CMPT 881: Pseudorandomness	Prof. Valentine Kabanets

Lecture 5: Random Walk on Graphs and Spectral Graph Theory

September 21, 2004 — *Scribe: Sasatte Prashant*

1 Random Walk on Graphs

1.1 Preliminaries

Definition 1 *A Markov chain is sequence of random variables $\{X_n\}$ where each X_i takes value from a finite set $[n] = \{1, 2, \ldots, n\}$ and has the following property*

$$\Pr(X_t = j | X_0 = i_0, X_1 = i_1 \ldots, X_{t-1} = i_{t-1}) = \Pr(X_t = j | X_{t-1} = i_{t-1}).$$

This property is called *Markov(memoryless)* property. It implies that the probability distribution of X_t is completely determined by X_{t-1}. Let p_{ij} denote $\Pr[X_t = j | X_{t-1} = i]$. Then, the $n \times n$ matrix P, with p_{ij} as entries, is called transition probability matrix of the Markov chain. The transition probability matrix together with initial distribution completely determines the Markov chain. If $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ is the initial probability distribution and A is the transition probability matrix of some **MC**(Markov chain) then πA is the probability distribution of that Markov chain after one step.

Definition 2 *A probability distribution π is called the stationary distribution of a Markov chain iff $\pi A = \pi$, where A is the transition probability matrix of the Markov chain.*

1.2 Spectral Theory of Graphs

Let G be an undirected multigraph on n vertices and A be the *normalized adjacency* matrix of G, that is, the $n \times n$ matrix whose ij^{th} entry equals the number of edges between i and j divided by $deg(i)$.

For example,

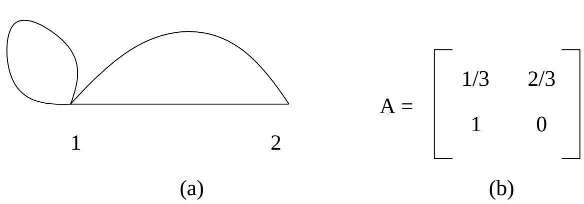

Figure 1: (a) Undirected Multigraph G (b) Normalized Adjacency matrix of G

The normalized adjacency matrix A of an undirected multigraph G is a real symmetric $n \times n$ matrix whose both rows and columns sum to 1. This matrix defines a **Markov Chain** which we interpret as a random walk on G.

Proposition 3 *If A is the normalized adjacency matrix of an undirected connected multigraph G then Markov chain defined by A has a unique stationary distribution π and $\forall i\ \pi_i > 0$. Morever, for any starting distribution*

$$\lim_{t \to \infty} \frac{N(i,t)}{t} = \pi_i \qquad \textit{Ergodic theorem of Markov chain}$$

Where, $N(i,t) = \#$ Number of visits to state i during t steps of ***MC***

Proposition 4 *In addition, If G is not a bipartite graph then for any initial distribution,*

$$\forall i \in [n] \lim_{t \to \infty} \Pr(X_t = i) = \pi_i.$$

Theorem 5 *if A is the normalized adjacency matrix of an undirected multigraph G then the stationary distribution of the Markov chain defined by A is given by $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ where $\pi_i = \frac{\deg(i)}{2e}$, where $e = |E(G)|$.*

Proof: We will show that the distribution π satisfies $A\pi = \pi$.

$$\begin{aligned}
(\pi A)_j &= \sum_i \pi_i a_{ij} \\
&= \sum_i \frac{\deg(i)}{D} \frac{\#\text{ edges from } j \text{ to } i}{\deg(i)} \text{ where } D = 2e \\
&= \frac{1}{D} \sum_i \#\text{ edges between } i \text{ and } j \\
&= \frac{\deg(j)}{D} \\
&= \pi_j.
\end{aligned}$$

Hence, By definition, π is the stationary distribution of the Markov chain defined by A and by proposition 3 it is unique. ■

Corollary 6 *If G is d-regular connected multigraph then the stationary distribution of the Markov chain defined by the normalized adjacency matrix of G is the uniform distribution $u = (1/n, 1, n, \dots, 1/n)$.*

From now on, we will consider only the d-regular, connected, undirected graphs. Thus, the stationary distribution of the Markov chains we will consider will be the uniform distribution u.

We state the following two theorems which we will use in the analysis of random walk on graphs.

Theorem 7 *If M is real symmetric $n \times n$ matrix then M has n real orthogonal eigenvectors $v_1, v_2, \dots, v_n$.(which form a basis in $\mathbb{R}^n$).*

Since, the normalized adjacency matrix A of an undirected, connected multigraph G is an $n \times n$ real symmetric matrix, it has n real orthogonal eigenvectors. Also, 1 is the eigenvalue of A and u is corresponding eigenvector.

Theorem 8 *If G is a d-regular undirected multigraph, then*

1. *All eigenvalues of A have absolute value at most 1. So, write $\lambda_1 = 1$ and $|\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$ (in descending order of absolute value). Also, Define $\lambda_i(G) = |\lambda_i|$.*

2. *G is connected iff the eigenvalue 1 has multiplicity 1.*

3. *if G is connected then G is bipartite iff -1 is also eigenvalue of A.*

4. *G is connected and nonbipartite then $|\lambda_2| \leq (1 - \frac{1}{n.d.\ diameter(\ G\ )})$.*