathcal{Y}}\Phi(R_{t-1}+r_t)$$

In the above definition we use the same notation as before:

$$R_{i,t-1} = \hat{L}_{t-1} - L_{i,t-1} = \sum_{s=1}^{t-1}(l(\hat{p}_s,y_s) - l(f_{i,s},y_s)), \quad r_{i,t} = l(p,y_t) - l(f_{i,t},y_t)$$

By using this strategy, the below theorem also holds the same as in the weighted average forecasters case.

$$\Phi(R_n) \leq \Phi(0) + \sum_{s=1}^{n} C_s$$

22.3 Limitation of Deterministic Decision

Consider a case of an adversarial setting.

Suppose there are two experts. One expert always predicts 0 and the other one always predicts 1. The nature can observe $\hat{p}_t$ and then choose $y_t = 1 - \hat{p}_t$ in an adversarial way.

Therefore the maximum regret $\sum_{t=1}^{n} \hat{L}_t - \inf_{i=0}^{1} f_{i,t} \geq n - \frac{n}{2} = \frac{n}{2}$. This is not sub-linear in n.

This result is sometimes called "Cover's Impossibility Result".

From this example, we get some basic motivation why we would like to pursue a randomized decision.

22.4 Randomization

In the case when both the decision space and the outcome space are $\mathcal{D} = \mathcal{Y} = \{0, 1\}$ and the loss function is $l(p, y) = \mathbb{I}(p \neq y)$. In this case for any deterministic strategy $p_t(y_{1:t-1}, f_{1:N,1:t-1})$, there exists an outcome $y_t = 1 - p_t$ such that the forecaster makes mistake every time, i.e. $\hat{L}_n = n$. However, if we let $f_{1,t} = 0, f_{2,t} = 1$ then $\min_i L_{i,n} \leq \frac{n}{2}$, which means $\hat{L}_n - \min_i L_{i,n} \geq \frac{n}{2}$. As we can see, these experts are actually useless but we still get large regret for any deterministic strategy. The way to deal with it is randomization.

Consider a game between a player (the forecaster) and the environment. There are some notations:

- Every round the player chooses one of the N actions $I_t \in \{1, \cdots, N\}$ with probability $\mathbf{p}_t = (P_{1,t}, \cdots, P_{N,t})$. i.e. $\mathbb{P}(I_t = i) = P_{i,t}$.
- Every round the environment chooses an action $y_t \in \mathcal{Y}$.
- The loss of choosing action i is $l(i, y) : \{1, \cdots, N\} \times \mathcal{Y} \to [0, 1]$.
- The player's goal is to minimize the cumulative regret

$$\hat{L}_n - \min_{i=1,\cdots,N} L_{i,n} = \sum_{t=1}^{n} l(I_t, y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, y_t)$$

.

- There are two kinds of environment:
  1. Oblivious Adversary: $\{y_t\}$ do not depend on $\{I_t\}$. For example, weather forecasting.
  2. Non-oblivious Adversary: $Y_t = g_t(I_1, \cdots, I_{t-1})$, where $g_t : [N]^{t-1} \to \mathcal{Y}$. For example, stock market, gambling.
- We define the "expected loss" as $\bar{l}(\mathbf{p}_t, Y_t) = \mathbb{E}_t l(I_t, Y_t) = \sum_{i=1}^{N} l(i, Y_t) p_{i,t}$. Here the expected value $\mathbb{E}_t$ is taken only with respect to the random variable I_t. It is actually a conditional expectation of $l(I_t, Y_t)$ given the past plays $I_1, \cdots, I_{t-1}$.

Note that here we use "actions" instead of "experts", but if we define the loss $l'(i, y_t) = l(f_{i,t}, y_t)$, then these two models are equivalent. So any bounds we will prove in this section also apply to the expert model.

Theorem 22.4 *Suppose for every $t = 1, \cdots, n$, p_t depends only on $\{y_s\}_{s=1}^{t-1}$ and not explicitly on the realization of $\{I_s\}_{s=1}^{t-1}$. Assume the forecaster's expected regret against an oblivious environment satisfies*

$$\sup_{y^n \in \mathcal{Y}^n} \mathbb{E}\left[\sum_{t=1}^{n} l(I_t, y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, y_t)\right] \leq B(n)$$

If the same forecaster is used against a non-oblivious environment, then

$$\sum_{t=1}^{n} \bar{l}(\mathbf{p}_t, Y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, Y_t) \leq B(n) \tag{22.1}$$

Moreover, $\forall \delta > 0$, with probability at least $1 - \delta$ the actual cumulative loss satisfies, for any non-oblivious environment,

$$\sum_{t=1}^{n} l(I_t, Y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, Y_t) \leq B(n) + \sqrt{\frac{n}{2} \ln \frac{1}{\delta}} \tag{22.2}$$

Proof: First, let's proof the Equation 22.1.

Observe first that if the environment is oblivious then the sequence $y_1, \cdots, y_n$ is fixed, so the $\mathbf{p}_t$ is also fixed. Therefore

$$B(n) \geq \mathbb{E}\left[\sum_{t=1}^{n} l(I_t, y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, y_t)\right] = \sum_{t=1}^{n} \bar{l}(\mathbf{p}_t, y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, y_t) \tag{22.3}$$

If the environment is non-oblivious, then for each $i = 1, \cdots, N$,

$$\begin{aligned}
&\sum_{t=1}^{n} \bar{l}(\mathbf{p}_t, Y_t) - \sum_{t=1}^{n} l(i, Y_t) \\
&\leq \sum_{t=1}^{n} \sup_{y_t \in \mathcal{Y}} \bar{l}(\mathbf{p}_t, y_t) - l(i, y_t) \\
&= \sup_{y_1 \in \mathcal{Y}} \mathbb{E}_1[l(I_1, y_1) - l(i, y_1)] + \cdots + \sup_{y_n \in \mathcal{Y}} \mathbb{E}_n[l(I_n, y_n) - l(i, y_n)] \\
&\overset{(i)}{=} \sup_{y^n \in \mathcal{Y}^n} \sum_{t=1}^{n} \mathbb{E}_t[l(I_t, y_t) - l(i, y_t)] \\
&= \sup_{y^n \in \mathcal{Y}^n} \sum_{t=1}^{n} \bar{l}(\mathbf{p}_t, y_t) - l(i, y_t) \\
&\overset{(ii)}{\leq} B(n) \\
\Rightarrow &\sum_{t=1}^{n} \bar{l}(\mathbf{p}_t, Y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, Y_t) \leq B(n)
\end{aligned}$$

The inequality (ii) results from the Equation 22.3. To see why equality (i) holds, lets look at the first two terms. Obviously, the optimal y_1 doesn't depend on I_2. Based on our assumption about $\mathbf{p}_t$, we know $\mathbf{p}_2$ doesn't depend on the realization of I_1, because the optimal y_2 only depends on $\mathbf{p}_2$, so the optimal y_2 doesn't depend on the realization of I_1. Therefore,

$$\begin{aligned}
&\sup_{y_1 \in \mathcal{Y}} \mathbb{E}_1[l(I_1, y_1) - l(i, y_1)] + \sup_{y_2 \in \mathcal{Y}} \mathbb{E}_2[l(I_2, y_2) - l(i, y_2)] \\
&= \sup_{y_1 \in \mathcal{Y}} \sup_{y_2 \in \mathcal{Y}} \{\mathbb{E}_1[l(I_1, y_1) - l(i, y_1)] + \mathbb{E}_2[l(I_2, y_2) - l(i, y_2)]\}
\end{aligned}$$

The proof for $n > 2$ follows by repeating the same argument.

For the Equation 22.2, observe that $\mathbb{E}_t[l(I_t, Y_t) - \bar{l}(\mathbf{p}_t, Y_t)] = 0$ and $l(I_t, Y_t) - \bar{l}(\mathbf{p}_t, Y_t) \in [0, 1]$, so $l(I_t, Y_t) - \bar{l}(\mathbf{p}_t, Y_t)$ is a bounded martingale difference sequence, based on Azuma-Hoeffding Inequality, we have

$$\mathbb{P}\left[\sum_{t=1}^{n} l(I_t, Y_t) - \sum_{t=1}^{n} \bar{l}(\mathbf{p}_t, Y_t) \geq \sqrt{\frac{n}{2} \ln \frac{1}{\delta}}\right] \leq \delta$$

Combining with Equation 22.1, we have Equation 22.2. ■

To understand the randomized prediction problem described above, we first consider a simplified version in which the player's goal is to minimize the difference between the cumulative expected loss and the loss of the pure strategy:

$$\bar{L}_n - \min_{i=1,\cdots,N} L_{i,n} \stackrel{def}{=} \sum_{t=1}^{n} \bar{l}(\mathbf{p}_t, y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, y_t)$$

Then this problem becomes a special case of the weighted average forecasters problem. Similarly to previous sections, we have

$$\begin{aligned}
\bar{l}(\mathbf{p}_t, y_t) &= \sum_{i=1}^{N} P_{i,t} l(i, y_t) \\
P_{i,t} &= \frac{\nabla_i \Phi(R_{t-1})}{\sum_{j=1}^{N} \nabla_j \Phi(R_{t-1})} \\
\Rightarrow &\sum_{t=1}^{N} r_{i,t} \nabla_i \Phi(R_{t-1}) \leq 0 \\
\Rightarrow &\Phi(R_n) \leq \Phi(R_0) + \sum_{t=1}^{n} C_t,
\end{aligned}$$

where $r_{i,t} = \bar{l}(\mathbf{p}_t, y_t) - l(i, y_t)$. Particularly, if we use the exponentially weighted average player

$$P_{i,t} = \frac{\exp(-\eta L_{i,t})}{\sum_{j=1}^{N} \exp(-\eta L_{j,t})}$$

we will get

$$\bar{L}_n - \min_{i=1,\cdots,N} L_{i,n} \leq \sqrt{\frac{n \ln N}{2}}$$

Definition 22.5 *Hannan consistency:*

$$\lim_{n\to\infty} \sup_{y^n \in \mathcal{Y}^n} \frac{1}{n} \left(\sum_{t=1}^{n} l(I_t, y_t) - \min_{i=1,\cdots,N} \sum_{t=1}^{n} l(i, y_t)\right) = 0,$$

with probability 1.

References

[BL] Nicolo Cesa-Bianchi and Gabor Lugosi, "Prediction, Learning, and Games" *Cambridge University Press*, ISBN-13978-0-521-84108-5.