401-4944-20L Mathematics of Data Science: Extra Problem Set 1

(Spring 2020)

Afonso S. Bandeira
`bandeira@math.ethz.ch`
`https://people.math.ethz.ch/~abandeira/`

July 20, 2020

This problem set is optional and it won't be graded. If you want to discuss a solution (to make sure it is correct) or want to ask questions about a problem stop by office hours or write on Piazza or a TA or myself an email and we can schedule a time to talk. Date is of last update (e.g. correction of typos)

If you need hints, feel free to write on Piazza or email or the TAs.

Problem 101.1 *Let G be a simple graph (no self-loops of multi-edges) that is connected and d-regular (all nodes have degree exactly d). Let A denote the adjacency matrix of G.*

Suppose $-d$ is an eigenvalue of A:

- *Prove that for every eigenvalue λ of A, $-\lambda$ is also an eigenvalue of A.*
- *A vertex coloring of G with k colors is a choice of one of k colors for each vertex of G such that no two neighboring vertices share an edge. $\chi(G)$ is the minimim number of colors with which it is possible to color all vertices of G this way. What can you say about $\chi(G)$?*

Connectivity of the Erdős-Rényi random graph

The Erdős-Rényi random graph $G(n,p)$ is a graph with n nodes, where each edge (i,j) appears (independently) with probability p. In this problem set, you will show a remarkable phase transition: if $\lambda < 1$, then $G(n, \frac{\lambda \ln n}{n})$ has, with high probability, isolated nodes while, if $\lambda > 1$, the graph is connected (with high probability).

Problem 101.2 *Let I_i be a random variable indicating whether node i is isolated: $I_i = 1$ if node i is isolated, and $I_i = 0$ otherwise. Let $X = \sum_{i=1}^n I_i$ be the number of isolated nodes.*

The goal is to show that $\Pr\{X = 0\}$ is small when $\lambda < 1$ (meaning that there are isolated notes, with high probability). In the proof you can use the approximation

$$(1 - \lambda/n)^n \approx e^{-\lambda} \quad \textit{(for large } n\textit{)}$$

1. *Show that* $\mathbb{E}[X] \approx n^{-\lambda+1}$*.* ***Note:*** *The fact that* $\mathbb{E}[X] \to \infty$ *is not sufficient to show* $\Pr\{X = 0\} \to 0$ *(**why? Can you give a counter-example?**). We need to ensure that* X *concentrates around its mean.*

2. *Use (a simple) concentration inequality derived in class to finish the proof. (The techinque you have just derived is known as the second moment method)*

Problem 101.3 *Prove that, if* $\lambda \geq 1$*,* $G(n, \frac{\lambda \ln n}{n})$ *is connected with high probability:*

1. *Derive the probability for a set of* k *nodes* ($k \leq n/2$) *being disconnected from the rest of the graph.*

2. *Prove the probability of graph* G *having a disconnected component goes to zero as n grows (hint: use union bound).*

Problem 101.4 (Problem 5.1. of "A Mathematical Introduction to Compressive Sensing") *The mutual coherence between two orthonormal bases* $U = (u_1, \ldots, u_n)$ *and* $V = (v_1, \ldots, v_n)$ *of* $\mathbb{C}^n$ *is defined as*

$$\mu(U, V) := \max_{i,j} |\langle u_i, v_j \rangle|.$$

Prove the following inequalities, and show they are sharp

$$\frac{1}{\sqrt{n}} \leq \mu(U, V) \leq 1.$$

Problem 101.5 ($*$**)** *The mutual coherence between two orthonormal bases* $U = (u_1, \ldots, u_n)$ *and* $V = (v_1, \ldots, v_n)$ *of* $\mathbb{C}^n$ *is defined as*

$$\mu(U, V) := \max_{i,j} |\langle u_i, v_j \rangle|.$$

A set of orthonormal bases $U_1, \ldots, U_k$ *of* $\mathbb{C}^n$ *are called mutually unbiased if for all* $a, b \in [k]$

$$\frac{1}{\sqrt{n}} = \mu(U_a, U_b).$$

- *Let* M_n *be the maximum number of mutually unbiased bases in* $\mathbb{C}^n$*. Give a lower and upper bound on* M_n*. (The exact number* M_n *is currently unknown)*