COMPSCI 614: Problem Set 5

Due: Tuesday, 5/13 by 11:59pm in Gradescope.

Note: This problem set is OPTIONAL. If you complete it, it can be used to replace your lowest score on the first four problem sets.

Instructions:

- You are allowed to, and highly encouraged to, work on this problem set in a group of up to three members.
- Each group should **submit a single solution set**: one member should upload a pdf to Gradescope, marking the other members as part of their group in Gradescope.
- You may talk to members of other groups at a high level about the problems but **not work through the solutions in detail together**.
- You must show your work/derive any answers as part of the solutions to receive full credit.

1. ℓ_1 Subspace Embedding via Sampling (6 points)

Given a matrix $A \in \mathbb{R}^{n\times d}$ we would like to find a sampling matrix $S \in \mathbb{R}^{m\times n}$ such that, with probability at least $1-\delta$, for all $x \in \mathbb{R}^d$, $(1-\epsilon)\|Ax\|_1 \leq \|SAx\|_1 \leq (1+\epsilon)\|Ax\|_1$, where $\|y\|_1 = \sum_{i=1}^n |y(i)|$ is the ℓ_1 norm.

To do so, we define the ℓ_1 *sensitivity* of row i as $\sigma_i = \max_{x\in\mathbb{R}^d} \frac{|[Ax](i)|}{\|Ax\|_1}$ and let $p_i = \frac{\sigma_i}{\sum_{j=1}^n \sigma_j}$. We pick each row of S independently, letting $S_{j,:} = \frac{1}{m\cdot p_i} e_i$, with probability p_i, where e_i is the i^{th} standard basis vector.

Let $T = \sum_{i=1}^n \sigma_i$. It can be shown that $T \leq d$ – this is maybe not surprising given the analogy to the ℓ_2 leverage scores, which sum to exactly d, but it is non-trivial to prove. You may use it as a fact going forward.

1. (4 points) Prove that for any fixed $x \in \mathbb{R}^d$, if $m = O\left(\frac{d\log(1/\delta)}{\epsilon^2}\right)$, then with probability at least $1-\delta$, $(1-\epsilon)\|Ax\|_1 \leq \|SAx\|_1 \leq (1+\epsilon)\|Ax\|_1$. **Hint:** Assume without loss of generality that $\|Ax\|_1 = 1$ and apply a Bernstein inequality. You'll want to target bounding the variance and maximum magnitude terms by T/m.
2. (2 points) Prove that for $m = O\left(\frac{d^2\log(1/\epsilon)+d\log(1/\delta)}{\epsilon^2}\right)$, with probability $1-\delta$, S satisfies: for all $x \in \mathbb{R}^d$, $(1-\epsilon)\|Ax\|_1 \leq \|SAx\|_1 \leq (1+\epsilon)\|Ax\|_1$. **Hint:** Follow the ϵ-net approach used in class for the ℓ_2 subspace embedding proof. You may use that the same net size bound of $(4/\epsilon)^d$ holds for the ℓ_1 norm.

2. Randomized Triangle Coloring (6 points)

A graph is k-colorable if there is an assignment of each node to one of k colors such that no two nodes with the same color are connected by an edge.

1. (2 points) Show that if a graph is 3-colorable then there is a coloring of the graph using 2 colors such that no triangle in monochromatic. I.e., for any three nodes u, v, w such that $(u, v), (v, w)$, and (u, w) are all edges, we do not have u, v, w all assigned to the same color (but we may have u, v assigned to the same color when (u, v) is an edge).

2. (4 points) Consider the following algorithm for coloring a 3-colorable graph with 2 colors so that no triangle is monochromatic. Start with an arbitrary 2-coloring (some triangles may be monochromatic, so it's not necessarily a valid coloring). While there are any monochromatic triangles, pick one arbitrarily and change the color of a randomly chosen vertex in that triangle. Give an upper bound on the expected number of steps of this process before a valid 2-coloring with all non-monochromatic triangles is found.

   Hint: Shoot for a polynomial, not an exponential number of steps here. Use the fact that part (1) actually implies the existence of *many* 2-colorings with non-monochromatic triangles.

3. Move to Top Shuffling (8 points)

Consider shuffling a deck of n unique cards by randomly picking a card and moving it to the top of the deck. Observe that with probability $1/n$, the top card is picked and so the order does not change from one step to the next.

1. (2 points) Prove that this Markov chain is irreducible and aperiodic.

2. (2 points) Prove that the chain converges to the the uniform distribution over all $n!$ possible permutations of the cards.

3. (2 points) In class, we argued that after $t = n \log(n/\epsilon)$ steps, the distribution of states q^t in this Markov chain satisfies $\|q^t - \pi\|_{TV} \leq \epsilon$. Say you are a casino, and you offer a game of pure chance where the customer must wager \$1. The game uses the shuffled deck of cards to determine a pay out somewhere between \$0 and \$1000. You have calculated that, when the deck is ordered according to a uniform random permutation (i.e., according to π), your expected winnings per game are \$0.1. How small must you set ϵ to ensure that your expected winnings are at least \$.09?

4. (2 points) Argue that our mixing time bound is essentially tight. In particular, show that if we run the Markov chain for $t \leq cn \log n$ steps for small enough constant c, then $\|q^t - \pi\|_{TV} \geq 99/100$. I.e., we are very far from a uniformly random permutation.

   Hint: Start by arguing that if $t \leq cn \log n$ for small enough c, with high probability there are $\sqrt{n}$ cards which are never swapped in the shuffle. Use the coupon collector analysis from Lecture 2. Then consider the probability that we have $\sqrt{n}$ consecutive cards in order after a uniform random shuffle, vs. after this shuffle starting from an ordered deck. Use the different in probabilities to lower bound the TV distance.