Lecture 6 (October 31, 2005)

prepared by Paul Constantine

1. Expander Graphs and Graph Spectra

First recall the definitions we stated last time. The expansion of a graph $G(V, E)$ is defined as

$$\rho(G) = \min_{S \subset V} \frac{C(S, \bar{S})}{\min(|S|, |\bar{S}|)}$$

Also, the volume of a subset S is defined as

$$\mathrm{vol}(S) = \sum_{i \in S} d_i$$

where d_i is degree i. Then define the conductance of a graph as

$$\phi(G) = \min_{S} \frac{C(S, \bar{S})}{\min(\mathrm{vol}(S), \mathrm{vol}(\bar{S}))}$$

From the definitions it follows that

$$\frac{\rho(G)}{d_{max}} \le \phi(G) \le \rho(G).$$

Also recall that the adjacency matrix of a graph of with n vertices is an $n \times n$ matrix where

$$A(i, j) = \begin{cases} 1 & \text{if } i \text{ is connected to } j \\ 0 & \text{otherwise} \end{cases}$$

Another useful matrix representation of a graph is called the *Laplacian*. The Laplacian of a graph $L(G)$ is defined as

$$L_G(i, j) = \begin{cases} d_i & \text{if } i = j \\ -1 & \text{if } i \text{ is connected to } j \\ 0 & \text{otherwise} \end{cases}$$

Last week we showed that if $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ are the eigenvalues of L_G, then $\lambda_1 = 0$ and G is connected if and only if $\lambda_2 > 0$.

Sometimes it's useful to use the normalized Laplacian given by

$$\mathcal{L}_G = D^{-1/2} L_G D^{-1/2}$$

or equivalently

$$\mathcal{L}_G(i,j) = \begin{cases} \frac{-1}{\sqrt{(d_i d_j)}} & \text{if } i \sim j \\ 1 & \text{if } i = j \\ 0 & \text{o.w.} \end{cases}$$

It is easy to see that for a regular graph, there is a 1-1 correspondence between the eigenvalues and eigenvectors of L_G and $\mathcal{L}_G$, i.e. x is an eigenvector of L_G if and only if $D^{1/2}x$ is an eigenvector of $\mathcal{L}_G$.

2. Cheeger's Inequality

This inequality is very useful for examining the expansion of a graph. In general, computing the expansion is NP-hard, but using Cheeger's, we only have to compute the second eigenvalue of the normalized Laplacian to get an estimate of the expansion of a graph. Cheeger's inequality is given by

$$\frac{\phi^2}{2} \leq \lambda_2(\mathcal{L}) \leq 2\phi.$$

If G is d-regular, then we can write this as

$$\frac{\rho^2}{2d^2} \leq \frac{\lambda_2(\mathcal{L})}{d} \leq \frac{2\rho}{d}$$

or equivalently

$$\frac{\rho^2}{2d} \leq \lambda_2(\mathcal{L}) \leq 2\rho$$

We now know that a d-regular graph is a constant expander if and only if $\lambda_2(\mathcal{L})$ is bouded away from zero; Cheeger's inequality gives us a way to determine if a d-regular graph is a constant expander without explicitly computing the expansion.

Proof: First we prove that $\lambda_2(\mathcal{L}) \leq 2\rho$. Observe

$$\begin{aligned} \lambda_2 &= \min_{x \perp e,\ \|x\|=1} \sum_{i \sim j} (x_i - x_j)^2 \\ &= \min_{x \perp e} \frac{\sum_{i \sim j}(x_i - x_j)^2}{\sum_i x_i^2} \end{aligned}$$

Note that $\sum_i x_i = 0$, then $\sum_{i,j}(x_i - x_j)^2 = 2n \sum_i x_i^2$, since

$$\sum_{i,j}(x_i - x_j)^2 \ = \ 2n \sum_i x_i^2 - 2 \sum_{i,j} x_i x_j \ = \ 2n \sum_i x_i^2 - 2 \sum_i x_i \sum_j x_i \ = \ 2n \sum_i x_i^2$$

This implies that

$$\begin{aligned} \lambda_2 &= \min_{x \perp e} 2n \frac{\sum_{i \sim j}(x_i - x_i)^2}{\sum_{i,j}(x_i - x_j)^2} \\ &= 2n \frac{\sum_{i \sim j}(x_i - x_i)^2}{\sum_{i,j}(x_i - x_j)^2}. \end{aligned}$$

Now examine ρ:

$$
\begin{aligned}
\rho(G) &= \min_{S \subset V} \frac{C(S,\bar{S})}{\min(|S|,|\bar{S}|)} \\
&\geq \min_{S \subset V} \frac{C(S,\bar{S})}{|S|,|\bar{S}|} \frac{n}{2} \\
&= \min_{S \subset V} \frac{\sum_{i \sim j}(x_i - x_j)^2}{\sum_{i<j}(x_i - x_j)^2} \frac{n}{2}
\end{aligned}
$$

where

$$
x_i = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \notin S \end{cases}
$$

From these two results, we deduce that $\lambda_2(\mathcal{L}) \leq 2\rho(G)$.

Next we prove the second inequality. Let $x_1, x_2, \ldots, x_n$ be the components of $\lambda_2(\mathcal{L})$. Recall that

$$
\lambda_1(\mathcal{L}) \ = \ 0 \ = \ e^T \mathcal{L} e
$$

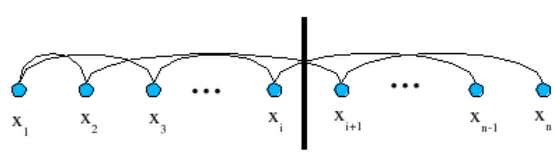
and

$$
\lambda_2(\mathcal{L}) \ = \ \min_{x \perp e, \|x\|=1} x^T \mathcal{L} x \ = \ \sum_{i \sim j} (x_i - x_j)^2
$$

Number the vertices of the graph such that $x_1 \geq x_2 \geq \ldots \geq x_n$ and the number of positive x_i's are less than or equal to the number of non-positive x_i's. Define

$$
y_i = \begin{cases} x_i & \text{if } x_i > 0 \\ 0 & \text{otherwise} \end{cases}
$$

Order the vertices in the order of the x_i's as shown in the figure, and cut the graph between two vertices i and $i+1$.



Let c_i be the set of edges that were cut, i.e.

$$
c_i = \{\{u, v\} \ : \ x_u \geq x_i > x_v\}
$$

Recall that we're trying to show that $\lambda_2 \geq \rho^2/2d$, which would mean that if we had a graph with a small eigenvalue gap, then I ought to be able to find a bad cut on that graph. So if λ_2 is small, then there is a way to cut the above representation that is a bad cut. We will show that

$$
\rho \leq \frac{|c_i|}{i}
$$

for some $i < n/2$. Motivated by this, we write

$$\begin{aligned}
\lambda_2(\mathcal{L}) &= \frac{\sum_{i\sim j}(x_i - x_j)^2}{\sum_i x_i^2} \\
&\geq \frac{\sum_{i\sim j}(y_i - y_j)^2}{\sum_i y_i^2} \\
&= \frac{\sum_{i\sim j}(y_i - y_j)^2}{\sum_i y_i^2} \times \frac{\sum_{i\sim j}(y_i + y_j)^2}{\sum_i (y_i + y_j)^2}
\end{aligned}$$

Now we employ the following useful inequality:

$$\sum_i \alpha_i^2 \times \sum_i \beta_i^2 \geq \sum (\alpha_i \beta_i)^2$$

to yield

$$\sum_{i\sim j}(y_i + y_j)^2 \leq 2d \sum_i y_i^2.$$

And this implies

$$\lambda_2(\mathcal{L}) \geq \frac{(\sum_{i\sim j}|y_i^2 - y_j^2|)^2}{2d(\sum_i y_i^2)^2}$$

Up to this point, everything we've done has been algebraic manipulation. Here is where we use insight. Assume $y_i > y_j$, then

$$\lambda_2(\mathcal{L}) \geq \frac{(\sum_{i\sim j, i<j} y_i^2 - y_j^2)^2}{2d(\sum y_i^2)^2}$$

Note we can rewrite the sum of the differnces as a telescoping sum since $y_i^2 - y_j^2 = (y_i^2 - y_{i+1}^2) + \ldots + (y_{j-1}^2 - y_j^2)$. Also, since we're only summing over the edges in the cut, this telescoping sum becomes

$$\begin{aligned}
\lambda_2(\mathcal{L}) &\geq \frac{(\sum_i |c_i|(y_{i+1}^2 - y_i^2)^2}{2d(\sum y_i^2)^2} \\
&\geq \frac{\rho^2(\sum_i i(y_{i+1}^2 - y_i^2))^2}{2d(\sum y_i^2)^2} \\
&\geq \frac{\rho^2}{2d} \times \left(\frac{\sum y_i^2}{\sum y_i^2}\right)^2 \\
&= \frac{\rho^2}{2d}
\end{aligned}$$

■

Remarks:

- It turns out that you don't have to compute the eigenvector. In fact, the proof works with any arbitrary vector.

- Consider the following algorithm: Given a graph,

  1. Compute the second eigenvector x of the normalized Laplacian.
  2. Order the vertices with respect to x_i's.
  3. Find i such that $|c_i|/\min(i, n-i)$ is minimized.
  4. Divide the graph into two subgraphs: $V_1 = \{v \ : \ x_v < x_i\}$ and $V_2 = \{v \ : \ x_v \geq x_i\}$.
  5. Repeat the process for the subgraphs on V_1 and V_2.

  This algorithm produces clusters in the graph.