AI3001: Advanced Topics in Machine Learning
Homework 1
October 3, 2022

Instructions:

- *The maximum number of points is 30. Each point is worth 0.5 marks.*
- *The last date of submission is end of day **Saturday, 15th Oct 2022**. You can either drop your handwritten homework in my office (C block, room:112/D) by sliding it under the door or hand it over after the class.*
- *If you are writing homework in latex/MS word, send the pdf document by email. Please do not scan and send handwritten pages.*
- *There will not be any extensions for homework submission.*
- *The Thompson Sampling algorithm will be covered in the class on Oct 10th.*

Problem 1. Let $R(x) = \sum_{i=1}^{n} x_i \log(x_i)$ for $x \in \mathbb{R}_+^n$ (with $0\log(0) = 0$). Show that the point $x^* \in \Delta_n$ with $x_i^* = \frac{y_i}{\sum_{i=1}^{n} y_i}$ is the Bregman projection of point $y \in \mathbb{R}_+^n$ i.e. $x^* = \arg\min_{x\in\Delta_n} B_R(x||y)$. [4 points]

Hint: Use Langrangean technique.

Problem 2. Show that executing FTRL on simplex Δ_n with scaled entropic regularizer $R_\eta(x) = \frac{1}{\eta}\sum_{i=1}^{n} x_i \log(x_i)$ and linear loss functions given by $(w_t)_{t\geq 1}$ is equivalent to running exponential weights algorithm with n experts with loss vectors $(w_t)_{t\geq 1}$ and parameter η. [3 points]

Hint: Use the result of Problem 1.

Problem 3. Let $f(x) = e^{-\eta\ell(x,y)}$ and consider the loss function $\ell(x, y) = ||x-y||_2^2$ over the closed ball of radius $r > 0$ and centered at origin i.e. $\mathcal{B}_r(0) \subseteq \mathbb{R}^d$. Find the range of values of η for which f is concave. [4 points]

Hint: First prove for $d = 1$ and then generalize for larger values of d.

Problem 4. For a stochastic MAB setting with K arms, show that $\mathcal{R}_T(\text{ALG}) = \sum_{i=1}^{K} \Delta_i \cdot \mathbb{E}[N_{i,T}]$. Recall from the class that $N_{i,T}$ is the number of time instances arm i is pulled till T, $\Delta_i = \mu_{i^*} - \mu_i$ and $\mathcal{R}(.)$ is pseudo-regret with respect to best arm i^*. [4 points]

Problem 5. Consider the stochastic multi-armed bandit problem with K arms with Bernoulli reward distributions. For $K \in \{10, 20, 50, 100\}$, set the arms' mean rewards as $\mu_i = 1 - \frac{i}{K+1}$ for arm $i = 1, 2, \cdots, K$. Assume that the time horizon T is known to each algorithm beforehand. Run the following algorithms.

a. EXP3-γ with $\gamma = \sqrt{\frac{K\log(K)}{2T}}$

b. Exploration-separated algorithm with $\epsilon = T^{-1/3}$

c. UCB1

d. Thompson Sampling algorithm with prior $Beta(1, 1)$

Send your code and plot (preferably Python, MATLAB, C, C++) via an email and your observations on paper.

1. Plot (in a single plot) the average cumulative regret for time horizons $T \in \{10^3, 10^4, 5 \times 10^4, 10^5, 5 \times 10^5, 10^6, 5 \times 10^6\}$ averaged over 10^5 runs for each algorithm. [10 points]

2. Write your observations. Do the theoretical results match with simulations? What happens for larger values of T and smaller values of Δ_i's? Why do you think some of these algorithms perform better than others? [5 points]