Lecture 10. Normalized Graph Laplacian and Cheeger's Inequality

Instructor: Yuan Yao, Peking University   *Scribe: Shaokun Li*

Introduction

In the last class, we give an introduction of Perron-Frobenius Theorem and unnormalized graph Laplacian.In this course, we will continue the discussion of last class. We will mainly discuss the normalized graph Laplacian and introduce the famous Cheeger's Inequality. We will also discuss the normalized graph Laplacian in directed graph in this class.

1 Normalized graph Laplacian

Definition (Normalized Graph Laplacian)**.**

$$
\mathcal{L}_{ij} = \begin{cases} 1 & i = j, \\ -\dfrac{1}{\sqrt{d_i d_j}} & i \sim j, \\ 0 & otherwise. \end{cases}
$$

In fact $\mathcal{L} = D^{-1/2}(D - A)D^{-1/2} = D^{-1/2}LD^{-1/2} = I - D^{-1/2}(D - A)D^{-1/2}$. From this one can see the relations between eigenvectors of normalized $\mathcal{L}$ and unnormalized L. For eigenvectors $\mathcal{L}v = \lambda v$, we have

$$
(I - D^{-1/2}LD^{-1/2})v = \lambda v \Leftrightarrow Lu = \lambda Du, \quad u = D^{-1/2}v,
$$

whence eigenvectors of $\mathcal{L}$, v after rescaling by $D^{-1/2}v$, become generalized eigenvectors of L.

We can also use the Rayleigh Quotient to calculate the eigenvalues of $\mathcal{L}$.

$$
\begin{aligned}
\frac{v^T \mathcal{L} v}{v^T v} &= \frac{v^T D^{-\frac{1}{2}}(D - A)D^{-\frac{1}{2}} v}{v^v} \\
&= \frac{u^T L u}{u^T D u} \\
&= \frac{\sum\limits_{i \sim j} (u_i - u_j)^2}{\sum\limits_{j} u_j{}^2 d_j}.
\end{aligned}
$$

Similarly we get the relations between eigenvalue and the connected components of the graph.

$$
\#\{\lambda_i(\mathcal{L}) = 0\} = \#\{connected\ components\ of\ G\}.
$$

Next we show that eigenvectors of $\mathcal{L}$ are related to random walks on graphs. This will show you why we choose this matrix to analysis the graph.

We can construct a random walk on G whose transition matrix is defined by

$$P_{ij} \sim \frac{A_{ij}}{\sum\limits_j A_{ij}} = \frac{1}{d_i}.$$

By easy calculation, we see the result below.

$$P = D^{-1}A = D^{-1/2}(I - \mathcal{L})D^{1/2}.$$

Hence P is similar to $I - \mathcal{L}$. So their eigenvalues satisfy $\lambda_i(P) = 1 - \lambda_i(\mathcal{L})$. Consider the right eigenvector ϕ and left eigenvector ψ of P.

$$u^T P = \lambda u,$$
$$Pv = \lambda v.$$

Due to the similarity between P and $\mathcal{L}$,

$$u^T P = \lambda u^T \Leftrightarrow u^T D^{-1/2}(I - \mathcal{L})D^{1/2} = \lambda u^T.$$

Let $\bar{u} = D^{-1/2}u$, we will get:

$$\bar{u}^T(I - \mathcal{L}) = \lambda \bar{u}^T$$
$$\Leftrightarrow \mathcal{L}\bar{u} = (1 - \lambda)\bar{u}.$$

You can see $\bar{u}$ is the eigenvector of $\mathcal{L}$, and we can get left eigenvectors of P from $\bar{u}$ by multiply it with $D^{1/2}$ on the left side. Similarly for the right eigenvectors $v = D^{-1/2}\bar{u}$.

If we choose $u_0 = \pi_i \sim \frac{d_i}{\sum d_i}$, then:

$$\bar{u}_0(i) \sim \sqrt{d_i},$$
$$\bar{u}_k^T \bar{u}_l = \delta_{kl},$$
$$u_k^T D v_l = \delta_{kl},$$
$$\pi_i P_{ij} = \pi_j P_{ji} \sim A_{ij} = A_{ji},$$

where the last identity says the Markov chain is time-reversible.

All the conclusions above show that the normalized graph Laplacian $\mathcal{L}$ keeps some connectivity measure of unnormalized graph Laplacian L. Furthermore, $\mathcal{L}$ is more related with random walks on graph, through which eigenvectors of P are easy to check and calculate. That's why we choose this matrix to analysis the graph.

2 A Glimpse of Laplacians for Directed Graphs

In this section, we will talk a little about the Laplacians for Directed Graphs. We will focus on establishing the Laplacians.

For undirected graph, we can define the normalized Laplacian $\mathcal{L} = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$. Suppose v is an eigenvector of $\mathcal{L}$, then we have

$$(I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}})v = \lambda v$$

Here, we define D as

$$d_i = \sum_{j=1}^{n} A_{ij}$$

$$D = diag(d_i)$$

But in directed graph, we have some problems in this definition. The first problem is that the in degree and out degree are not the same in directed graph. So we have that

$$D_{in} = diag(\sum_i^n A_{ij})$$

$$D_{out} = diag(\sum_j^n A_{ij})$$

We define the transition probability matrix as

$$P = D_{out}^{-1} A$$

If the graph is primitive(indeed we do not need such a strong assumption, the graph is strong connected is enough), then we know that there exist a stationary distribution ϕ. According to Perron-Frobenius theorem, we have $\phi \geq 0, \phi^T P = 1\phi^T$. Define $\Phi = diag(\phi(i))$, Φ will help us to establish the Laplacians for directed graphs.

Now we give a definition on those balanced weights.

Definition (circulation).
$$F: \quad E \quad \rightarrow \quad \mathbb{R}_{\geq 0}$$

If F satisfies
$$\sum_{u,u \rightarrow v} F(u,v) = \sum_{w,v \rightarrow w} F(v,w), \forall v,$$

then F is called a circulation.

Note. A circulation is a flow with no source or sink.

Example. For a directed graph, $F_\phi(u,v) = \phi(u)P(u,v)$ is a circulation, for

$$\sum_{u,u \rightarrow v} F_\phi(u,v) = \phi(v) = \sum_{w,v \rightarrow w} F_\phi(v,w).$$

In the end of this section, we establish the Laplacian in directed graph.

In undirected graph, we have

$$
\begin{aligned}
\mathcal{L} &= I - D^{-1/2}AD^{-1/2} \\
&= I - D^{1/2}PD^{-1/2} \\
&= I - \Phi^{1/2}P\Phi^{-1/2}
\end{aligned}
$$

We generalize it to the directed graph. We will find the most important problem is that P is not symmetric in directed graph. So we use this following definition to promise that the normalized Laplacian is symmetric.

Definition (Laplacian for Directed Graph)**.**

$$\mathcal{L} = I - \frac{1}{2}(\Phi^{1/2}P\Phi^{-1/2} + \Phi^{-1/2}P^T\Phi^{1/2}).$$

This is the normalized Laplacian in directed graph. We can see easily that the Laplacian is symmetric.

3   Cheeger Inequality

3.1   Cheeger constant

Now we are ready to introduce the Cheeger's inequality with normalized graph Laplacian.

Let G be a graph, $G = (V, E)$ and S is a subset of V whose complement is $\bar{S} = V - S$. We define $Vol(S)$, ∂S and $NCUT(S)$ as below.

$$Vol(S) = \sum_{i \in S} d_i.$$

$$\partial S = \sum_{i \in S, j \in \bar{S}} A_{ij}.$$

$$NCUT(S) = \frac{\partial S}{\min(Vol(S), Vol(\bar{S}))}.$$

$NCUT(S)$ is called normalized-cut. We define the Cheeger constant

Definition (Cheeger Constant)**.**

$$h_G = \min_S NCUT(S).$$

Finding minimal normalized graph cut is NP-hard. It is often defined that

$$\text{Cheeger ratio (expander): } h_S := \frac{\partial(S)}{Vol(S)}$$

and

$$\text{Cheeger constant: } h_G := \min_S \max\{h_S, h_{\bar{S}}\}.$$

3.2   Cheeger Inequality

Cheeger Inequality says the second smallest eigenvalue provides both upper and lower bounds on the minimal normalized graph cut. Its proof gives us a constructive polynomial algorithm to achieve such bounds.

Theorem 3.1 (Cheeger Inequality)**.** For every undirected graph G,

$$\frac{h_G^2}{2} \le \lambda_1(\mathcal{L}) \le 2h_G.$$

Proof. (1) Upper bound:

Assume the following function f realizes the optimal normalized graph cut,

$$f(i) = \begin{cases} \frac{1}{Vol(S)} & i \in S, \\ \frac{-1}{Vol(\bar{S})} & i \in \bar{S}, \end{cases}$$

By using the Rayleigh Quotient, we get

$$\lambda_1 = \inf_{g \perp D^{1/2}e} \frac{g^T \mathcal{L} g}{g^T g}$$
$$\leq \frac{\sum_{i \sim j}(f_i - f_j)^2}{\sum f_i^2 d_i}$$
$$= \frac{(\frac{1}{Vol(S)} + \frac{1}{Vol(\bar{S})})^2 \partial S}{Vol(S)\frac{1}{Vol(S)^2} + Vol(\bar{S})\frac{1}{Vol(\bar{S})^2}} \qquad \text{span}$$
$$= (\frac{1}{Vol(S)} + \frac{1}{Vol(\bar{S})})\partial S$$
$$\leq \frac{2\partial S}{\min(Vol(S), Vol(\bar{S}))} =: 2h_G.$$

which gives the upper bound.

(2) Lower bound: the proof of lower bound actually gives a constructive algorithm to compute an approximate optimal cut as follows, which is called Cheeger Sweep.

Let v be the second eigenvector, i.e. $\mathcal{L}v = \lambda_1 v$, and $f = D^{-1/2}v$. Then we reorder node set V such that $f_1 \leq f_2 \leq ... \leq f_n$). Denote $V_- = \{i; v_i < 0\}, V_+ = \{i; v_i \geq v_r\}$. Without Loss of generality, we can assume

$$\sum_{i \in V_-} d_v \geq \sum_{i \in V_+} d_v$$

Define new functions f^+ to be the magnitudes of f on V_+.

$$f_i^+ = \begin{cases} f_i & i \in V_+, \\ 0 & otherwise, \end{cases}$$

Now consider a series of particular subsets of V,

$$S_i = \{v_1, v_2, ...v_i\},$$

and define

$$\widetilde{Vol}(S) = \min(Vol(S), Vol(\bar{S})).$$
$$\alpha_G = \min_i NCUT(S_i).$$

Clearly finding the optimal value α just requires comparison over $n - 1$ NCUT values.

Below we shall show that

$$\frac{h_G^2}{2} \leq \frac{\alpha_G^2}{2} \leq \lambda_1.$$

First, we have $Lf = \lambda_1 Df$, so we must have

$$\sum_{j:j \sim i} f_i(f_i - f_j) = \lambda_1 d_i f_i^2. \tag{1}$$

From this we will get the following results,

$$
\begin{aligned}
\lambda_1 &= \frac{\sum_{i \in V_+} f_i \sum_{j:j \sim i}(f_i - f_j)}{\sum_{i \in V_+} d_i f_i^2}, \\
&= \frac{\sum_{i \sim j \ i,j \in V_+}(f_i - f_j)^2 + \sum_{i \in V_+} f_i \sum_{j \sim i \ j \in V_-}(f_i - f_j)}{\sum_{i \in V_+} d_i f_i^2}, (f_i - f_j)^2 = f_i(f_i - f_j) + f_j(f_j - f_i) \\
&> \frac{\sum_{i \sim j \ i,j \in V_+}(f_i - f_j)^2 + \sum_{i \in V_+} f_i \sum_{j \sim i \ j \in V_-}(f_i)}{\sum_{i \in V_+} d_i f_i^2}, \\
&= \frac{\sum_{i \sim j}(f_i^+ - f_j^+)^2}{\sum_{i \in V} d_i f_i^{+2}}, \\
&= \frac{(\sum_{i \sim j}(f_i^+ - f_j^+)^2)(\sum_{i \sim j}(f_i^+ + f_j^+)^2)}{(\sum_{i \in V} f_i^{+2} d_i)(\sum_{i \sim j}(f_i^+ + f_j^+)^2)} \\
&\geq \frac{(\sum_{i \sim j} f_i^{+2} - f_j^{+2})^2}{(\sum_{i \in V} f_i^{+2} d_i)(\sum_{i \sim j}(f_i^+ + f_j^+)^2)}, \quad \text{Cauchy-Schwartz Inequality} \\
&\geq \frac{(\sum_{i \sim j} f_i^{+2} - f_j^{+2})^2}{2(\sum_{i \in V} f_i^{+2} d_i)^2},
\end{aligned}
$$

where the second last step is due to the Cauchy-Schwartz inequality $|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle$, and the last step is due to $\sum_{i \sim j \in V}(f_i^+ + f_j^+)^2 = \sum_{i \sim j \in V}(f_i^{+2} + f_j^{+2} + 2f_i^+ f_j^+) \leq 2\sum_{i \sim j \in V}(f_i^{+2} + f_j^{+2}) \leq 2\sum_{i \in V} f_i^{+2} d_i$. Continued from the last inequality,

$$
\begin{aligned}
\lambda_1 &\geq \frac{(\sum_{i \sim j} f_i^{+2} - f_j^{+2})^2}{2(\sum_{i \in V} f_i^{+2} d_i)^2}, \\
&\geq \frac{(\sum_{i \in V}(f_i^{+2} - f_{i-1}^{+2})CUT(S_{i-1}))^2}{2(\sum_{i \in V} f_i^{+2} d_i)^2}, \quad \text{since } f_1 \leq f_2 \leq \ldots \leq f_n \\
&\geq \frac{(\sum_{i \in V}(f_i^{+2} - f_{i-1}^{+2})\alpha_G \widetilde{Vol}(S_{i-1}))^2}{2(\sum_{i \in V} f_i^{+2} d_i)^2} \\
&= \frac{\alpha_G^2}{2} \cdot \frac{(\sum_{i \in V} f_i^{+2}(\widetilde{Vol}(S_{i-1}) - \widetilde{Vol}(S_i)))^2}{(\sum_{i \in V} f_i^{+2} d_i)^2}, \\
&= \frac{\alpha_G^2}{2} \frac{(\sum_{i \in V} f_i^{+2} d_i)^2}{(\sum_{i \in V} f_i^{+2} d_i)^2} = \frac{\alpha_G^2}{2}.
\end{aligned}
$$

where the last inequality is due to the assumption $Vol(V_-) \geq Vol(V_+)$, whence $\widetilde{Vol}(S_i) = Vol(\bar{S}_i)$ for $i \in V_+$.

This completes the proof. $\square$

3.3   A Short Proof of Cheeger Inequality

Fan Chung gives a short proof of the lower bound in Simons Institute workshop, 2014.

Short Proof. The proof is based on the fact that

$$h_G = \inf_{f \neq 0} \sup_{c \in \mathbb{R}} \frac{\sum_{i \sim j} |f(i) - f(j)|}{\sum_i |f(i) - c| d_i}$$

where the supreme over c is reached at $c^* = median(f(x) : x \in V)$.

$$
\begin{aligned}
\lambda_1 &= R(f) = \sup_c \frac{\sum_{i \sim j} (f(i) - f(j))^2}{\sum_i (f(i) - c)^2 d_i}, \\
&\geq \frac{\sum_{i \sim j} (g(i) - g(j))^2}{\sum_i g(i)^2 d_i}, \quad g(i) = f(i) - c \\
&= \frac{(\sum_{i \sim j} (g(i) - g(j))^2)(\sum_{i \sim j} (g(i) + g(j))^2)}{(\sum_{i \in V} g^2(i) d_i)((\sum_{i \sim j} (g(i) + g(j))^2)} \\
&\geq \frac{(\sum_{i \sim j} |g^2(i) - g^2(j)|)^2}{(\sum_{i \in V} g^2(i) d_i)((\sum_{i \sim j} (g(i) + g(j))^2)}, \quad \text{(Cauchy-Schwartz Inequality)} \\
&\geq \frac{(\sum_{i \sim j} |g^2(i) - g^2(j)|)^2}{2(\sum_{i \in V} g^2(i) d_i)^2}, \quad (g(i) + g(j))^2 \leq 2(g^2(i) + g^2(j)) \\
&\geq \frac{h_G^2}{2}.
\end{aligned}
$$

$\square$

Remark: In fact, this proof is wrong with a obvious mistake in the last step. I think the proof might be total wrong since that $g^2(i) = g^2(j)$ is possible even if $g(i) \neq g(j)$.

4 Appendix: Laplacians and the Cheeger inequality for directed graphs

The following section is mainly contained in [**?**], which described the following results:

1. Define Laplacians on directed graphs.

2. Define Cheeger constants on directed graphs.

3. Give an example of the singularity of Cheeger constant on directed graph.

4. Use the eigenvalue of Lapacian and the Cheeger constant to estimate the convergence rate of random walk on a directed graph.

Another good reference is [**?**].

4.1 Definition of Laplacians on directed graphs

On a finite and strong connected directed graph $G = (V, E)$ (A directed graph is strong connected if there is a path between any pair of vertices), a weight is a function

$$w : \quad E \quad \to \quad \mathbb{R}_{\geq 0}$$

The in-degree and out-degree of a vertex are defined as

$$\begin{aligned}
d^{in} : \quad V \quad &\to \quad \mathbb{R}_{\geq 0} \\
d_i^{in} \quad &= \quad \textstyle\sum_{j \in V} w_{ji} \\
d^{out} : \quad V \quad &\to \quad \mathbb{R}_{\geq 0} \\
d_i^{out} \quad &= \quad \textstyle\sum_{j \in V} w_{ij}
\end{aligned}$$

Note that d_i^{in} may be different from d_i^{out}.

A random walk on the weighted G is a Markov chain with transition probability

$$P_{ij} = \frac{w_{ij}}{d_i^{out}}.$$

Since G is strong connected, P is irreducible, and consequently there is a unique stationary distribution ϕ. (And the distribution of the Markov chain will converge to it if and only if P is aperiodic.)

Example (undirected graph).

$$\phi(x) = \frac{d_x}{\sum_y d_y}.$$

Example (Eulerian graph). If $d_x^{in} = d_x^{out}$ for every vertex x, then $\phi(x) = \frac{d_x^{out}}{\sum_y d_y^{out}}$.

This is because d_x^{out} is an unchanged measure with

$$\sum_x d_x^{out} P_{xy} = \sum_x w_{xy} = d_y^{in} = d_y^{out}.$$

Example (exponentially small stationary dist.). G is a directed graph with $n + 1$ vertices formed by the union of a directed circle $v_0 \to v_1 \to \cdots \to v_n$ and edges $v_i \to v_0$ for $i = 1, 2, \cdots, n$. The weight on any edge is 1. Checking from v_n to v_0 with the prerequisite of stationary distribution that the inward probability flow equals to the outward probability flow, we can see that

$$\phi(v_0) = 2^n \phi(v_n), \, i.e. \phi(v_n) = 2^{-n} \phi(v_0).$$

This exponentially small stationary distribution cannot occur in undirected graph cases for then

$$\phi(i) = \frac{d_i}{\sum_j d_j} \geq \frac{1}{n(n-1)}.$$

However, the stationary dist. can be no smaller than exponential, because we have

Theorem 4.1. If G is a strong connected directed graph with $w \equiv 1$, and $d_x^{out} \leq k, \forall x$, then $\max\{\phi(x) : x \in V\} \leq k^D \min\{\phi(y) : y \in V\}$, where D is the diameter of G.

It can be easily proved using induction on the path connecting x and y.

Now we give a definition on those balanced weights.

Definition (circulation).

$$F : \quad E \quad \to \quad \mathbb{R}_{\geq 0}$$

If F satisfies

$$\sum_{u, u \to v} F(u, v) = \sum_{w, v \to w} F(v, w), \forall v,$$

then F is called a circulation.

Note. A circulation is a flow with no source or sink.

Example. For a directed graph, $F_\phi(u, v) = \phi(u)P(u, v)$ is a circulation, for

$$\sum_{u, u \to v} F_\phi(u, v) = \phi(v) = \sum_{w, v \to w} F_\phi(v, w).$$

Definition (Rayleigh quotient)**.** For a directed graph G with transition probability matrix P and stationary distribution ϕ, the Rayleigh quotient for any $f : V \to$ is defined as

$$R(f) = \frac{\sum_{u \to v} \mid f(u) - f(v) \mid^2 \phi(u)P(u, v)}{\sum_v \mid f(v) \mid^2 \phi(v)}.$$

Note. Compare with the undirected graph condition where

$$R(f) = \frac{\sum_{u \sim v} \mid f(u) - f(v) \mid^2 w_{uv}}{\sum_v \mid f(v) \mid^2 d(v)}.$$

If we look on every undirected edge (u, v) as two directed edges $u \to v, v \to u$, then we get a Eulerian directed graph. So $\phi(u) \sim d_u^{out}$ and $d_u^{out}P(u, v) = w_{uv}$, as a result $R(f)(directed) = 2R(f)(undirected)$. The factor 2 is the result of looking on every edge as two edges.

The next step is to extend the definition of Laplacian to directed graphs. First we give a review on Lapalcian on undirected graphs. On an undirected graph, adjacent matrix is

$$A_{ij} = \begin{cases} 1, & i \sim j; \\ 0, & i \not\sim j. \end{cases}$$

$$D = \text{diag}(d(i)),$$

$$\mathcal{L} = D^{-1/2}(D - A)D^{-1/2}.$$

On a directed graph, however, there are two degrees on a vertex which are generally inequivalent. Notice that on an undirected graph, stationary distribution $\phi(i) \sim d(i)$, so $D = c\Phi$, where c is a constant and $\Phi = \text{diag}(\phi(i))$.

$$\begin{aligned} \mathcal{L} &= I - D^{-1/2}AD^{-1/2} \\ &= I - D^{1/2}PD^{-1/2} \\ &= I - c^{1/2}\Phi^{1/2}Pc^{-1/2}\Phi^{-1/2} \\ &= I - \Phi^{1/2}P\Phi^{-1/2} \end{aligned}$$

Extending and symmetrizing it, we define Laplacian on a directed graph

Definition (Laplacian)**.**

$$\mathcal{L} = I - \frac{1}{2}(\Phi^{1/2}P\Phi^{-1/2} + \Phi^{-1/2}P^*\Phi^{1/2}).$$

Suppose the eigenvalues of $\mathcal{L}$ are $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1}$. Like the undirected case, we can calculate λ_1 with the Rayleigh quotient.

Theorem 4.2.

$$\lambda_1 = \inf_{\sum f(x)\phi(x)=0} \frac{R(f)}{2}.$$

Before proving that, we need

Lemma 4.3.

$$R(f) = 2\frac{g\mathcal{L}g^*}{\parallel g \parallel^2}, \text{ where } g = f\Phi^{1/2}.$$

Proof.

$$
\begin{aligned}
R(f) &= \frac{\sum_{u \to v} \mid f(u) - f(v) \mid^2 \phi(u)P(u,v)}{\sum_v \mid f(v) \mid^2 \phi(v)} \\
&= \frac{\sum_{u \to v} \mid f(u) \mid^2 \phi(u)P(u,v) + \sum_v \mid f(v) \mid^2 \phi(v) - \sum_{u \to v}(\overline{f(u)}f(v) + f(u)\overline{f(v)})\phi(u)P(u,v)}{f\Phi f^*} \\
&= \frac{\sum_u \mid f(u) \mid^2 \phi(u) + \sum_v \mid f(v) \mid^2 \phi(v) - (f^*\Phi Pf + f\Phi Pf^*}{f\Phi f^*} \\
&= 2 - \frac{f(P^*\Phi + \Phi P)f^*}{f\Phi f^*} \\
&= 2 - \frac{(g\Phi^{-1/2})(P^*\Phi + \Phi P)(\Phi^{-1/2}g^*)}{(g\Phi^{-1/2})\Phi(\Phi^{-1/2}g^*)} \\
&= 2 - \frac{g(\Phi^{-1/2}P^*\Phi^{1/2} + \Phi^{1/2}P\Phi^{-1/2})g^*}{gg^*} \\
&= 2 \cdot \frac{g\mathcal{L}g^*}{\parallel g \parallel^2}
\end{aligned}
$$

$\square$

Proof of Theorem 4.2. With Lemma 4.3 and $\mathcal{L}(\phi(x)^{1/2})_{n \times 1} = 0$, we have

$$
\begin{aligned}
\lambda_1 &= \inf_{\sum g(x)\phi(x)^{1/2}=0} \frac{R(f)}{2} \\
&= \inf_{\sum f(x)\phi(x)=0} \frac{R(f)}{2}.
\end{aligned}
$$

$\square$

Note.

$$
\begin{aligned}
\lambda_1 &= \inf_{f,\sum f(x)\phi(x)=0} \frac{R(f)}{2} \\
&= \inf_{f,\sum f(x)\phi(x)=0} \frac{\sum_{u \to v} \mid f(u) - f(v) \mid^2 \phi(u)P(u,v)}{2\sum_v \mid f(v) \mid^2 \phi(v)} \\
&= \inf_{f,\sum f(x)\phi(x)=0} \sup_c \frac{\sum_{u \to v} \mid f(u) - f(v) \mid^2 \phi(u)P(u,v)}{2\sum_v \mid f(v) - c \mid^2 \phi(v)}
\end{aligned}
$$

Theorem 4.4. Suppose the eigenvalues of P are $\rho_0, \cdots, \rho_{n-1}$ with $\rho_0 = 1$, then

$$\lambda_1 \leq \min_{i \neq 0}(1 - Re\rho_i).$$

4.2 Definition of Cheeger constants on directed graphs

We have a circulation $F(u, v) = \phi(u)P(u, v)$. Define

$$F(\partial S) = \sum_{u \in S, v \notin S} F(u, v), F(v) = \sum_{u, u \to v} F(u, v) = \sum_{w, v \to w} F(v, w), F(S) = \sum_{v \in S} F(v),$$

then $F(\partial S) = F(\partial \bar{S})$.

Definition (Cheeger constant). The Cheeger constant of a graph G is defined as

$$h(G) = \inf_{S \subset V} \frac{F(\partial S)}{\min\left(F(S), F(\bar{S})\right)}$$

Note. Compare with the undirected graph condition where

$$h_G = \inf_{S \subset V} \frac{|\partial S|}{\min\left(|S|, |\bar{S}|\right)}.$$

Similarly, we have

$$
\begin{aligned}
h_G(\text{undirected}) &= \inf_{S \subset V} \frac{|\partial S|}{\min\left(|S|, |\bar{S}|\right)} \\
&= \inf_{S \subset V} \frac{\sum_{u \in S, v \in \bar{S}} w_{uv}}{\min\left(\sum_{u \in S} d(u), \sum_{u \in \bar{S}} d(u)\right)} \\
h_G(\text{directed}) &= \inf_{S \subset V} \frac{\sum_{u \in S, v \in \bar{S}} \phi(u)P(u, v)}{\min\left(\sum_{u \in S} \phi(u), \sum_{u \in \bar{S}} \phi(u)\right)} \\
&= \inf_{S \subset V} \frac{F(\partial S)}{\min\left(F(S), F(\bar{S})\right)}.
\end{aligned}
$$

Theorem 4.5. For every directed graph G,

$$\frac{h^2(G)}{2} \leq \lambda_1 \leq 2h(G).$$

The proof is similar to the undirected case using Rayleigh quotient and Theorem 4.2.

4.3 An example of the singularity of Cheeger constant on a directed graph

We have already given an example of a directed graph with $n + 1$ vertices and stationary distribution ϕ satisfying $\phi(v_n) = 2^{-n}\phi(v_0)$. Now we make a copy of this graph and denote the new $n + 1$ vertices $u_0, \ldots, u_n$. Joining the two graphs together by two edges $v_n \to u_n$ and $u_n \to v_n$, we get a bigger directed graph. Let $S = (v_0, \cdots, v_n)$, we have $h(G) \sim 2^{-n}$. In comparison, $h(G) \geq \frac{2}{n(n-1)}$ for undirected graph.

4.4 Estimate the convergence rate of random walks on directed graphs

Define the distance of P after s steps and ϕ as

$$\Delta(s) = \max_{y \in V} \left(\sum_{x \in V} \frac{(P^s(y, x) - \phi(x))^2}{\phi(x)}\right)^{1/2}.$$

Modify the random walk into a lazy random walk $\tilde{P} = \frac{I + P}{2}$, so that it is aperiodic.

Theorem 4.6.
$$\Delta(t)^2 \le C(1 - \frac{\lambda_1}{2})^t.$$

4.5 Random Walks on Digraphs, The Generalized Digraph Laplacian, and The Degree of Asymmetry

In this paper the following have been discussed:

1. Define an asymmetric Laplacian L on directed graph;

2. Use L to estimate the hitting time and commute time of the corresponding Markov chain;

3. Introduce a metric to measure the asymmetry of L and use this measure to give a tighter bound on the Markov chain mixing rate and a bound on the Cheeger constant.

Let P be the transition matrix of Markov chain, and $\pi = (\pi_1, \ldots, \pi_n)^T$(column vector) denote its stationary distribution (which is unique if the Markov chain is irreducible, or if the directed graph is strongly connected). Let $\Pi = \text{diag}\{\pi_1, \ldots, \pi_n\}$, then we define the normalized Laplacian L on directed graph:

$$L = I - \Pi^f P \Pi^f \tag{2}$$

4.5.1 Hitting time, commute time and fundamental matrix

We establish the relations between L and the hitting time and commute time of random walk on directed graph through the fundamental matrix $Z = [z_{ij}]$, which is defined as:

$$z_{ij} = \sum_{t=0}^{\infty} (p_{ij}^t - \pi_j), \ 1 \le i, j \le n \tag{3}$$

or alternatively as an infinite sum of matrix series:

$$Z = \sum_{t=0}^{\infty} (P^t - 1\pi^T) \tag{4}$$

With the fundamental matrix, the hitting time and commute time can be expressed as follows:

$$H_{ij} = \frac{z_{jj} - z_{ij}}{\pi_j} \tag{5}$$

$$C_{ij} = H_{ij} + H_{ji} = \frac{z_{jj} - z_{ij}}{\pi_j} + \frac{z_{ii} - z_{ji}}{\pi_i} \tag{6}$$

Using (4), we can write the fundamental matrix Z in a more explicit form. Notice that

$$(P - 1\pi^T)(P - 1\pi^T) = P^2 - 1\pi^T P - P1\pi^T + 1\pi^T 1\pi^T = P^2 - 1\pi^T \tag{7}$$

We use the fact that 1 and π are the right and left eigenvector of the transition matrix P with eigenvalue 1, and that $\pi^T 1 = 1$ since π is a distribution. Then

$$Z + 1\pi^T = \sum_{t=0}^{\infty} (P - 1\pi^T)^t = (I - P + 1\pi^T)^{-1} \tag{8}$$

4.5.2 Green's function and Laplacian for directed graph

If we treat the directed graph Laplacian L as an asymmetric operator on a directed graph G, then we can define the Green's Function G (without boundary condition) for directed graph. The entries of G satisfy the conditions:

$$(GL)_{ij} = \delta_{ij} - \sqrt{\pi_i \pi_j} \tag{9}$$

or in the matrix form

$$GL = I - \pi^f {\pi^f}^T \tag{10}$$

The central theorem in the second paper associate the Green's Function G, the fundamental matrix Z and the normalize directed graph Laplacian L:

Theorem 4.7. Let $\tilde{\mathcal{Z}} = \Pi^f Z \Pi^f$ and $L^\dagger$ denote the Moore-Penrose pseudo-inverse L, then

$$G = \tilde{\mathcal{Z}} = L^\dagger \tag{11}$$

4.6 measure of asymmetric and its relation to Cheeger constant and mixing rate

To measure the asymmetry in directed graph, we write the L into the sum of a symmetric part and a skew-symmetric part:

$$L = f(L + L^T) + f(L - L^T) \tag{12}$$

$f(L + L^T) = \mathcal{L}$ is the symmetrized Laplacian introduced in the first paper. Let $\Delta = f(\mathcal{L} - \mathcal{L}^T)$, the Δ captures the difference between $\mathcal{L}$ and its transpose. Let σ_i, λ_i and $\delta_i (1 \le i \le n)$ denotes the i-th singular value of $\mathcal{L}$, $\mathcal{L}$, Δ in ascending order ($\sigma_1 = \lambda_1 = \delta_1 = 0$). Then the relation $\mathcal{L} = \mathcal{L} + \Delta$ implies

$$\lambda_i \le \sigma_i \le \lambda_i + \delta_n \tag{13}$$

Therefore $\delta_n = \|\Delta\|_2$ is used to measure the degree of asymmetry in the directed graph.

The following two theorems are application of this measure.

Theorem 4.8. The second singular of $\mathcal{L}$ has bounds :

$$\frac{h(G)^2}{2} \le \sigma_2 \le (1 + \frac{\delta_n}{\lambda_2}) \cdot 2h(G) \tag{14}$$

where $h(G)$ is the Cheeger constant of graph G

Theorem 4.9. For a aperiodic Markov chain P,

$$\delta_n^2 \le \max\{\frac{\|\tilde{P}f\|^2}{\|f\|^2} : f \perp \pi^f\} \le (1 - \lambda_2)^2 + 2\delta_n \lambda_n + \delta_n^2 \tag{15}$$

where $\tilde{P} = \Pi^f P \Pi^f$