CS 170 DIS 12

Released on 2019-4-15

1 Perfect Expert

In this question, we will analyze a simplified version of multiplicative weights in the presence of a perfect expert. In this version, a player chooses in each round, one of two actions, a_1 or a_2. The player is aided by a set of n experts $e_1, \ldots, e_n$ where the i^{th} expert suggests action $a_i^{(t)}$ in the t^{th} round. Based on this choice, the player is rewarded with either 1 or 0. As always, the reward for the player at the end of the process is the sum of the rewards at the end of T rounds. Suppose that the player makes choices $c_1, \ldots, c_T$ at the T rounds, then the total reward of the player is $\sum_{i=1}^{T} r_{c_t,t}$ where $r_{c_t,t}$ is the reward at round t for choice c_t. Suppose that there is an expert whose choice always produces a reward of 1, analyze the following instance of the multiplicative weights strategy.

1. First assign each expert, e_i, with the weight $w_i^{(1)} = 1$.
2. At each time step, choose $c_t = \arg\max_{i\in\{1,2\}} \sum_{j=1}^{n} w_i^{(t)} 1(c_t = i)$
3. If the player receives 0 reward, update $w_i^{(t+1)} = 0$ if $a_i^{(t)} = c_t$.

(a) Show that the weight for the perfect expert, e_{i^*} remains 1. 2cm

(b) Let $W_t = \sum_{i=1}^{n} w_i^{(t)}$. Show that $W_t \leq n\left(\frac{1}{2}\right)^{(t-1)-\sum_{i=1}^{t-1} r_{c_i,i}}$. 4cm

(c) Put the previous two bounds together to bound the expected regret of the algorithm described. 3cm

Solution:

(a) Note that the weight for an expert only changes if the expert made a mistake. Since, e_{i^*} never makes a mistake, their weight never changes.

(b) We will prove the statement via induction. The statement trivially holds true when $t = 1$. Let $t = k + 1$. In this scenario, we assume two cases:

Case 1: $r_{c_i,i} = 1$. In this case, none of the experts change their weights and the statement follows from the inductive hypothesis.

Case 2: $r_{c_i,i} = 0$. In this case, let $B_t = \{i : a_i^{(t)} = c_t\}$. We know that $W_{t+1} = \sum_{i\notin B_t} w_i^{(t)}$. Furthermore, we also know from the definition of c_t that $\sum_{i\in B_t} w_i^{(t)} \geq \sum_{i\notin B_t} w_i^{(t)}$. Therefore, we have:

$$W_{t+1} = \sum_{i\notin B_t} w_i^{(t)} \leq \frac{1}{2}\left(\sum_{i\notin B_t} w_i^{(t)} + \sum_{i\in B_t} w_i^{(t)}\right) = \frac{1}{2} W_t$$

This verifies the inductive hypothesis in the second case, proving the theorem.

(c) By applying previous statement with $t = T + 1$, we have:

$$1 = w_{i^*}^{(T+1)} \leq W_{T+1} \leq n \left(\frac{1}{2}\right)^{T - \sum_{i=1}^{T} r_{c_i,i}} \implies \sum_{i=1}^{T} r_{c_i,i} \geq T - \log n$$

2 Multiplicative Weights

Recall from the notes that in the experts problem, if there are n experts and the best expert has cost m, the randomized multiplicative weights algorithm has expected cost at most $(1+\epsilon)m + \frac{\ln n}{\epsilon}$.

(a) We run the randomized multiplicative weights algorithm with two experts and believe the best expert will have cost 10000. What choice of ϵ should we use to minimize the bound on the cost of the algorithm? 3cm

(b) We run the randomized multiplicative weights algorithm with two experts. In all of the first 140 days, Expert 1 has cost 0 and Expert 2 has cost 1. If we chose $\epsilon = 0.01$, on the 141st day with what probability will we play Expert 1? (Hint: You can assume that $0.99^{70} = \frac{1}{2}$) 3cm

Solution:

(a) Taking the derivative of $(1+\epsilon)m + \frac{\ln n}{\epsilon}$ with respect to ϵ gives $m - \frac{\ln n}{\epsilon^2}$. Setting this equal to 0 and solving for ϵ gives $\epsilon = \sqrt{\frac{\ln n}{m}}$. Plugging in $n = 2, m = 10000$ gives $\epsilon = \sqrt{\frac{\ln 2}{10000}} \approx .008326$.

(b) The weight assigned to expert 1 is $.99^{0 \cdot 140} = 1$, while the weight assigned to expert 2 is $.99^{1 \cdot 140} \approx 1/4$. So, the probability we play expert 1 is $\frac{1}{1+1/4} = 4/5$.

3 Multiplicative Rewards

Recall that in the classical experts scenario, the player first picks an action, a_t to play in each time step and observes a loss, l_t corresponding to the action they chose. We will now consider a different scenario where the losses are multiplicative instead. We will show that this is equivalent to the setting where one is given additive losses instead. You start off with 1$ to trade on a stock market with two stocks. In each time step, t, you are given the choice of investing your money in one of the two stocks. At each time step, stock i rises by $r_{t,i}\%$ (which is negative in the scenario that the stock falls in value). Show that this scenario can be transformed to the additive losses setting (You may assume that $|r_{t,i}| \leq 10$).

Solution: Suppose that we have x_t\$ at time step, t and we choose to invest in stock 1. Then at time $t+1$, stock 1, we are left with $(1 + r_{t,1}/100)x_t$. Suppose that we chose $c_1, \ldots, c_T$ at each of the time steps between 1 and T. At the end of this process, the amount of money we are left with is $\prod_{t=1}^{T}(1 + r_{t,c_t})$. By taking logarithms we see that this is equivalent to maximizing the following quantity:

$$\sum_{t=1}^{T} \log(1 + r_{t,c_t})$$

This is equivalent to the experts scenario where the rewards are $\log(1 + r_{t,c_t})$ instead.