

Notes on Spectral Graph Theory

Mufei Li

October 2019

This notes is based on the reading group on spectral graph theory at AWS Shanghai AI Lab, Fall 2019. Our primary reference are [1] and [2], accompanied by [3]. Some sections have not been covered directly in the reading group and I mark the titles of them with an *.

Contents

1 Background	1
1.1 Linear Algebra	1
2 Introduction	2
2.1 Laplacian Matrices	2

1 Background

This section is mainly for a preparation and a refresh of the memory. You can skip it first and go back if necessary.

1.1 Linear Algebra

Theorem 1.1 (Variational Characterization of Eigenvalues). Let $M \in \mathbb{R}^{n \times n}$ be a symmetric matrix, and $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the eigenvalues of M . Then

$$\lambda_k = \min_{k\text{-dim } V} \max_{\mathbf{x} \in V \setminus \{\mathbf{0}\}} \frac{\mathbf{x}^T M \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \quad (1)$$

The quantity $\frac{\mathbf{x}^T M \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$ is called the *Rayleigh quotient* of \mathbf{x} with respect to M , and we will denote it by $R_M(\mathbf{x})$.

Proof. By the spectral theorem, let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be eigenvectors of M corresponding to eigenvalues $\lambda_1, \dots, \lambda_n$ that form an orthonormal basis of \mathbb{R}^n . $\forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}, \mathbf{x} = \sum_{i=1}^n a_i \mathbf{v}_i, \mathbf{x}^T \mathbf{x} = \sum_{i=1}^n a_i^2, \mathbf{x}^T M \mathbf{x} = \sum_{i=1}^n \lambda_i a_i^2$.

Consider V a space spanned by $\mathbf{v}_1, \dots, \mathbf{v}_k$. It follows that $\mathbf{x}^T M \mathbf{x} \leq \lambda_k \sum_{i=1}^n a_i^2$ and $R_M(x) \leq \lambda_k$. In particular λ_k is obtained when $\mathbf{x} = a_k \mathbf{v}_k$. Thus $\lambda_k \geq \min_{k\text{-dim } V} \max_{\mathbf{x} \in V \setminus \{\mathbf{0}\}} \frac{\mathbf{x}^T M \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$.

For the other direction, let V be any k -dim space. Let S be the span of $\mathbf{v}_k, \dots, \mathbf{v}_n$. Then $\exists \mathbf{x} \in V \cap S, \mathbf{x} \neq \mathbf{0}$. Since $\mathbf{x} = \sum_{i=k}^n b_i \mathbf{v}_i, R_M(\mathbf{x}) \geq \lambda_k$. \square

Remark 1.1. Several facts can be immediately obtained from Theorem 1.1:

1. If λ_1 is the smallest eigenvalue of a real symmetric matrix M , then $\lambda_1 = \min_{\mathbf{x} \neq \mathbf{0}} R_M(\mathbf{x})$. Furthermore, every minimizer is an eigenvector of λ_1 .
2. If λ_n is the largest eigenvalue of a real symmetric matrix M , then $\lambda_n = \max_{\mathbf{x} \neq \mathbf{0}} R_M(\mathbf{x})$. Furthermore, every maximizer is an eigenvector of λ_n .
3. If λ_1 is the smallest eigenvalue of a real symmetric matrix M , and \mathbf{x}_1 is an eigenvector of λ_1 , then $\lambda_2 = \min_{\mathbf{x} \neq \mathbf{0}, \mathbf{x} \perp \mathbf{x}_1} R_M(\mathbf{x})$. Furthermore, every minimizer is an eigenvector of λ_2 .

2 Introduction

2.1 Laplacian Matrices

A graph G is specified by its vertex set V and edge set E . Given an undirected graph $G = (V, E)$, the approach of spectral graph theory is to associate a symmetric real-valued matrix to G , and to relate the eigenvalues of the matrix to combinatorial properties of G .

The most natural matrix associated to G is the weighted adjacency matrix A such that $A_{ij} = A_{ji} > 0$ if $(i, j) \in E$ and $A_{ij} = 0$ otherwise. Note that unless otherwise specified, the edges are unique up to symmetry in E , i.e. (i, j) and (j, i) are considered to be the same edge in E . We define the **Laplacian matrix** of G to be

$$L = D - A, \quad (2)$$

where the degree matrix D is a diagonal matrix with $D_{ii} = A_i \mathbf{1}$.

The Laplacian matrix gives rise to a useful quadratic form associated with a graph:

$$\mathbf{x}^T L \mathbf{x} = \sum_{(u,v) \in E} A_{uv} (x_u - x_v)^2. \quad (3)$$

If we consider \mathbf{x} to be a function $\mathbf{x} : V \rightarrow \mathbb{R}$, then this form measures the smoothness of the function and the evaluated result is small if \mathbf{x} does not change much over any edge (with large positive weights). From the perspective of graph cut, if $\mathbf{x} \in \{0, 1\}^V$ is a Boolean vector representing a cut in the graph and $A_{ij} = A_{ji} = 1, \forall (i, j) \in E$, then the evaluated result counts the number of edges that cross the cut. It's also quite often that we consider instead normalized Laplacian matrices $D^{-1}L$ or $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$.

Remark 2.1 (Properties of Laplacian Matrices). Laplacian matrices have several properties:

1. They are positive semidefinite. It follows that their eigenvalues are non-negative.
2. They are additive. For a graph $G = (V, E)$, denote its Laplacian matrix by L_G . An edge (i, j) with $i, j \in V$ can be viewed as a graph $\Delta G = (V, \{(i, j)\})$. Let \tilde{G} be the graph obtained by adding the edge (i, j) to G , we have $L_{\tilde{G}} = L_G + L_{\Delta G}$.
3. Their eigenvalues are invariant under graph isomorphism. If we permute the vertices then the eigenvectors are similarly permuted. Let P be a permutation matrix, then

$$L\mathbf{v} = \lambda\mathbf{v} \text{ iff } (PLP^T)(P\mathbf{v}) = \lambda P\mathbf{v}. \quad (4)$$

Proof.

1. That Laplacian matrices are positive semi-definite comes from the quadratic form we derived above. The existence of eigenvalues is justified by the real Spectral Theorem. Since the matrices are positive semi-definite, the eigenvalues are non-negative.
2. This can be verified by applying the definition.

3. Note that $P^T P = I$.

□

Below We prove several relations between the Laplacian matrices and certain purely combinatorial properties of the corresponding graphs. We will start with unnormalized Laplacian matrices and then extend the results to the normalized cases.

Theorem 2.1. Let $G = (V, E)$ be a weighted undirected graph and denote its unnormalized Laplacian matrix by L . Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the real eigenvalues of L with multiplicities in nondecreasing order. Then

1. $\lambda_1 = 0$.
2. $\lambda_k = 0$ iff G has at least k connected components.
3. Let D denote the degree matrix of G and let $D_{\max} = \max_{v \in V} D_{vv}$. Then $\lambda_n \leq 2D_{\max}$. In particular, the equality holds when all nodes have the same degree (i.e. the graph is a D_{\max} -regular graph) and at least one of the connected components of G is bipartite.

Note that the second result suggests that the multiplicity of 0 as an eigenvalue is precisely the number of connected components of G .

Proof.

1. By the variational characterization of eigenvalues, we have

$$\lambda_1 = \min_{\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} R_L(\mathbf{x}) = \min_{\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{\sum_{(u,v) \in E} A_{uv}(x_u - x_v)^2}{\sum_{v \in V} x_v^2} \geq 0. \quad (5)$$

If we take \mathbf{x} to be an all-one vector $\mathbf{1} = (1, \dots, 1)$, then $R_L(\mathbf{x}) = 0$. Hence $\lambda_1 = 0$.

2. By the variational characterization of eigenvalues, we have

$$\lambda_k = \min_{k\text{-dim } S} \max_{\mathbf{x} \in S \setminus \{\mathbf{0}\}} R_L(\mathbf{x}) \quad (6)$$

If $\lambda_k = 0$, then $\exists S$ a k -dim subspace such that $\forall \mathbf{x} \in S$, it is constant within each connected component. Therefore the graph has at least k connected components.

Conversely, assume G has at least k connected components. Let $\mathbf{1}^{(i)} \in \mathbb{R}^n$ be a vector such that $\mathbf{1}_v^{(i)} = 1$ if v belongs to the i -th connected component and 0 otherwise. Consider S to be the span of $\mathbf{1}^{(1)}, \dots, \mathbf{1}^{(k)}$, then $R_L(\mathbf{x}) = 0 \forall \mathbf{x} \in S$. Hence $\lambda_k = 0$.

- 3.

$$\mathbf{x}^T L \mathbf{x} = \sum_{(u,v) \in E} A_{uv}(x_u - x_v)^2 = \sum_{v \in V} D_{vv} x_v^2 - 2 \sum_{(u,v) \in E} A_{uv} x_u x_v \quad (7)$$

$$= 2 \sum_{v \in V} D_{vv} x_v^2 - \left(\sum_{v \in V} D_{vv} x_v^2 + 2 \sum_{(u,v) \in E} A_{uv} x_u x_v \right) \quad (8)$$

$$= 2 \sum_{v \in V} D_{vv} x_v^2 - \sum_{(u,v) \in E} A_{uv} (x_u + x_v)^2 \quad (9)$$

$$\leq 2D_{\max} \sum_{v \in V} x_v^2 - \sum_{(u,v) \in E} A_{uv} (x_u + x_v)^2 \quad (10)$$

By the variational characterization of eigenvalues,

$$\lambda_n = \max_{\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} R_L(\mathbf{x}) \leq \max_{\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} 2D_{\max} - \frac{\sum_{(u,v) \in E} A_{uv} (x_u + x_v)^2}{\sum_{v \in V} x_v^2} \leq 2D_{\max}. \quad (11)$$

The equality holds when $D_{vv} = D_{\max}, \forall v \in V$ and $\sum_{(u,v) \in E} A_{uv}(x_u + x_v)^2 = 0$. Since $\mathbf{x} \neq \mathbf{0}$, $\exists U = \{v | v \in V, x_v = c\}$ and $V = \{v | v \in V, x_v = -c\}$ for some $c \neq 0$ s.t. the nodes in U are only connected to some nodes in V and vice versa. In other words, they form a connected component that is bipartite.

□

Theorem 2.2. Let $G = (V, E)$ be a weighted undirected graph. Let $\hat{L} = D^{-1}L$ and $\tilde{L} = D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$ be its two normalized Laplacian matrices. Then

1.

$$\mathbf{x}^T \tilde{L} \mathbf{x} = \sum_{(u,v) \in E} A_{uv} \left(\frac{x_u}{\sqrt{D_{uu}}} - \frac{x_v}{\sqrt{D_{vv}}} \right)^2 \quad (12)$$

2. λ is an eigenvalue of \hat{L} with eigenvector u iff $Lu = \lambda Du$.

3. λ is an eigenvalue of \hat{L} with eigenvector u iff λ is an eigenvalue of \tilde{L} with eigenvector $w = D^{\frac{1}{2}}u$.

4. 0 is an eigenvalue of \hat{L} with eigenvector $\mathbf{1}$ and an eigenvalue of \tilde{L} with eigenvector $D^{\frac{1}{2}}\mathbf{1}$.

5. \hat{L} and \tilde{L} are positive semi-definite and have n non-negative real-valued eigenvalues $0 = \lambda_1 \leq \dots \leq \lambda_n$.

6. The multiplicity of the eigenvalue 0 of both \hat{L} and \tilde{L} equals the number of connected components in G .

7. $\lambda_n \leq 2$. In particular, the equality holds when at least one of the connected components of G is bipartite.

Proof.

1.

$$\mathbf{x}^T \tilde{L} \mathbf{x} = \mathbf{x}^T (D^{-\frac{1}{2}} L D^{-\frac{1}{2}}) \mathbf{x} = \mathbf{x}^T (D^{-\frac{1}{2}} (D - A) D^{-\frac{1}{2}}) \mathbf{x} = \mathbf{x}^T \mathbf{x} - \mathbf{x}^T D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \mathbf{x} \quad (13)$$

$$= \mathbf{x}^T \mathbf{x} - 2 \sum_{(u,v) \in E} \frac{A_{uv}}{\sqrt{D_{uu}} \sqrt{D_{vv}}} x_u x_v \quad (14)$$

$$= \sum_{(u,v) \in E} \frac{A_{uv}}{D_{uu}} x_u^2 + \frac{A_{uv}}{D_{vv}} x_v^2 - 2 \frac{A_{uv}}{\sqrt{D_{uu}} \sqrt{D_{vv}}} x_u x_v \quad (15)$$

$$= \sum_{(u,v) \in E} A_{uv} \left(\frac{x_u}{\sqrt{D_{uu}}} - \frac{x_v}{\sqrt{D_{vv}}} \right)^2 \quad (16)$$

2. The verification takes a direct computation.

3.

$$\hat{L}u = D^{-1}Lu = \lambda u \iff D^{-1}LD^{-\frac{1}{2}}w = \lambda D^{-\frac{1}{2}}w \iff \tilde{L}w = \lambda w \quad (17)$$

4. $\hat{L}\mathbf{1} = D^{-1}L\mathbf{1} = \mathbf{1} - D^{-1}A\mathbf{1} = 0$. Hence $\mathbf{1}$ is an eigenvector of \hat{L} with eigenvalue 0. Apply the previous result we have that $D^{\frac{1}{2}}\mathbf{1}$ is an eigenvector of \tilde{L} with eigenvalue 0.

5. That \tilde{L} is positive semi-definite is from the first result. The existence of its n eigenvalues is justified by the spectral theorem. Since it's positive semi-definite, all of its eigenvalues are non-negative. By the third result, the eigenvalues of \hat{L} are also non-negative, hence \hat{L} is positive semi-definite.

6. We only need to prove the case for \tilde{L} and the case for \hat{L} will follow from the third result.

By the variational characterization of eigenvalues, we have

$$\lambda_k = \min_{k-\dim S} \max_{\mathbf{x} \in S \setminus \{\mathbf{0}\}} R_{\tilde{L}}(\mathbf{x}) = \min_{k-\dim S} \max_{\mathbf{x} \in S \setminus \{\mathbf{0}\}} \frac{\sum_{(u,v) \in E} A_{uv} \left(\frac{x_u}{\sqrt{D_{uu}}} - \frac{x_v}{\sqrt{D_{vv}}} \right)^2}{\sum_{v \in V} x_v^2} \quad (18)$$

If $\lambda_k = 0$, then $\exists S$ a k -dim subspace such that $\forall \mathbf{x} \in S$, $D^{-\frac{1}{2}}\mathbf{x}$ is constant within each connected component. Therefore the graph has at least k connected components.

Conversely, assume G has at least k connected components. Let $D^{\frac{1}{2}}\mathbf{1}^{(i)} \in \mathbb{R}^n$ be a vector such that $(D^{\frac{1}{2}}\mathbf{1}^{(i)})_v = \sqrt{D_{vv}}$ if v belongs to the i -th connected component and 0 otherwise. Consider S to be the span of $D^{\frac{1}{2}}\mathbf{1}^{(1)}, \dots, D^{\frac{1}{2}}\mathbf{1}^{(k)}$, then $R_{\tilde{L}}(\mathbf{x}) = 0 \ \forall \mathbf{x} \in S$. Hence $\lambda_k = 0$.

It follows that $\lambda_k = 0$ iff G has at least k connected components. Hence the multiplicity of the eigenvalue 0 of \tilde{L} equals the number of connected components in G .

7.

$$\mathbf{x}^T \tilde{L} \mathbf{x} = \sum_{(u,v) \in E} A_{uv} \left(\frac{x_u}{\sqrt{D_{uu}}} - \frac{x_v}{\sqrt{D_{vv}}} \right)^2 = \sum_{v \in V} x_v^2 - 2 \sum_{(u,v) \in E} \frac{A_{uv}}{\sqrt{D_{uu}}\sqrt{D_{vv}}} x_u x_v \quad (19)$$

$$= 2 \sum_{v \in V} x_v^2 - \left(\sum_{v \in V} x_v^2 + 2 \sum_{(u,v) \in E} \frac{A_{uv}}{\sqrt{D_{uu}}\sqrt{D_{vv}}} x_u x_v \right) \quad (20)$$

$$= 2 \sum_{v \in V} x_v^2 - \sum_{(u,v) \in E} A_{uv} \left(\frac{x_u}{\sqrt{D_{uu}}} + \frac{x_v}{\sqrt{D_{vv}}} \right)^2 \quad (21)$$

By the variational characterization of eigenvalues,

$$\lambda_n = \max_{\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} R_{\tilde{L}}(\mathbf{x}) \leq \max_{\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} 2 - \frac{\sum_{(u,v) \in E} A_{uv} \left(\frac{x_u}{\sqrt{D_{uu}}} + \frac{x_v}{\sqrt{D_{vv}}} \right)^2}{\sum_{v \in V} x_v^2} \leq 2 \quad (22)$$

The equality holds when $\sum_{(u,v) \in E} A_{uv} \left(\frac{x_u}{\sqrt{D_{uu}}} + \frac{x_v}{\sqrt{D_{vv}}} \right)^2 = 0$. Since $\mathbf{x} \neq \mathbf{0}$, $\exists U = \{v | v \in V, \frac{x_v}{\sqrt{D_{vv}}} = c\}$ and $V = \{v | v \in V, \frac{x_v}{\sqrt{D_{vv}}} = -c\}$ for some $c \neq 0$ s.t. the nodes in U are only connected to some nodes in V , vice versa. In other words, they form a connected component that is bipartite.

□

References

- [1] Luca Trevisan. Lecture notes on graph partitioning, expanders and spectral methods, 2016.
- [2] Daniel A. Spielman. Lecture notes on spectral graph theory, 2018.
- [3] Ulrike von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007.