October 26, 2007
Prepared by Adam Guetz

Random Walks and Spectral Graph Theory

Let $G(V, E)$ be a simple, undirected, connected graph, with $|V| = n$, $|E| = m$.

P_t is the vector indicating the probability that the particle is in each vertex at time t. Let $M_{n\times n} = D_{n\times n}A_{n\times n}$, where A is the adjacency matrix, and D is the diagonal matrix with $D_{ii} = 1/d_i$. Then $P_{t+1} = M^T P_t$.

The *hitting time* $H(i, j)$ is the expected time for a row starting at i to visit j for the first time.

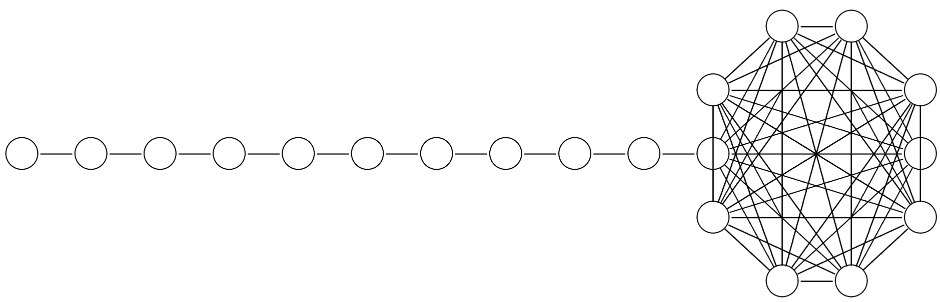
Remark: Even if G is undirected, $H(i, j)$ is not necessarily symmetric, i.e. $H(i, j) \neq H(j, i)$ in general.

Definition 7.1: $C(i, j) = H(i, j) + H(j, i)$

Remark: There is an interesting connection between hitting time and the conductance of an electrical network. See the link on the webpage to a book-length treatment of this connection.

The *cover time* $C(i)$ is the expected time for a random walk started at i to visit all the vertices.

What is the graph with the longest cover time? The lollipop graph:



What is the graph with the shortest cover time? This is actually an open problem:

Conjecture 7.1 *The graph with the shortest cover time on n vertices is K_n.*

It is easy to verify that $P = \left(\frac{d_1}{2m}, \frac{d_2}{2m}, \ldots, \frac{d_n}{2m}\right)$ satisfies $P = M^T P$.

It is often easier to work with a symmetrized version of M,

$$N = D^{1/2}AD^{1/2} = D^{-1/2}MD^{1/2}$$

N is symmetric, so it can be written

$$N = \sum_{k=1}^{n} \lambda_k v_k v_k^T$$

with the first eigenvector $v_1 = \left(\sqrt{\frac{d_1}{2m}}, \sqrt{\frac{d_2}{2m}}, \ldots, \sqrt{\frac{d_n}{2m}}\right)$, and the first eigenvalue $\lambda_1 = 1$. Therefore, $|\lambda_k| \leq 1$ for $1 \leq k \leq n$. It can be verified that if G is bipartite, then $\lambda = -1$ and otherwise $\lambda_n > -1$.

$$M^T = D^{1/2}N^T D^{-1/2} = D^{1/2}v_1 v_1^T D^{-1/2} + \sigma_{k=1}^{n} \lambda_k^T D^{1/2} v_k v_k^T D^{-1/2}$$

The first term in the above equation, $QD^{1/2}v_1v_1^TD^{-1/2}$ is such that $Q : Q_{ij} = d_j/2m$. If G is not bipartite, then

$$P_t \to= P = \left(\frac{d_1}{2m}, \frac{d_2}{2m}, \ldots, \frac{d_n}{2m}\right)$$

The rate of convergence will depend on $\lambda = \max(\lambda_2, -\lambda)n$. For a random walk starting from vertex i,

$$|P_t(j) - P(j)| \leq \lambda^t \sqrt{\frac{d_j}{d_i}}$$

What this says is that the principle non-stationary eigenvalues of the matrix need to be well separated from 1 to convergy quickly to the stationary distribution.

Recall that the graph Laplacian is defined as

$$L_{ij}(G) = \begin{cases} d_i & : i = j \\ -1 & : i \sim j \\ 0 & : \text{otherwise} \end{cases}$$

It has eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ and $\lambda_2 = 0$ iff G is connected.

It is often easier to work with the *normalized Laplacian*,

$$\mathcal{L}_{ij}(G) = I - N = \begin{cases} d_i & : i = j \\ -\frac{1}{\sqrt{d_i d_j}} & : i \sim j \\ 0 & : otherwise \end{cases},$$

with $0 = \lambda_1(\mathcal{L}) \leq \lambda_2(\mathcal{L}) \leq \ldots \leq \lambda_n(\mathcal{L}) \leq 2$.

Definition 7.2: Expansion The *expansion* $\rho(G)$ is the minimum edge cut between two sets divided by the size of the smaller set,

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

Definition 7.3: Sparsity The *sparsity* of a cut $sp(S, \bar{S})$ is

$$sp(S, \bar{S}) = \frac{C(S, \bar{S})}{\min(vol(S), vol(\bar{S}))},$$

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

Definition 7.4: Conductance The *conductance* of a graph $\phi(G)$ is

$$\phi(G) = \min_{s \subset V} sp(S, \bar{S})$$

Unfortunately, these measures are NP hard to compute. Not only that, they are difficult to approximate.

There is, however, an interesting and useful relationship between these conductance, expansion and the spectral gap.

Theorem 7.1 (Cheeger's inequality)

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

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

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

Given the vector x that minimizes this quantity, label the vertices of the graph so that $x_1 \geq x_2 \geq \cdots \geq x_n$ and the number of positive x_i's are less than or equal to the number of non-positive x_i's. Define

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

We define c_i the cut $C(\{x_1,\ldots,x_i\},\{x_{i+1},\ldots,x_n\})$. Note 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

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

And this implies

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

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

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

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

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

■