Algorithm Theory

Jiayi Wu

January 22, 2023

Contents

Chapter 1

Graph Matrix

Let G(V, E) be a graph, where $|V| = n$ and $|E| = m$. Assume G is unweighted and undirected, with no self loops or multiple edges.

1.1 Laplacian

1.1.1 Adjacency Matrix

For graph G, let the adjacency matrix $A = A_G$ be defined as

$$A_G = \begin{cases} 1 & (i, j) \in E \\ 0 & (i, j) \notin E \end{cases}$$

1.1.2 Laplacian

The Laplacian of the graph, $L = L_G$ is defined as

$$L_G = \begin{cases} -1 & (i, j) \in E \\ d_i & i = j \\ 0 & otherwise \end{cases}$$

$$L_G = D_G - A_G$$

Let some vector **v** be thought of as a function that maps to the index to some value. ie.

$$v_i = X(i)$$

$L_G\mathbf{v}$ can then be though as:

$$[L_G\mathbf{v}]_i = d_i[X(i) - (\overline{X(j)} \mid (i,j) \in E)]$$

Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of L_G with corresponding eigenvectors $\mathbf{v_1}, \ldots \mathbf{v_n}$. Then

- $\mathbf{v_1} = \mathbf{1}$, $\lambda_1 = 0$
- All $\lambda_i \geq 0$
- Much information can be obtained by the first few non-trival eigenvectors.

1.1.3 Properties of the Laplacian

Edge Union

If G and H share the same vertex set have **disjoint** edge sets,

$$L_{G \cup H} = L_G + L_H$$

Isolated Vertex

If $d_i = 0$, then $[L_G]_{i,j} = 0$ and $[L_G]_{j,i} = 0$

Disjoint Union

If G and H have disjoint vertex sets, then

$$L_{G \sqcup H} = L_G \oplus L_H = \begin{pmatrix} L_G & 0 \\ 0 & L_H \end{pmatrix}$$

If L_G had eigenvectors $\mathbf{v_i}$ with eigenvalues λ_i and L_H had eigenvectors $\mathbf{u_i}$ with eigenvalues μ_i, then $L_{G \sqcup H}$ has eigenvectors

$$\mathbf{v_i} \oplus \mathbf{0}, \mathbf{0} \oplus \mathbf{u_i}$$

with the eigenvalues λ_i and μ_i each.

Laplacian of Edge

The Laplacian of an edge L_e has the form $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} -1 & 1 \end{pmatrix}$.

The graph can be written as $\sum L_e$. Hence for vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbf{x^T} L_e \mathbf{x} = (x_1 - x_2)^2 \geq 0$. Hence the Laplacian is **positive semidefinite**. This implies that all eigenvalues $\lambda_i \geq 0$ and that the matrix can be diagonalized as $L_e = Q^T \Lambda Q$.

$$\mathbf{x^T} L \mathbf{x} = \sum_{(i,j) \in E} (x_i - x_j)^2$$

In addition, because L is positive semidefinite, all eigenvalues of L are non-negative, $L_e = A^T A$ (for some A).
A need not be square or unique.

Factoring the Laplacian

The **Incidence Matrix** $\nabla_G = \nabla_{e,v} = \begin{cases} 1 & e = (v, w), v < w \\ -1 & e = (v, w), v > w \\ 0 & otherwise \end{cases}$

Then $L_G = \nabla^T \nabla$.

Dimension of Nullspace

If G is connected, then the nullspace is $span\{\mathbf{1}\}$.

Thus, for connected graphs, $\lambda_2 \geq 0$

The dimensions of the nullspace is equal to the number of connected components of G.

Spectra of Common Graphs

Complete Graph has eigenvalue 0 with multiplicity 1 and eigenvalue n with multiplicity $(n - 1)$
eigenvectors of $\mathbf{1}$ and $\{\mathbf{x} \mid \mathbf{x} \cdot \mathbf{1} = 0\}$

Ring Graph with n vertices has eigenvectors

$$x_k(u) = \sin\frac{2\pi ku}{n}$$
$$y_k(u) = \cos\frac{2\pi ku}{n}$$

where the eigenvalue x_k and y_k is $2 - 2\cos\frac{2\pi k}{n}$
Path Graph with n vertices has the eigenvectors

$$v_k(u) = \sin(\frac{\pi ku}{n} + \frac{\pi}{2n})$$

That has the same Eigenvalues as R_{2n}

Graph Products

Definition Let G=(V,E) and H=(W,F). The Graph $G \times H$ has vertex set $V \times W$ and edge set

$$((v_1, w), (v_2, w)), \forall (v_1, v_2) \in E, w \in W$$
$$((v, w_1), (v, w_2)), \forall v \in V, (w_1, w_2) \in F$$

The product of 2 paths graph is the grid graph.
Eigenvales If L_G has the eigenvalues λ_i with the eigenvectors v_i and L_H has the eigenvalues μ_i with the eigenvectors w_i, then the eigenvectors of $L_{G\times H}$ are listed

$$z_{ij}(v, w) = x_i(v)y_j(w)$$

with eigenvalues $\lambda_i + \mu_u$.

1.1.4 Why it is Called the Laplacian

If one is to discrete the derivative on graphs, as in treat the vertex as a function, then the "Laplacian" would be $-L_G$

Laplacian Eigenvalues

Sum of Eigenvalues Given an n-vertex graph, where $d_{max} = \max d_i$,

$$\sum \lambda_i = \sum d_i \leq d_{max} n$$

This is obtained from the trace. If one removes $\lambda_1 = 0$, then

$$\lambda_2 \leq \frac{\sum d_i}{n-1}$$
$$\lambda_n \geq \frac{\sum d_i}{n-1}$$

Bounds of λ_2 and λ_{max}

Courant-Fischer Formula For any $n \times n$ symmetric matrix A,

$$\lambda_1 = \min_{||x||=1} \mathbf{x^T} A \mathbf{x} = \min_{||x|| \neq 0} \frac{\mathbf{x^T} A \mathbf{x}}{\mathbf{x^T x}}$$
$$\lambda_{max} = \max_{||x||=1} \mathbf{x^T} A \mathbf{x} = \max_{x \neq \vec{0}} \frac{\mathbf{x^T} A \mathbf{x}}{\mathbf{x^T x}}$$

Let S_k denote the span of $v_1 \cdots v_k$, and let $S_k^{\perp}$ denote the orthogonal complement of S_k

$$\lambda_k = \min_{|x|=1, x \in S_k^{\perp}} x^T A x = \min_{x \neq 0, x \in S_k^{\perp}} \frac{x^T A x}{x^T x}$$

Rayleigh Quotient Let G=(V,E) be a graph with Laplacian L_G, then

$$\begin{aligned}
\lambda_1 &= 0 & \mathbf{v_1} &= \mathbf{1} \\
\lambda_2 &= \min_{x \perp v_1, x \neq 0} \frac{x^T L_G x}{x^T x} &= \min_{\sum x = 0, x \neq 0} &\frac{\sum_{(i,j) \in E} (x_i - x_j)^2}{\sum_{i \in V} x_i^2} \\
\lambda_{max} &= \max_{x \neq 0} \frac{x^T L_G x}{x^T x} &= \max_{x \neq 0} &\frac{\sum_{(i,j) \in E} (x_i - x_j)^2}{\sum_{i \in V} x_i^2}
\end{aligned}$$

One can use this to bound the eigenvalues of graphs easily.

1.2 Cut Graphs

1.2.1 Cut Ratio

Let $e(S)$ of a cut $S - \overline{S}$ denote the number of edges between S and $\overline{S}$
Cut Ratio ϕ of a cut $S - \overline{S}$ is $\phi = \frac{e(S)}{\min(|S, |\overline{S}|)}$
The **cut of minimum ratio** is the cut that minimizes $\phi(S)$.
The **isoperimetric number** of a graph is the value of minimum cut.

Integer Function of the Cut Ratio

Associate every cut $S - \overline{S}$ with a vector $x \in \{-1, 1\}^n$, such that

$$x_i = \begin{cases} 1 & i \in S \\ -1 & i \in \overline{S} \end{cases}$$

Thus,

$$e(S) = \frac{1}{4} \sum_{(i,j) \in E} (x_i - x_j)^2$$

$$|S||\overline{S}| = \sum_{i,j \in V} [i \in S, j \in \overline{S}] = \frac{1}{4} \sum_{i<j} (x_i - x_j)^2$$
$$\frac{n}{2} \min(|S|, |\overline{S}|) \leq |S||\overline{S}| \leq n \min(|S|, |\overline{S}|)$$

Hence:

$$\frac{1}{n}\phi(G) \leq \min_{x \in \{1,-1\}^n} \frac{\sum_{(i,j)\in E}(x_i - x_j)^2}{\sum_{i<j}(x_i - x_j)^2} \leq \frac{2}{n}\phi(G)$$

This allows approximation of $\phi(G)$ to a factor of 2, but NP-hard. Can change the range of x to be $[-1, 1]^n$, this is solvable.

Relaxation

To find the min of function $f(x)$ on C (NP hard), relax the constraints $x \in C' \supseteq C$, let $f(p)$ and $f(q)$ be the min with constraints C and C'. $f(p) \geq f(q)$. Round q to $q' \in C$ such that $f(q') \leq \gamma f(q)$. Thus $f(q') \leq \gamma f(q) \leq \gamma f(p)$

Bounding Min Cut Relax $x \in \{-1, 1\}$ to $x \in R^n, \sum x_i = 0$

$$\sum_{i<j}(x_i - x_j)^2 = n \sum x_i^2$$

$$\min_{x \in R^n} \frac{\sum_{(i,j)\in E}(x_i - x_j)^2}{\sum_{i<j}(x_i - x_j)^2} = \min_{x \in R^n, x \perp 1} \frac{\sum_{(i,j)\in E}(x_i - x_j)^2}{n \sum x_i^2} = \frac{\lambda_2}{n}$$

By definition of Courant-Fischer formula of the second eigenvalue. So,

$$\phi(G) \geq \frac{\lambda_2}{2}$$

1.2.2 Cheeger's Inequality

Upper bound on the minimum cut. These bounds are tight.

$$\frac{\phi(G)^2}{2d_{max}} \le \lambda_2 \le 2\phi(G)$$

Proof

For any $x \perp \mathbf{1}, x_1 \le x_2 \cdots x_n$, $\exists i$

$$\frac{x^T Lx}{x^T x} \ge \frac{\phi(\{1 \cdots i\})}{2d_{max}}$$

Let

$$m = \frac{n+1}{2}$$
$$y_i = x_i - x_m$$

Then

$$\frac{x^T Lx}{x^T x} \ge \frac{y^T Ly}{y^T y}$$

Next, if $(i,j) \in E, i < m < j$, then replace (i, j) with (i, m) and (m, j) in E'.

$$\frac{\sum_{(i,j)\in E}(y_i - y_j)^2}{\sum_{i\in V} y_i^2} \ge \frac{\sum_{(i,j)\in E'}(y_i - y_j)^2}{\sum_{i\in V} y_i^2}$$

Split the latter on m,

$$\frac{\sum_{(i,j)\in E'_-}(y_i - y_j)^2 + \sum_{(i,j)\in E'_+}(y_i - y_j)^2}{\sum_{i\in[1,m]} y_i^2 + \sum_{i\in[m,n]} y_i^2}$$

which has lower bound

$$\frac{\sum_{(i,j)\in E'_-}(y_i - y_j)^2}{\sum_{i\in[1,m]} y_i^2}$$

Lemma for $z_1 \le z_2 \cdots \le z_n = 0$,

$$\sum_{(i,j)\in E'_-} |z_i - z_j| \ge \phi \sum_{i=1}^{m} |z_i|$$

Put it together

1. Normalize: $\sum_{i=1}^{m} y_i^2 = 1$
2. Let $z_i = -y_i^2$, apply lemma to gather

$$\sum_{(i,j)\in E'_{-}} |y_i^2 - y_j^2| \geq \phi \sum_{i=1}^{m} |y_i|^2 = \phi$$

3. Cauchy-Schwartz:

$$\sum_{(i,j)\in E'_{-}} |y_i^2 - y_j^2| \leq \Big(\sum_{(i,j)\in E'_{-}} (y_i - y_j)^2 \Big)^{1/2} \Big(\sum_{(i,j)\in E'_{-}} (y_i + y_j)^2 \Big)^{1/2}$$

4.

$$\sum_{(i,j)\in E'_{-}} (y_i + y_j)^2 \leq 2 \sum_{(i,j)\in E'_{-}} (y_i^2 + y_j^2) \leq 2 \sum_{i=1}^{m} d_{max} y_i^2 \leq 2 d_{max}$$

5.

$$\frac{\sum_{(i,j)\in E'_{-}} (y_i - y_j)^2}{\sum_{i=0}^{m} y_i^2} \geq \frac{\big(\sum_{(i,j)\in E'_{-}} |y_i^2 - y_j^2|\big)^2}{\sum_{(i,j)\in E'_{-}} (y_i + y_j)^2} \geq \frac{\phi^2}{2 d_{max}}$$

6.

$$\frac{x^T L x}{x^T x} \geq \frac{y^T L y}{y^T y} \geq \frac{\phi^2}{2 d_{max}}$$

1.2.3 Random Walks

Definitions

Consider starting at some vertex $v \in V(G)$ and repeatedly moving to a neighboring vertex uniformly at random
Let $p_t(u)$ denote the probability of being at a vertex at time t.

$$\sum_{u \in V} p_t(u) = 1$$

$$p_t(u) = \sum_{(u,v)\in E} p_{t-1}(v) \cdot \frac{1}{d(v)}$$

Define a matrix W_G as follows:

$$[W]_{i,j} = \begin{cases} \frac{1}{d(j)} & (i,j) \in E \\ 0 & otherwise \end{cases}$$
$$W = A \cdot D^{-1}$$

Where A is the adjacency matrix and D is diagonal matrix with degrees.

Stationary Distribution

Let π be defined as a probability vector *stationary distribution* of the random walk:

$$\pi(u) = \frac{d(u)}{\sum_{v \in V} d(v)}$$

π is a probability distribution

$$W \cdot \pi = \pi$$

Hence is at one point the random walk follows distribution π, it shall follow π for all future steps. Or π is W's eigenvector with eigenvalue 1. Note that the probability need not converge to π or at all.

Lazry Random Walks

Lazy Random Walks

- With probability 1/2 stay at the current node
- With probability 1/2 take a step of the original walk

This breaks the periodicity of the walk in the following way

$$W' = (W + I)/2 = (I + A \cdot I)/2$$

However, since W and W' are not symmetric, it is difficult to analyze. *Normalized Random Walk Matrix*

$$N = D^{-1/2} \cdot W \cdot D^{1/2} \qquad = D^{-1/2} \cdot A \cdot D^{-1/2}$$
$$N' = D^{-1/2} \cdot W' \cdot D^{1/2} \qquad = (I + D^{-1/2} \cdot A \cdot D^{-1/2})/2$$

N and W share eigenvalues and their eigenvectors are related.

If v is eigenvector of N, then $q = D^{1/2}v$ is eigenvector of W

therefore because W has eigenvector $D \cdot \mathbf{1}$ with eigenvalue 1, N has eigenvector $D^{1/2} \cdot \mathbf{1}$ with eigenvalue 1

Connection to Laplacians

The normalized laplacian $\mathcal{L}$ is defined as

$$\mathcal{L} = D^{-1/2} \cdot L \cdot D^{-1/2}$$

and

$$N = I - \mathcal{L}$$

Eigenvalues of N are given by $\mu = 1 - \lambda$ of the eigenvalues of $\mathcal{L}$, reorder them such that

$$1 = \mu_1 \geq \mu_2 \cdots \geq \mu_n$$

Theorems

- for all i, $\mu_i \in [-1, 1]$
- if G is connected, $\lambda_2 < 1$
- the eigenvalue of -1 only occurs for bipartite Graphs

For eigenvalues of N', μ_i'

- for all i, $\mu_i' \in [0, 1]$
- if G is connected, then $\mu_2' < 1$

ℓ_2 Convergence

The spectral gap is defined as $\lambda = 1 - \mu_2$
ℓ_2 **Distance** for probability distribution p, q is defined to be $||p - q||_2 = \sqrt{\sum_i (p(i) - q(i))^2}$, basically the Euclidean Distance. **Theorem** Let p_0 denote some initial probability distribution, and p_t the distribution after t steps of lazy random walks. Then

$$||p_t - \pi||_2 \leq (1 - \lambda_2)^2 \sqrt{\frac{max_x d(x)}{min_y d(y)}}$$

This is easily shown for regular graphs.

Conductance

Conductance controls how well-knit a graph is: how fast a random walk converges to the stationary distribution. Or how fast one can move from one part of a graph to another **Definition** For $S \subset V$, let

$$\Phi(S) = \frac{e(S)}{min(\sum_{v \in S} d(v), \sum_{v \in \overline{S}} d(v)}$$

The conductance

$$\Phi(G) = \min_{S \subset V} \Phi(S)$$
$$\Theta(1) \cdot \Phi^2(G) \leq 1 - \mu_2' \leq \Theta(1) \cdot \Phi(G)$$

The above is analogous to Cheeger's Inequality.

Monte Carlo Methods

For ranomd variable $r \in 0, 1$ such that $\mathbf{Pr}[r = 1] = p$ and $\mathbf{Pr}[r = 0] = 1 - p$ Assuming draw n independent $r_1, r_2 \ldots r_n$. Let $R = \sum_i r_i$. By linearity of expectations, $E[R] = n \cdot p$
We say R ϵ-*Approximates* $\mathbf{E}[R]$ if

$$(1 - \epsilon)\mathbf{E}[R] \leq R \leq (1 + \epsilon)\mathbf{E}[R]$$

This error measure is multiplicative.
Chernoff Bound The Probability that R fails to ϵ approximate **E**[R] is

$$\mathbf{Pr}[|R - \mathbf{E}[R]| \geq \epsilon \mathbf{E}[R]] \leq 2e^{-np\epsilon^2/12} = 2e^{-\mathbf{E}[R]\epsilon^2/12}$$

- The bound is nearly tight
- The trials need to be independent for the bound to hold
- The guarentee is multiplicative but not additive
- For fiexed ϵ, it falls off exponentially in n. So one can improve from $1/2$ to $1/2^k$ by doing k trials
- Therefore, smaller n requires more trials
- If a ϵ-approximation is required with probability $1 - \delta$,

$$N \geq \Theta(\frac{log(1/\delta)}{p\epsilon^2})$$

Chapter 2

Combinatorics

2.1 Expanders and Eigenvalues

2.1.1 Expander Graphs

The **edge expansion** or **Cheeger Constant** is

$$h(G) = \min_{|S| \le n/2} \frac{e(S, \overline{S})}{|S|} \tag{2.1}$$

A graph is **(d, ϵ)-expander** if it is d-regular and $h(G) \ge \epsilon$

2.1.2 Random Graphs

A bipartite graph with $n + n$ vertices $L \cup R$ is a **(d, β)-expander** if the degrees in L are d and any set of vertex $S \subset L$ of size $|S| \le n/d$ has at least $\beta|S|$ neighbors in R.
Let $d \ge 4$ and G be a bipartite graph obtained by randomly choosing d edges for each vertex in L, then G is a (d, d/4)-expander with constant positive probability.

$$Pr[\exists S, T; |S| \le n/d, T \le \beta|S|] < \sum_{s=1}^{n/d} \binom{n}{s} \binom{n}{\beta s} (\frac{\beta s}{n})^{ds}$$

2.1.3 Eigenvalue Bounds on Expansion

$$h(G) \geq \frac{1}{2}(d - \lambda_2)$$

Proof For any subset of the vertices S let $\mathbf{x} = (n-s)\mathbf{1}_S - s\mathbf{1}_{\overline{S}}$. Then $\mathbf{x^T x} = (n-s)^2 s + s^2(n-s) = s(n-s)n$.

$$\begin{aligned}
\mathbf{x^T} A\mathbf{x} &= 2\sum_{(i,j)\in E} x_i x_j = 2(n-s)^2 e(S) - 2s(n-s)e(S,\overline{S}) + 2s^2 e(\overline{S}) \\
&= (n-s)^2(ds - e(S,\overline{S})) - 2s(n-s)e(S,\overline{S}) + s^2(d(n-s) - e(S,\overline{S})) \\
&= dns(n-s) - n^2 e(S,\overline{S})
\end{aligned}$$

$$\lambda_2 \geq \frac{x^T A x}{x^T x} = \frac{dns(n-s) - n^2 e(S,\overline{S})}{s(n-s)n} = d - \frac{ne(S,\overline{S})}{s(n-s)}$$

For $s \leq \frac{n}{2}$,

$$\frac{e(S,\overline{S})}{|S|} \geq \frac{n-s}{n}(d - \lambda_2) \geq \frac{1}{2}(d - \lambda_2)$$

$(d - \lambda_2)$ is called the **spectral gap**

$$h(G) \leq \sqrt{d(d - \lambda_2)}$$