CSC 696H Homework 1

Chicheng Zhang

September 2024

- This homework is due on Oct 4 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 sequential 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 sequantial 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 $\lambda \in \mathbb{R}$,*

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

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. (Hint: somewhere in your proof you can use the independence property of X and Y.)

Based on the above, why do you think "variance proxy of X" is a good name for b^2?

Problem 3 (10pts)

Suppose we have a coin represented by a random variable X with unknown bias p. We would like to estimate p by flipping the coin and computing $\bar{X}_n$, the sample average of its iid outcomes $X_1, \ldots, X_n$. Additionally, suppose we know that $p \le 0.05$ (i.e., it is heavily biased towards tail).

- Prove that $\mathrm{Var}(X) \le 0.05(1 - 0.05)$.
- Suppose we would like to ensure that $\bar{X}_n$ is within 0.01 of p, with probability at least 0.999. How many samples n do we need, according to Chebyshev, Hoeffding, and Bernstein's inequalities, respectively? Which inequality suggests the smallest sample size?
- Suppose instead, we only get to flip the coin $n = 1000$ times. According to Chebyshev, Hoeffding, and Bernstein's inequalities respectively, how close do you expect $\bar{X}_n$ to be within p, with probability 0.99? Which inequality gives the best result?

Problem 4 (7pts)

In the lecture we saw a useful lemma:

Lemma 2. *Suppose events E_1, E_2 happen with probabilities at least $1 - \delta_1$ and $1 - \delta_2$, respectively. Then with probability $1 - \delta_1 - \delta_2$, E_1 and E_2 happen simultaneously.*

- How was this fact used in our lecture?
- Suppose F_1 and F_2 are independent and have probabilities 0.999 and 0.998, respectively. Calculate a lower bound of $P(F_1 \cap F_2)$ using Lemma 2, and compare it against the exact value of $P(F_1 \cap F_2)$. In this case, do you find the lower bound provided by the lemma to be good, and why?
- Suppose we have a collection of events $\{E_n\}_{n=1}^N$, with each $P(E_n) \ge 1 - \delta_n$ for $\delta_n > 0$, prove that:

$$P\left(\bigcap_{n=1}^{N} E_n\right) \ge 1 - \sum_{n=1}^{N} \delta_n.$$

Problem 5 (10pts)

Suppose loss function $\ell(\hat{y}, y)$ is the absolute loss ($\hat{y} \in [-1, 1], y \in \{-1, 1\}$) and we are given a finite predictor 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 $f^* = \mathrm{argmin}_{f \in \mathcal{F}} L_D(f)$. Define the expected regret of the learner as:

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

Answer the following questions:

- Show that $\mathrm{Reg}(T)$ is equal to $\mathbb{E}[\mathrm{PReg}(T)]$, where $\mathrm{PReg}(T) := \sum_{t=1}^{T} L_D(\hat{f}_t) - L_D(f^*)$ is called the pseudo-regret of the learner. (Hint: use the law of total expectation)

- 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. Show that, for every t, there exists an event E_t such that $P(E_t) \geq 1-\delta/T$, and when E_t happens,

$$L_D(\hat{f}_t) - L_D(f^*) \leq \min\left(2, 4\sqrt{\frac{\ln \frac{2|\mathcal{F}|T}{\delta}}{t-1}}\right)$$

- Show that, there exists some constant $c > 0$ and an event E, such that $P(E) \geq 1 - \delta$, and when E happens,

$$\mathrm{PReg}(T) \leq c\sqrt{T \ln \frac{2|\mathcal{F}|T}{\delta}}$$

  You may find the inequality $\sum_{t=1}^{n} \frac{1}{\sqrt{t}} \leq 2\sqrt{n}$ useful. (Hint: use results from Problem 4.)

- Provide a sublinear regret bound on $\mathrm{Reg}(T)$. (Hint: Try to decompose $\mathbb{E}[\mathrm{PReg}(T)]$ to $\mathbb{E}[\mathrm{PReg}(T)I(E)]+\mathbb{E}\left[\mathrm{PReg}(T)I(E^C)\right]$, and bound each part respectively. Finally we can tune the free parameter δ to optimize our bound.)

Problem 6 (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)?