

COMS 4773 Fall 2026 Problem Set 1

Problem 1. Consider the expert advice scenario with the realizability assumption, except now the outcome to predict is ternary: Rain, Shine, or Blizzard. (Assume these outcomes are mutually exclusive!)

- (a) Give an algorithm that ensures the learner makes no more than $C \log_2 N$ mistakes, where $C > 0$ is a positive constant that you should try to make as small as possible. What value C are you able to achieve? Give a proof that this mistake bound holds for your algorithm.
- (b) Let us now change the the expert advice scenario again; it is the same as above, with ternary outcomes, but at the end of the day, the learner does not actually get to observe the actual outcome. Instead, the learner is trapped inside an underground news-station, and only learns whether or not its prediction was correct. Now re-do Part (a) for this new scenario.

Problem 2. Consider the following modification of Weighted Majority (WM) for the expert advice scenario (with binary outcomes):

- Initialize $w_{1,i} = 1$ for all $i \in [N]$.
- For $t = 1, 2, \dots$:
 - Observe experts' predictions $b_{t,1}, \dots, b_{t,N} \in \{0, 1\}$.
 - For each possible outcome $y \in \{0, 1\}$, let

$$Z_{t,y} := \sum_{i=1}^N \mathbf{1}\{b_{t,i} = y\} w_{t,i}$$

be the total weight of experts that predict y .

- Predict

$$a_t := \begin{cases} 1 & \text{if } Z_{t,1} > Z_{t,0}, \\ 0 & \text{otherwise.} \end{cases}$$

- Observe outcome $y_t \in \{0, 1\}$.
- Update weights: for all $i \in [N]$,

$$w_{t+1,i} = \begin{cases} w_{t,i} & \text{if } y_t = b_{t,i} \text{ or } w_{t,i} < \frac{1}{3} \cdot \frac{1}{N} \sum_{j=1}^N w_{t,j}; \\ \frac{1}{2} w_{t,i} & \text{if } y_t \neq b_{t,i} \text{ and } w_{t,i} \geq \frac{1}{3} \cdot \frac{1}{N} \sum_{j=1}^N w_{t,j}. \end{cases}$$

This is like WM (with $\beta = 1/2$), except that a mistaken expert's weight is halved only if its weight was at least $1/3$ of the average weight in that round.

- (a) Let $Z_t = \sum_{i=1}^N w_{t,i}$ for all rounds t . Explain why, at the start of any round t , every expert i has weight $w_{t,i}$ satisfying

$$w_{t,i} \geq \frac{Z_t}{6N}.$$

(The modification to how weights are updated is important for this part.)

- (b) Prove that this modified version of WM has the following guarantee. Pick any segment of consecutive rounds, say, $a, a+1, a+2, \dots, b$, with $a \leq b$. Let $K_{a,b}$ be the smallest number of mistakes committed by an expert in the rounds $a, a+1, a+2, \dots, b$. Then the learner makes at most

$$8K_{a,b} + 12 \ln N + 21$$

mistakes in the rounds $a, a+1, a+2, \dots, b$.

The guarantee in part (b) is useful if the “good expert” changes over time. For example, suppose Expert 1 always predicts 1, and Expert 2 always predicts 0. Further, suppose the outcomes in the first 1000 rounds are all 1’s, so $K_{1,1000} = 0$ (achieved by Expert 1), but the outcomes in the next 1000 rounds are all 0’s, so $K_{1001,2000} = 0$ (achieved by Expert 2). Each of Expert 1 and Expert 2 makes 1000 mistakes over all 2000 rounds; neither is particularly good for all 2000 rounds. (Think about how well the original WM, with $\beta = 1/2$, would do in this example. . .) The learner, using this modified algorithm, will make at most a small number of mistakes overall:

- at most $8K_{1,1000} + 12 \ln 2 + 21 \approx 30$ mistakes in rounds $1, \dots, 1000$;
- at most $8K_{1001,2000} + 12 \ln 2 + 21 \approx 30$ mistakes in rounds $1001, \dots, 2000$.

The modification in how the weights are updated ensures that the learner never stops anticipating that some currently poorly-performing expert may become good later on.

Problem 3. Consider the following alternative algorithm for Online Allocation (assuming $\ell_t \in [0, 1]^N$ for all rounds t):¹

- Initialize $w_{1,i} = 1$ for all $i \in [N]$.
- For $t = 1, 2, \dots$:
 - Choose allocation vector $p_t = (p_{t,1}, \dots, p_{t,N}) \in \Delta^{N-1}$, where

$$p_{t,i} = \frac{w_{t,i}}{Z_t} \quad \text{for all } i \in [N]$$

and $Z_t = \sum_{i=1}^N w_{t,i}$.

- Get loss vector $\ell_t = (\ell_{t,1}, \dots, \ell_{t,N}) \in [0, 1]^N$.
- If $R_{t,i} \leq 0$ for all $i \in [N]$, then set $w_{t+1,i} = 1$ for all $i \in [N]$; otherwise set $w_{t+1,i} = \max\{R_{t,i}, 0\}$ for all $i \in [N]$.
 (Recall that $R_{t,i} = \sum_{s=1}^t \langle p_s, \ell_s \rangle - \sum_{s=1}^t \ell_{s,i}$.)

In this problem, you will prove that this algorithm guarantees $R_T \leq \sqrt{TN}$. (Yes, this has quite a bit worse dependence on N than what Hedge achieves, but at least the dependence on T is sublinear. And the algorithm is simple and has no hyperparameters!)

- (a) Define $r_{t,i} = \langle p_t, \ell_t \rangle - \ell_{t,i}$ for all $i \in [N]$. Prove that, for every round t , and every $i \in [N]$,

$$(\max\{R_{t,i}, 0\})^2 \leq (\max\{R_{t-1,i}, 0\})^2 + 2 \max\{R_{t-1,i}, 0\} r_{t,i} + r_{t,i}^2,$$

where we regard $R_{0,i} = 0$ for all $i \in [N]$.

- (b) Using the result from part (a) and the way the algorithm chooses allocation vectors, prove that

$$\sum_{i=1}^N (\max\{R_{t,i}, 0\})^2 \leq \sum_{i=1}^N (\max\{R_{t-1,i}, 0\})^2 + \sum_{i=1}^N r_{t,i}^2.$$

- (c) Use the previous two parts to complete the proof that, for any T , the regret of the learner after T rounds satisfies $R_T \leq \sqrt{TN}$.

(Remember that the loss vectors are chosen from $[0, 1]^N$.)

¹This algorithm is related to—but not the same as—the “NormalHedge” algorithm from <https://arxiv.org/abs/0903.2851>.