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

Lecture 23: April 23

Lecturer: Pradeep Ravikumar *Scribes: Cheng Cheng, Ge Huang, Guoxi Zhang*

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

Recall that in last lecture we studied randomized prediction, and we looked at:

$$\mathbb{E}\left[\frac{1}{n}\left(\sum_{t=1}^{n} l(I_t, y_t) - \min_{i=1,\cdots,N}\sum_{t=1}^{n} l(i, y_t)\right)\right]$$

and we want this to go to 0 as n goes to infinity with probability 1. Today we looked at some other randomized prediction strategy.

23.1 Follow the leader strategy (FTL)

FTL tries to predict at time t according to the action i whose cumulative Loss $L_{i,t-1}$ up to that time is minimal. However, it may not achieve Hannan consistency. Below is a counterexample.

Example Suppose there are only $N = 2$ actions so that for the first action $l(1, y_t) = (1/2, 0, 1, 0, 1, 0, 1, ...)$ and the for the second actcion $l(2, y_t) = (1/2, 0, 1, 0, 1, 0, 1, ...)$. Then we have $\sum_{t=1}^{n} l(p^{(FTL)}, y_t) = n$ but $L_{i,n} = n/2$. In order to improve the performance, one may add a small random perturbation to the cumulative losses and follow the "perturbed" leader.

23.2 Follow the perturbed leader strategy (FTPL)

Let $Z_1, Z_2, ...$ be independent, identically distributed random N-vectors with components $Z_{i,t} (i = 1, ..., N)$. $Z_{i,t} \sim Q(.)$At time t, an action:

$$I_t = \arg\min_{i=1,...N}(L_{i,t-1} + Z_{i,t})$$

Example $Z_{i,t} \sim Q(.)$ with density $f_z(z) = (\frac{\eta}{2})^N \exp\left(-\eta ||z||_1\right)$. By setting η appropriately, we have:

$$\sum_{t=1}^{n} \bar{l}(p_t, y_t) - \min_{i=1,\cdots,N}\sum_{t=1}^{n} l(i, y_t) \leq \sqrt{L_n^* lnN} + clnN$$

and further:

$$\sum_{t=1}^{n} l(I_t, y_t) - \min_{i=1,\cdots,N}\sum_{t=1}^{n} l(i, y_t) \leq \sqrt{L_n^* lnN} + clnN + \sqrt{\frac{n}{2} lnN}$$

with respect to probability.

23.3 Prediction with limited feedback

We now move on to study prediction with limited feedback. The setting is defined as follows:

For each round $t = 1, 2, ...$

- the environment chooses y_t;
- the forecaster chooses $I_t \in \{1, ..., N\}$;
- the forecaster incurs loss $l(I_t, y_t)$, and the experts (actions) incur loss $l(i, y_t)_{i=1}^N$ where none of these values is necessarily revealed to the forecaster;
- if less than $\mu(t)$ queries have been issued so far, the forecaster may issue a new query to obtain the outcome Y_t; if no query is issued, then Y_t remains unknown.

The goal of the forecaster is to minimize its regret:

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

We can use a variant of weighted average forecasting to solve this problem. At step t, the forecaster chooses an action randomly according to $P_{i,t} = \frac{w_{i,t-1}}{\sum_{j=1}^N w_{j,t-1}}$. He then flips a coin $Z_t \sim \text{Bernoulli}(\epsilon)$ and updates weights as the following:

$$w_{i,t} = \begin{cases} w_{i,t-1}, & Z = 0; \\ w_{i,t-1} \exp -\eta \dfrac{\ell(i, y_t)}{\epsilon}, & Z = 1, I_t = i. \end{cases}$$

In this procedure the forecaster estimates his loss as:

$$\tilde{\ell}(i, y_t) = \begin{cases} 0, & Z = 0; \\ \dfrac{\ell(i, y_t)}{\epsilon} & Z = 1. \end{cases}$$

This is an unbiased estimate of his loss: $\mathbb{E}[\tilde{\ell}(i, y_t)] = \ell(i, y_t)$.

Theorem 23.1 *Let* $\epsilon = \frac{m}{n}$ *and* $\eta = \sqrt{2 \ln N \frac{m}{n}}$, *with the variant of weighted average forecasting* $\mathbb{E}\hat{L}_n - \min_{i=1,\ldots,N} L_{i,n} \leq \sqrt{2 \ln N} \frac{n}{\sqrt{m}}$, *where* m *is the number queries the forecaster is allowed to make.*

23.4 Partial Monitoring

We now consider a different setting in which the forecaster only receives indirect information. Consider the following example. A seller is trying to sell products to customers in a stream. At each step t, the seller chooses a price $I_t \in \{1, \ldots, N\}$ with which he will sell his products to the customer at step t. The customer at step t has an internal value $y_t \in \{1, 2, \ldots, M\}$, and she will buy the product only if $y_t \geq I_t$. In other words, about y_t the seller only receives partial information $\mathbb{I}(y_t \geq I_t)$ not the exact value y_t. We assume that at each step t the seller incurs loss $\ell(I_t, y_t) = \mathbb{I}(y_t \geq I_t)(y_t - I_t) + \mathbb{I}(y_t < I_t)c$. That is, if he manages to sell the product, he will earn $I_t - y_t$; if he fails to sell the product he will loose c, a constant. Note that the seller is not informed with the value of his loss, otherwise he will be able to deduce y_t from that.

Formally, let $L = (\ell(i,j)) \in \mathbb{R}^{N\times M}$ and $H = (h(i,j)) \in \mathbb{R}^{N\times M}$ be two known matrices. Let $\mathcal{Y} = \{1, 2, \ldots, M\}$ be the domain of sequence, and $\mathcal{D} = \{1, \ldots, N\}$ be the decision space. Consider the following protocol: For each round $t = 1, 2, \ldots$

1. the environment chooses $y_t \in \mathcal{Y}$;
2. the forecaster chooses $I_t \in \mathcal{D}$;
3. the forecaster incurs loos $\ell(I_t, y_t)$, the experts (actions) incur $\ell(i, y_t), i = 1\ldots, N$.
4. $h(I_t, y_t)$ is revealed to the forecaster

23.4.1 Multi-arm Bandit

Multi-arm Bandit is a specific partial monitoring setting. Suppose we have N slot machines. Each of them has a lever. At each time one may choose one machine and pull the lever. With some probability the machine being pulled will return a large amount of money, but it may also return noting. This probability is a machine specific parameter.

It turns out that we could model this problem as sequence forecasting with partial monitoring, and utilize a variant of weighted average forecasting to it. Assume there exists an known invertable matrix $K \in \mathbb{R}^{N\times N}$ such that $L = KH$where $L \in R^{N\times M}$ and $H \in R^{N\times M}$. In other words, $\ell(i,j) = \sum_r k(i,r)h(r,j)$. We don't know $\ell(i,j)$, one way is to use $\tilde{\ell}(i, y_t) = \frac{k(i,I_t)h(I_t,y_t)}{P_{I,t}}$ as an estimator for lose, where $P_{I,t}$ is the probability for choosing action I at time t. The expectation $E_{T_t}[\tilde{\ell}](i, y_t) = \sum_{r=1}^{N} P_{r,t} \frac{k(i,r)h(r,y_t)}{P_{r,t}} = \ell(i, y_t)$. There three things we need to know. One is the probability taking action I_t, two is that $\forall i$, we need to $K(i, I+t)$, and $h(I_t, y_t)$ is provided. Now we can do forecasting, but it won't work. The reason things don't work is the loss in this case is not enough for the loss to have an expectation that's equal to this loss. But I can always have a variable with the mean equals to this but with a huge variance, which is essentially junk. So do you guarantee the vairance of $\tilde{\ell}$ is not too large? It turns out that if you use a typical weighted average forward casting probabilities, the variance is too large. The reason variance comes into play is that the RHS is essentially $1/P$, and if P is extremely small, the variance scales quadratically with $1/P$, hence variance is extremely large. For the weighted average forecasting, you want weights to be small for most guys except for the best expert. So it can't guarantee the variance bounded and you get into trouble. So all you need to make sure is that $P_{i,t}$ dosen't decay to 0 for any action. So one can pick action $I_t \in [N]$ according to $P_{i,t} = (1-\gamma)\frac{w_{i,t-1}}{\sum_{j=1}^{N} w_{j,t-1}} + \gamma\frac{1}{N}$. The extra term ensures the probability is at least $\gamma\frac{1}{N}$ and isn't too small even when the best expect has most of the mass and everybody else has probability 0. $w_{i,t} = w_{i,t-1}exp(-\eta\ell(i, y_t))$.

Theorem 23.2 *$L = KH$ for some $K \in R^{N\times N}$ with $k^* = max1, max|k(i,j)|$. If $= \frac{1}{c}()^{\frac{lnN}{Nn}}2/3$ and $\gamma = c(\frac{N^2 lnN}{n})^{1/3}$, then $E[\sum_{t=1}^{N} \ell(I_t, y_t) - min_{i=1}^{N} \sum_{t=1}^{n} \ell(i, y_t)] \leq c(Nn)^{2/3}(lnN)^{1/3}$* This is a tight bound. These constant depends on k^*. It's worse than $\sqrt{n}$, which is not too bad since you are only given access to partial information.

We can actually do better. Specifically for the multi-arm Bandit problem, the trick is surprisingly simple. We still do weighted averaged forecasting, but in a different way to look at gains instead of losses. $g(i, y_t) = 1 - \ell(i, y_t)$. This gives you the $\sqrt{n}$ because if you have losses with a suddenly bad action, the probability or weight of the action is gonna to sharply decay, but for gains, you are just going to stop increasing when having a suddenly bad action. We still need the variance of the gains. The estimated gains is

$$\tilde{g}(i, y_t) = \begin{cases} \dfrac{g(i, y_t)}{p_{i,t}}, & I_t = i; \\ 0 & I_t = 0. \end{cases}$$

$P_{i,t} = \frac{w_{i,t-1}}{\sum_{j=1}^{N} w_{j,t-1}} + \frac{\gamma}{N}$ Even with this additional restriction, the variance is not guaranteed to be bounded sufficiently enough. We need to bound the variance of $\tilde{g}$ even more to get the $\sqrt{n}$. In the fully observed case, there is no variance because you see exactly value of the loss, but here you are seeing a stochastic estimator of the gain, so the variance is large. One way for you to control the variance is to calculate the estimated gains by $g'(i, y_t) = \tilde{g}(i, y_t) + \frac{\beta}{P_{i,t}}$ which also modifying the $\tilde{g}$ to ensure the estimated gain is not too small. This is no longer unbiased, but the small deviation from g is the price to pay for a smaller variance. Again, the weights will still be $w_{i,t} = w_{i,t-1} exp(\eta g'(i, y_t))$. There are fancier ways to reduce the variance, but this simply trick is suffices.