Prediction of Online Sequences

Nicolas CLOAREC
Tom ROUSSEAU

Homework

1 Theory - Sleeping expert

The classic definition of regret compares the performance of an algorithm with the performance of the best "constant" action. But in some applications, some actions may be sometimes unavailable. The purpose of this exercise is to deal with this issue.

We consider the following full-information setting with a finite set $\mathcal{X} = \{1, \ldots, K\}$. At each time $t \geq 1$, a subset of active decisions $A_t \subseteq \mathcal{X}$ is available, the other decisions are sleeping (or inactive) and cannot be chosen; the player chooses a distribution p_t over active decisions A_t (i.e., $\sum_{j \in A_t} p_t(j) = 1$ and $p_t(k) = 0$ for $k \notin A_t$) and observes the loss $\ell_t(k) \in [0,1]$ of all decisions in A_t. The **sleeping regret** is defined

$$R_T(k) := \sum_{t=1}^{T} (p_t \cdot \ell_t - \ell_t(k)) \mathbb{1}\{k \in A_t\} \tag{1}$$

with respect to decision $k \in \mathcal{X}$, where $p_t \cdot \ell_t = \sum_{j \in A_t} p_t(j)\ell_t(j)$ is the loss of the player.

1. **The prod algorithm** Here, we consider the case where all experts are active $A_t = \mathcal{X}$ for all $t \geq 1$. Let $\eta(1), \ldots, \eta(K) \leq 1/2$ be K parameters. We define the weights

$$p_t(k) = \frac{\eta(k) w_t(k)}{\sum_{j=1}^{K} \eta(j) w_t(j)} \text{ where } w_t(k) = \prod_{s=1}^{t-1} (1 + \eta(k)(p_s \cdot \ell_s - \ell_s(k))) \text{ if } t \geq 2 \text{ and } w_1(k) = 1, \tag{2}$$

for all $k \in \mathcal{X}$ and $t \geq 1$.

Question 1. Prove that $\log(1+x) \geq x - x^2$ for $x \geq -1/2$.

Answer 1. Denoting $\phi :]-1;+\infty] \to \mathbb{R}, x \mapsto \log(1+x) - x + x^2$. ϕ is clearly differentiable on its domain, and we have

$$\forall x > -1, \qquad \phi'(x) = \frac{1}{1+x} - 1 + 2x = \frac{x(2x+1)}{1+x}.$$

For $x \geq -1/2$, we have ϕ decreasing on $[\frac{-1}{2};0]$ and increasing on $[0;+\infty[$. Then, for all $x \geq -1/2$

$$\phi(x) = \log(1+x) - x + x^2 \geq \phi(0) = 0,$$

which shows the desired result.

Question 2. Denoting $W_t = \sum_{k=1}^{K} w_t(k)$. Prove that for all $k \in \mathcal{X}$

$$\log W_{T+1} \geq \eta(k) \sum_{t=1}^{T} (p_t \cdot \ell_t - \ell_t(k)) - (\eta(k))^2 \sum_{t=1}^{T} (p_t \cdot \ell_t - \ell_t(k))^2 \tag{3}$$

Answer 2. First of all, we will assume that all parameters $\eta(1),\dots\eta(K)$ to be (*strictly*) greater than 0. If it is not the case, weights $p_t(k)$ are ill-defined as the denominator can be equal to zero and weights can take negative values, which is, of course, not what we want as p_t denotes probability distribution over $\mathcal{X}$.

Assuming $\eta(k) \in]0;-1/2]$ and since $\ell_t(k) \in [0;1]$, we have

$$\eta(k)\,(\,p_s \cdot \ell_s - \ell_s(k)\,) \geq \frac{1}{2}\,(\,0-1\,) = -1/2,$$

which implies that $w_t(k) \geq 0$ for all $k \in \mathcal{X}$.

Using the previous inequality with $x = \eta(k)\,(\,p_s \cdot \ell_s - \ell_s(k)\,) \geq -1/2$, we have for all $k \in \mathcal{X}$

$$\begin{aligned}\log W_{T+1} = \log \sum_{j=1}^{K} w_{T+1}(j) \geq \log w_{T+1}(k) \\ &\geq \sum_{s=1}^{T} \log\{\,1+\eta(k)\,(p_s \cdot \ell_s - \ell_s(k))\,\} \\ &\geq \sum_{s=1}^{T} \eta(k)\,(p_t \cdot \ell_s - \ell_s(k)) - \eta(k)^2\,(p_s \cdot \ell_s - \ell_s(k))^2\,.\end{aligned}$$

Question 3. Show that $W_{t+1} = W_t$ for all $t \geq 1$. What is the value of $\log(W_{t+1})$?

Answer 3. Using the definition of W_{t+1}, we have

$$\begin{aligned}W_{t+1} &= \sum_{k\in[K]} w_{t+1}(k) \\ &= \sum_{k\in[K]} w_t(k)\{\,1+\eta(k)\{\,p_t \cdot \ell_t - \ell_t(k)\}\,\} \\ &= W_t + \sum_{k\in[K]} w_t(k)\eta(k)\{\,p_t \cdot \ell_t - \ell_t(k)\} \\ &= W_t + \left(\sum_{j\in[K]} \eta(j)w_t(j)\right) \sum_{k\in[K]} p_t(k)\{\,p_t \cdot \ell_t - \ell_t(k)\}.\end{aligned}$$

A small computation shows that

$$\sum_{k\in[K]} p_t(k)\{\,p_t \cdot \ell_t - \ell_t(k)\} = (\,p_t \cdot \ell_t\,) \underbrace{\sum_{k\in[K]} p_t(k)}_{=1} - p_t \cdot \ell_t = 0.$$

Using the initial values for $w_1(k)$ with $k \in \mathcal{X}$ and the previous result, we have

$$\log W_{T+1} = \log W_1 = \log \sum_{k\in[K]} w_1(k) = \log \sum_{k\in[K]} 1 = \log K.$$

Question 4. Assuming $\eta(k)$ are well-optimized, show the regret bound for all arms $k \in [K]$

$$\sum_{t=1}^{T} p_t \cdot \ell_t - \ell_t(k) \le 2\sqrt{(\log K) \sum_{t=1}^{T} (p_t \cdot \ell_t - \ell_t(k))^2}. \tag{4}$$

Answer 4. By taking in (3)

$$\eta(k) = \sqrt{\frac{\log W_{T+1}}{\sum_{t=1}^{T} (p_t \cdot \ell_t - \ell_t(k))^2}},$$

and replacing $\log W_{T+1}$ with $\log K$ yields to the desired inequality (4).

2. **Sleeping experts** Now we assume that some decisions are sometimes not possible (sleeping), i.e., $A_t \subsetneq \mathcal{X}$ for some $t \ge 1$. The idea is to use Algorithm above with past modified losses

$$\tilde{\ell}_t(k) := \begin{cases} \ell_t(k) & \text{if } k \in A_t \\ p_t \cdot \ell_t = \sum_{k \in A_t} p_t(k)\ell_t(k) & \text{if } k \notin A_t \end{cases} \tag{5}$$

i.e., by assigning the loss of the algorithm $p_t \cdot \ell_t$ to all inactive decisions $k \notin A_t$. The algorithm outputs weights $\tilde{p}_t(k)$ and $\tilde{w}_t(k)$ obtained by replacing $\ell_t(k)$ with $\tilde{\ell}_t(k)$ in (2). This vector is then used to form another weight vector

$$p_t(k) = \frac{\tilde{p}_t(k)\mathbb{1}_{k \in A_t}}{\sum_{j=1}^{K} \tilde{p}_t(j)\mathbb{1}_{j \in A_t}} \tag{6}$$

which has non zero weights only on active arms A_t.

Question 5. Show that the instantaneous regret on the modified losses equals the sleeping regret on the original rewards; i.e. forall $t \ge 1$, and all $k \in \mathcal{X}$

$$\tilde{p}_t \cdot \tilde{\ell}_t - \tilde{\ell}_t(k) = (p_t \cdot \ell_t - \ell_t(k))\, \mathbb{1}_{k \in A_t} \tag{7}$$

Answer 5. Let $k \in \mathcal{X}$. Using definitions of $\tilde{\ell}_t(k)$ and $\tilde{p}_t(k)$, we get

$$\begin{aligned}
\tilde{p}_t \cdot \tilde{\ell}_t - p_t \cdot \ell_t &= \sum_{k \in [K]} \tilde{p}_t(k) \left\{ \tilde{\ell}_t(k) - p_t \cdot \ell_t \right\} \\
&= \sum_{k \in A_t} \tilde{p}_t(k) \left\{ \tilde{\ell}_t(k) - p_t \cdot \ell_t \right\} + \sum_{k \notin A_t} \tilde{p}_t(k) \underbrace{\left\{ \tilde{\ell}_t(k) - p_t \cdot \ell_t \right\}}_{=0} \\
&= \left(\sum_{j \in A_t} \tilde{p}_t(j) \right) \underbrace{\sum_{k \in A_t} p_t(k) \{ \ell_t(k) - p_t \cdot \ell_t \}}_{=0 \text{ as shown in Q3.}} = 0.
\end{aligned}$$

- If $k \notin \mathcal{A}_t$, $\tilde{\ell}_t(k) = p_t \cdot \ell_t$ and (7) is reduced to $\tilde{p}_t \cdot \tilde{\ell}_t - p_t \cdot \ell_t = 0$ which is true as shown above.
- If $k \in A_t$, $\tilde{\ell}_t(k) = \ell_t(k)$ and (7) is reduced to $\tilde{p}_t \cdot \tilde{\ell}_t - \tilde{\ell}_t(k) = p_t \cdot \ell_t - \ell_t(k)$ which is also true.

Question 6. Conclude that $R_T(k) \leq 2\sqrt{(\log K) T_k}$ where $T_k = \sum_{t=1}^{T} \mathbb{1}\{k \in A_t\}$ is the number of times arm k is active.

Answer 6. The sleeping regret with losses ℓ_t [1] and weights p_t defined in (1) is equal to usual regret with losses $\tilde{\ell}_t$ and weights $\tilde{p}_t$. Applying the results of the first question with corresponding losses and weights [2], we have

$$\sum_{t=1}^{T} \tilde{p}_t \cdot \tilde{\ell}_t - \tilde{\ell}_t(k) \leq 2\sqrt{(\log K) \sum_{t=1}^{T} \left(\tilde{p}_t \cdot \tilde{\ell}_t - \tilde{\ell}_t(k)\right)^2}.$$

Using the formula (7), we get an upper bound of the right term in the above inequality

$$\sum_{t=1}^{T} \left(\tilde{p}_t \cdot \tilde{\ell}_t - \tilde{\ell}_t(k)\right)^2 = \sum_{t=1}^{T} (p_t \cdot \ell_t - \ell_t(k))^2 \, \mathbb{1}_{k \in A_t} \leq \sum_{t=1}^{T} \mathbb{1}_{k \in A_t}$$

where we have used the fact that $\ell_t(k) \in [0;1]$ for all $k \in \mathcal{A}_t$.

1. we only observe ℓ_t on the subset of active decisions A_t.
2. we note that while for ℓ_t we get a partial feedback on $A_t \subsetneq \mathcal{X}$, we have a full-information setting for $\tilde{\ell}_t$.