CSC 696H Homework 1

Chicheng Zhang

September 2023

- This homework is due on Sep 26 in class.
- Your solutions to these problems will be graded based on both correctness and clarity. Your arguments should be clear: there should be no room for interpretation about what you are writing. Otherwise, I will assume that they are wrong, and grade accordingly.
- If you feel hard to make progress on any of the questions, you can post your questions on Piazza. Try posing your questions to be as general as possible, so that it can promote discussion among the class.
- You are encouraged to discuss the homework questions with your classmates, but the discussions should only be at a high level, and you should write your solutions in your own words. For every question you have had discussions on, please mention explicitly whom you have discussed with; otherwise it may be counted as academic integrity violation.
- Feel free to use existing theorems from the course notes / the textbook.

Problem 1 (10pts)

- Describe one real-world decision making problem that can be modeled as contextual bandits but not multi-armed bandits, and explain why. What is this problem's context space, action space, and reward function?
- Describe one real-world decision making problem that can be modeled as reinforcement learning but not contextual bandits, and explain why. What is this problem's state space, action space, reward function, and transition probability?

Problem 2 (10pts)

Recall that in the class, we defined subgaussian random variables:

Definition 1. *A random variable X is said to be subgaussian with variance proxy b^2 (abbrev. b^2-subgaussian), if for any $\lambda \in \mathbb{R}$,*

$$\mathbb{E}\left[e^{\lambda(X-\mathbb{E}[X])}\right] \le e^{\frac{\lambda^2 b^2}{2}}.$$

Verify that:

- If X is v^2-subgaussian, then for any $a \in \mathbb{R}$, aX is a^2v^2-subgaussian.
- If X is v^2-subgaussian and Y is w^2-subgaussian, and X, Y are independent, then $X + Y$ is $(v^2 + w^2)$-subgaussian.

Based on the above, why do you think b^2 has the name "variance proxy of X"?

Problem 3 (10pts)

Verify the following two useful facts in the lectures on online learning:

- Suppose we are in the online square loss regression setting, where $\mathcal{Y} = \{-1, +1\}$, and $\hat{\mathcal{Y}} = [-1, +1]$, and $\ell(\hat{y}, y) = (\hat{y} - y)^2$. Verify that $\ell(\hat{y}, y)$ is $\frac{1}{8}$-exp-concave as a function of $\hat{y} \in \hat{\mathcal{Y}}$.

  You can use the following fact about concave functions:

  Fact 2. *Given a function $f : [a, b] \to \mathbb{R}$; f is concave if and only if for all $x \in [a, b]$, $f''(x) \leq 0$.*

- Suppose we are in the online density estimation with log loss setting, where $\mathcal{Y} = \{1, \ldots, K\}$, and $\hat{\mathcal{Y}} = \Delta(\mathcal{Y}) = \left\{(p_1, \ldots, p_K) : \forall i, p_i \geq 0, \sum_{i=1}^{K} p_i = 1\right\}$. Verify that $\ell(p, y) = \ln \frac{1}{p_y}$ is 1-exp-concave as a function of $p \in \hat{\mathcal{Y}}$.

Problem 4 (10pts)

Suppose loss function ℓ is the absolute loss and we are given a hypothesis class $\mathcal{F}$. Consider the following "online learning with iid data" protocol:

for $t = 1, 2, \ldots, T$: **do**
  Learner chooses a predictor $\hat{f}_t$.
  Example (x_t, y_t) is drawn from distribution D and is shown to the learner.
  Learner incurs loss $\ell(\hat{f}_t(x_t), y_t)$.
end for

Let $f^* = \text{argmin}_{f \in \mathcal{F}} L_D(f)$. Define the expected regret of the learner as:

$$\text{Reg}(T) = \mathbb{E}\left[\sum_{t=1}^{T} \ell(\hat{f}_t(x_t, y_t)) - \ell(f^*(x_t), y_t)\right].$$

Answer the following questions:

- Show that $\text{Reg}(T)$ is equal to $\mathbb{E}[\text{PReg}(T)]$, where $\text{PReg}(T) := \sum_{t=1}^{T} L_D(\hat{f}_t) - L_D(f^*)$ is called the pseudo-regret of the learner.
- Suppose that for every t, $\hat{f}_t$ is chosen by the Follow-the-Leader (FTL) algorithm, i.e. it is the ERM over $\mathcal{F}$ on the first $t - 1$ examples. Provide an upper bound on $\text{PReg}(T)$ that holds with probability $1 - \frac{1}{T}$, and use it to give an upper bound on $\text{Reg}(T)$. (You may find the inequality $\sum_{t=1}^{n} \frac{1}{\sqrt{t}} \leq 2\sqrt{n}$ useful.)
- Suppose instead we run the Exponential Weight Algorithm. What is a valid upper bound on $\text{Reg}(T)$, and why? Is this bound better or worse than that of FTL you got from solving the previous question?

Problem 5 (2pts)

- How much time did it take you to complete this homework?
- Do you have any suggestions for this course so far (e.g. homework length, lecture pace, etc)?