PROBLEM SET 2

MARK DITSWORTH

1. Problem 2.1

Given graph $G = (V, E, W)$ consider a random walk on V with transition probabilities $M_{ij} = PX(t+1) = j|X(t) = i = \frac{w_{ij}}{\deg(i)}$.

Partition the vertex set as $V = V_+ \cup V_- \cup V_*$. Suppose that every node in V_* is connected to at least one node in either V_+ or V_-. Given a node $i \subset V$ let $g(i)$ be the probability that a random walker starting at i reaches a node in V_+ before reaching one in V_-. If $i \in V_+$, then $g(i) = 1$ and if $i \in V_-$, then $g(i) = 0$. Find $g(i)$ for $i \in V_*$.

Solution: There are two scenarios in which the random walker will reach a node in V_+ before one in V_-: Moving from V_* to V_+ immediately, or moving from V_* to V_* repeatedly until moving to V_+. We essentially view leaving V_* as an absorbing barrier, and want the probability that we end in V_+, which is the sum of all probabilities that will end in that scenario:

$$\sum_{j \in V_+} M_{ij} + \sum_{j \in V_*} \sum_{k \in V_+} M_{ij} M_{jk} + \sum_{j \in V_*} \sum_{k \in V_*} \sum_{l \in V_+} M_{ij} M_{jk} M_{kl} + \dots$$

Define $\nu_* \in \mathbb{R}^n$, $n = |V|$ where the jth element is 1 if $j \in V_*$, and 0 otherwise. Similarly define $\nu_+ \in \mathbb{R}^n$ for the partition V_+. Let $\Psi_i \in \mathbb{R}^n$, $n = |V|$ have elements everywhere equal 0 except for the ith node that is the starting node. The probability that a random walker starting at node i is absorbed in the V_+ partition at time-step k is expressed as

$$\Psi_i^T \left[M \operatorname{diag}(V_*) \right]^{(k-1)} M V_+$$

The $\left[M \operatorname{diag}(V_*) \right]^{(k-1)}$ matrix expresses the probability of starting at node i, and ending at node j after $k-1$ steps, discounting the intermediate nodes that would absorb the random walker earlier (nodes belonging to V_+ or V_-) For convenience, we will denote this matrix as M_*, as in only accounting for transitions within partition V_*. The resulting matrix is dotted with M for the last transition time step, and then dotted with V_+ to attain the vector of probabilities end in partition V_+. Ψ_i^T dotted with this vector selects the probability stemming from starting at node i. Thus, summing the probabilities across all k yields

$$\Psi_i^T \left[\sum_{k=0}^{\infty} (M_*)^k \right] M V_+$$

With each element of M_* less than 0, and the summation across each row bounded above by 1, it is clear that $\lim_{n \to \infty} (M_*)^n = \mathbf{0}$. Thus, the infinite sum converges to

$(I - M_*)^{-1}$. In conclusion, the probability that a random walker starting at node $i \in V_*$ will reach a node in V_+ before a node in V_- is calculated by

$$g(i) = \Psi_i^T \left(I - M_*\right)^{-1} M V_+.$$

Note that for graphs with large $|V|$, the inverse operation can be quite expensive. Thus $g(i)$ should be approximated either via pseudo-inverse operations or monte carlo simulations. See the attached notebook ps2.ipynb for justification of this closed-form expression for $g(i)$.

2. Problem 2.2

For a graph G let $h(G)$ denote its Cheeger constant and $\lambda_2(\mathcal{L}_G)$ the second-smallest eigenvalue of its normalized graph Laplacian ($\mathcal{L}_G = D - W$). The Cheeger inequality guarantees that

$$\frac{1}{2}\lambda_2(\mathcal{L}_G) \leq h_G \leq \sqrt{2\lambda_2(\mathcal{L}_G)}$$

This excercise shows that this inequality is tight (at least up to constants).

1. Construct a family of graphs $\mathcal{G}$ for which $\lambda_2(\mathcal{L}_G) \to 0$ and for which there exists a constant $C > 0$ for which

$$\forall G \in \mathcal{G}, h_G \leq C\lambda_2(\mathcal{L}_G)$$

2. Construct a family of graphs $\mathcal{G}$ for which $\lambda_2(\mathcal{L}_G) \to 0$ and for which there exists a constant $c > 0$ for which

$$\forall G \in \mathcal{G}, h_G \geq c\sqrt{\lambda_2(\mathcal{L}_G)}$$

3. Problem 2.3

Given a graph G show that the dimension of the nullspace of $\mathcal{L}_G$ corresponds to the number of connected components of G.

Solution: The nullspace of $\mathcal{L}_G$ is the set of linearly independent vectors $\mathcal{X}$ such that $\forall \mathbf{x} \in \mathcal{X}$, $\mathcal{L}_G\mathbf{x} = \mathbf{0}$. By expanding $\mathcal{L}_G$ we can see that both D and W transform each $\mathbf{x}$ to equal vectors.

$$\mathcal{L}_G\mathbf{x} = (D - W)\mathbf{x} = D\mathbf{x} - W\mathbf{x} = \mathbf{0}$$

$$D\mathbf{x} = W\mathbf{x}$$

Consider that graph $G = (V, E, W)$, $(|V| = n)$ has k connected components. Let $\mathbf{x}^i$ be a vector where

$$x_j^i = \begin{cases} 1 & j \in \text{ component } i \\ 0 & \text{otherwise} \end{cases}$$

Clearly, without loss of generality, $D\mathbf{x}^i$ is equal to a vector containing the degree of each node in G that is in component i. $W\mathbf{x}^i$ is clearly the same since for each node, it will sum all weights of edges that connect to nodes in component i. Thus, there will be k linearly independent vectors that are transformed to $\mathbf{0}$ by $\mathcal{L}_G$. All other vectors that do this will be linear combinations of $\mathcal{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^k\}$. Therefor, the cardinality of the nullspace of $\mathcal{L}_G$ is equal to the number of conneted components in graph G.

4. Problem 2.4

Given a connected unweighted graph $G = (V, E)$, its diameter is equal to

$$\operatorname{diam}(G) = \max_{u,v \in V} \min_{path p, u \to v} \operatorname{len}(p)$$

Show that

$$\operatorname{diam}(G) \geq \frac{1}{\operatorname{vol}(G)\lambda_2(\mathcal{L}_G)}$$

5. Problem 2.5

Prove the Courant Fisher Theorem: Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$ with eigenvalues $\lambda_1 \leq \dots \lambda_n$, for $k \leq n$,

$$\lambda_k(A) = \min_{U:\dim(U)=k} \left[\max_{x \in U} \frac{x^T A x}{x^T x} \right].$$

Also show that

$$\lambda_2(A) = \max_{y \in \mathbb{R}^n} \left[\min_{x \in \mathbb{R}^n : x \perp y} \frac{x^T A x}{x^T x} \right]$$

6. Problem 2.6

Given a set of points $x_1, \dots, x_n \in \mathbb{R}^p$ and a partition of them in k clusters $S_1, \dots, S_k$ reca the k-means objective

$$\min_{S_1,\dots,S_k} \min_{\mu_1,\dots,\mu_k} \sum_{l=1}^{k} \sum_{i \in S_l} \|x_i - \mu_l\|^2.$$

Show that this is equivalent to

$$\min_{S_1,\dots,S_k} \sum_{l=1}^{k} \frac{1}{S_l} \sum_{i,j \in S_l} \|x_i - x_j\|^2.$$

Solution:

$$\min_{S_1,\dots,S_k} \min_{\mu_1,\dots,\mu_k} \sum_{l=1}^{k} \sum_{i \in S_l} \|x_i - \mu_l\|^2 =$$

$$\min_{S_1,\ldots,S_k} \sum_{l=1}^{k} \sum_{i \in S_l} \|x_i - \frac{1}{|S_l|} \sum_{j \in S_l} x_j\|^2 =$$

$$\min_{S_1,\ldots,S_k} \sum_{l=1}^{k} \frac{1}{|S_l|} \sum_{i \in S_l} \||S_l| x_i - \sum_{j \in S_l} x_j\|^2 =$$

$$\min_{S_1,\ldots,S_k} \sum_{l=1}^{k} \frac{1}{|S_l|} \sum_{i \in S_l} \sum_{j \in S_l} \|x_i - x_j\|^2 =$$

$$\min_{S_1,\ldots,S_k} \sum_{l=1}^{k} \frac{1}{|S_l|} \sum_{i,j \in S_l} \|x_i - x_j\|^2.$$