You are allowed to discuss with others but not allowed to use any references except the course notes. Please list your collaborators for each question. This will not affect your marks. In any case, you must write your own solutions.

There are totally 90 marks, and the full mark is 70. This homework is counted 14% of the course.

Some hints will be provided on May 4 in a separate file.

---

1. **Bipartite Graphs**

   (10 marks) Consider the adjacency matrix A of an undirected connected graph G. Let $\alpha_1 \geq \ldots \geq \alpha_n$ be the eigenvalues of A. Prove that $\alpha_1 = -\alpha_n$ if and only if G is bipartite.

   (You may use the Perron-Frobenius theorem to assume that all the entries of the first eigenvector is positive.)

2. **Spanning Trees**

   (15 marks) Let $G = (V, E)$ be an undirected graph.

   (a) Let $V = \{1, \ldots, n\}$, $e = ij$, and b_e be the n-dimensional vector with $+1$ in the i-th entry and -1 in the j-th entry and 0 otherwise. Let B be an $n \times m$ matrix where the columns are b_e and m is the number of edges in G. Prove that the determinant of any $(n-1) \times (n-1)$ submatrix of B is ± 1 if and only if the $n-1$ edges corresponding to the columns form a spanning tree of G.

   (b) Let L be the Laplacian matrix of G and let L' be the matrix obtained from L by deleting the last row and last column. Use (a) to prove that $\det(L')$ is equal to the number of spanning trees in G.

   (You can use the Cauchy-Binet formula (see wikipedia) to solve this problem.)

3. **Local Cheeger's Inequality**

   (10 marks) In this question, we study the relation between "local eigenvalues" and "local conductance". 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 $\phi^*(S) \geq \lambda_S \geq (\phi^*(S))^2/2$.

4. **Page Ranking**

(15 marks) 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.)

5. **Graph Partitioning by Random Walks**

(15 marks) Let $G = (V, E)$ be an undirected d-regular graph and $S \subseteq V$ be a subset of vertices with $|S| \leq |V|/2$. Let $p_t = W^t p_0$ where $W = AD^{-1}$ is the random walk matrix (D is the diagonal degree matrix and A is the adjacency matrix) and p_0 is the initial distribution.

(a) Prove that there is a probability distribution q on V such that if $p_0 = q$ then $\sum_{i \in S} p_t(i) \geq (1 - \phi(S))^t$.

(b) Use part (a) to argue that the random walk algorithm has the following performance guarantee when there is a small set S of small conductance. More precisely, say $|S| = |V|^{0.99}$ and $\phi(S) = \phi$, prove that we can use the random walk algorithm in L24 to find a set with conductance $O(\sqrt{\phi})$ and $|S| \leq |V|/2$.

6. **Hitting Time**

(10 marks) 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.

7. **Effective Resistances and Spanning Trees**

(15 marks) 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.

(a) 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 $x \in \mathbb{R}^n$.)

(b) 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.