MATHEMATICAL TOOLS - PROBLEM SET 9

Due Sunday, January 15th, 23:55, either in the course mailbox or through the Moodle.

General guidelines:

- All graphs mentioned are finite.
- **Do not** use the Perron-Frobenius theorem in this problem set. It provides a "solution by magic" to some questions, but the point is to deepen understanding of the underlying concepts.

Spectral Properties of Graph Adjacency Matrices.

Problem 1. Let $G = ([n], E)$ be a d-regular graph. Let A be G's adjacency matrix.

(1) Prove that A is orthogonally diagonalizable.
(2) Let $\lambda_1 \geq \ldots \geq \lambda_n$ be A's eigenvalues (with multiplicities).
   (a) Show that $\lambda_1 = d$.
   (b) Show that for all $1 \leq i \leq n$, $|\lambda_i| \leq d$.
   (c) Let k be the number of connected components in G. Show that $\lambda_1 = \ldots = \lambda_k = d$, and $\lambda_{k+1} < d$.
   (d) Show that if G is bipartite, then for all $1 \leq i \leq n$, $\lambda_i = -\lambda_{n+1-i}$.
   (e) Show that G has a bipartite connected component iff $\lambda_n = -d$.

Hint/general thoughts:

- We can think of $\mathbb{R}^n$ as the space of functions from $V = [n]$ to $\mathbb{R}$. Then A is a linear operator where for each $i \in [n]$, $(Af)(i) = \sum_{j:\{i,j\}\in E} f(j)$. That is, $Af : V \to \mathbb{R}$ is the function that gives each vertex the sum of f over its neighbors.
- If f is an eigenvector of A , it's useful to consider $i \in [n]$ that maximizes $|f(i)|$.

Total Variation Distance.

Definition. Let $p = (p_1, \ldots, p_n)$ and $q = (q_1, \ldots, q_n)$ be probability distributions on $[n]$. The **total variation distance** between p and q is $\delta(p,q) = \frac{1}{2}\|p-q\|_1 = \frac{1}{2}\sum_{i=1}^{n}|p_i - q_i|$.

Problem 2. Let $p = (p_1, \ldots, p_n)$ and $q = (q_1, \ldots, q_n)$ be probability distributions on $[n]$. For $A \subseteq [n]$, let $P(A) = \sum_{i\in A} p_i, Q(A) = \sum_{i\in A} q_i$. Show that

$$\delta(p,q) = \max_{A\subseteq[n]} |P(A) - Q(A)|$$

Random Walks on Graphs.

Problem 3. Let $G = (V, E)$ be a connected non-bipartite graph with at least two vertices. What is the stationary distribution of the random walk on G? Note that G isn't necessarily regular.

Definition. For a graph $G = (V, E)$ with no isolated vertices, the **lazy random walk** on G is a Markov chain $X_0, X_1, \ldots$ with each X_t taking values in V. The transition probabilities are:

$$\forall t \in \mathbb{N}, \forall u, v \in V, \mathbb{P}\left[X_t = v | X_{t-1} = u\right] = \begin{cases} \frac{1}{2} & u = v \\ \frac{1}{2\deg(u)} & \{u, v\} \in E \\ 0 & otherwise \end{cases}$$

In other words, with probability $\frac{1}{2}$ the walk stays in place, and with probability $\frac{1}{2}$ it moves to a uniformly random neighbor. The lazy random walk is useful in avoiding issues of periodicity in bipartite graphs.

Problem 4. Let $G = ([n], E)$ be a d-regular connected graph ($d \geq 1$). Let $x^0 = \left(x^0_1, \ldots, x^0_n\right)$ be some distribution on $[n]$, and let $X_0, X_1, \ldots$ be the lazy random walk on G where $X_0 \sim x^0$. We'll prove that for all $k \in [n]$:

$$\lim_{t\to\infty} \mathbb{P}\left[X_t = k\right] = \frac{1}{n}$$

Let A be the transition matrix for the Markov chain. Explicitly:

$$A_{i,j} = \begin{cases} \frac{1}{2} & i = j \\ \frac{1}{2d} & \{i, j\} \in E \\ 0 & otherwise \end{cases}$$

Show that $\lim_{t\to\infty} \left\| x^0 A^t - \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)\right\|_2 = 0$, and explain why this implies that $\lim_{t\to\infty} \mathbb{P}\left[X_t = k\right] = \frac{1}{n}$ and $\lim_{t\to\infty} \delta\left(x^0 A^t, \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)\right) = 0$.

The general idea is to follow the steps in the analogous proof for connected non-bipartite graphs:

- Show that A is orthogonally diagonalizable, and has eigenvalue 1 with multiplicity 1. Show that all other eigenvalues are strictly smaller than 1 in absolute value.
- It might help to prove the following claim: Let B be G's normalized adjacency matrix. Then $A = \frac{1}{2}(I + B)$. Let $\lambda_1, \ldots, \lambda_n$ be B's eigenvalues (with multiplicity). Then A's eigenvalues (with multiplicity) are $\frac{1}{2}(1 + \lambda_1), \ldots, \frac{1}{2}(1 + \lambda_n)$.
- Express x^0 as the linear sum of orthonormal eigenvectors. This implies an explicit expression for $x^0 A^t$, from which the behaviour in the limit can be deduced.

Problem 5. Let $G = \left(\{0, 1\}^n, E\right)$, where there is an edge between two vertices iff they differ by a single coordinate (G is known as the **discrete cube**).

(1) Show that G is bipartite.
(2) Consider the lazy random walk on G, where $X_0 = (0, 0, \ldots, 0)$. Convince yourself that it is equivalent to the following process: At each time step a coordinate is chosen uniformly at random and set to 0 or 1 with probability $\frac{1}{2}$ each.
Let A_t be the event that by time t, all coordinates have been chosen. Show that for all $t \geq n$ and all $x \in \{0, 1\}^n$, $\mathbb{P}\left[X_t = x | A_t\right] = \frac{1}{2^n}$.
(3) Show that the for all t, $\mathbb{P}\left[A_t\right] \geq 1 - n\left(1 - \frac{1}{n}\right)^t$.
(4) Let p_t be the distribution of X_t. Show that there exists some $C > 0$ that doesn't depend on n s.t. if $t \geq Cn^2$ then $\delta\left(p_t, \left(2^{-n}, \ldots, 2^{-n}\right)\right) \leq \frac{1}{4}$.