Lecture 17, 18, 19

Agenda
* Multiplicative weights update for prediction and multiarmed bandit
* Lower bounds

A Competing with the best coin

For simplicity, we will restrict to a simple, yet canonical formulation.

Prediction Problem (PP)

There is an unknown sequence $y^n = (y_1, \ldots, y_n) \in \{0,1\}^n$.

Our goal is to predict the next entry by looking at the previous entries.

→ Clearly, in the worst-case we can be wrong every time.

But we only seek to be competitive with a set of k experts, which we can denote by k coins:

Coin i outputs $x_i(t) \in \{0,1\}$ at time t, $1 \le t \le n$

Reward: # of correct predictions

Observation: $\underline{x}^t = \big((x_i(1), \ldots, x_i(t)),\ 1 \le i \le k\big)$

$$O_t = (y^{t-1}, \underline{x}^t) \equiv \Big(y^{t-1}, \underbrace{\big(\mathbb{1}_{\{x_i(1) \ne y_1\}}, \ldots, \mathbb{1}_{\{x_i(t-1) \ne y_{t-1}\}}\big),\ 1 \le i \le k,}_{\text{losses of the experts}} \underbrace{(x_1(t), \ldots, x_k(t))}_{\text{prediction of the future}} \Big)$$

$\supseteq$ (losses of the past, prediction) → the algo. we cover only requires this

Action: Choose an expert I_t based on O_t:

$$I_t = \hat{\theta}_t(O_t) \in [k]$$

Regret Formulation:

$$r^{PP}(k,n) = \min_{\hat{\theta}} \max_{y^n \in \{0,1\}^n} \left(\sum_{t=1}^{n} \mathbb{1}_{\{x_{I_t}(t) \neq y_t\}} - \min_{1 \le i \le k} \sum_{t=1}^{n} \mathbb{1}_{\{x_i(t) \neq y_t\}} \right)$$

Multiarmed Bandit Problem (MAB)

Observations: We only observe the output of the selected coin, $O_t = (x_{I_j}(j),\ 1 \le j \le t-1)$

We can define $r^{MAB}(k,n)$ similarly as above.

It is easy to see $r^{MAB}(k,n) \ge r^{PP}(k,n)$.

But as we saw earlier, in the worst case the predictor can be wrong every time, and therefore $r^{PP}(k,n) = O(n)$.

Thus, these definitions are not very useful. Surprisingly, the multiplicative regret can be handled. We define

$$R^{PP}(k,n) = \min_{\hat{\theta}} \max_{y^n \in \{0,1\}^n} \max_{1 \le i \le k} \log \frac{\sum_{t=1}^{n} \mathbb{1}_{\{x_{I_t}(t) \neq y_t\}}}{\sum_{t=1}^{n} \mathbb{1}_{\{x_i(t) \neq y_t\}}}$$

→ We will see that if the best expert incurs a loss of $\ge \frac{1}{4}\log k$, then

$$R^{PP}(k,n) \le 2$$

(B) **Weighted Majority Algorithm (WM)** ("The multiplicative weight update method", Arora, Hazan, Kale)

3

Initialize: $w_i^1 = 1, \quad 1 \le i \le k$

For $t = 1, \ldots, n$, update

$$w_i^{t+1} = w_i^t \left(1 - \varepsilon \mathbb{1}_{\{x_i(t) \neq y(t)\}}\right), \quad 1 \le i \le k$$

Output: $\hat{\theta}_t$ = least $i \in [k]$ s.t. $x_i(t)$ is the weighted majority, i.e.,

$$\sum_{j=1}^{k} w_j^t \mathbb{1}_{\{x_j(t) = x_i(t)\}} \geq \sum_{j=1}^{k} w_j^t / 2$$

Denote: $\hat{x}(t)$ = prediction $x_{I_t}(t)$ at time t

l_{WM}^t = total loss incurred till time t

$$= \sum_{j=1}^{t} \mathbb{1}_{\{\hat{x}(j) \neq y_j\}} \longrightarrow l_{WM}(j)$$

= total loss of expert i till time t

$$l_i^t = \sum_{j=1}^{t} \mathbb{1}_{\{x(j) \neq y_j\}} \longrightarrow l_i(j)$$

Theorem For every $\varepsilon \le 1/2$,

$$l_{WM}^t \le \frac{2}{\varepsilon} \log k + 2(1+\varepsilon)\, l_i^t, \quad \forall\, i \in [k].$$

Proof. We follow a potential function approach.

Let $\Phi^t = \sum_{i=1}^{k} w_i^t$ be the potential at time t.

Thus,

$$\Phi^1 = k,$$

and $w_i^t = (1-\varepsilon)^{l_i(t)}$

The key observation is the growth rate (4)

$$\Phi^{t+1} = \sum_{i=1}^{k} \omega_i^t - \varepsilon \sum_{i=1}^{k} \omega_i^t \mathbb{1}_{\{x_i(t) \neq y_t\}}$$

When $\hat{x}(t) \neq y_t$,

$$\Phi_{t+1} = \sum_{i=1}^{k} \omega_i^t - \varepsilon \sum_{i=1}^{k} \omega_i^t \mathbb{1}_{\{x_i(t) = \hat{x}(t)\}}$$

$$\leq \sum_{i=1}^{k} \omega_i^t - \varepsilon \sum_{i=1}^{k} \omega_i^t / 2 = \Phi_t (1 - \varepsilon/2).$$

Therefore, $\Phi^t = \Phi^1 (1-\varepsilon/2)^{\ell_{WM}^t}$

and $\Phi^t \geq \omega_i^t \geq (1-\varepsilon)^{\ell_i^t}$

Thus,

$$\ell_i^t \log(1-\varepsilon) \leq \ell_{WM}^t \log(1-\varepsilon/2) + \log k$$

$$\Rightarrow \quad \ell_{WM}^t \, \varepsilon/2 \leq \ell_i^t \log\frac{1}{1-\varepsilon} + \log k$$

$$\leq \ell_i^t (\varepsilon + \varepsilon^2) + \log k,$$

$$\left(\text{since } \log\frac{1}{1-\varepsilon} \leq \varepsilon + \varepsilon^2 \text{ if } \varepsilon \leq 1/2\right)$$

when $\varepsilon < 1/2$. Thus,

$$\ell_{WM}^t \leq 2(1+\varepsilon)\ell_i^t + \frac{2}{\varepsilon}\log k. \quad \blacksquare$$

Corollary. If $\ell_{min}^t := \min_{i \in [k]} \ell_i^t \geq (\log k)/4$, then

$$R^{PP}(k,n) \leq 2.$$

Proof. Using the previous theorem, we get

$$\frac{l_{WM}^t}{l_{min}^t} \le 2\left(\varepsilon + \frac{1}{\varepsilon} \cdot \log k / l_{min}^t\right) + 2$$

$$\le 4\sqrt{\log k / l_{min}^t} + 2$$

where we used assumption of the corollary to set $\varepsilon \le 1/2$.

If $l_{WM}^t \le \log k$, by assumption, $l_{WM}^t / l_{min}^t \le 4$.

Otherwise we have,

$$\frac{l_{WM}^t}{l_{min}^t} \le 4\sqrt{\frac{l_{WM}^t}{l_{min}^t}} + 2 \Rightarrow \left(\sqrt{\frac{l_{WM}^t}{l_{min}^t}} - 2\right)^2 \le 2$$

$$\Rightarrow \frac{l_{WM}^t}{l_{min}^t} \le 4.$$

$$\Rightarrow R^{PP}(k,n) \le \log 4 = 2. \quad \blacksquare$$

$\boxed{C}$ Multiplicative weight update

While weighted majority allowed us to come multiplicatively close to the least loss l_{min}^t, and this is the best we can do, on average we can get a sublinear regret. We now allow randomized $\hat{\theta}$ that choose expert $I_t = i$ with probability $\theta_t(i \mid O_t)$. Define

$$\bar{R}^{PP}(k,n) = \min_{\hat{\theta}} \max_{\underline{y}} \sum_{t=1}^{n} \mathbb{E}\left[l_{I_t}(t)\right] - \min_{i \in (k)} \sum_{t=1}^{n} l_i(t).$$

Multiplicative Weight Update (MWU)

- Update weight as before.
- Select $I_t = i$ with probability $p_t(i) = w_i^t / \Phi^t$, $i \in [k]$.

Theorem For $\varepsilon \le 1/2$,

$$\mathbb{E}[\ell^n_{MWU}] - \ell^n_{min} \le \frac{\log k}{\varepsilon} + \varepsilon \, \ell^n_{min}.$$

Proof.

$$\begin{aligned}
\Phi^{t+1} &= \sum_{i=1}^{k} w_i^t (1-\varepsilon)^{\ell_i(t)} \\
&= \Phi^t \sum_{i=1}^{k} p_t(i) \, (1-\varepsilon)^{\ell_i(t)} \\
&= \Phi^t \sum_{i=1}^{k} p_t(i) \, (1-\varepsilon \, \ell_i(t)) \\
&= \Phi^t \Big(1 - \varepsilon \underbrace{\sum_{i=1}^{k} p_t(i) \, \ell_i(t)}_{\mathbb{E}[\ell_{I_t}(t)]}\Big) \\
&\le \Phi^t e^{-\varepsilon \mathbb{E}[\ell_{I_t}(t)]} \\
&\le k \cdot e^{-\varepsilon \sum_{j=1}^{t} \mathbb{E}[\ell_{I_j}(j)]} \\
&= k \cdot e^{-\varepsilon \mathbb{E}[\ell^t_{MWU}]}
\end{aligned}$$

$$\Phi^t \ge w_i^t = (1-\varepsilon)^{\ell_i^t}$$

$$\Rightarrow \quad \ell_i^n \log(1-\varepsilon) \le \log k - \varepsilon \, \mathbb{E}[\ell^n_{MWU}]$$

$$\Rightarrow \quad \mathbb{E}[\ell^n_{MWU}] \le \frac{1}{\varepsilon} \log k + \varepsilon \, \ell_i^n + \ell_i^n$$

(using $-\log(1-\varepsilon) \le \varepsilon + \varepsilon^2$, $\varepsilon \le 1/2$) ∎

__Corollary__

$$\mathbb{E}[\ell^n_{MWU}] - \ell^n_{min} \le \frac{1}{\varepsilon}\log k + \varepsilon n$$

$$\le \sqrt{n \log k} \quad \text{if} \quad \underbrace{\frac{\log k}{n} \le \frac{1}{4}}_{1}$$

where we used

$$\varepsilon = \sqrt{\frac{\log k}{n}}.$$

__Remark 1.__ The condition 1 above is unnecessary. We can get rid of it using a more careful analysis as follows: Denote $e^{-\eta} = (1-\varepsilon)$. Then,

$$\frac{\Phi^{t+1}}{\Phi^t} = \frac{\sum_{i=1}^{k} e^{-\eta(\ell_i^t + \ell_i(t+1))}}{\sum_{i=1}^{k} e^{-\eta \ell_i^t}}$$

$$\le 1 - \eta \frac{\sum_{i=1}^{k} e^{-\eta \ell_i^t} \cdot \ell_i(t+1)}{\sum_{i=1}^{k} e^{-\eta \ell_i^t}} + \frac{\eta^2}{2} \frac{\sum_{i=1}^{k} \ell_i^2(t+1)\, e^{-\eta \ell_i^t}}{\sum_{i=1}^{k} e^{-\eta \ell_i^t}},$$

(using $e^{-x} \le 1 - x + x^2/2 \ \forall x \ge 0$)

$$\le 1 - \eta\, \mathbb{E}[\ell_{I_{t+1}}(t+1)] + \frac{\eta^2}{2} \mathbb{E}[\ell_{I_{t+1}}(t+1)],$$

where we used $\ell_i(t+1) \le 1$ (in our case $\ell_i^2(t+1) = \ell_i(t)$).

Thus, proceeding as before,

$$-\eta\, \ell^n_{min} \le \log k - \eta\, \mathbb{E}[\ell^n_{MWU}] + \frac{\eta^2 n}{2}$$

$$\Rightarrow \mathbb{E}\left[\ell^n_{MWU} \right] - \ell^n_{min} \leq \frac{\log k}{\eta} + \frac{\eta n}{2}, \quad \forall \eta > 0$$

which gives the corollary without assuming ①.

Remark 2. The scheme above requires the knowledge of n to set η. An adaptive η can be used to get the same result without knowing n.

We have shown:

$$\bar{\mathfrak{R}}^{PP}(k,n) \leq \frac{3}{2}\sqrt{n \log k}$$

$\boxed{D}$ Multiarmed Bandit

We now move to the MAB problem and consider

$$\bar{\mathfrak{R}}^{MAB}(k,n) = \min_{\hat{\theta}} \max_{y^n} \sum_{t=1}^{n} \mathbb{E}\left[\ell_{I_t}(t) \right] - \min_{i \in [k]} \sum_{t=1}^{n} \ell_i(t)$$

where

$$\ell_i(t) \equiv \ell\left(x_i(t), y_t\right).$$

* A simple extension of MWU to the MAB will simply update the weight of the selected arm. In this case,

$$\Phi^{t+1} = \sum_{i=1}^{k} \omega_i^t \, e^{-\eta \, \ell_i(t) \mathbb{1}_{\{I_t = i\}}}$$

$$\leq \sum_{i=1}^{k} \omega_i^t \left(1 - \eta \, \ell_i(t) \mathbb{1}_{\{I_t=i\}} + \frac{\eta^2}{2} \ell_i(t) \mathbb{1}_{\{I_t=i\}}\right)$$

from where we don't know how to proceed. ⑨

→ Fix 1. Denote by $p_t(i)$ the probability of selecting the i^{th} arm.

Let $\hat{\ell}_i(t) = \frac{\ell_i(t)}{p_t(i)} \mathbb{1}_{\{I_t = i\}}$,

and update the weights as

$w_i^{t+1} = w_i^t \cdot e^{-\eta \hat{\ell}_i(t)}$.

Then,
$$\Phi^{t+1} \leq \sum_{i=1}^{k} w_i^t \left(1 - \eta \hat{\ell}_i(t) + \frac{\eta^2}{2} \cdot \hat{\ell}_i^2(t)\right)$$
$$= \Phi^t \left(1 - \eta \sum_{i=1}^{k} p_t(i) \frac{\ell_i(t)}{p_t(i)} \cdot \mathbb{1}_{\{I_t=i\}} + \frac{\eta^2}{2} \cdot \sum_{i=1}^{k} p_t(i) \cdot \frac{\ell_i^2(t)}{p_t^2(i)} \cdot \mathbb{1}_{\{I_t=i\}}\right)$$
$$\leq \Phi^t \left(1 - \eta\, \ell_{I_t}(t) + \frac{\eta^2}{2} \frac{\ell_{I_t}(t)}{p_t(I_t)}\right),$$

but we are stuck again since $p_t(I_t)$ can be arbitrarily small.

→ Fix 2. Add a constant $p_t(i)$.

Let $p_t(i) = (1-\eta) \frac{w_i^t}{\sum_{i=1}^{k} w_i^t} + \eta/k$

and update $w_i^{t+1} = w_i^t \cdot e^{-\frac{\eta}{k} \hat{\ell}_i(t)}$. → This is the popular EXP_3 algo.

For this algo., we have

$$\Phi^{t+1} = \sum_{i=1}^{k} w_i^t \cdot e^{-\frac{\eta}{k}\hat{\ell}_i(t)}$$

$$\leq \sum_{i=1}^{k} w_i^t \cdot \left(1 - \eta\,\frac{\hat{\ell}_i(t)}{k} + \frac{\eta^2}{2}\,\frac{\hat{\ell}_i^2(t)}{k^2}\right)$$

$$= \Phi^t \sum_{i=1}^{k} \left(\frac{p_t(i) - \eta/k}{1-\eta}\right) \cdot \left(1 - \frac{\eta}{k}\hat{\ell}_i(t) + \frac{\eta^2}{2k^2}\hat{\ell}_i^2(t)\right)$$

$$\Rightarrow \frac{\Phi^{t+1}}{\Phi^t} \leq 1 - \frac{\eta}{k(1-\eta)} \cdot \sum_{i=1}^{k} p_t(i)\,\hat{\ell}_i(t) + \frac{\eta^2}{k^2(1-\eta)} \sum_{i=1}^{k} \hat{\ell}_i(t)$$

$$+ \frac{\eta^2}{2k^2(1-\eta)} \sum_{i=1}^{k} p_t(i)\,\hat{\ell}_i^2(t),$$

where

$$\hat{\ell}_i(t) = \frac{\ell_i(t)}{p_t(i)} \mathbb{1}_{\{I_t = i\}}.$$

Note that

$$\frac{\eta^2}{k^2} \sum_{i=1}^{k} p_t(i)\,\hat{\ell}_i^2(t) = \frac{\eta^2}{k^2} \sum_{i=1}^{k} \ell_i(t)\,\hat{\ell}_i(t) \leq \frac{\eta^2}{k^2} \sum_{i=1}^{k} \hat{\ell}_i(t).$$

Thus,

$$\frac{\Phi^{t+1}}{\Phi^t} \leq 1 - \frac{\eta}{k(1-\eta)}\,\ell_{I_t}(t) + \frac{3\eta^2}{2k^2(1-\eta)} \sum_{i=1}^{k} \hat{\ell}_i(t)$$

Therefore, $\Phi^n \leq \Phi^1 \cdot \prod_{t=1}^{n} \left(1 - \frac{\eta}{k(1-\eta)}\ell_{I_t}(t) + \frac{\eta}{k(1-\eta)} + \frac{\eta^2}{2k^2(1-\eta)} \cdot \sum_{i=1}^{k} \hat{\ell}_i(t)\right)$

$$\leq \Phi^1 \cdot \exp\left(-\frac{\eta}{k(1-\eta)} \sum_{t=1}^{n} l_{I_t}(t) + \frac{\eta}{k(1-\eta)} \cdot n + \frac{\eta^2}{2k^2(1-\eta)} \cdot \sum_{i=1}^{k} \sum_{t=1}^{n} \hat{l}_i(t) \right)$$

On the other hand,

$$\Phi^n \geq w_i^n = \exp\left(-\frac{\eta}{k} \sum_{t=1}^{n} \hat{l}_i(t)\right)$$

Combining the bounds above, we get

$$-\frac{\eta}{k} \sum_{t=1}^{n} \hat{l}_i(t) \leq \log k - \frac{\eta}{k(1-\eta)} \sum_{t=1}^{n} l_{I_t}(t) + \frac{3}{2} \frac{\eta^2}{k^2(1-\eta)} \sum_{i=1}^{k} \sum_{t=1}^{n} \hat{l}_i(t)$$

Rearranging the terms and taking expectation on both sides

$$\frac{\eta}{k(1-\eta)} \mathbb{E}\left[\sum_{t=1}^{n} l_{I_t}(t)\right] \leq \log k + \frac{\eta}{k(1-\eta)} \sum_{t=1}^{n} \mathbb{E}[\hat{l}_i(t)] + \frac{3}{2} \cdot \frac{\eta^2}{k^2(1-\eta)} \cdot \sum_{t=1}^{n} \sum_{i=1}^{k} \mathbb{E}[\hat{l}_i(t)]$$

Claim: $\mathbb{E}[\hat{l}_i(t)] = l_i(t)$

Pf. $\mathbb{E}[\hat{l}_i(t) \mid I_1, \ldots, I_t] = \frac{l_i(t)}{P_t(i)} \cdot \mathbb{1}_{\{I_t = i\}}$

$$\Rightarrow \mathbb{E}[\hat{l}_i(t)] = l_i(t)$$

Therefore,

$$\mathbb{E}[l^n_{EXP_3}] \leq \frac{k \log k}{\eta} + l^n_{min} + \frac{3}{2} \cdot \eta \cdot \sum_{t=1}^{n} \frac{1}{k} \sum_{i=1}^{k} l_i(t)$$

$$\Rightarrow \mathbb{E}\left[\ell^n_{EXP_3}\right] - \ell^n_{min} \le \frac{k \log k}{\eta} + \frac{3}{2}\eta n$$

$$= \frac{5}{2}\sqrt{n k \log k} \quad \text{for } \eta = \sqrt{\frac{k \log k}{n}}.$$

E Lower Bounds

Both $\bar{r}^{PP}(k,n)$ and $\bar{r}^{MAB}(k,n)$ have the following common form:

$$\bar{r}(k,n) = \min_{\hat{\theta}} \max_{y^n \in \{0,1\}^n} \max_{1 \le i \le k} \mathbb{E}\left[\sum_{t=1}^n \mathbb{1}_{\{x_{I_t}(t) \ne y_t\}}\right] - \sum_{t=1}^n \mathbb{1}_{\{x_i(t) \ne y_t\}}$$

The two problems only differ in their information structure. We follow the same approach to derive lower bounds for both problems:

$$\bar{r}(k,n) \ge \bar{R}(k,n) = \min_{\hat{\theta}} \max_{\substack{p_1,\dots,p_k \\ \text{where} \\ Z_i^n \sim \text{iid } Ber(p_i)}} \mathbb{E}\left[\sum_{t=1}^n Z_{I_t}(t) - \min_{i \in [k]} \sum_{t=1}^n Z_i(t)\right]$$

To derive our lower bound, we relate the regret problem above to that of finding the i minimizing p_i, namely the "best arm identification" problem.

The particular example we work with is the following:

Generate $J \sim unif[k]$ and let $p_J = \frac{1}{2} - \varepsilon$ and $p_i = \frac{1}{2}$ $\forall i \ne J$.

Denote by $\mathbb{P}_j$ and $\mathbb{E}_j$ the probabilities under $J=j$, and $\mathbb{E}_*$ the overall expectation. Then, for any $\hat{\theta}$,

$$\mathbb{E}_*\left[\sum_{t=1}^{n} Z_{I_t}(t) - \min_{i\in[k]} \sum_{t=1}^{n} Z_i(t)\right]$$

$$= \frac{1}{k}\sum_{j=1}^{k} \mathbb{E}_j\left[\sum_{t=1}^{n} Z_{I_t}(t) - \min_{i\in[k]} \sum_{t=1}^{n} Z_i(t)\right]$$

$$\geq \frac{1}{k}\sum_{j=1}^{k} \mathbb{E}_j\left[\sum_{t=1}^{n} Z_{I_t}(t)\right] - \min_{i\in[k]} \mathbb{E}_j\left[\sum_{t=1}^{n} Z_i(t)\right]$$

$$= \frac{1}{k}\sum_{j=1}^{k}\sum_{t=1}^{n} \mathbb{E}_j\left[Z_{I_t}(t)\right] - n\left(\frac{1}{2}-\varepsilon\right)$$

Note that $\mathbb{E}_j\left[Z_{I_t}(t)\right] = \underbrace{\sum_{i=1}^{k} \mathbb{P}_j(I_t=i)\, \mathbb{E}_j\left[Z_i(t)\right]}$

This uses an important property of our construction, namely that random losses Z_i^n are generated indep. of $I_1,\dots,I_n$.

$$= \frac{1}{2} - \varepsilon \sum_{i=1}^{k} \mathbb{P}_j(I_t=i)\, \mathbb{1}_{\{i=j\}} = \frac{1}{2} - \varepsilon\, \mathbb{P}_j(I_t=j)$$

Denote by N_j the number of times $I_t=j$ is chosen in $t=1,\dots,n$, i.e., $N_j = \sum_{t=1}^{n} \mathbb{1}_{\{I_t=j\}}$.

On combining the bounds above, we get

$$\mathbb{E}_*\left[\sum_{t=1}^{n} Z_{I_t}(t) - \min_{i\in[k]} \sum_{t=1}^{n} Z_i(t)\right]$$

$$\geq \varepsilon\left(\underbrace{n - \mathbb{E}_*[N_J]}\right)$$

average # of times the best arm is not selected

Our goal now is to bound $\mathbb{E}_*[N_J]$ for PP and MAB. The idea is that $\mathbb{E}_*[\cdot]$ is not very different from $\mathbb{E}_0$ where each $Z_i^n \sim \mathrm{Ber}(1/2)$. Formally, we will show:

<u>Lemma</u>

$$\mathbb{E}_*[N_J] \le \frac{1}{k}\sum_{j=1}^{k} \mathbb{E}_0[N_j] + \varepsilon \cdot n\sqrt{n} \cdot g(k),$$

where

$$g(k) = \begin{cases} c/\sqrt{k} & \text{for MAB} \\ c & \text{for PP.} \end{cases}$$

Using this lemma, and noting that

$$\sum_{j=1}^{k} N_j = n,$$

we get

$$\bar{R}(k,n) \ge \varepsilon\, n\left(1-\frac{1}{k}\right) - c\, n\sqrt{n}\, \varepsilon^2 g(k)$$
$$\ge n\left[\varepsilon - \varepsilon^2 \sqrt{n}\, g(k)\right]$$

which for $\boxed{\varepsilon = \frac{1}{2\sqrt{n}\, g(k)}}$ gives

(2) Can be selected when $\sqrt{\frac{k}{n}} \le 1$.

$$\bar{R}(k,n) \ge \frac{\sqrt{n}}{4g(k)} = \frac{1}{4c} \cdot \begin{cases} \sqrt{nk}, & \text{for MAB,} \\ \sqrt{n}, & \text{for PP.} \end{cases}$$

We have:

$$\Omega(\sqrt{n}) \le \bar{\mathcal{R}}^{PP}(k,n) \le O(\sqrt{n \log k})$$
$$\Omega(\sqrt{nk}) \le \bar{\mathcal{R}}^{MAB}(k,n) \le O(\sqrt{nk\log k}), \quad \boxed{\text{when } k \le c^2 n}$$

(The case $k > c^2 n$ will be assigned as HW)

Proof of Lemma

Note that N_j depends only on $I_1, \dots, I_n$ which in turn depends on $O_1, \dots, O_n$. Recall that O_t includes which arm was selected at time j for $1 \le j \le t-1$. Denote by P the distribution of O^n under $\mathbb{P}_0$ and Q its distribution under $\mathbb{P}_j$. Then,

$$\mathbb{E}_j[N_j] - \mathbb{E}_0[N_j] \le 2n\, d(P, Q)$$
$$\le \sqrt{2 \ln 2} \cdot n \cdot \sqrt{D(P \| Q)}$$

Consider

$$D(P_{O^n} \| Q_{O^n}) = \sum_{t=0}^{n-1} D(P_{O_{t+1}|O^t} \| Q_{O_{t+1}|O^t} \,|\, P_{O^t})$$

$\boxed{\text{For MAB}}$

Note that $O_{t+1} = (I_{t+1}, Z_{I_{t+1}}(t+1))$. Furthermore, given a fixed $O^t = o^t$

$$P_{I_{t+1}|O^t}(i | o^t) = Q_{I_{t+1}|O^t}(i | o^t).$$

Therefore, $D(P_{O_{t+1}|O^t} \| Q_{O_{t+1}|O^t} | P_{O^t})$

(3) $$= D(P_{Z_{I_{t+1}}(t+1)|O^t, I_{t+1}} \| Q_{Z_{I_{t+1}}(t+1)|O^t, I_{t+1}} \,|\, P_{O^t, I_{t+1}})$$

$$= \mathbb{E}_0\left[\mathbb{1}_{\{I_{t+1} = j\}} D\left(\tfrac{1}{2} \,\|\, \tfrac{1}{2} - \varepsilon\right)\right]$$

$$\leq c' \cdot \varepsilon^2 \; \mathbb{E}_0\left[\mathbb{1}_{\{I_{t+1}=j\}}\right]$$

Thus,

$$\mathbb{E}_j[N_j] - \mathbb{E}_0[N_j]$$

$$\leq \sqrt{c' \cdot 2\ln 2} \cdot n \cdot \sqrt{\sum_{t=1}^{n} \mathbb{E}_0\left[\mathbb{1}_{\{I_{t+1}=j\}}\right]} .$$

$$\Rightarrow \mathbb{E}_*[N_J] - \frac{1}{k}\sum_{j=1}^{k} \mathbb{E}_0[N_j]$$

$$\leq c \cdot \varepsilon \cdot n \cdot \frac{1}{k}\sum_{j=1}^{k} \sqrt{\sum_{t=1}^{n} \mathbb{E}_0\left[\mathbb{1}_{\{I_t=j\}}\right]}$$

$$\leq c \cdot \varepsilon \cdot n \cdot \sqrt{\frac{1}{k}\sum_{t=1}^{n} \mathbb{E}_0\left[\sum_{j=1}^{k} \mathbb{1}_{\{I_t=j\}}\right]}$$

$$= \varepsilon \cdot n\sqrt{n} \cdot \frac{c}{\sqrt{k}},$$

which completes the proof for MAB.

$\boxed{\text{For PP}}$

Now $O_t = (Z_1(t), \ldots, Z_k(t), I_t)$. Only (3) in the previous analysis changes to

$$D\left(P_{O_{t+1}|O^t} \,\|\, Q_{O_{t+1}|O^t} \,\middle|\, P_{O^t}\right)$$

$$= D\left(P_{Z(t+1)|O^t, I_{t+1}} \,\|\, Q_{Z(t+1)|O^t, I_{t+1}} \,\middle|\, P_{O^t, I_{t+1}}\right)$$

$= D\left(\frac{1}{2} \,\|\, \frac{1}{2} - \varepsilon\right)$. Following the remaining steps, we have the result.

Lecture 20 (contd.)

F Universal portfolio and prediction using MWU

Recall that our goal was to find $\hat{\theta}$ such that $\hat{\theta}_{t+1}(\cdot \mid \underline{x}^t) \in \mathcal{P}_k$ and minimize

$$\max_{P \in \mathcal{P}_k} \log \frac{\prod_{t=1}^{n} \left(\sum_{i=1}^{k} P_i \, x_i(t) \right)}{\prod_{t=1}^{n} \left(\sum_{i=1}^{k} \hat{\theta}_t(i \mid \underline{x}^{t-1}) \, x_i(t) \right)}$$

Using Cover's universal portfolio, we established that the problem is equivalent to probability assignment, and the minimum regret equals

$$r(k,n) = \frac{k-1}{2} \log n + O(1)$$

But the proposed algorithm had computational complexity that depends exponentially on k.

We now show that this problem is related closely to the prediction problem, but with time dependent loss function.

Specifically,

$$\log \frac{\prod_{t=1}^{n} \left(\sum_{i=1}^{k} P_i x_i(t) \right)}{\prod_{t=1}^{n} \sum_{i=1}^{k} \hat{\theta}_t(i|\underline{x}^{t-1})\, x_i(t)}$$

$$= \sum_{t=1}^{n} \log \frac{\sum_{i=1}^{k} P_i x_i(t)}{\sum_{i=1}^{k} \hat{\theta}_t(i|\underline{x}^{t-1})\, x_i(t)}$$

$$\leq \sum_{t=1}^{n} \left[\frac{\sum_{i=1}^{k} P_i x_i(t)}{\sum_{i=1}^{k} \hat{\theta}_t(i|\underline{x}^{t-1})\, x_i(t)} - \frac{\sum_{i=1}^{k} \hat{\theta}(i|\underline{x}^{t-1})\, x_i(t)}{\sum_{i=1}^{k} \hat{\theta}(i|\underline{x}^{t-1})\, x_i(t)} \right]$$

Let

$$l_i(t) = B - \frac{x_i(t)}{\sum_{j=1}^{k} \hat{\theta}_t(j|\underline{x}^{t-1})\, x_j(t)}$$

where by assuming $x_i(t) \in [c, C]$, $B = \frac{C}{c} > 1$.

Thus, the regret can be bounded by

$$\sum_{t=1}^{n} \mathbb{E}\left[l_{I_t}(t) \right] - \sum_{i=1}^{k} P_i \sum_{t=1}^{n} l_i(t),$$

where $l_i(t) \leq B$.

Using MWU, we can bound the right-side by $B\sqrt{n \log k}$

for every $P \in \mathcal{P}_k$.

Remark: (i) The algorithm above is easier to implement in comparison with Cover's universal portfolio, but has regret depends on n as $\sqrt{n}$. Surprisingly, the dependence of regret on k is only logarithmic.

(ii) While the use of MWU is semantically similar to our earlier use, but there is a subtle difference: The loss function now depends on time (and observations till that time). The power of MWU stems from the fact that it can handle such loss functions.

$\boxed{F}$ Probability assignment and prediction

Note that our MWU algorithm actually assigned probabilities to all sequences $i_1, \ldots, i_n \in [k]^n$. Specifically,

$$Q^{(n)}(\underline{i}) = \frac{\exp\left(-\eta \sum_{t=1}^{n} l_{i_t}(t)\right)}{\sum_{\underline{i}'} \exp\left(-\eta \sum_{t=1}^{n} l_{i'_t}(t)\right)}$$

where the denominator equals

$$\sum_{\underline{i}'} \prod_{t=1}^{n} \exp\left(-\eta \, l_{i'_t}(t)\right)$$

$$= \prod_{t=1}^{n} \sum_{i \in [k]} \exp\left(-\eta \, l_i(t)\right)$$

ignore

This is closer to Cover's assignment

$$Q^n(i_1, \ldots, i_n) = \prod_{t=1}^{n} Q_t(i_t)$$

where $Q_{t+1}(i) \propto \exp\left(-\eta \sum_{j=1}^{t} l_i(j)\right)$

$$= \exp\left(-\eta \sum_{j=1}^{t} l_i(j)\right) \Big/ \Phi^t$$

A log-moment generating function perspective

Let $L_t = \ell_{I_t}(t)$. Then, for $L_1, \dots, L_n$,

$$\log \mathbb{E}\left[e^{-\eta \sum_{t=1}^{n} L_t}\right]$$

$$\leq \mathbb{E}\left[-\eta \sum_{t=1}^{n} L_t + \frac{\eta^2}{2}\left(\sum_{t=1}^{n} L_t\right)^2\right]$$

in general. But for our independent choice,

$$\log \mathbb{E}\left[e^{-\eta \sum_{t=1}^{n} L_t}\right] \leq$$

$$= \sum_{t=1}^{n} \log \mathbb{E}\left[e^{-\eta L_t}\right]$$

$$\leq \sum_{t=1}^{n}\left[-\eta \mathbb{E}[L_t] + \frac{\eta^2}{2}\mathbb{E}[L_t^2]\right]$$

$$\leq -\eta \mathbb{E}\left[\sum_{t=1}^{n} L_t\right] + \frac{\eta^2}{2} n.$$

On the other hand,

$$\log \mathbb{E}\left[e^{-\eta L_{t+1}}\right] = \log \sum_i p_{t+1}(i) e^{-\eta L_{t+1}}.$$

$$= \log \frac{\Phi^{t+1}}{\Phi^{t}}$$

$$\Rightarrow \sum_{t=1}^{n} \log \mathbb{E}\left[e^{-\eta L_t}\right] = \sum_{t=1}^{n} \log \frac{\Phi^{t}}{\Phi^{t-1}}$$

$$= \log \frac{\Phi^{n}}{\Phi^{0}} = \log \frac{1}{k} \sum_{i=1}^{k} \exp\left[-\eta \sum_{t=1}^{n} l_i(t)\right]$$

$$\geq \frac{1}{k} \sum_{i=1}^{k} \left(-\eta \sum_{i=1}^{n} l_i(t)\right)$$

(by concavity of log)

Therefore,

$$\mathbb{E}\left[\sum_{t=1}^{n} L_t\right] - \underbrace{\frac{1}{k} \sum_{i=1}^{k} \left(\sum_{t=1}^{n} l_i(t)\right)}_{\text{average loss of experts}}$$

$$\leq \frac{\eta n}{2},$$

which can be made arbitrarily small!

Also,

$$\max_{i} -\eta \sum_{t=1}^{n} l_i(t) \leq -\eta \mathbb{E}\left[\sum_{t=1}^{n} L_t\right] + \log k + \eta^2 n/2;$$

$$\Rightarrow \quad \mathbb{E}\left[\sum_{t=1}^{n} L_t\right] - \sum_{t=1}^{n} \ell_i(t)$$

$$\leq \frac{\log k}{\eta} + \frac{\eta n}{2} \quad \forall \eta > 0.$$

which recovers our original regret bound.

Thus, what we have shown is the following:

for $I \sim \text{unif}[k]$,

$$\log \frac{\mathbb{E}\left[e^{-\eta \sum_{t=1}^{n} L_t}\right]}{\mathbb{E}\left[e^{-\eta \sum_{t=1}^{n} \ell_I(t)}\right]} \leq -\eta\, \mathbb{E}\left[\sum_{t=1}^{n} L_t\right] + \frac{\eta^2 n}{2}$$