# CSC 696H Homework 2

Chicheng Zhang

October 2023

- This homework is due on Oct 20, 11:59pm.
- 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 (16pts)

**Linear bandits and elliptical potential.** Suppose we are in the stochastic linear bandit setting. The sequence of (context, action)'s are represented by feature vectors $\{\phi_t\}_{t=1}^T$ such that

$$\phi_t = \begin{cases} (1,2) & t \text{ is odd} \\ (3,1) & t \text{ is even} \end{cases}$$

and we see rewards $\{r_t\}_{t=1}^T$ such that all $r_t = 0$.

- Recall the definition of $V_t(\lambda) = \sum_{s=1}^t \phi_s \phi_s^\top + \lambda I$ is the (scaled) regularized data covariance matrix up to round $t$; Compute $V_6(1), \hat{\theta}_7(1)$ and $V_{12}(1), \hat{\theta}_{13}(1)$. (You may find it useful to use matrix computation tools such as numpy or matlab to calculate these.)
- Define the confidence set at time step $t$ for $\theta^*$, $\Theta_t := \left\{\theta : \|\theta - \hat{\theta}^t(1)\|_{V_{t-1}(1)} \leq 2\right\}$. (Note that the norm bound here is changed from $\beta_t(1)$ in the original lecture to 2 for simplicity.) In one plot, draw the graphs of $\Theta_7$ and $\Theta_{13}$ and label them. You can use any plotting software you like; I recommend `https://www.desmos.com/calculator`.
- For $x = (1,1)$ and $y = (1,-1)$, calculate $[\min_{\theta\in\Theta_{13}} \langle\theta, x\rangle, \max_{\theta\in\Theta_{13}} \langle\theta, x\rangle]$ and $[\min_{\theta\in\Theta_{13}} \langle\theta, y\rangle, \max_{\theta\in\Theta_{13}} \langle\theta, y\rangle]$ respectively. These are confidence intervals for $\langle\theta^*, x\rangle$, $\langle\theta^*, y\rangle$ respectively. Which confidence interval has a higher uncertainty? Does this match your intuition?

  (Hint: we have an alternative expression for $\max_{\theta\in\Theta_{13}} \langle\theta, x\rangle$ in class that is simpler to calculate.)
- calculate the cumulative elliptical potential $\sum_{s=1}^t \|\phi_s\|^2_{V_s(1)^{-1}}$ for $t = 1, \ldots, 100$ and plot it as a function of $t$. Is your result consistent with what the elliptical potential lemma predicts? Why?

## Problem 2 (8pts)

In this exercise, we aim to show the useful fact that given any symmetric positive definite matrix $M \succ 0$ and vector $\phi$, $\max_{x:\|x\|_M \le \beta} \langle x, \phi \rangle = \beta \|\phi\|_{M^{-1}}$.

- Where is this fact used in our lectures?

We now set up some notations. Let $M = U \begin{bmatrix} \lambda_1 & & \\ & \dots & \\ & & \lambda_d \end{bmatrix} U^\top$ be $M$'s eigendecomposition, where $UU^\top = I_d$. Define $M^{\frac{1}{2}} = U \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \dots & \\ & & \sqrt{\lambda_d} \end{bmatrix} U^\top$ and $M^{-\frac{1}{2}} = U \begin{bmatrix} \sqrt{\lambda_1^{-1}} & & \\ & \dots & \\ & & \sqrt{\lambda_d^{-1}} \end{bmatrix} U^\top$.

- Verify that for any vector $x$, $\|x\|_M = \|M^{\frac{1}{2}} x\|$.
- Using the fact that $\langle x, \phi \rangle = \left\langle M^{\frac{1}{2}} x, M^{-\frac{1}{2}} \right\rangle$ and Cauchy-Schwarz inequality, prove that $\max_{x:\|x\|_M \le \beta} \langle x, \phi \rangle \le \beta \|\phi\|_{M^{-1}}$.
- Prove that $\max_{x:\|x\|_M \le \beta} \langle x, \phi \rangle \ge \beta \|\phi\|_{M^{-1}}$. (Hint: consider $x^* = \beta \frac{M^{-1}\phi}{\|\phi\|_{M^{-1}}}$)

## Problem 4 (8pts)

Answer the following questions related to nonlinear bandits:

- Suppose $f_1, f_2$ and $g$ and a dataset $S$ are such that $\|f_1 - g\|_S^2 \le \beta$ and $\|f_1 - g\|_S^2 \le \beta$. Prove that $\|f_1 - f_2\|_S^2 \le 4\beta$. Where was this result used?
- Suppose we try to estimate some $f^*$ in function class $\mathcal{F}$ by making queries to it, by repeating the following process with $\alpha = 0$ until we cannot do so:

  (In the "politician" analogue, think of this as a way for the reporter to ask questions to the politician to nail down their political position $f^*$.)

> Choose $u_1$, query $f^*(u_1)$
> Choose $u_2$ $\alpha$-independent of $u_1$, query $f^*(u_2)$.
> ...
> Choose $u_N$ $\alpha$-independent of $u_1, \dots, u_{N-1}$, query $f^*(u_N)$.

  - Prove that $N \le \mathrm{Edim}(\mathcal{F}, 0)$.
  - Define $\mathcal{F}_N = \{f \in \mathcal{F} : f(u_1) = f^*(u_1), \dots, f(u_N) = f^*(u_N)\}$. What functions are there in $\mathcal{F}_N$? Why?

## Problem 3 (10pts)

For multi-armed bandits problem with $A$ actions, define event $E$ as

$$E = \left\{ \text{for all time steps } t = 1, \dots, T, \left| f^*(a) - \hat{f}_t(a) \right| \le b_t(a) \right\}.$$

Suppose $E$ happens. Instead of UCB, consider the following alternative action selection rule:

At timestep $t$:

- For every action $a$, define $\mathrm{UCB}_t(a) = \hat{f}_t(a) + b_t(a)$, $\mathrm{LCB}_t(a) = \hat{f}_t(a) - b_t(a)$.
- Define $\mathcal{A}_t = \{a \in \mathcal{A} : \mathrm{UCB}_t(a) \geq \max_{b \in \mathcal{A}} \mathrm{LCB}_t(b)\}$.
- Choose $a_t \in \mathrm{argmax}_{a \in \mathcal{A}_t} w_t(a)$.

Prove that:

- For every $t$, the optimal action, $a^* = \mathrm{argmax}_{a \in \mathcal{A}} f^*(a)$, is in $\mathcal{A}_t$.
- The pseudoregret of the learner $\mathrm{PReg}(T) \leq 4 \sum_{t=1}^{T} b(a_t)$.
- If $b(a_t)$ are defined according to our MAB lecture, what is the order of the above bound in terms of $A$ and $T$? Why?

Hints:

- Try to draw a picture to visualize e.g. $\hat{f}_t(a)$, $f^*(a)$, $\mathrm{UCB}_t(a)$, $\mathrm{LCB}_t(a)$, for say, a 3 arm bandit setting.
- You may find the following observation useful: $\mathrm{UCB}_t(a) \leq f^*(a) + 2b_t(a)$. (Can you see why?)

# Problem 5 (2pts)

- How much time did it take you to complete this homework?
- What paper are you planning to present?

# Bonus question

If you made mistakes in HW1, submit your updated solutions here; for each subproblem, your final score will be the maximum of: (1) your initial score and (2) 0.75× your score at the second trial. (In other words, for any subproblem, if you already get more than 75% score on it, your score will not increase by making a second trial. I am always happy to provide feedback on your updated solutions though.)