Assignment-2

Topics in ML (CSE975)
Total Marks : 30
Submission Deadline : 03/10/2017 11:55 PM

September 22, 2017

General Instructions

1. Assignment can be implemented in Matlab/Octave, Python, C/C++, R.
2. Ensure that submitted assignment is your original work. Please do not copy any part from any source including your friends, seniors and/or the internet. If any such attempt is caught then serious actions including an F grade in the course is possible.
3. A single zip file needs to be uploaded to the moodle course portal. The file should contain (1)pdf report file containing the experiment findings (2)the code you have written. Include the assignment number, your name and roll number at the top of the first page of the pdf report.
4. Your grade will depend on the correctness of answers and output. In addition, due consideration will be given to the clarity and details of your answers and the legibility and structure of your code.
5. Assignments submitted through moodle.iiit.ac.in only will be considered for evaluation. No assignment submitted through the mail will be considered.
6. There will not be any deadline extension given.
7. Please make sure that you submit the assignment well ahead of the time to avoid last minute power/internet issues.

Problem 1. [10 Points] **EXP3 MAB:** Consider the problem of multi-armed bandit with $n = 10$ arms. Assume that the losses assigned to each expert are generated according to independent Bernoulii distributions. The adversary generates loss for experts 1 to 8 according to Bernoulii(0.5) in each round. For the 9th expert, loss is generated according to Bernoulii$(.5 - \Delta)$ in each round. The losses for the 10th expert are generated according to different Bernoulii random variable in each round: for the first $T/2$ rounds, they are generated according to Bernoulii$(0.5 + \Delta)$ and the remaining $T/2$ rounds they are generated according to Bernoulii$(0.5 - 2\Delta)$. $\Delta = 0.1$.

Run EXP3 algorithm with $\eta = \sqrt{\frac{2 \ln n}{nT}}$. Take different values of T as $T = 1000, 10000$ and 100000. For each value of T, do the following.

- Generate 200 different runs (sample paths) by varying the random seed.
- Create a plot with the time horizon on the x axis and pseudo regret on the y axis. Each point must be the average over all runs for the corresponding horizon.
- Create a plot where x axis is the arm index and y axis is the number of times that arm has been pulled.

Problem 2. [10 Points] **ϵ-Greedy MAB:** Consider the problem of multi-armed bandit with $n = 10$ arms. Let i_t be the arm pulled at time t and $x_{i_t}^t$ be the corresponding reward. Reward for ith arm is distributed as $x_i^t \sim \mathcal{N}(\mu_i, 1)$, where μ_i is actual mean reward of ith arm. Let $\mu_i = 5+i,\ i = 1 \ldots 10$. Let $T = 100000$. Run ϵ-greedy algorithm for $\epsilon = 0, 0.01, 0.1$. For each value of ϵ, do the following.

- Generate 200 different runs (sample paths) by varying the random seed.
- Create a plot with the time horizon on the x axis and pseudo regret on the y axis. Each point must be the average over all runs for the corresponding horizon and ϵ value.
- Create a plot with the $\log(t)$ on the x axis and pseudo regret on the y axis. Each point must be the average over all runs for the corresponding horizon and ϵ value.
- Create a plot where x axis is the arm index and y axis is the number of times that arm has been pulled. Create this plot for different values of ϵ.

Problem 3. [10 Points] **Upper Confidence Bound MAB:** Consider the problem of multi-armed bandit with $n = 10$ arms. Let i_t be the arm pulled at time t and $x_{i_t}^t$ be the corresponding reward. Reward for ith arm is distributed as $x_i^t \sim \mathcal{N}(\mu_i, 1)$, where μ_i is actual mean reward of ith arm. Let $\mu_i = 5 + i,\ i = 1 \ldots 10$. Let $T = 100000$. Run UCB algorithm. Do the following.

- Generate 200 different runs (sample paths) by varying the random seed.
- Create a plot with the time horizon on the x axis and pseudo regret on the y axis. Each point must be the average over all runs for the corresponding horizon.
- Create a plot with the $\log(t)$ on the x axis and pseudo regret on the y axis. Each point must be the average over all runs for the corresponding horizon.
- Create a plot where x axis is the arm index and y axis is the number of times that arm has been pulled.