CSC 588: Homework 4

Chicheng Zhang

May 4, 2021

- This homework is due on May 4, 11:59pm on gradescope. As stated in the syllabus, no late homeworks will be accepted.
- If you feel unable 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

In this problem we investigate the p-norm algorithm for online convex optimization and its relationship to statistical learning in greater detail. Consider conjugate exponents p, q such that $1 < p \leq 2 \leq q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Recall that for vector $z \in \mathbb{R}^d$, $\|z\|_c = (\sum_{i=1}^d |z_i|^c)^{\frac{1}{c}}$ denotes its ℓ_c-norm.

1. Show that for $a, b \geq 0$, $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$. For any a, find a corresponding b such that this inequality is tight.
2. Show that for vectors $x, y \in \mathbb{R}^d$, $\langle x, y \rangle \leq \frac{\|x\|_p^p}{p} + \frac{\|y\|_q^q}{q}$. For any x, find a corresponding y such that this inequality is tight.
3. Show that for vectors $x, y \in \mathbb{R}^d$, $\langle x, y \rangle \leq \|x\|_p \|y\|_q$. For any x, find a corresponding y such that this inequality is tight.
4. Define $\psi(w) = \frac{1}{2}\|w\|_p^2$. Show that its Fenchel conjugate $\psi^\star(\theta) = \max_{w \in \mathbb{R}^d} \langle w, \theta \rangle - \psi(w)$ is $\psi^\star(\theta) = \frac{1}{2}\|\theta\|_q^2$.
5. Consider a set of examples $S = \{x_t\}_{t=1}^T \subset \left\{x \in \mathbb{R}^d : \|x\|_q \leq R_q\right\}$. Define $\mathcal{F} = \left\{m_w(x) := \langle w, x \rangle : \|w\|_p \leq B_p\right\}$. Prove that $\mathrm{Rad}_S(\mathcal{F}) \leq B_p R_q \sqrt{\frac{1}{(p-1)T}}$. (Hint: given iid Rademacher random variables $\{\sigma_t\}_{t=1}^T$, consider running the p-norm algorithm over linear functions $\left\{f_t(w) = \langle w, -\sigma_t x_t \rangle\right\}_{t=1}^T$; write down its regret guarantee, and take expectation on both sides.)

Problem 2

Suppose we have a sequence of regression examples $(x_t, y_t)_{t=1}^T$, where for all t, $\|x_t\|_2 \leq R$ and $|y_t| \leq Y$. Define constraint set $\Omega = \left\{w : \|w\|_2 \leq B\right\}$. For every t, define loss function $f_t(w) = \frac{1}{2}(\langle w, x_t \rangle - y_t)^2$.

1. Show that for every t, f_t is $\frac{1}{(BR+Y)^2}$-exp-concave over Ω.

2. Show that for every t, f_t is $R(BR+Y)$-Lipschitz with respect to $\|\cdot\|_2$ over Ω.

3. Show that Online Newton Step, with appropriate tuning of its parameters $\tilde{\alpha}$, λ, has a regret of $O((BR+Y)^2 \cdot d \ln T)$ with respect to $\{f_t\}_{t=1}^T$ against Ω. What are your choices of $\tilde{\alpha}$, λ?

Problem 3

Consider a K-armed MAB environment for $K = 5$, where the losses of arm a, $\left\{\ell_t(a)\right\}_{t=1}^T$ are drawn iid from Bernoulli($\ell(a)$). Let the expected loss vector $\ell = (\ell(1), \ldots, \ell(5)) = (0.7, 0.6, 0.5, 0.4, 0.3)$.

1. Implement the UCB(α) algorithm, where at time step t, the lower confidence bound for arm a is defined as $\mathrm{LCB}_t(a) = \bar{\ell}_{t-1}(a) - \alpha\sqrt{\frac{\ln T}{m_{t-1}(a)+1}}$. (The UCB algorithm we discussed in class is UCB($\alpha = 2$).) Let $A = \{0.01, 0.03, 0.1, 0.3, 1, 3, 10\}$.

   For each $\alpha \in A$:

   (a) Run UCB(α) on the above-defined MAB environment for a time horizon of $T = 1000$, and define its *regret learning curve* as: the pseudo-regret $R_t = \sum_{s=1}^t \Delta(a_s)$ as a function of round number t;

   (b) Repeat the above simulation 2000 times and plot the averaged learning curve of the 2000 obtained learning curves.

   Plot the 7 averaged learning curves on one graph with legends indicating the values of α used. Out of the 7 choices of α's, which one achieves the best average pseudo-regret at $T = 1000$?

2. Given an MAB algorithm and an environment, define its *optimal action learning curve* as: the fraction of pulls to the optimal arm $\frac{1}{t}\sum_{s=1}^t \mathbf{1}\left[a_s = a^\star\right]$ as a function of t. Repeat the experiments in item 1 and plot the 7 optimal action learning curves. Out of the 7 choices of α's, which one plays the highest (average) portion of the optimal arm at $T = 1000$?

3. Implement the ϵ-greedy algorithm. Repeat the above two experiments for the ϵ-greedy algorithm for each $\epsilon \in E = \{0, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3\}$ and plot the respective average learning curves.

   Out of the 7 choices of ϵ's, which one achieves the best average pseudo-regret / plays the highest (average) portion of the optimal arm at $T = 1000$?

Problem 4

How much time did it take you to complete this homework?