CS 170 DIS 13

Released on 2018-04-24

1 Follow the regularized leader

(a) **Follow the leader.** You are playing T rounds of the following game: At round t you pick one of n strategies; your payoff for picking strategy i is $A(t, i) \in [0, 1]$. You try the following algorithm: at each iteration pick the strategy which gave the highest average payoff so far (on the first iteration, you pick strategy 1).

Give an example of payoffs for $T = 100$ and $n = 2$, where your algorithm obtains a payoff of 0, but sticking to either $i = 1$ or $i = 2$ would have given you a payoff of almost 50.

Solution: Let the payoffs be at $t = 1$, $(0, 1 - \varepsilon)$ and then for every other odd t be $(0, 1)$ and for each even t be $(1, 0)$.

Prior to every even numbered round, the strategy with the higher average payoff is strategy 2. Similarly, prior to ever odd numbered round, the strategy with the higher average payoff is strategy 1. But, by construction, this will yield an overall payoff of 0 as the strategies alternate in success in dissonance with the alterantion of average payoff.

If you stuck to strategy 1, you would obtain a payoff of 50 and if you stuck to strategy 2, you would obtain a payoff of $50 - \varepsilon$.

(b) **Follow the randomized leader.** The reason the algorithm above didn't do so well, is because when we deterministically jump from one strategy to another, an adversarially chosen set of strategies can be designed to thwart the algorithm.

To trick such adversaries, we want to use a *randomized* strategy; at time t we pick our strategy i at random from distribution D_t. Let $p_t(i) \geq 0$ denote the probability that we assign to strategy i (i.e. $\sum_{i=1}^{n} p_t(i) = 1$).

The previous algorithm ("Follow the leader") corresponds to setting D_t that maximizes

$$\sum_{i=1}^{n} \left(p_t(i) \cdot \sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau, i)] \right).$$

Why is this no better? (Hint: convexity).

Solution: A randomized solution is a linear combination of deterministic solutions. By the convexity of linear combinations, it can do no better than a deterministic solution.

(c) **Follow the regularized leader.** Instead, it is common to add a /regularized term that favors smoother distributions. A commonly used regularizer is the entropy function, i.e. we want to use pick i from the distribution that maximizes

$$\sum_{i=1}^{n} \left(p_t(i) \cdot \sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau, i)] - \eta p_t(i) \ln p_t(i) \right). \tag{1}$$

(Here, $\eta > 0$ is a parameter that we can tweak to balance exploration and exploitation. Notice also that $\ln p_t(i) \le 0$.)

In this exercise you will show that following the regularized leader with the entropy regularizer is the same as Multiplicative Weights Update!

Show that for any distribution p_t, (1) is at most

$$\eta \cdot \ln \left(\sum_{i=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} \right) \tag{2}$$

(Hint: you may use the inequality $\sum_{i=1}^{n} p_t(i) \cdot \ln(y_i) \le \ln\left(\sum_{i=1}^{n} p_t(i) \cdot y_i\right)$ for any vector $\vec{y}$.)

Solution:

$$\begin{aligned}
(1) &= \eta \sum_{i=1}^{n} p_t(i) \cdot \left(\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta - \ln p_t(i) \right) \\
&= \eta \sum_{i=1}^{n} p_t(i) \cdot \ln \left(e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} / p_t(i) \right) \\
&\le \eta \cdot \ln \left(\sum_{i=1}^{n} \left(p_t(i) \cdot e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} / p_t(i) \right) \right) \\
&= \eta \cdot \ln \left(\sum_{i=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} \right).
\end{aligned}$$

(d) Show that for some choice of ϵ (which depends on η), when computing p_t using Multiplicative Weight Update, (1) is equal to (2). What is the dependence of ϵ on η?

Solution: When using the MWU algorithm, we have:

$$\begin{aligned}
p_t(i) &= \frac{w_t(i)}{\sum_{j=1}^{n} w_t(j)} \\
&= \frac{(1-\epsilon)^{-\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]}}{\sum_{j=1}^{n} (1-\epsilon)^{-\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,j)]}}.
\end{aligned}$$

If we set ϵ such that $(1-\epsilon) = e^{-1/\eta}$, and plugin into the last equation, we have:

$$p_t(i) = \frac{e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta}}{\sum_{j=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,j)]/\eta}}.$$

Therefore,

$$
\begin{aligned}
(1) &= \eta \sum_{i=1}^{n} p_t(i) \cdot \ln\left(e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta} / p_t(i)\right) \\
&= \eta \sum_{i=1}^{n} p_t(i) \cdot \ln\left(\sum_{i=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta}\right) \\
&= \eta \cdot \ln\left(\sum_{i=1}^{n} e^{\sum_{\tau \in \{1,\ldots,t-1\}} [A(\tau,i)]/\eta}\right).
\end{aligned}
$$