Homework 5

Topics in Graph Algorithms
CS598CCI, Spring 2020
Due: 5/08/2020 10am

Instructions and Policy: Each person should write up their own solutions independently. You need to indicate the names of the people you discussed a problem with. Solutions to most of these problems can be found from one source or the other. Try to solve on your own first, and cite your sources if you do use them.

Please write clearly and concisely. Refer to known facts. You should try to convince me that you know the solution, as quickly as possible.

Do as many problems as you can. I expect you to do at least 3 with at least 2 from Problems 3 - 7.

Most of these problems are directly from Lap Chi Lau's home works from his course algorithmic spectral graph theory.

Problem 1. Let $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ be two edge-disjoint graphs on the same vertex set. Let $G = (V, E_1 \cup E_2)$.

- Prove that the algebraic connectivity of the two graphs is superadditive; that is, $\lambda_2(G_1) + \lambda_2(G_2) \leq \lambda_2(G)$.
- For any graph G, let H be a spanning subgraph of G. Infer that $\lambda_2(H) \leq \lambda_2(G)$.

Problem 2. A hypercube of n-dimension is an undirected graph with 2^n vertices. Each vertex corresponds to a string of n bits. Two vertices have an edge if and only if their corresponding strings differ by exactly one bit.

1. Given two undirected graphs $G = (V, E)$ and $H = (U, F)$, we define $G \times H$ as the undirected graph with vertex set $V \times U$ and two vertices (v_1, u_1), (v_2, u_2) have an edge if and only if either (1) $v_1 = v_2$ and $u_1 u_2 \in F$ or (2) $u_1 = u_2$ and $v_1 v_2 \in E$. Let x be an eigenvector of the Laplacian of G with eigenvalue α, and let y be an eigenvector of the Laplacian of H with eigenvalue β. Prove that we can use x and y to construct an eigenvector of the Laplacian of $G \times H$ with eigenvalue $\alpha + \beta$.
2. Use previous part to compute the spectrum of the hypercube of n dimension.
3. **Extra credit:** Show that the spectral partitioning algorithm may return a set S of conductance $\Omega(\sqrt{\phi(S^*)})$ in a hypercube, where S^* is the set with minimum conductance in the hypercube and $\phi(S^*)$ is the conductance of S^*.

Problem 3. In this question, we study the relation between "local eigenvalues" and "local conductance" similar to how Cheeger's inequality relates the global quantities for the entire graph. Let $G = (V, E)$ be an undirected d-regular graph and $\mathcal{L}$ be its normalized Laplacian matrix. Let $S \subseteq V$ be a subset of vertices with $|S| \leq |V|/2$.

First we define local eigenvalues. Let $\mathcal{L}_S$ be the $|S| \times |S|$ submatrix of $\mathcal{L}$ with rows and columns restricted to those indexed by vertices in S. Let λ_S be the smallest eigenvalue of $\mathcal{L}_S$. We say λ_S is the smallest local eigenvalue of S.

Next we define local conductance. Let $\phi(S)$ be the conductance of S in G, and let $\phi^*(S) = \min_{S' \subseteq S} \phi(S')$. We say $\phi^*(S)$ is the local conductance of S.

Prove that $\phi^*(S) \geq \lambda_S \geq (\phi^*(S))^2/2$.

Problem 4. This problem explores a Cheeger like inequality for checking how close a graph is to being bipartite. This is connected to the Max Cut problem.

Let $G = (V, E)$ be a general (non-regular) undirected graph, and $\mathcal{A} = D^{-1/2}AD^{-1/2}$ be its normalized adjacency matrix. Let

$$\beta_n = \min_x \frac{x^T(I + \mathcal{A})x}{x^T x}$$

and let

$$\beta'(G) = \min_{y \in \{-1,0,+1\}^n} \frac{\sum_{ij \in E}(y_i + y_j)^2}{\sum_{i \in V} d_i y_i^2}.$$

Prove that $\beta_n \leq \beta'(G) \leq \sqrt{2\beta_n}$.

Problem 5. This problem is on the use of random walks for PageRank.

Suppose someone searches a keyword (like "car") and we would like to identify the webpages that are the most relevant for this keyword and the webpages that are the most reliable sources for this keyword (a page is a reliable source if it points to many most relevant pages). First we identify the pages with this keyword and ignore all other pages. Then we run the following ranking algorithm on the remaining pages. Each vertex corresponds to a remaining page, and there is a directed edge from page i to page j if there is a link from page i to page j. Call this directed graph $G = (V, E)$. For each vertex i, we have two values $s(i)$ and $r(i)$, where intentionally $r(i)$ represents how relevant is this page and $s(i)$ represents how reliable it is as a source (the larger the values the better). We start from some arbitrary initial values, say $s(i) = 1/|V|$ for all i, as we have no ideas at the beginning. At each step, we update s and r (where s and r are vectors of $s(i)$ and $r(i)$ values) as follows: First we update $r(i) = \sum_{j:ji \in E} s(j)$ for all i, as a page is more relevant if it is linked by many reliable sources. Then we update $s(i) = \sum_{j:ij \in E} r(j)$ for all i (using the just updated values $r(j)$), as a page is a more reliable source if it points to many relevant pages. To keep the values small, we let $R = \sum_{i=1}^{|V|} r(i)$ and $S = \sum_{i=1}^{|V|} s(i)$, and divide each $s(i)$ by S and divide each $r(i)$ by R. We repeat this step for many times to refine the values.

Let $s, r \in \mathbb{R}^{|V|}$ be the vectors of the s and r values. Give a matrix formulation for computing s and r. Suppose G is weakly connected (when we ignore the direction of the edges the underlying undirected graph is connected) and there is a self-loop at each vertex. Prove that there is a unique limiting s and a unique limiting r for any initial s as long as $s \geq 0$ and $s \neq 0$.

(You may use the Perror-Frobenius theorem which states that for any aperiodic irreducible matrix, there is a unique positive eigenvalue with maximum absolute value and the entries of the corresponding eigenvector are all positive.)

Problem 6. Consider a random walk on a graph $G = (V, E)$ that starts at a vertex $v \in V$, and stops when it reaches s or t. Let $p(v)$ be the probability that if the random walk starts at v then it reaches s before t. Establish a connection between these probabilities and some parameters of an appropriate electric flow problem.

Problem 7. Let $G = (V, E)$ be an undirected graph where each edge e has an integral weight w_e. The weight of a spanning tree T is defined as $w_T := \prod_{e \in T} w_e$. Let $p_T = w_T / \sum_{T'} w_{T'}$, where the sum is over all the spanning trees T' of G. Let T^* be a random spanning tree sampled from the distribution p.

1. Prove that $\Pr(e \in T^*) = w_e R_{\text{eff}}(e)$ for any $e \in E$, where $R_{\text{eff}}(e)$ is the effective resistance of e when every edge e' has resistance $1/w_{e'}$.

   (You may use the equation $\det(M + xx^T) = (1 + x^T M^{-1} x) \det(M)$, for any non-singular matrix $M \in \mathbb{R}^{n \times n}$ and $x \in \mathbb{R}^n$.)

2. Prove that
$$\Pr(e \in T^* \mid f \in T^*) \leq \Pr(e \in T^*),$$
   for any two edges $e, f \in E$. In words, conditioned on the event $f \in T^*$, the probability of the event $e \in T^*$ could not increase (i.e. the events $e \in T^*$ and $f \in T^*$ are negatively correlated).