CS395T Pseudorandomness and Combinatorial Constructions
Lecture 12, February 22, 2001
Lecturer: Prof. David Zuckerman, Scribe: K. J. Balakrishnan

1 Random Walks on Graphs

A random walk on a graph $G = (V, E)$ is a sequence of random variables $X_0, X_1, X_2 \ldots \in V$ such that X_{t+1} is a uniformly random neighbour of X_t.

1.1 Properties

If G is connected and not bipartite, then there exists a unique limiting distribution π called the stationary distribution, i.e. $p_t \to \pi$, where p_t is the probability distribution of X_t. For a regular graph $G = (V, E)$, $|V| = n$, $|E| = m$, $\pi_i = 1/n$. In general, $\pi_i = \frac{d_i}{2m}$, d_i is the degree of vertex i. Let A be the transition matrix of the random walk, $A_{ij} = \Pr[X_{t+1} = j / X_t = i] = \frac{1}{d_i}$. If G is d-regular, then $A = \frac{1}{d}\hat{A}$, where $\hat{A}$ is the adjacency matrix of the graph.

Now $p_{t+1}^T = p_t^T A$. For non-bipartite graphs, $p_t \to \pi$ and hence $\pi^T = \pi^T A$ i.e. π is the left eigenvector for eigenvalue 1.

1.2 Relation between Convergence and 2^{nd} eigenvalue

Consider a d-regular graph. For such a graph, $A = A^T$. Hence $p_{t+1} = Ap_t$. A has eigenvalues given by $\lambda_i = \frac{\hat{\lambda}_i}{d}$, where $\hat{\lambda}_i$ are the eigenvalues of the adjacency matrix of the graph. Hence,

$$1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n \geq -1$$

The last equality holds only when G is bipartite. Let us assume that G is not bipartite. Let

$$\lambda = \max(\lambda_2, -\lambda_n) = \max_{i \neq 1} |\lambda_i|.$$

We can write $p_t = \pi + q_t$ where q_t can be considered as an error term at time t. Now $q_t \perp \pi$ (Reason: $\pi = \frac{1}{n}\mathbf{1}$, so $\mathbf{1} \cdot q_t = \mathbf{1} \cdot p_t - \mathbf{1} \cdot \pi = 1 - 1 = 0$). Therefore,

$$\begin{aligned} p_{t+1} &= Ap_t \\ &= A\pi + Aq_t \\ &= \pi + Aq_t \end{aligned}$$

$\|Aq_t\|$ has components only in eigenvectors corresponding to eigenvalues not equal to 1. Therefore, $\|Aq_t\| \leq \lambda \|q_t\|$, or $\|q_t\| \leq \lambda^t \|q_0\|$. Hence the random walk converges quickly if the 2^{nd} eigenvalue is small.

1.2.1 Measurement of Closeness to Uniform Distribution

In order to have a correct measurement of the closeness of a given distribution to the uniform distribution, we need to define the variation distance between two distributions.

Definition 1 *Variation Distance : If D_1 and D_2 are distributions on X, then the variation distance $|D_1 - D_2|$ is*

$$\max_{S \subseteq X} |D_1(S) - D_2(S)| = \frac{1}{2}|D_1 - D_2|_1 = \frac{1}{2} \sum_{x \in X} |D_1(x) - D_2(x)|$$

Now for $t \geq \frac{c \ln n}{1-\lambda}$,

$$\begin{aligned} |p_t - \pi| &= \frac{1}{2}|q_t|_1 \\ &\leq \frac{1}{2}\sqrt{n}\|q_t\| \\ &\leq \sqrt{n}\lambda^t \\ &\leq n^{1/2-c} \end{aligned}$$

Definition 2 *Mixing Time is defined as*

$$\tau = \max_{P_0} \min_t \{t : |p_t - \pi| \leq 1/2\}.$$

With this definition, we can show that: $\frac{1}{1-\lambda} \leq \tau \leq \frac{\ln n}{1-\lambda}$. Hence the mixing time depends on the 2^{nd} eigenvalue. The following table shows the values of the 2^{nd} eigenvalue and mixing time for some examples.

Graph	Vert Expansion	Edge Expansion	λ	τ
Cycle	$\Theta(1/n)$	$\Theta(1/n)$	1 - $\Theta(1/n^2)$	$\Theta(n^2)$
2-D grid	$\Theta(1/\sqrt{n})$	$\Theta(1/\sqrt{n})$	$1 - \Theta(1/n)$	$\Theta(n)$
Hypercube	$\Theta(1/\sqrt{\log(n)})$	$\Theta(1)$	$1 - \Theta(1/\log(n))$	$\Theta(\log(n)\log(\log n))$
Const. deg exp.	$\Theta(1)$	$\Theta(1)$	$1 - \Theta(1)$	$\Theta(\log(n))$

Definition 3 *A lazy random walk is a random walk in which at each time step the walks stays where it is with probability $1/2$ and otherwise makes a step according to the usual random walk.*

Thus $A_{Lazy} = \frac{1}{2}(I + A)$ and $\lambda_{i_{Lazy}} = \frac{1}{2}(1 + \lambda_i)$. Hence none of the eigenvalues stay negative.

2 Application to Pseudorandomness

Random walks on graphs can be used to amplify the success probability of a **BPP** algorithm. Here we consider the **RP** case. Suppose we have an **RP** algorithm using r random bits. We will construct a pseudorandom generator based on a constant degree expander enabling the algorithm to use $r + O(s)$ bits and have an error $\leq 2^{-s}$.

Algorithm(*Ajtai-Komlos-Szemeredi*) G is d-regular with $V = \{0,1\}^r$ and has 2^{nd} eigenvalue in absolute value at most λ (add self-loops if necessary). Pick $X \in_u \{0,1\}^r = V$. Pick

$y_1, y_2, \dots y_s \in_u [d]$. Perform a random walk starting at X and following the edges $y_1, y_2, \dots y_s$, say $RW = \{X = X_0, X_1, \dots X_s\}, X_i \in V$. Now we run the **RP** algorithm with random strings $X_0, \dots, X_s$.

Proof: Let $W \subseteq \{0,1\}^r$ be the set of witnesses for the **RP** algorithm. We find the probability $\Pr[\forall i X_i \notin W]$. Let N be the projection matrix into $\bar{W}$, i.e. N is a diagonal matrix with entries 1 corresponding to elements of $\bar{W}$ and 0 corresponding to elements of W. In other words, for any column vector v,

$$(Nv)_i = \begin{cases} v_i, i \in \bar{W} \\ 0, o.w. \end{cases}$$

We start with $X_0 = \pi$. $N\pi$ zeroes out all the elements of W from the distribution. Now $(NA)N\pi$ zeroes out all elements of W after one step of the walk. In general, the i^{th} component of the column vector $(NA)^t N\pi$ represents the probability of being at vertex i and not having visited any single witness vertex on the walk, ie $\Pr[\forall j, j \le t, X_j \notin W]$. Hence the required probability is:

$$\begin{aligned} \Pr[\forall i, X_i \notin W] &= |NA(NA(NA \dots N\pi))|_1 \\ &\le \sqrt{2^r} \|NA\|^s \|N\pi\| \\ &\le \|NA\|^s \end{aligned}$$

since $\|N\pi\| \le \|\pi\| = \sqrt{2^r 2^{-2r}} = \sqrt{2^{-r}}$. We now prove that $\|NA\| \le c$ for some constant $c < 1$. Consider $NA\mathbf{1} = N\mathbf{1}$. This is a column matrix with elements 1 corresponding to $\bar{W}$ and 0 for W. Hence

$$\|NA\mathbf{1}\| = \sqrt{|\bar{W}|}$$

or

$$\frac{\|NA\mathbf{1}\|}{\|\mathbf{1}\|} = \sqrt{\frac{|\bar{W}|}{2^r}} = \sqrt{\mu(\bar{W})}$$

For any $v \perp \mathbf{1}$,

$$\|NAv\| \le \|Av\| \le \lambda \|v\|$$

or

$$\frac{\|NAv\|}{\|v\|} \le \lambda.$$

Therefore

$$\|NA\| \le \lambda + \sqrt{\mu(\bar{W})}$$

Remember that the norm of a matrix M is defined as

$$\|M\| = \max_{x \ne 0} \frac{\|Mx\|}{\|x\|}$$

So, we have

$$\Pr[\forall i, X_i \notin W] \le \|NA\|^s \le \left(\sqrt{\mu(\bar{W})} + \lambda\right)^s \le \left(\sqrt{1/2} + 1/4\right)^s = 2^{-\Theta(s)}$$

for $\lambda \le 1/4$. Hence we need choose our graph such that $\lambda \le 1/4$. □