# Homework #3
# CMSC 35401: The Interplay of Economics and Machine Learning (Winter'24)

Due Saturday 02/17, 9:00 pm

**General Instructions** The assignment is meant to practice your understanding of course materials, and some of them are challenging. You are allowed to discuss with fellow students, however please write up your solutions independently (e.g., start writing solutions after a few hours of any discussion) and, equally importantly, acknowledge everyone you discussed the homework with on your writeup. All course materials are available on the course webiste here https://www.haifeng-xu.com/cmsc35401win24. You may refer to any materials covered in our class. However, any attempt to consult outside sources, on the Internet or otherwise, for solutions to any of these homework problems is *not* allowed.

Whenever a question asks you to "show" or "prove" a claim, please provide a formal mathematical proof. These problems have been labeled based on their difficulties. `Short` problems are intended to take you 5-15 minutes each and `medium` problems are intended to take 15-30 minutes each. `Long` problems may take anywhere between 30 minutes to several hours depending on whether inspiration strikes. Note that, the total score is meant to *not* be normalized to 100 (for instance, this HW has 30 in total for regular students and additional 15 points for those who take it as elective).

Finally, please write your solutions in latex — **hand written solutions will not be accepted.** Hope you enjoy the homework!

## Problem 1 (`Short, 5 points`)

When we argue that pseudo-regret is at most the (external) regret, we used the following fact: for any random vector $C \in \mathbb{R}^n$, we have $\min_{j\in[n]} \mathbb{E}[C(j)] \geq \mathbb{E}\big[\min_{j\in[n]} C(j)\big]$. Prove this claim.

(For this HW problem, you can assume $C$ has finite support, i.e., value of $C$ is from a finite set of vectors, though the stated conclusion above holds in general.)

## Problem 2: Exercising the Regret Analysis for Exponential-Weight Update

During lecture we mentioned a simple variant of multiplicative weight update algorithm, called the *Exponential-Weight (EW) update*, but omitted its regret analysis in the *full information* setting (i.e., the learner can see the whole cost vector). In this problem, you are asked to give a complete proof of EW's regret upper bound, with the following steps. Recall the notations: (1) there are $n$ actions in set $[n] = \{1, \cdots, n\}$; (2) $c_t \geq 0$ is the cost *vector* the learner observes at round $t$; (3) $W_t = \sum_{i=1}^{n} w_t(i)$ is the total weight at round $t$; (4) the update rule in EW is as follows: at time $t$, for any action $i$ we set $w_{t+1}(i) = w_t(i)e^{-\epsilon c_t(i)}$.

1. (`Short, 5 points`) Prove that $W_{t+1}/W_t = \sum_{i=1}^n p_t(i)e^{-\epsilon c_t(i)}$ where $p_t(i) = w_t(i)/W_t$.

2. (**Regret Bound of Exponential-Weight Update**, `Medium, 10 points`) Using the above conclusion, together with the fact we proved in class

$$\sum_{t=1}^{T} \log\big(\sum_{i=1}^{n} p_t(i)e^{-\epsilon c_t(i)}\big) \leq \sum_{t=1}^{T}\sum_{i=1}^{n} p_t(i)\big(-\epsilon c_t(i) + \frac{\epsilon^2}{2}[c_t(i)]^2\big),$$

prove the following regret bound for EW

$$R_T \leq \frac{\ln n}{\epsilon} + \frac{\epsilon}{2}\sum_{t=1}^{T}\sum_{i=1}^{n} p_t(i)[c_t(i)]^2.$$

# Additional Problems for those who take the course as Elective

These following problems are designed specifically for the very few *UChicago CS PhD* student who want to take the course as elective, hence will need these to fulfill their elective requirement.

Notes for regular students: *if you are interested*, you can try to solve these problems as well, but this is *not* a requirement for you. We will grade your solution just *by courtesy*, but your grades on these problems will NOT count towards your final grade.

## Problem 3: The Experts' Advice Problem

The *experts' advice* problem is a slight variant of the online learning problem. Here there are $n$ experts, each making a prediction about a *binary* event (e.g., the stock market will go up or down tomorrow). For round $t = 1, \cdots, T$ ($T > n$), the following occurs in order: (1) each expert $i$ makes a binary prediction $a_t(i) \in \{0, 1\}$; (2) after observing these predictions, the learner comes up with his own prediction $\overline{a_t}$; (3) the binary event is realized; (4) The learner observes whether she made a correct prediction as well as whether each expert made a correct prediction at this round. The learner's goal is to design an online learning algorithm that makes as few mistakes as possible for any set of experts and any event realization.

1. (`Short, 5 points`) Formalize the definition of regret in this setting.

2. (`Short, 5 points`) Assume that one of the expert is perfect, i.e., all his predictions are correct. Show that in this case, there exists a learning algorithm that has regret at most $O(\ln n)$.

3. (`Medium, 10 points`) If none of the experts are perfect, one natural algorithm to make predictions is to use a *weighted majority voting* rule. In particular, consider the following algorithm, parameterized by $\epsilon$.

   (a) Initialize $w_1(i) = 1$ for all expert $i$.

   (b) At round $t = 1 \cdots, T$

      i. After observing each expert's prediction, the learner computes the total weight for prediction 1 and 0, as $W_t^1 = \sum_{i=1}^n w_t(i) \cdot \mathbb{I}(a_t(i) = 1)$ and $W_t^0 = \sum_{i=1}^n w_t(i) \cdot \mathbb{I}(a_t(i) = 0)$, respectively, and predict 1 if and only if $W_t^1 \geq W_t^0$. Here, $\mathbb{I}(A)$ is the indicator function, which equals 1 if and only if $A$ is true and 0 otherwise.

ii. After the binary event is realized, update expert $i$'s weight as follows: $w_{t+1}(i) = w_t(i)(1 - \epsilon)$ if $i$ made a wrong prediction and $w_{t+1}(i) = w_t(i)$ if $i$ made a correct prediction

Derive a regret upper bound for this online learning algorithm, as a function of the parameter $\epsilon, T, n$. Is your regret bound sublinear in $T$ for any problem instance (note: in regret minimization, a sublinear regret in $T$ is typically what we pursue)?

4. (`Short, 5 points`) Show that there exists an online learning algorithm for the experts' advice problem which has sublinear regret in $T$ for any problem instance.