COL866: Special Topics in Algorithms
Concentration inequalities and their Applications in Computer Science
II semester, 2022-23.
Minor 4 exam.
22 April 2023, Maximum Marks: 25.

Instructions: Please handwrite your solutions clearly and concisely. Try to fit the solution for each problem in one side on an A4 sized page, definitely no more than 2. Latexed solutions are highly encouraged. Upload the paper on gradescope by the deadline (no extensions will be given under any circumstances). Make sure you mark the pages related to each problem on gradescope.

Problem 1 ($6 + 5 + 6 = 17$ **marks)**
A *signed graph*, $\Gamma = (G, \sigma)$ is a simple undirected graph $G = (V, E)$ and mapping $\sigma : E \to \{-1, +1\}$, called its *edge labelling*, that assigns a sign to each edge. A signed graph is said to be *balanced* if there is a partition (V_1, V_2) of V such that for every $(u, v) \in E$ with $\sigma(u, v) = -1$, $u \in V_1$ iff $v \in V_2$. The Laplacian L of a signed graph $\Gamma = (G = (V, E), \sigma)$, is a symmetric $|V| \times |V|$ matrix with $L_{ii} = d_i$, and $L_{ij} = -\sigma(i, j)$ if $(i, j) \in E$ and 0 otherwise for $i \neq j$.

Given a $p \in (0, 1)$ we define a random signed graph $\Gamma_{n,p}$ by taking n vertices and including each possible undirected edge with probability p independent of all other edges. Further, the sign of each edge that is included is uniformly chosen from $\{-1, 1\}$ independent of all other random choices. In this problem we will study the conditions on p under which $\Gamma_{n,p}$ is or is not balanced.

Problem 1.1 (6 **marks)**
Given any set E' of edges of Γ, the sign of E' is $\prod_{e \in E'} \sigma(e)$. It is known that a signed graph is balanced iff the sign of each cycle of the graph is positive. It is an interesting graph theoretic exercise to prove this but for this exam you can take it as given. Use this fact and the first moment method to find $f(n)$ such that $\Pr\{\Gamma_{n,p} \text{ is balanced}\}$ tends to 1 as $n \to \infty$ whenever $p \leq f(n)$ (asymptotic bound is also okay, but try for the best possible result).

Problem 1.2 (5 **marks)**
Prove that L for any signed graph is positive semidefinite. Also show that if Γ is balanced then L has eigenvalue 0.

Problem 1.3 (6 **marks)**
Use the result of Problem 1.2 and Matrix Chernoff bounds to show a lower bound on $g(n)$ such that $\Pr\{\Gamma_{n,p} \text{ is balanced}\}$ tends to 0 as $n \to \infty$ whenever $p \geq f(n)$ (asymptotic bound is also okay, but try for the best possible result).

Problem 2 (8 **marks)**
Suppose $\boldsymbol{X}$ is a random vector chosen from $\mathbb{R}^d$ with the property that $\|\boldsymbol{X}\| = 1$ and $\mathrm{E}[\boldsymbol{X}(i)] = 0$ for each $i \in [d]$. Clearly, $\boldsymbol{X}(i)$ and $\boldsymbol{X}(j)$ may be correlated whenever $i \neq j$. Let Σ be their covariance matrix. We will approximate Σ by considering n i.i.d. copies of $\boldsymbol{X}$, $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ and computing

$$\widehat{\Sigma}_n = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_i \boldsymbol{X}_i^\top.$$

First argue that $\widehat{\Sigma}_n$ is an unbiased estimator of Σ (i.e. $\mathrm{E}\left[\widehat{\Sigma}_n\right] = \Sigma$) and then use the Matrix Bernstein inequality to find a value n_ε such that,

$$\frac{\|\Sigma - \widehat{\Sigma}_n\|}{\|\Sigma\|} \leq \varepsilon$$

with high probability for all $n \geq n_\varepsilon$.