College of Computer & Information Science
Northeastern University
CS 7880: Network Algorithms and Analysis
Spring 2012
3 February 2012
Scribe: Rajmohan Rajaraman

- Normalized random walk and Laplacian matrices
- Random walks in nonregular undirected graphs
- Cover time

1 Normalized random walk and Laplacian matrices

Thus far, we have focused on d-regular graphs. The associated random walk matrix in that case is symmetric and we can draw results from the spectral decomposition of such graphs to obtain bounds on the rate of convergence. What can be done in the more general setting of nonregular undirected graphs?

For a nonregular undirected graph, the adjacency matrix is still symmetric; however, the random walk matrix M, where M_{ij} equals $1/d(i)$ with $d(i)$ being the degree of i, is not. Note that $M = AD^{-1}$. Define the *normalized random walk matrix* $N = D^{-1/2}MD^{1/2}$, where D is the diagonal matrix with the ith diagonal entry being $d(i)$. We thus have

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

which is a symmetric matrix.

Lemma 1. *The matrices M and N have the same eigenvalues and closely related eigenvectors.*

Proof: Let v be an eigenvector of N, with eigenvalue λ. Then, we have

$$D^{-1/2}MD^{1/2}v = Nv = \lambda v.$$

Multiplying both sides by $D^{1/2}$ we obtain

$$MD^{1/2}v = \lambda D^{1/2}v.$$

Thus $D^{1/2}v$ is an eigenvector of M with eigenvalue λ/ □

2 Random walks in nonregular undirected graphs

What is the stationary distribution of a random walk in an arbitrary undirected graph? It is easy to verify that the stationary distribution is given by π:

$$\pi(v) = \frac{d(v)}{\sum_{u \in V} d(u)} = \frac{d(v)}{2m},$$

where m is the number of edges. We now study the convergence rate of a random walk to the stationary distribution.

Let $|\mu_1| \geq |\mu_2| \geq \ldots \geq |\mu_n|$ denote the eigenvalues of the normalized walk matrix N with corresponding eigenvectors $v_1, \ldots, v_n$. Consider probability vector p_t of the random walk at time t, with p_0 being the initial vector. Since $N = \sum_k \mu_k v_k^T v_k$, we obtain

$$M^t = (D^{1/2} N D^{-1/2})^t = D^{1/2} N^t D^{-1/2} = \sum_{k=1}^{n} \mu_k^t D^{1/2} v_k^T v_k D^{-1/2} = \pi + \sum_{k=2}^{n} \mu_k^t D^{1/2} v_k^T v_k D^{-1/2}.$$

Suppose we start the random walk from vertex i, then p_0 has its ith component as 1 and the others being zero. Then, we obtain

$$p_t(j) = \pi(j) + \sum_{k=2}^{n} \mu_k^t v_k(i) v_k(j) \sqrt{\frac{d(j)}{d(i)}}.$$

We now can bound $p_t(j)$ as follows.

$$\begin{aligned} |p_t(j) - \pi(j)| &= |\sum_{k=2}^{n} \mu_k^t v_k(i) v_k(j) \sqrt{\frac{d(j)}{d(i)}}| \\ &\leq |\mu_2|^t |\sum_{k=2}^{n} v_k(i) v_k(j) \sqrt{\frac{d(j)}{d(i)}}| \\ &\leq |\mu_2|^t \sqrt{\frac{d(j)}{d(i)}}|. \end{aligned}$$

In the last step, we use the fact that $\sum_{k=2}^{n} v_k(i) v_k(j) \leq 1$. This follows from the fact that the matrix Q formed by the eigenvectors $v_1, \ldots, v_n$ is orthonomal – so $Q^T Q = I$ – so every row and column is a unit vector.

3 Cover time of random walks

In addition to the convergence rate of random walks, measures of interest include *hitting time*, the expected time that it takes for the random walk to visit a particular vertex and *cover time*, the expected time it takes to visit every vertex in the graph.

Let $h_{u,u}$ denote the expected time it takes for a random walk starting from u to return to u. Suppose X_i denote the number of steps between the ith visit and the $(i+1)$st visit to u of the walk. Then, $h_{u,u}$ is simply the limit, as s tends to infinity of $(X_1 + X_2 + \ldots + X_s)/s$. (Note that both the mean and variance of X_i can be shown to be finite.) But as s tends to infinity, $s/(X_1 + X_2 + \ldots + X_s)$ is simply the fraction of time the walk is at s, which equals the visit probability in the stationary distribution. So we have:

$$\pi(v) = \frac{d(v)}{2m} = \frac{1}{h_{v,v}}.$$

We thus know that $h_{v,v} = 2m/d(v)$. Note that $\pi(v)/d(v) = 1/2m$. Therefore, the probability that the walk is going from vertex u to v in stationary distribution is $1/(2m)$ for any edge (u, v). From this, we can calculate the expected number of steps for the walk to traverse the edge from v to u, if it is presently traversing the edge from u to v. It is simply $h_{v,v} d(v) = 2m$. We can now place an upper bound on the cover time by simply adding the above traversal time bound over the Eulerian walk edges of a spanning tree to obtain a cover time bound of at most $2m(n-1)$.

Theorem 1. *The cover time for any undirected graph with n vertices and m edges is at most $2m(n-1)$.*