Convergence of Random Walks in Undirected Graphs

1 Introduction

Let $G = (V, E)$ be a connected graph with $|V| = n$. We have seen various properties of random walks on connected graphs:

1. The stationary distribution depends only on degrees.
2. Cover time is $O(mn)$.
3. $h_{i,j} \leq 2m$ (hitting time).
4. Connection to electrical networks.

We are interested in the convergence of the random walk to its stationary distribution. To make this formal, we define the following.

Definition 1. Given two distributions π and $\tilde{\pi}$ on V, their *total variational distance* is

$$\|\pi - \tilde{\pi}\|_{TV} = \frac{1}{2} \sum_{v} |\pi(v) - \tilde{\pi}(v)|.$$

We say that a Markov chain converges to its stationary distribution if $\|\pi^t - \pi\|_{TV} \to 0$ as $t \to \infty$, where π is the stationary distribution.

We will be interested in the rate of convergence: starting from some arbitrary distribution π^0, how long does it take to get ϵ-close to π?

Note: A random walk may be periodic, in which case it does not converge. We will work with the lazy walk instead.

We will focus on random walks in undirected graphs using spectral analysis, relating spectral properties to combinatorial properties of the graph.

2 Setting up the Matrices

It is useful to switch from row vectors to column vectors. So far we have been working with distributions over states as row vectors. Since we will work with eigenvalues and eigenvectors, it is useful to switch to plot p^t as column vectors. Recall P was the probability transition matrix of the chain. In order to work with column vectors, we use P^T. Thus $p^t = P^T p^{t-1}$ where p is the distribution of the chain at time step t.

2.1 Graph Setup

Consider random walk on undirected graph $G = (V, E)$ with $|V| = n$. Let A be the adjacency matrix of G. Let D be the $n \times n$ diagonal matrix with $D_{ii} = \deg(i)$. Then D^{-1} is the matrix with $(D^{-1})_{ii} = \frac{1}{\deg(i)}$.

Recall the transition matrix P was defined as $P_{ij} = \frac{1}{\deg(i)}$ if $(i, j) \in E$. Hence $P = D^{-1}A$.

We can see that $P^T = AD^{-1}$.

We call $W = AD^{-1}$ the *walk matrix*.

For the lazy random walk, the walk matrix is $W = \frac{1}{2}I + \frac{1}{2}AD^{-1}$.

Suppose G is d-regular. Then $W = \frac{1}{d}A$ is symmetric. Symmetric matrices have substantial structure and there is a beautiful and powerful spectral theory for them with many applications.

3 Review of Linear Algebra

3.1 Eigenvalues and Eigenvectors

For an $n \times n$ matrix $A \in \mathbb{R}^{n\times n}$, a vector $v \in \mathbb{R}^n$ is an eigenvector of A if there exists $\lambda \in \mathbb{R}$ such that $Av = \lambda v$, where λ is an eigenvalue.

Eigenvalues are solutions to the polynomial $\det(A - \lambda I) = 0$.

In general, this polynomial may not have real roots, and hence A may not have real eigenvalues and eigenvectors. However, if A is viewed as a complex matrix, then we have complex eigenvalues and eigenvectors, since every univariate polynomial can be factorized.

Symmetric matrices are, however, special and have substantial structure. This is captured by the Spectral Theorem.

Theorem 2 (Spectral Theorem)**.** Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then A has n real eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding eigenvectors $v_1, \ldots, v_n$ which form an orthonormal basis of $\mathbb{R}^n$. Moreover, $A = VDV^T$ where D is a diagonal matrix with $D_{ii} = \lambda_i$ and $V = [v_1 \cdots v_n]$.

3.2 Rayleigh Quotient and Extremal Eigenvalues

A useful characterization of the eigenvectors is obtained via the Rayleigh quotient. Given A and $x \in \mathbb{R}^n$, consider $x^T Ax$, which is the quadratic form induced by A.

Note that

$$x^T Ax = \sum_{i,j} A_{ij} x_i x_j.$$

When A is a symmetric matrix, we get

$$x^T Ax = 2 \sum_{i<j} A_{ij} x_i x_j.$$

Consider the problem of

$$\max_{\|x\|=1} x^T Ax,$$

which is the same as

$$\max \min \frac{x^T A x}{\|x\|^2}.$$

It turns out that if A is symmetric, then

$$\lambda_1 = \max_{\|x\|=1} x^T A x, \quad \lambda_n = \min_{\|x\|=1} x^T A x.$$

In fact, one can characterize all eigenvalues via:

$$\lambda_k = \max_{U_k} \min_{\substack{x \in U_k \\ \|x\|=1}} x^T A x$$

where U_k is a k-dimensional subspace of $\mathbb{R}^n$.

3.3 Proof of Extremal Characterization

One can derive this characterization from the spectral theorem or directly. Suppose $v_1, \ldots, v_n$ are orthonormal unit vectors that are eigenvectors of A and let $\lambda_1 \geq \cdots \geq \lambda_n$.

Let $v \in \mathbb{R}^n$ and $\|v\| = 1$ be any unit vector. Then

$$v = \sum_{i=1}^{n} c_i v_i \text{ where } \sum_{i=1}^{n} c_i^2 = 1.$$

Consider

$$\max_{\|v\|=1} \frac{x^T A x}{\|x\|^2}, \quad \min_{\|v\|=1} \frac{x^T A x}{\|x\|^2}.$$

We have

$$x^T A x = v^T A v = v^T \left(\sum_i \lambda_i c_i v_i v_i^T \right) v = \sum_i \lambda_i c_i^2,$$

where $v = \sum_i c_i v_i$ with $\sum_i c_i^2 = 1$.

Hence

$$\max_{\|v\|=1} v^T A v = \lambda_1, \quad \min_{\|v\|=1} v^T A v = \lambda_n.$$

Similarly, $\min_{\|v\|=1} v^T A v = \lambda_n$, etc.

3.4 Positive Semidefinite Matrices

Definition 3. A real symmetric matrix A is positive semidefinite (PSD) if $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$.

Theorem 4. Let A be a real symmetric matrix of size $n \times n$. A is PSD if and only if one of the following conditions holds:

1. $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$.
2. All eigenvalues of A are non-negative: $\lambda_1, \ldots, \lambda_n \geq 0$.

3. $A = W^T W$ for some $W \in \mathbb{R}^{n\times n}$.

Proof of (3) $\Rightarrow$ (1):

$$v^T A v = v^T W^T W v = (Wv)^T (Wv) = \|Wv\|^2 \geq 0.$$

Proof of (1) $\Rightarrow$ (2): Recall for symmetric A,

$$\lambda_{\min} = \min_{\|v\|=1} v^T A v.$$

Here $\lambda_{\min} \geq 0$.
A is symmetric, so $A = VDV^T$ where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_i \geq 0$.
Then $D = \sqrt{D}\sqrt{D}$ where $\sqrt{D} = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$.
Hence $A = V\sqrt{D}\sqrt{D}V^T = W^T W$ where $W = \sqrt{D}V^T$.

4 Convergence Analysis for Regular Graphs

For d-regular undirected graphs, $W = \frac{1}{d}A$ is symmetric. Note that W is doubly stochastic. By spectral theorem, all eigenvalues are real. Let $\alpha_1, \ldots, \alpha_n$ be the eigenvalues of W.

Claim 5. $\alpha_1 = 1$ and $\alpha_n \geq -1$.

Proof: This is an exercise.

Claim 6. Show that $\alpha_n = -1$ if G is bipartite. Otherwise, $\alpha_n > -1$.

Claim 7. Consider $W = \frac{1}{2}I + \frac{1}{2}AD^{-1}$. The eigenvalues are $1, 1-\lambda_2, 1-\lambda_2, \ldots$ where

$$0 \leq \lambda_2, \ldots, \lambda_n \leq 1.$$

4.1 Spectral Decomposition and Convergence

By spectral theorem, W can be written as

$$W = \sum_{i=1}^{n} \alpha_i v_i v_i^T$$

where $v_1, \ldots, v_n$ are the eigenvectors of W normalized to be unit vectors, and $\cdot$ denotes outer product. We have $v_1, \ldots, v_n$ are orthonormal.

We want to know $W^t p^0$ where p^0 is the starting distribution.
Since $v_1, \ldots, v_n$ are orthonormal, we can write $p^0 = \sum_i c_i v_i$ where $c_i = p^0 \cdot v_i$.
Then

$$W^t p^0 = \sum_i c_i \alpha_i^t v_i.$$

Recall $\alpha_1 = 1$ and $|\alpha_i| \leq 1$ for all i. Since we normalize to unit vectors, we have $c_i \in \mathbb{R}$.
The spectral gap is defined as

$$\beta = \min_{i\geq 2}(1 - \alpha_i).$$

Note that $0 \leq \beta \leq 1$.

Suppose $\beta > 0$, $\alpha_2 < 1$, and $\alpha_n > -1$. Then $\alpha_1 = 1 > \alpha_i$ for $i \geq 2$.

We have

$$W^t p^0 = c_1 \alpha_1^t v_1 + \sum_{i \geq 2} c_i \alpha_i^t v_i = c_1 v_1 + \sum_{i \geq 2} c_i \alpha_i^t v_i.$$

Since $\alpha_i < 1 - \beta$ for $i \geq 2$, as $t \to \infty$, all $\alpha_i^t \to 0$.

Thus $W^t p^0 \to c_1 v_1$.

Since $\sum_i c_i v_i = p^0$ and $v_1, \ldots, v_n$ are orthonormal, we have $c_1 = p^0 \cdot v_1$.

Recall that v_1 is the first eigenvector. Since $\pi = \nVdash/n$ (uniform distribution) and the stationary distribution for a d-regular graph, and $W v_1 = v_1$, we know that π is proportional to $\nVdash$.

Actually, for d-regular graphs, $\pi(i) = 1/n$ for all i, and $\nVdash = (1, 1, \ldots, 1)^T$, so $v_1 = \nVdash/\sqrt{n}$ (normalized).

Hence $c_1 v_1 = (p^0 \cdot \nVdash/\sqrt{n}) \cdot \nVdash/\sqrt{n} = \frac{1}{n}\nVdash = \pi$ (the uniform/stationary distribution).

4.2 Rate of Convergence

Claim 8. Mixing time is $O\left(\frac{1}{\beta} \log \frac{1}{\epsilon}\right)$.

Proof: Why is $\sum_i c_i \alpha_i^t v_i$ small?

Hence

$$\left\| W^t p^0 - \pi \right\|_{TV} = \left\| \sum_{i \geq 2} c_i \alpha_i^t v_i \right\|.$$

By Cauchy-Schwarz,

$$\left\| \sum_{i \geq 2} c_i \alpha_i^t v_i \right\| \leq \sqrt{\sum_i c_i^2} \sqrt{\sum_i |\alpha_i|^{2t}}.$$

Full expansion:

$$\|W^t p^0 - \pi\|_{TV} \leq \sqrt{\sum_{i \geq 2} c_i^2} \cdot \sqrt{\sum_{i \geq 2} |\alpha_i|^{2t}}.$$

Since v_i are orthonormal, $\sum_i c_i^2 = \|p^0\|^2 = 1$.

Hence $\sqrt{\sum_{i \geq 2} c_i^2} \leq 1$.

Also, $|\alpha_i|^{2t} \leq (1 - \beta)^{2t}$ for $i \geq 2$.

Therefore,

$$\|W^t p^0 - \pi\|_{TV} \leq \sqrt{n}(1 - \beta)^t.$$

Want $\|W^t p^0 - \pi\|_{TV} \leq \epsilon$. Need $\sqrt{n}(1 - \beta)^t \leq \epsilon$.

Thus $(1 - \beta)^t \leq \epsilon/\sqrt{n}$.

Taking logarithms: $t \log(1 - \beta) \leq \log(\epsilon/\sqrt{n})$.

Since $\log(1 - \beta) \approx -\beta$ for small β, we get $t \geq \frac{1}{\beta} \log \frac{\sqrt{n}}{\epsilon}$.

Thus $t = O\left(\frac{1}{\beta} \log \frac{1}{\epsilon}\right)$ suffices (absorbing the $\log n$ factor).

4.3 Lazy Random Walk

For the lazy random walk, $W = \frac{1}{2}I + \frac{1}{2}AD^{-1}$.

Recall that $\alpha_1 = 1$, $\alpha_n \leq 1$ with $\alpha_n < -1$ if bipartite, else $\alpha_n > -1$.

And $\alpha_i' = 1 - \frac{1}{2}\lambda_i$ where λ_i are eigenvalues of $\frac{1}{d}A$ (or AD^{-1} for regular graphs).

Thus $\alpha_1' = 1 - \frac{1}{2} = 1/2$ and $\alpha_2' = 1 - \frac{1}{2}(1-\beta) = 1/2 + \beta/2$.

Wait, let me reconsider. For the lazy walk with $W = \frac{1}{2}I + \frac{1}{2}W_0$ where $W_0 = AD^{-1}$ has eigenvalues $\alpha_1 = 1, \alpha_2, \ldots, \alpha_n$, the eigenvalues of W are

$$\alpha_i' = \frac{1}{2} + \frac{1}{2}\alpha_i.$$

So $\alpha_1' = 1$, and the second eigenvalue is $\alpha_2' = \frac{1}{2}(1 + \alpha_2)$.

Spectral gap: $\beta' = 1 - \alpha_2' = \frac{1}{2}(1 - \alpha_2)$.

5 Example: Cycle Graph

Example: $G = C_n$, the n-cycle.

What are eigenvalues of AD^{-1}?

Can show that they are

$$\lambda_k = \cos\left(\frac{2\pi k}{n}\right), \quad k = 0, 1, \ldots, n-1.$$

If n is even, $\alpha_n = -1$. The spectral gap is $\beta = 0$. Hence need to use lazy walk.

What about α_1, α_2?

We have

$$\lambda_1 = \cos\left(\frac{2\pi}{n}\right) = 1 - \frac{1}{2}\left(\frac{2\pi}{n}\right)^2 + O(n^{-4}) = 1 - \frac{2\pi^2}{n^2} + O(n^{-4}).$$

For large n,

$$\beta = 1 - \lambda_1 \approx \frac{2\pi^2}{n^2}.$$

Convergence time is $O(1/\beta) = O(n^2)$.

Not surprising!

6 General Non-Regular Graphs

For graphs that are not necessarily regular, we use the lazy walk.

$W = \frac{1}{2}I + \frac{1}{2}AD^{-1}$ is the walk matrix. W is not symmetric, so we cannot use the spectral theorem directly.

Consider the normalized adjacency matrix

$$\tilde{A} = D^{-1/2}AD^{-1/2}.$$

Note that $\tilde{A}_{ij} = \frac{A_{ij}}{\sqrt{\deg(i)\deg(j)}}$, which is symmetric.

We can write

$$W = D^{-1/2}\left(\frac{1}{2}I + \frac{1}{2}D^{-1/2}AD^{-1/2}\right)D^{1/2} = D^{-1/2}\left(\frac{1}{2}I + \frac{1}{2}\tilde{A}\right)D^{1/2}.$$

So W is similar to the symmetric matrix $\frac{1}{2}I + \frac{1}{2}\tilde{A}$.

W is similar to a symmetric matrix.

Definition 9. If X and Y are $n \times n$ matrices, we say X is similar to Y if there exists a non-singular matrix B such that $X = BYB^{-1}$.

The action of X can be understood via action of Y.

Claim 10. Eigenvalues of X and Y are the same. Hence same spectrum. Eigenvectors may be different.

Proof: Suppose $Yv = \lambda v$. Let $u = B^{-T}v$ (i.e., $u = (B^T)^{-1}v$). Then

$$Xu = BYB^{-1}u = BY(B^{-1}u) = BY\frac{1}{B^{-T}v}Bv = BYv = B(\lambda v) = \lambda(Bv) = \lambda(B(B^T)^{-1}u) = \lambda u.$$

Wait, let me redo this. If $Yv = \lambda v$ and $u = B^T v$, then

$$Xu = BYB^{-1}u = BYB^{-1}(B^T)^{-1}v = BYv = \lambda Bv = \lambda(B^T)^{-1}u = \lambda u.$$

Hmm, this doesn't quite work. Let me reconsider.

If $Yv = \lambda v$, let $u = B^{-1}v$. Then $v = Bu$ and

$$X(B^{-T}v) = BYB^{-1}(B^{-T}v) = BYB^{-1}B^{-T}v.$$

Actually, let me use the standard approach. If $Yv = \lambda v$, define $u = Bv$. Then

$$Xu = BYB^{-1}u = BYB^{-1}Bv = BYv = \lambda Bv = \lambda u.$$

So u is an eigenvector of X with eigenvalue λ. Thus X and Y have the same eigenvalues.

Corollary 11. If X is similar to a symmetric matrix, then all eigenvalues of X are real. Eigenvectors span $\mathbb{R}^n$ even though they may not be orthonormal.

Now back to W.

We have $W = D^{-1/2}\left(\frac{1}{2}I + \frac{1}{2}\tilde{A}\right)D^{1/2}$ where

$$\tilde{A} = D^{-1/2}AD^{-1/2}$$

is the normalized adjacency matrix.

Let the eigenvalues of $\tilde{A}$ be $\lambda_1, \lambda_2, \ldots, \lambda_n$ with $\lambda_1 = 1$ and $|\lambda_i| \leq 1$ for $i \geq 2$.

The eigenvalues of W are

$$\alpha_i = \frac{1}{2} + \frac{1}{2}\lambda_i, \quad i = 1, \ldots, n.$$

If $v_1, v_2, \ldots, v_n$ are eigenvectors of $\tilde{A}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then eigenvectors of W are

$$u_i = D^{-1/2} v_i, \quad i = 1, \ldots, n.$$

Note that $v_1, \ldots, v_n$ are orthonormal (from spectral theorem for $\tilde{A}$). Thus

$$D^{-1/2} v_i, \quad i = 1, \ldots, n$$

are independent and span $\mathbb{R}^n$.

Now we want to understand $W^t p^0$ where p^0 is the starting distribution.

We can write $p^0 = \sum_i c_i u_i$ where $u_i = D^{-1/2} v_i$ are eigenvectors of W.

Since the eigenvectors of W span $\mathbb{R}^n$, therefore

$$W^t p^0 = \sum_{i=1}^{n} c_i \alpha_i^t u_i = \sum_{i=1}^{n} c_i \alpha_i^t D^{-1/2} v_i.$$

Recall $D^{-1/2} v_i$ is eigenvector of W with eigenvalue α_i.

We have

$$W^t D^{-1/2} v_i = \alpha_i^t D^{-1/2} v_i.$$

Thus

$$W^t p^0 = \sum_{i=1}^{n} c_i \alpha_i^t D^{-1/2} v_i.$$

Recall $\alpha_1 = 1$, $\alpha_2, \ldots, \alpha_n$ satisfy $|\alpha_i| < 1$ for $i \geq 2$ (since λ_i are eigenvalues of a normalized adjacency matrix and $|\lambda_i| \leq 1$ with $\lambda_1 = 1$).

If p^0 is a unit vector (a point mass), then

$$W^t p^0 = c_1 D^{-1/2} v_1 + \sum_{i=2}^{n} c_i \alpha_i^t D^{-1/2} v_i.$$

Since $\alpha_i < 1$ for $i \geq 2$, as $t \to \infty$,

$$W^t p^0 \to c_1 D^{-1/2} v_1.$$

Converges to $c_1 D^{-1/2} v_1 =$ stationary distribution.

What is $c_1 D^{-1/2} v_1$?

Recall $p^0 = \sum_{i=1}^{n} c_i D^{-1/2} v_i$. Thus

$$c_i D^{-1/2} v_i = ?$$

Wait, we need to be more careful. We have $p^0 = \sum_i c_i u_i$ where u_i are eigenvectors of W.

Now, $u_1 = D^{-1/2} v_1$ where v_1 is the first eigenvector of $\tilde{A}$.

For the normalized adjacency matrix $\tilde{A} = D^{-1/2} A D^{-1/2}$, the first eigenvector (corresponding to eigenvalue 1) is proportional to $D^{-1/2} \mathbb{1} / \| D^{-1/2} \mathbb{1} \|$.

Actually, for a connected graph, the first eigenvector of $\tilde{A}$ is $v_1 = D^{1/2} \mathbb{1} / \| D^{1/2} \mathbb{1} \|$.

Let me reconsider. We have $\tilde{A} = D^{-1/2} A D^{-1/2}$.

Note that $\tilde{A}(D^{1/2}\mathbb{1}) = D^{-1/2}A\mathbb{1} = D^{-1/2}d$ where d is the degree vector. So $\tilde{A}(D^{1/2}\mathbb{1}) = D^{1/2}\mathbb{1}$ (since $d = D\mathbb{1}$ and thus $D^{-1/2}D\mathbb{1} = D^{1/2}\mathbb{1}$).

Thus $v_1 = D^{1/2}\mathbb{1}/\|D^{1/2}\mathbb{1}\|$ is the first eigenvector of $\tilde{A}$ with eigenvalue 1.

Then $u_1 = D^{-1/2}v_1 = D^{-1/2} \cdot D^{1/2}\mathbb{1}/\|D^{1/2}\mathbb{1}\| = \mathbb{1}/\|D^{1/2}\mathbb{1}\|$.

Hmm, this is proportional to $\mathbb{1}$, but normalized differently.

Actually, the stationary distribution is $\pi = \frac{d}{\mathrm{vol}(G)}$ where $\mathrm{vol}(G) = \sum_i d_i = 2m$.

So $\pi = \frac{D\mathbb{1}}{2m}$ (where $D\mathbb{1}$ is the degree vector).

Now, $c_1 u_1 = c_1 D^{-1/2} v_1 = c_1 D^{-1/2} \cdot D^{1/2}\mathbb{1}/\|D^{1/2}\mathbb{1}\| = c_1\mathbb{1}/\|D^{1/2}\mathbb{1}\|$.

Recall $p^0 = \sum_i c_i u_i$. In particular, $c_1 = p^0 \cdot u_1/\|u_1\|^2 = p^0 \cdot u_1$ (if u_i are normalized).

Hmm, let me think about this more carefully.

Actually, let's use the following: u_i are eigenvectors of W, and they may or may not be orthonormal. But v_i are orthonormal eigenvectors of $\tilde{A}$.

If $p^0 = \sum_i c_i u_i$ and $v_i = D^{1/2}u_i$, then... hmm, this is getting messy.

Let me try a different approach. The claim is that $u_1^T = \frac{d^T}{\|d\|}$ where d is the degree vector. More precisely, the stationary distribution is $\pi^T = \frac{d^T}{\mathrm{vol}(G)}$.

Therefore, $c_1 D^{-1/2} v_1 = \pi$.

So $W^t p^0 \to \pi$ as $t \to \infty$.

Thus the lazy random walk converges to the stationary distribution whatever the starting distribution is. Note that the lazy walk only assumes connectivity (i.e., does not require regularity).

Since $\beta \geq \lambda_2$ which is true if G is connected, the mixing time analysis applies.

6.1 Mixing Time

Let $\pi = \frac{d}{\mathrm{vol}(G)}$ be the stationary distribution. We have $W^t p^0 = \pi +$ (lower order terms).

It turns out that $\|W^t p^0 - \pi\|_{TV} \leq \|p^0 - \pi\|_{TV}$.

More precisely, $\|W^t p^0 - \pi\|_{TV} = \left\|\sum_{i\geq 2} c_i \alpha_i^t D^{-1/2} v_i\right\|$.

By Cauchy-Schwarz, $\left\|\sum_{i\geq 2} c_i \alpha_i^t D^{-1/2} v_i\right\| \leq \sqrt{\sum_i c_i^2} \cdot \sqrt{\sum_i |\alpha_i|^{2t}} \leq \|p^0\|\beta^t$ where $\beta = \min_i(1 - |\alpha_i|)$ for $i \geq 2$.

Actually, more carefully: $\|W^t p^0 - \pi\|_{TV} \leq C \cdot \beta^t$ where C depends on the starting distribution and graph structure.

In particular, if p^0 is a point mass at vertex v, then $\|W^t p^0 - \pi\|_{TV} \leq \frac{1}{\sqrt{d_{\min}}}\beta^t$ where $d_{\min}$ is the minimum degree.

Also, $\|p^0\| \leq \sqrt{\max_i d_i/d_{\min}}$ (roughly).

Therefore, to get $\|W^t p^0 - \pi\|_{TV} \leq \epsilon$, we need $t \geq \frac{1}{\beta}\log\frac{1}{\epsilon}$.

7 Conclusion

To summarize:

- For d-regular graphs, the walk matrix $W = \frac{1}{d}A$ is symmetric, and we can use spectral analysis directly.

- For general graphs, the lazy walk $W = \frac{1}{2}I + \frac{1}{2}AD^{-1}$ is similar to a symmetric matrix, so spectral analysis still applies.
- The convergence rate depends on the spectral gap $\beta = 1 - \alpha_2$, where α_2 is the second-largest eigenvalue.
- The mixing time is $O\left(\frac{1}{\beta}\log\frac{1}{\epsilon}\right)$.

8 Next Lecture

Which graphs have spectral gap that is a constant? (i.e., $\beta = \Omega(1)$). This will ensure that random walk converges in $O(\log n)$ steps.