Sample Solutions to Problem Set 1

Problem 1 (10 points) λ_2 for Laplacian of $G(n,p)$

We have seen in class that the connectivity threshold in the $G(n,p)$ model is $p = \Theta((\ln n)/n)$. In particular, we had shown that when $p < (\ln n)/2n$, the graph is disconnected with high probability (i.e., with probability $1 - o(1)$), and when $p > (4 \ln n)/n$, the graph is connected with high probability. Note that much tighter bounds on the connectivity threshold are known.

Give tight (high probability) bounds on the second smallest eigenvalue of the Laplacian of $G(n,p)$ for the corresponding cases: (a) $p < (\ln n)/2n$ and (b) $p \geq (4 \ln n)/n$.

Answer:

(a) When $p < (\ln n)/2n$, we showed in class (see lecture notes) that the graph is disconnected with high probability. We know that the number of connected components of a graph equals the multiplicity of the eigenvalue 0 of the Laplacian; therefore, with high probability, the second smallest eigenvalue is 0 with high probability.

(b) When $p \geq (4 \ln n)/n$, the graph is connected with high probability (as we showed in class). By Cheeger's inequality, we know that the second smallest eigenvalue λ_2 can be bounded as follows:

$$\frac{\phi^2}{2d_{\max}} \leq \lambda_2 \leq 2\phi,$$

where ϕ is the isoperimetric number of the graph. Let us try to bound $\phi(S)$ for a given set S of vertices, where $|S| \leq n/2$. (Recall that n is the number of vertices.) Since each edge is selected with probability p, the expected number of edges in the cut $(S, V-S)$ equals $p|S|(|V| - S|)$. It follows that the expected number of edges in the cut is $(n - |S|)|S|p$. We now use a Chernoff bound argument (which you can find on Wikipedia and all randomized algorithms textbooks). The version we use is the following. The sum X of n independent Bernoulli random variables, each taking the value 1 with probability q and 0 with probability $1-q$ satisfies the following inequality for any $0 < \delta < 1$.

$$\Pr[X < (1-\delta)E[X]] < \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{E[X]} \tag{1}$$

By choosing $\varepsilon = (1-\delta)$ sufficiently close to 0, we can make the right hand side above strictly smaller than $e^{-E[X]/2}$. We now invoke the above bound and derive that the probability that the number of edges in the cut is less than $\varepsilon(n - |S|)|S|p$ is at most

$$e^{-(n-|S|)|S|p/2} \leq \frac{1}{n^{|S|}},$$

since $|S| \leq n/2$ and $p \geq 4(\ln n)/n$. Since the number of sets of size k is at most $\binom{n}{k}$, it follows that the probability that every cut has at least $\varepsilon(n-|S|)|S|p$ edges crossing is at least

$$1 - \sum_{k=1}^{n/2} \binom{n}{k} \frac{1}{n^k} \geq 1 - \sum_{k=1}^{n/2} \left(\frac{e}{k}\right)^k$$

Unfortunately, the above bound is too weak since $(e/k)^k$ exceeds 1 for small k. The problem is that for small $k = |S|$, we are settling on a much weaker bound of $1/n^k$ than can be shown. We revisit this now. For $k = |S| \leq n/4$, we have

$$e^{-(n-|S|)|S|p/2} \leq e^{-(3n/4)kp/2} = e^{-3k\ln n/2} = \frac{1}{n^{1.5k}},$$

So we now have the desired probability bounded as

$$1 - \left(\sum_{k=1}^{n/4} \binom{n}{k} \frac{1}{n^{1.5k}}\right) - \left(\sum_{k=n/4+1}^{n/2} \binom{n}{k} \frac{1}{n^{1.5k}}\right) \geq 1 - \left(\sum_{k=1}^{n/4} \left(\frac{e}{k}\right)^k \frac{1}{n^{1.5k}}\right) - \left(\sum_{k=n/4+1}^{n/2} \left(\frac{e}{k}\right)^k \frac{1}{n^k}\right).$$

This bound tends to 1 as $n \to \infty$. We thus have that with high probability, $\phi(S)$ is $\Theta(n|S|p/|S|) = \Theta(np)$ for all S, implying that ϕ is $\Theta(np)$.

We next bound the maximum degree. The expected degree of each node is $(n-1)p$. Again, by a Chernoff bound, the degree of each node is $\Theta(np)$ with high probability. Applying these bounds to Cheeger's inequality we obtain that λ_2 is $\Theta(np)$ with high probability.

Problem 2 (10 points) Eigenvalues of the hypercube

The n-dimensional hypercube is the graph $G = (V, E)$ with $V = \{0, 1, \ldots, 2^n - 1\}$ and the set E given as follows: (i, j) is in E if the binary representations of i and j differ in exactly one bit. Find all the eigenvalues of the Laplacian of the hypercube graph.

Answer: Consider the 1-dimensional hypercube, which is simply the line with 2 nodes. The Laplacian L_1 is

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

The eigenvalues for L_1 are 0 and 2.

The $d+1$-dimensional hypercube is composed by taking two d-dimensional hypercubes, and adding an edge between corresponding nodes in each hypercube. The Laplacian L_{d+1} is given by

$$\begin{pmatrix} I_d + L_d & -I_d \\ -I_d & I_d + L_d \end{pmatrix},$$

where I_d denotes the identity matrix of dimension 2^d. Note that the d-dimensional hypercube has 2^d vertices.

If v_d is an eigenvector of L_d with eigenvalue λ, then it is easy to verify that the vector $(v_d v_d)^T$ is also an eigenvector of L_{d+1} with eigenvalue of λ. Furthermore, the vector $(v_d - v_d)^T$ is a different eigenvector of L_{d+1} with eigenvalue $\lambda + 2$. It is easy to confirm that that these vectors are all orthogonal with one another.

We use the above observations to prove by induction that for any d and $0 \le i \le d$, $2i$ is an eigenvalue of L_d with multiplicity $\binom{d}{i}$. The induction base case is immediate. Suppose the claim is true for dimension d. Consider dimension $d+1$ and $0 \le i \le d$. The number $2i$ is an eigenvalue with multiplicity $\binom{d}{i} + \binom{d}{i-1} = \binom{d+1}{i}$. Finally, $2(d+1)$ is an eigenvalue with multiplicity $\binom{d}{d} = \binom{d+1}{d+1}$.

Problem 3 (10 points) Graph properties and the Laplacian spectrum

The independence number of a graph G is the size of the largest set S of vertices such that there are no edges between any two vertices in S. The chromatic number of a graph G is the smallest number of colors needed for coloring the vertices such that no two adjacent vertices are of the same color. The diameter of the graph is the maximum pairwise distance in the graph, where the distance between any two vertices is the length of the shortest path between the vertices. Below, λ_n is the largest eigenvalue and λ_2 is the second smallest eigenvalue.

(a) Prove that the independence number of a d-regular graph is at most $n(1 - d/\lambda_n)$.

Answer: Let I denote a set of vertices that forms an independent set in G. Recall that λ_n equals

$$\max_{\|x\|=1} x^T Lx = \max_{x \neq 0} \frac{x^T Lx}{x^T x} = \max_{x \neq 0} \frac{\sum_{(i,j)\in E}(x_i - x_j)^2}{\sum_{i \in V} x_i^2}.$$

Let x denote the characteristic vector of I; $x_i = 1$ for each $i \in I$ and $x_i = 0$ for each $i \in V - I$. Then, we have

$$x^T Lx = \sum_{(i,j)\in E} (x_i - x_j)^2 = |I|d.$$

Also $x^T x = |I|$. By Courant-Fischer, we know that

$$\lambda_n = \max_{y \perp \mathbf{1}} \frac{y^T Ly}{y^T y}.$$

Unfortunately, the above ratio, taking $y = x$, evaluates as d yielding the lower bound $\lambda_n \ge d$, which is true but not useful for us.

We present a new matrix M that has the same largest eigenvalue as L, but does not encounter the same issue as above.

$$M = L + \frac{\lambda_n}{n} J,$$

where J is the all-1s $n \times n$ matrix (i.e., every entry in J is 1). Consider an eigenvector $v \neq \mathbf{1}$ of L. Since v is orthogonal to $\mathbf{1}$, it follows that $Jv = 0$. Thus, $Mv = Lv$, implying that all but the smallest eigenvalue of L are also eigenvalues of M, with exactly the same eigenvectors. Finally, Since $J\mathbf{1} = n\mathbf{1}$, we obtain that $M\mathbf{1} = L\mathbf{1} + \lambda_n \mathbf{1} = \lambda_n \mathbf{1}$, implying that $\mathbf{1}$ is an eigenvector of M with eigenvalue λ_n. Therefore, we have

$$\lambda_n = \max_{y \perp \mathbf{1}} \frac{y^T My}{y^T y} \ge \frac{x^T Mx}{x^T x} = \frac{d|I| + \lambda_n |I|^2/n}{|I|} = d + \frac{\lambda_n |I|}{n}.$$

Rewriting the above inequality, we obtain the following as desired.

$$|I| \le n(1 - \frac{d}{\lambda_n}.$$

(b) Prove that the chromatic number is at least $\lambda_n/(\lambda_n - d)$.

Answer: This follows from part (a). If χ is the chromatic number, then the independence number is at least n/χ since each of the colors identifies an independent set. From (a), we thus have

$$\frac{n}{\chi} \leq n\left(1 - \frac{d}{\lambda_n}\right).$$

Rearranging the terms yields the desired inequality.

(c) Prove that the diameter of G is at most $4d(\log n)/\lambda_2)$.

Answer: We will prove a slightly weaker bound using the constant 4.16 instead of 4; one can get it down to 4 or even better through more painful calculations. By Cheeger's inequality, we know that the isoperimetric number $\phi \geq \lambda_2/2$. Thus, for every set S of size at most $n/2$, the number of edges going out of $|S|$ is at least $\phi|S|$. Thus, the number of nodes within one hop of S (not counting the nodes in S) is at least $\phi|S|/d$.

Consider the breadth first search tree from an arbitrary node v. Using the above bound, the number of nodes at distance at most k from v is at least $(1 + \phi/d)^k$. Thus, in $k = \log(n/2)/\log(1 + \phi/d)$ steps, the number of nodes reached is at least $n/2$. Since this is true for every node, we find that the diameter of the graph is at most

$$\begin{aligned}
2\frac{\log(n/2)}{\log(1+\phi/d)} &\leq \frac{2\ln 2 \log n}{\ln(1+\phi/d)} \\
&\leq \frac{2d \ln 2 \log n(1+\phi/d)}{\phi} \\
&= 2d \ln 2 \log n(1/\phi + 1/d) \\
&\leq 2d \ln 2 \log n(2/\lambda_2 + 1/d) \\
&= \frac{4d \log n \ln 2}{\lambda_2}\left(1 + \frac{\lambda_2}{2d}\right) \\
&\leq \frac{6d \ln 2 \log n}{\lambda_2} \\
&\leq \frac{4.16 \log n}{\lambda_2}
\end{aligned}$$

(The second step follows from the inequality $\ln(1+x) > x/(1+x)$ for all $x > 0$. The second last step follows from the fact that $\lambda_2 \leq dn/(n-1)$.)

Problem 4 (10 points) Laplacian spectrum and k-partition

Let $G = (V, E)$ be an undirected graph and let $V_1, \ldots, V_k$ denote a partition of the vertex set with $|V_i| = n/k$ (assume n is divisible by k). Let e_k denote the number of edges with endpoints in two different sets of the partition. Let $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ denote the eigenvalues of the Laplacian of G. Prove that

$$\frac{n}{k}\sum_{i=1}^{k} \lambda_i \leq 2e_k \leq \frac{n}{k}\sum_{i=n-k+1}^{n} \lambda_i.$$

Answer: The above inequality is the revised version of the inequality originally posed ($n-k+2$ has been replaced by $n-k+1$). The original inequality, taken by the text by Brouwer and Haemers, seems to require a result such as the Interlacing theorem in linear algebra, which we did not cover in class. So we will settle for the weaker result.

We first establish the upper bound on $2e_k$. It is known from Schur's inequality (from linear algebra) that the sum of the k largest eigenvalues of the Laplacian is at least the sum of the k largest degrees. One can also derive Schur's inequality through a straightforward inductive argument. So we have

$$\sum_{i=n-k+1}^{n} \lambda_i \geq \sum_{i=n-k+1}^{n} d_i,$$

where d_i is the ith largest degree. Thus we have

$$\frac{n}{k} \sum_{i=n-k+1}^{n} \lambda_i \geq \frac{n}{k} \sum_{i=n-k+1}^{n} d_i \geq 2e_k.$$

We now establish the lower bound on $2e_k$. Let $v_1, v_2, \ldots, v_k$ denote k n-dimensional real vectors that are orthogonal to one another. Then, we claim that

$$\sum_{i=1}^{k} \lambda_i \leq \sum_{i=1}^{k} \frac{v_i^T L v_i}{v_i^T v_i}. \tag{2}$$

Equation 2 follows from a straightforward induction on k and Courant-Fischer's Theorem, which says the following.

$$\lambda_i = \min_{x \perp u_1, \ldots, u_{i-1}} \frac{x^T L x}{x^T x},$$

where u_j is the unit eigenvector associated with the jth smallest eigenvalue. The point is that for a given $v_1, \ldots, v_{k-1}$ the vector v_k that minimizes the right hand side of Equation 2 is the one the vector x that achieves the following minimum.

$$\min_{x \perp v_1, \ldots, v_{i-1}} \frac{x^T L x}{x^T x}.$$

The rest of the proof follows by induction.

We now use Equation 2 to establish the lower bound on $2e_k$. Let v_i be the vector that assigns 1 to all vertices in V_i and 0 to vertices not in V_i. Then, we have

$$\frac{v_i^T L v_i}{v_i^T v_i} = \frac{e(V_i, V - V_i)}{n/k}.$$

Adding the above equation over i, and applying Equation 2 since v_i's are orthogonal to one another, we obtain

$$\frac{n}{k} \sum_{i=1}^{k} \lambda_i \leq \frac{n}{k} \sum_{i=1}^{k} \frac{v_i^T L v_i}{v_i^T v_i} \leq \frac{n}{k} \sum_{i=1}^{k} \frac{e(V_i, V - V_i)}{n/k} = 2e_k.$$