

Random Walk

Lecturer: Sushant Sachdeva

Scribe: Junwei Su

HW: The proof of statement is left as exercise for the student

1 Remark from last class

If at step t , your distribution is given by \mathbf{p}_t , then the next distribution \mathbf{p}_{t+1} is given by:

$$\mathbf{p}_{t+1} = \mathbf{A}\mathbf{D}^{-1}\mathbf{p}_t, \quad (1)$$

where:

A: weighted adjacency matrix

D: diagonal weighted degree matrix

The state transition can also be expressed as:

$$\mathbf{p}_{t+1}(x) = \sum_{y:(x,y) \in E} \frac{w(x,y)}{\mathbf{D}(y)} * \mathbf{p}_t(y)$$

2 Stationary State

Definition 2.1. If distribution $\pi \in \mathbb{R}^v$ is said to be stationary distribution for G if $\mathbf{A}\mathbf{D}^{-1}\pi = \pi$

Lemma 2.2. Any undirected graph has a stationary distribution

Proof. Given any undirected graph G , let

$$\pi = \frac{1}{\mathbf{1}^\top \mathbf{D} \mathbf{1}} \mathbf{D} \mathbf{1}$$

. This is the stationary distribution for G since

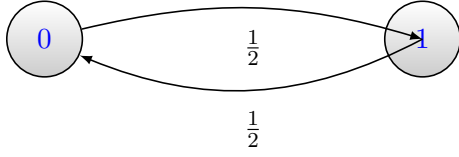
$$\mathbf{A}\mathbf{D}^{-1}\pi = \frac{1}{\mathbf{1}^\top \mathbf{D} \mathbf{1}} \mathbf{A} \mathbf{1} = \frac{1}{\mathbf{1}^\top \mathbf{D} \mathbf{1}} \mathbf{D} \mathbf{1} = \pi$$

□

Claim: If G is connected, π is unique

Remark: Even if G is connected, it is not true that for any \mathbf{p}_0 , $\mathbf{p}_t \rightarrow \pi$

Example:



The stationary state $\pi = (\frac{1}{2}, \frac{1}{2})$ but the random walk will alternate between Vertex 0 and Vertex 1

3 Positive Semi-Definite(PSD)

Definition 3.1. A symmetric matrix \mathbf{M} is positive semi-definite(psd) if $\forall \mathbf{x}, \mathbf{x}^\top \mathbf{M} \mathbf{x} \geq 0$

Theorem 3.2. the following statements are equivalent:

- (1) \mathbf{M} is psd
- (2) All eigenvalues of \mathbf{M} are non-negative
- (3) There exist an matrix \mathbf{A} such that $\mathbf{M} = \mathbf{A} \mathbf{A}^\top$

Lemma 3.3. If \mathbf{M} is psd, then for all matrices \mathbf{C} , $\mathbf{C}^\top \mathbf{M} \mathbf{C}$ is psd

Proof. $\forall \mathbf{x}, \mathbf{x}^\top \mathbf{C}^\top \mathbf{M} \mathbf{C} \mathbf{x} = (\mathbf{C} \mathbf{x})^\top \mathbf{M} (\mathbf{C} \mathbf{x}) \geq 0$ since \mathbf{M} is psd

□

Notation: \mathbf{M} is psd $\Leftrightarrow \mathbf{M} \succeq 0$

Lemma 3.4. $\mathbf{L} \succeq 0$ where \mathbf{L} is the laplacian matrix of some graph

Remark: $\mathbf{L} \succeq 0$ implies $\mathbf{N} \succeq 0$ since $\mathbf{N} = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}}$

Lemma 3.5. $\mathbf{L} \preceq 2\mathbf{D} \Leftrightarrow \mathbf{N} \preceq 2\mathbf{I} \Leftrightarrow \lambda_i(\mathbf{N}), \nu_i \leq 2$

HW: If $\mathbf{A} \succeq \mathbf{B}$, then $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B})$

4 Lazy random walk

4.1 Lazy random walk matrix

At each step, the lazy random walk will do the following

$$\begin{cases} \text{with probability } \frac{1}{2} & \text{stay at the current vertex} \\ \text{with probability } \frac{1}{2} & \text{take a usual random step} \end{cases}$$

Lazy Random Walk Transition Matrix $\mathbf{W} = \frac{1}{2}(\mathbf{I} + \mathbf{A}\mathbf{D}^{-1})$

We know that the normalized Laplacian(\mathbf{N}) can be expressed as:

$$\begin{aligned}\mathbf{N} &= \mathbf{D}^{-\frac{1}{2}}\mathbf{L}\mathbf{D}^{-\frac{1}{2}} \\ &= \mathbf{D}^{-\frac{1}{2}}(\mathbf{D} - \mathbf{A})\mathbf{D}^{-\frac{1}{2}} \\ &= \mathbf{I} - \mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}\end{aligned}$$

applied this result to the lazy walk transition matrix

$$\begin{aligned}\mathbf{W} &= \frac{1}{2}\mathbf{I} + \frac{1}{2}\mathbf{A}\mathbf{D}^{-1} \\ &= \frac{1}{2}\mathbf{I} + \frac{1}{2}\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}\mathbf{D}^{-\frac{1}{2}} \\ &= \frac{1}{2}\mathbf{I} + \frac{1}{2}\mathbf{D}^{\frac{1}{2}}(\mathbf{I} - \mathbf{N})\mathbf{D}^{-\frac{1}{2}} \\ &= \mathbf{I} - \frac{1}{2}\mathbf{D}^{\frac{1}{2}}\mathbf{N}\mathbf{D}^{-\frac{1}{2}}\end{aligned}$$

Thus, we can express the lazy random walk transition matrix as:

$$\mathbf{W} = \mathbf{I} - \frac{1}{2}\mathbf{D}^{\frac{1}{2}}\mathbf{N}\mathbf{D}^{-\frac{1}{2}} \quad (2)$$

4.2 Eigenpair for lazy random walk matrix

Lemma 4.1. *If (ν_i, ψ_i) is an eigenpair for \mathbf{N} , i.e $\mathbf{N}\psi_i = \nu_i\psi_i \Leftrightarrow (1 - \frac{1}{2}\nu_i, \mathbf{D}^{\frac{1}{2}}\psi_i)$ is an eigenpair for \mathbf{W}*

Proof.

$$\begin{aligned}\mathbf{W}\mathbf{D}^{\frac{1}{2}}\psi_i &= (\mathbf{I} - \frac{1}{2}\mathbf{D}^{\frac{1}{2}}\mathbf{N}\mathbf{D}^{-\frac{1}{2}})\mathbf{D}^{\frac{1}{2}}\psi_i \\ &= \mathbf{D}^{\frac{1}{2}}\psi_i - \frac{1}{2}\nu_i\mathbf{D}^{\frac{1}{2}}\psi_i\end{aligned}$$

□

Because of lemma 4.1 and lemma 3.5, we can obtain the following corollary
corollary: $0 \leq \lambda_i(\mathbf{W}) \leq 1$

Warning: \mathbf{W} is not symmetric. Thus, its eigenvector need not be orthogonal

5 Convergence of Lazy Random Walk

5.1 Finding an expression for \mathbf{p}_t

State transition from \mathbf{p}_t to \mathbf{p}_{t+1} in a lazy random is given by :

$$\mathbf{p}_{t+1} = \mathbf{W}\mathbf{p}_t$$

When $t = 0$:

$$\mathbf{p}_1 = \mathbf{W}\mathbf{p}_0$$

We know that

$$\mathbf{D}^{-\frac{1}{2}}\mathbf{p}_0 = \sum_{i=1}^n \alpha_i \psi_i \Leftrightarrow \mathbf{p}_0 = \sum_{i=1}^n \alpha_i \mathbf{D}^{\frac{1}{2}} \psi_i$$

Thus, we can express \mathbf{p}_1 as:

$$\mathbf{p}_1 = \mathbf{W}\mathbf{p}_0 = \sum_{i=1}^n \alpha_i (\mathbf{W}\mathbf{D}^{\frac{1}{2}} \psi_i) = \sum_{i=1}^n \alpha_i (1 - \frac{\nu_i}{2}) \mathbf{D}^{\frac{1}{2}} \psi_i$$

Iterating the process above, we obtain:

$$\mathbf{p}_t = \sum_{i=1}^n \alpha_i (1 - \frac{\nu_i}{2})^t \mathbf{D}^{\frac{1}{2}} \psi_i$$

Claim: If G is connected $\Leftrightarrow \nu_2 > 0$

Remark: the claim above implies the following:

$$\forall i \neq 1, 0 \leq 1 - \frac{\nu_i}{2} < 1$$

5.2 Given ϵ , finding step t such that \mathbf{p}_t is ϵ closed to the stationary distribution

At arbitrary vertex u , we have the following:

$$\begin{aligned} \mathbf{1}_u^\top \mathbf{p}_t - \mathbf{1}_u^\top \pi &= \mathbf{1}_u^\top \mathbf{p}_t - \frac{\mathbf{1}_u^\top \mathbf{D} \mathbf{1}}{\mathbf{1}^\top \mathbf{D} \mathbf{1}} \\ &= \sum_{i=1}^n \alpha_i (1 - \frac{\nu_i}{2})^t \mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i - \frac{\mathbf{1}_u^\top \mathbf{D} \mathbf{1}}{\mathbf{1}^\top \mathbf{D} \mathbf{1}} \end{aligned} \tag{3}$$

We know that:

$$\psi_1 = \frac{(\mathbf{D}^{\frac{1}{2}} \mathbf{1})}{\|\mathbf{D}^{\frac{1}{2}} \mathbf{1}\|} = \frac{(\mathbf{D}^{\frac{1}{2}} \mathbf{1})}{\sqrt{\mathbf{1}^\top \mathbf{D} \mathbf{1}}}$$

multiplied both side with $\mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}}$, we get

$$\begin{aligned}\mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_1 &= \mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \frac{(\mathbf{D}^{\frac{1}{2}} \mathbf{1})}{\|\mathbf{D}^{\frac{1}{2}} \mathbf{1}\|} \\ &= \frac{(\mathbf{1}_u^\top \mathbf{D} \mathbf{1})}{\sqrt{\mathbf{1}^\top \mathbf{D} \mathbf{1}}}\end{aligned}\tag{4}$$

We can express $\mathbf{D}^{-\frac{1}{2}} \mathbf{p}_0$ as following

$$\mathbf{D}^{-\frac{1}{2}} \mathbf{p}_0 = \sum_{i=1}^n \alpha_i \psi_i$$

multiply each side with ψ_1^\top

$$\psi_1^\top \mathbf{D}^{-\frac{1}{2}} \mathbf{p}_0 = \alpha_1$$

This gives us:

$$\alpha_1 = \frac{(\mathbf{1}^\top \mathbf{D}^{\frac{1}{2}}) \mathbf{D}^{-\frac{1}{2}} \mathbf{p}_0}{\sqrt{\mathbf{1}^\top \mathbf{D} \mathbf{1}}}$$

because \mathbf{p}_0 is a probability vector and sum up to 1, we have,

$$= \frac{1}{\sqrt{\mathbf{1}^\top \mathbf{D} \mathbf{1}}}\tag{5}$$

Now, we can further simplify (3) as:

$$\mathbf{1}_u^\top \mathbf{p}_t - \mathbf{1}_u^\top \pi = \alpha_1 (1 - \frac{\nu_1}{2})^\top \psi_1 + \sum_{i=2}^n \alpha_i (1 - \frac{\nu_i}{2})^\top \mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i - \frac{\mathbf{1}_u^\top \mathbf{D} \mathbf{1}}{\mathbf{1}^\top \mathbf{D} \mathbf{1}}$$

with equation(4) and (5) we obtained

$$= \sum_{i=2}^n \alpha_i (1 - \frac{\nu_i}{2})^\top \mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i$$

Combining the result above, we have the following

$$|\mathbf{1}_u^\top \mathbf{p}_t - \mathbf{1}_u^\top \pi| \leq \sum_{i=2}^n |\alpha_i \mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i| (1 - \frac{\nu_i}{2})^\top$$

because $\nu_2 \leq \nu_3 \leq \dots$, we have

$$\begin{aligned}&\leq (1 - \frac{\nu_2}{2})^\top \sum_{i=2}^n |\alpha_i \mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i| \\ &\leq (1 - \frac{\nu_2}{2})^\top \sqrt{\sum_{i=2}^n \alpha_i^2 \sum_{i=2}^n (\mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i)^2}\end{aligned}\tag{6}$$

Now let's try to simplify the term inside the square root, starting with $\sum_{i=2}^n \alpha_i^2$

$$\sum_{i=2}^n \alpha_i^2 \leq \|\mathbf{D}^{-\frac{1}{2}} \mathbf{p}_0\|_2^2$$

Assuming we start at vertex v

$$\begin{aligned} &\leq \mathbf{1}_v^\top \mathbf{D}^{-1} \mathbf{1}_v \\ &= \frac{1}{D(v)} \end{aligned} \tag{7}$$

Now let's simplify $\sum_{i=2}^n (\mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i)^2$

We shall start with finding an expression for $\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u$

Claim:

$$\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u = \sum_{i=1}^n (\mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i) \psi_i$$

Proof. Let's express $\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u$ with eigenvector and eigenvalue

$$\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u = \sum_{i=1}^n \beta_i \psi_i$$

multiplied both sides with ψ_j , we obtain

$$\psi_j^\top \mathbf{D}^{\frac{1}{2}} \mathbf{1}_u = \beta_j$$

take the transpose of left hand side

$$(\psi_j^\top \mathbf{D}^{\frac{1}{2}} \mathbf{1}_u)^\top = \mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_j^\top$$

$$\begin{aligned} \|\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u\|_2^2 &= (\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u)^\top (\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u) \\ &= \left(\sum_{i=1}^n \beta_i \psi_i \right)^\top \left(\sum_{j=1}^n \beta_j \psi_j \right) \\ &= \sum_{i=1}^n \beta_i^2 \\ &= \sum_{i=1}^n (\mathbf{1}_u^\top \mathbf{D}^{\frac{1}{2}} \psi_i)^2 \end{aligned} \tag{8}$$

□

With the results above, (6) can be further simplified as following:

$$\begin{aligned} |\mathbf{1}_u^\top \mathbf{P}_t - \mathbf{1}_u^\top \pi| &\leq (1 - \nu_2)^\top \sqrt{\|\mathbf{D}^{-\frac{1}{2}} \mathbf{p}_0\|_2^2 \|\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u\|_2^2} \\ &= \|\mathbf{D}^{\frac{1}{2}} \mathbf{1}_u\| \cdot \|\mathbf{D}^{-\frac{1}{2}} \mathbf{p}_0\| (1 - \frac{\nu_2}{2})^\top \end{aligned}$$

with (7) and assuming the walk starts at v , then

$$= \sqrt{\frac{D_u}{D_v}} (1 - \frac{\nu_2}{2})^\top \quad (9)$$

Notice that $(\mathbf{1} - \mathbf{x}) \leq \mathbf{e}^{-\mathbf{x}}$

$$\leq \sqrt{\frac{D_u}{D_v}} e^{-\nu_2 t/2}$$

Theorem 5.1. *Given an undirected unweighted graph G and ϵ , it suffices for the lazy random walk to take t steps to get ϵ closed to stationary distribution.*

In other word, if

$$t \geq \frac{2}{\nu_2} \log\left(\frac{n}{\epsilon}\right)$$

then

$$|\mathbf{1}_u^\top \mathbf{P}_t - \mathbf{1}_u^\top \pi| \leq \epsilon$$

Interpretation: On a graph with n vertex, if the lazy random walk want to be $\frac{\epsilon}{2}$ closed to the stationary distribution, then it would satisfy the following relation.

$$\begin{aligned} T\left(\frac{\epsilon}{2}\right) &= \frac{2}{\nu_2} \log\left(\frac{2n}{\epsilon}\right) \\ &= \Theta\left(\frac{1}{\nu_2}\right) + T(\epsilon) \end{aligned}$$

5.3 Application of the theorem on some examples

K_n : complete graph with n vertices

Lazy random walk mixes in $\Omega(\log n)$ steps

$$\begin{aligned} L_{K_n} &= n\mathbf{I} - \mathbf{1}\mathbf{1}^\top \\ \lambda_i(L) &= \begin{cases} 0 & i = 1 \\ n - 1 & O.W \end{cases} \\ \nu_2(N_{K_n}) &= 1 \end{aligned}$$

the theorem gives that the lazy random walk would mix in $O(\log n)$. In this case, the bound given by the theorem is tight

R_n : n -ring graph

The second eigenvalue of n -ring graph is given by:

$$\nu_2(N_{K_n}) = \theta\left(\frac{1}{n^2}\right)$$

thus, the theorem gives that the lazy random walk will mixed in $O(n^2 \log n)$
the lazy random walk on the n-ring graph can be defined as following:

$$x_{i+1} = \begin{cases} x_i & \text{with probability } \frac{1}{2} \\ x_i + 1 & \text{with probability } \frac{1}{4} \\ x_i - 1 & \text{with probability } \frac{1}{4} \end{cases}$$

$$E[x_{i+1}|x_i] = x_i$$

$$E[x_{i+1}^2|x_i] = x_i^2 + \frac{1}{2}$$

thus, the Lazy Random Walk is mixed in $\Omega(n^2)$ steps. In this case, the bound given by the theorem is off up to $\log n$