PROPERTIES AND APPLICATIONS OF GRAPH LAPLACIANS

HANCHEN LI

Abstract. Laplacian matrices are widely studied in spectral graph theory to gain understanding of graphs with results from linear algebra. This paper aims to introduce properties of the graph Laplacian and show how these properties can be utilized to help generate insights about graphs with respect to the applications of graph partitioning and more.

Contents

1. Introduction and Basic Background

Graphs are mathematical structures that depict the pairwise relationship between objects. They are very helpful tools when modeling computer networks, protein structures, electric circuits, etc. The idea of spectral graph theory is to first represent a graph as a matrix, and then study the properties of the graph by studying the spectrum of this matrix representation. In spectral graph theory, the central object of study is the graph Laplacian matrix. In this paper, we will introduce the Laplacian matrix and explain how its properties can be utilized to provide insights about the structure of a graph.

We will begin by giving some basic definitions, then we will show several important properties of the Laplacian. After this, we will utilize the spectrum to study a very important application: graph partitioning. In this paper, background in basic linear algebra is assumed. We recommend [1] as a reference.

2. Basic Definitions of Graph and Laplacian Matrix

In this section, we will define some basic concepts and introduce two equivalent definitions of the graph Laplacian.

Definition 2.1. A *graph* is an unordered pair $G = (V, E)$ of two sets. The set V denotes the *vertices*, represented as an ordered set:

$$V = \{v_1, v_2, ..., v_n\}$$

while the set E denotes the set of *edges*, which is an ordered set of edges, represented as:

$$E = \{E_1, E_2, ..., E_m.\}$$

Depends on whether a graph is directed or undirected, the edges have different definitions. The edges in undirected graphs connect pairs of vertices without orientation, thus every edge E_k is defined by a set of two vertices $\{v_i, v_j\}$. However, in a directed graph, there are two possible edges between two vertices by a difference of orientation. In this case, each edge E_k is defined with an ordered pair (v_i, v_j)

We will continue to present the definition of the weight of the edges.

Definition 2.2. The *weight* of an edge E_k is a real number associated with the edge E_k. We represent the weight in $Weight(E_K)$. In undirected graphs, it can also be represented as $Weight(v_i, v_j)$.

We now define some basic concepts.

Definition 2.3. A *path* is a non-empty graph $P = (V', E')$ of the form

$$V' = \{v_{a_0}, v_{a_1}, ..., v_{a_p}\} \subset V,$$

$$E' = \{\{v_{a_0}, v_{a_1}\}, \{v_{a_1}, v_{a_2}\}, ..., \{v_{a_{p-1}}, v_{a_p}\}\} \subset E,$$

where $\{a_1, a_2, ..., a_p\}$ is a distinct sequence of integers from 1 to n.

When there exists a path that starts with x_i and ends with x_j, we say x_i and x_j are connected by a path. With this definition, we can define a connected component.

Definition 2.4. A *connected component* of an undirected graph is a maximal set of nodes such that each pair of nodes is connected by a path.

Now we are going to get into "spectral" part of the spectral graph theory. First, we define concepts of adjacency, degree and incidence matrices.

Definition 2.5. In a graph $G = (V, E)$, two vertices x, y are *adjacent* if $\{x, y\} \in E$. The *adjacency matrix* A is a matrix of dimension $n(V) \times n(V)$. It is defined as

$$A_{ij} = \begin{cases} 1, & \{i, j\} \in E \\ 0, & \text{otherwise.} \end{cases}$$

Definition 2.6. The *degree* of a vertex x is the number of vertices that are adjacent to it. Denote this as $d(x)$. The *degree matrix* D is a diagonal matrix of dimension $n(V) \times n(V)$, defined by

$$D_{ij} = \begin{cases} d(i), & i = j \\ 0, & \text{otherwise.} \end{cases}$$

Definition 2.7. The *incidence matrix* B of a graph $G = (V, E)$ is a matrix of dimension $n(v) \times n(E)$. The definition is

$$B_{ik} = \begin{cases} 1, & (i, j) \in E_k \text{ for some } j \\ -1, & (j, i) \in E_k \text{ for some } j \\ 0, & \text{otherwise.} \end{cases}$$

After defining these three matrices, we finally come to define the Laplacian. We will first start with the more typical definition.

Definition 2.8. The *Laplacian matrix* of $G = (V, E)$ is defined as

$$L = D - A.$$

Note that the adjacency matrix has all zeros on its diagonal entries and the degree matrix is diagonal.

Proposition 2.9. *The Laplacian matrix can also be written as*

$$L = BB^T,$$

where B is the incidence matrix.

Proof. Note:

$$(BB^T)_{ij} = \sum_{p=1}^{k} B_{ik}B_{jk} \tag{2.10}$$

Applying the definition of incidence matrix, we have the equation that:

$$B_{ik}B_{jk} = \begin{cases} 1, & i = j, (i,j) \in E_k \\ -1, & i \neq j, (j,i) \in E_k \\ 0, & \text{otherwise.} \end{cases}$$

So

$$(BB^T)_{ij} = \begin{cases} \sum_{k|i \in E_k} 1, & i = j \\ -1, & i \neq j, (i,j) \in E \\ 0, & \text{otherwise.} \end{cases}$$

Note that $\sum_{k|i\in E_k} 1$ is just the degree of the i^{th} vertex, so

$$BB^T = D - A = L.$$

□

Because of this equivalence, we can utilize both definitions based on our needs later.

3. Properties of The Graph Laplacian Matrix

In this section, we are going to introduce several properties of the Laplacian matrix. The first thing we want to show is that the Laplacian matrix L is symmetric and positive semi-definite.

The symmetry follows from Definition 2.8. Since D and A are both symmetric, we get that L is symmetric.

Theorem 3.1. *The Laplacian matrix L is a positive semi-definite matrix.*

Proof. We want to show that for all $x \in \mathbb{R}^n, x^T Lx \geq 0$.

Recall from Definition 2.8 that $L = BB^T$, so

$$\begin{aligned} x^T Lx &= x^T(BB^T)x \\ &= (B^T x)^T(B^T x). \end{aligned} \tag{3.2}$$

Now following the definition of the incidence matrix, we then see that the kth element of the vector $B^T x$ is equal to $x_i - x_j$ where (i, j) is E_k. Then

$$\begin{aligned} x^T Lx &= (B^T x)^T(B^T x) \\ &= \sum_{(i,j)\in E} (x_i - x_j)^2 \\ &\geq 0. \end{aligned} \tag{3.3}$$

□

Now we will utilize some linear algebra results. Since L is real and symmetric, all its eigenvalues are real and its eigenvectors are orthogonal. Since L is also positive semi-definite, all its eigenvalues are non-negative too.

The eigenvalues of L are very useful objects to study since they contain information about the structures of a graph. We are going to show several examples in the rest of this section.

First, we consider the cases where one of the eigenvalues is zero.

Theorem 3.4. *There always exists an eigenvalue of the Laplacian matrix that is equal to 0.*

Proof. Consider the vector $v = (1, 1, ..., 1)$. We have that the kth entry of the vector Lv is equal to

$$\begin{aligned}\sum_{i=1}^{n} L_{ki} &= \sum_{i=1}^{n} D_{ki} - A_{ki} \\ &= D_{kk} - \sum_{i=1}^{n} A_{ki}.\end{aligned} \tag{3.5}$$

This is equal to 0, since D_{kk} is defined to be the degree of the kth vertex and A_{ki} is equal to 1 if and only if vertex i is adjunct to vertex k. Thus, we have that $Lv = 0v$ □

We now present a more general and powerful form of this theorem.

Theorem 3.6. *The number of zero eigenvalues of the Laplacian L is the same as the number of connected components of the graph.*

Proof. First, we try to show that the multiplicity of zero eigenvalues is greater or equal to the number of connected components. Suppose there are k connected components. Call them $S_1, ..., S_k$. Define k vectors $v_1, ..., v_k$ such that

$$v_i(j) = \begin{cases} \frac{1}{\sqrt{|S_i|}}, & j \in S_i \\ 0, & \text{otherwise.} \end{cases}$$

For $i = 1, ..., k$, it holds that $\|v_i\| = 1$. Additionally, for $i \neq j$, since S_i, S_j are disjoint, we have $\langle v_i, v_j \rangle = 0$. Finally, note

$$(Lv_i)_k = \sum_{i=1}^{n} L_{ki} v_i(k).$$

For $k \in S_i$, $L_{ki} = 0$ when $i \notin S_i$. For $k \notin S_i$, we have $v_i(k) = 0$.

Therefore, L_{ki} and $V_i(k)$ have disjoint support over k, and hence Lv_i is equal to zero. This means that there are at least k orthogonal eigenvectors with eigenvalue 0 that are eigenvectors of L.

Now we prove the multiplicity of zero eigenvalues is less or equal to the number of connected components. Recall from Equation 3.3 that:

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

This is 0 only if x is constant on every connected component. Then suppose there exists a $(k+1)th$ vector that satisfies this condition. The vector must be non-zero on some entries. Thus, the vector need to be non-zero and constant on all entries that belong to the corresponding connected component. Thus, the vector cannot be orthogonal to the group of vectors v_i. Then we have a contradiction. So the multiplicity of zero eigenvalues is less or equal to the number of connected components. $\square$

4. Graph Partitioning

We will now show that we can use the graph Laplacian to measure how well a graph can be separated into two parts. Again, let us first define some new concepts related to graph partitioning.

Let S be a vertex subset of a graph. To measure how well S can be separated from the graph, it is natural to think about the number of edges connecting S to the rest of the graph. However, we should also take the size of S into consideration. For example, the number of edges connecting a single point with the rest of the graph is the same for both the point and the rest, but the "difficulty" of separating these two from the rest of the graph should not be the same.

This motivates us to define the following concept to measure how well a subset S can be separated from the rest of a graph G as the following:

Definition 4.1. Given a subset S of the vertices of a graph, the *isoperimetric ratio* of S is defined by:

$$\theta(S) = \frac{|(i,j) \in E, i \in S, j \notin S|}{|S|}$$

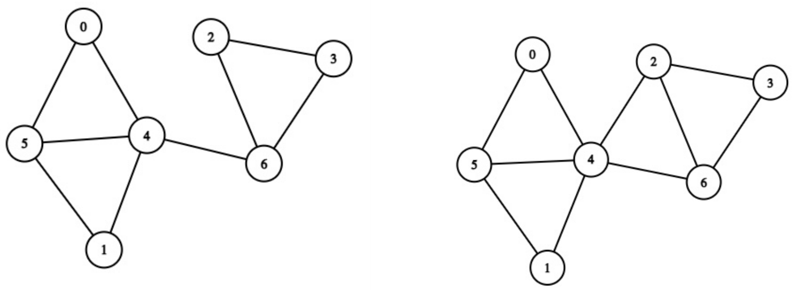
Note that the numerator calculates the number of edges connecting S from its complement, while the denominator is normalized by the size of the subset.

We can also define the *isoperimetric number* for a graph.

Definition 4.2. The *isoperimetric number* of a graph G is the minimum *isoperimetric ratio* over all sets of at most half of the vertices.

$$\theta_G = \min_{|S| \leq \frac{n}{2}} \theta(S)$$

Notice that we only consider subsets with element number less or equal to half of the total degree n. This is because for subset S_1 with $|S_1| \leq n/2$ and its complement, $|(i,j) \in E, i \in S, j \notin S|$ are the same, but the denominators in the isoperimetric ratio are different. Choosing S_1^C always gives a smaller isoperimetric ratio, so we only consider subsets of size less or equal to $n/2$ when taking the minimum.



(A) Graph A (B) Graph B

Figure 1. Two Graph Examples

We are going to present an example of these ideas with respect to the graphs in Figure 1. Consider graph A and the subset of $\{2, 3, 6\}$, the isoperimetric ratio of this set should be $\frac{1}{3}$ since there is only one edge connecting the subset with its complement, and there are 3 elements in the subset. The isoperimetric number of the graph is equal to the isoperimetric ratio of this subset, since the ratio of this subset is the minimum among all other subsets with element number less or equal to half of the total number.

We can compare this graph to the graph B, which has an isoperimetric number of 2/3, taking the same subset $\{2, 3, 6\}$. Our visual intuition also suggests that graph B is harder to separate than graph A.

Now we have a mathematical measure of how easily a graph can be separated into two parts. We will then relate it with the spectrum of the Laplacian matrix. It turns out that θ_G is closely related to the second-smallest eigenvalue. (Recall that there always exist a zero eigenvalue.)

Let us first derive a lower bound on θ_G.

Theorem 4.3. *Let λ_2 be the second-smallest eigenvalue of the Laplacian matrix L. We have the following lower bound for the isoperimetric ratio of the graph given a subset S:*

$$\theta(S) \geq \lambda_2 / \left(1 - \frac{|S|}{n}\right)$$

Before we start proving this theorem, we will have to utilize a result from linear algebra. The name of the full theorem is Courant-Fischer. The version we are using will only be a special case.

Theorem 4.4.

$$\lambda_2 = \min_{x : x^T \mathbf{1} = 0} \frac{x^T L x}{x^T x}$$

The Courant-Fischer theorem is a famous theorem in linear algebra with multiple proofs. One recommended proof is on page 211 of [1].

Now with this result, we are going to prove Theorem 4.3.

Proof. From Theorem 4.4, for all non-zero x orthogonal to $\mathbf{1}$ we know

$$\lambda_2 \leq \frac{x^T L x}{x^T x}. \tag{4.5}$$

Now given an arbitrary set of vertices S, we create a vector v_s defined as follows:

$$v_s(i) = \begin{cases} 1, & i \in S \\ 0, & \text{otherwise.} \end{cases}$$

We can compute

$$\begin{aligned} v_s^T L v_s &= \sum_{(i,j)\in E} (v_s(i) - v_s(j))^2 \\ &= |(i,j) \in E : i \in S, j \notin S| \\ &= \theta(S)|S|. \end{aligned} \tag{4.6}$$

Note that the theorem requires vectors to be orthogonal to $\mathbf{1}$. Thus, we create a vector x_s from v_s such that

$$x_s(i) = \begin{cases} 1 - \frac{|S|}{n}, & i \in S \\ -\frac{|S|}{n}, & \text{otherwise.} \end{cases} = v_s(i) - \frac{|S|}{n}$$

Notice that $x_s^T \mathbf{1} = 0$, and we have from Equation 3.3

$$\begin{aligned} x_s^T L x_s &= \sum_{(i,j)\in E} \left(\left(v_s(i) - \frac{|S|}{n} \right) - \left(v_s(j) - \frac{|S|}{n} \right) \right)^2 \\ &= \sum_{(i,j)\in E} (v_s(i) - v_s(j))^2 \\ &= \theta(S)|S|. \end{aligned} \tag{4.7}$$

Note that:

$$\begin{aligned} x_s^T x_s &= \left(1 - \frac{|S|}{n}\right)^2 |S| + \frac{|S|^2}{n^2}(n - |S|) \\ &= n\left(\frac{|S|}{n} - \left(\frac{|S|}{n}\right)^2 \right). \end{aligned} \tag{4.8}$$

Applying this to (4.5) gives us

$$\lambda_2 \leq \frac{x_s^T L x_s}{x_s^T x_s} = \frac{\theta(S)}{1 - \frac{|S|}{n}}. \tag{4.9}$$

Moving the denominator to the left side and using the fact $|S| \leq n/2$, we get that:

$$\theta(S) \geq \lambda_2 \cdot \left(1 - \frac{|S|}{n}\right) \geq \lambda_2/2. \tag{4.10}$$

This holds for all S, thus we get a lower bound for the isoperimetric number of the graph:

$$\theta_G \geq \lambda_2/2. \tag{4.11}$$

□

This lower bound gives us a relationship between the second-smallest eigenvalue and the separateness of the graph.

This result is consistent with Theorem 3.5. Notice if we have two connected components in the graph, then the second-smallest eigenvalue of the Laplacian matrix is 0. This conforms to the fact that we could have a subset of vertices that are completely disconnected with its complement.

However, there is a problem with the isoperimetric number. Going back to the definition of isoperimetric ratio of a vertex subset, we see that $\theta(S)$ and $\theta(S^c)$ are not equal even when both capture the same information of "how much the graph is connected" after we separate in this way.

We now introduce a notion of graph connectivity of edge-weighted graphs that is symmetric for subsets and complements. The definition is called the *conductance*:

Definition 4.12. We define the *conductance* of a subset S to be:

$$\phi(s) = \frac{\sum\limits_{(v_i,v_j)|(v_i,v_j)\in E, v_i\in S, v_j\in S^c} \text{Weight}(v_i, v_j)}{\min(d(S), d(S^c))}$$

This definition is similar to Definition 4.1. Both fractions have a value measuring the connectedness of the set from its complement as numerator and a value measuring the size of the sets on the denominator. However, the numerator also captures the weight of the edges and the denominator is the smaller one of S and S^C. This solves the potential symmetric problem we talked about.

Definition 4.13. The *conductance* of a graph is defined as:

$$\phi_G = \min_{S\subset V} \phi(S).$$

Generally speaking, the definition of conductance is more useful when we are considering problems when the relationships among vertices are not equivalent since they take into account of the edge weights, while the isoperimetric number is more useful when we only consider vertices because it only cares about whether vertices are connected but not how.

The conductance is also closely related with the spectrum of Laplacian. However, we need to use the normalized Laplacian N.

Definition 4.14. We define the *normalized Laplacian* N with respect to the degree matrix D as:

$$N = D^{-1/2} L D^{-1/2},$$

where $D^{-1/2}$ represents the matrix whose diagonal entries are square roots of the degree matrix and zero elsewhere.

We will refer to the eigenvalues of N as λ_{N1}, λ_{N2}, ...

The normalized Laplacian matrix is often used when the graph is not regular, which means the vertices have different degrees.

For the conductance of the graph and the normalized Laplacian, there are both an upper bound and lower bound:

Theorem 4.15 (Cheeger's Inequality)**.** *For all subset $S \subset V$, we have the following inequality for $\phi(S)$:*

$$\frac{\lambda_{N2}}{2} \le \phi(S) \le \sqrt{2\lambda_{N2}}.$$

The left half of the inequality is proven using the same strategy as used in Theorem 4.4: We modify v_s with a constant to make it orthogonal to $\mathbf{1}$ in order to use the Courant-Fischer theorem.

The right half can be proved with the help of the Rayleigh quotient. In fact, it is a discrete version of the original form of Cheeger's inequality, which was initially proven in the context of Riemannian Geometry. And there are some other variations (proved for the isoperimetric number instead of conductance or use eigenvalues of walk matrix instead of the normalized Laplacian). Interested readers can refer to Chapter 20, 21 of [3]

With this inequality, we are able to have a rough sense of the conductance number of the graph by only calculating λ_{N2}

This inequality can be applied to many graph clustering problems including local clustering, which involves finding small clusters of small conductance near an input vertex. The normalized Laplacian is also widely involved in other graph partitioning results, examples are studied in the first few chapters of [2].

5. More Applications of the Laplacian

Graph Laplacians have more applications. The spectrum of a graph can be used to compute Fourier transforms on graphs, so they can transform graph signals ($f : V \mapsto \mathbb{R}$) to signals in spectral space. This allows us to classify graphs and perform other operations on this informative structure. The Laplacian can also be used to group similar graphs and transform graph structure to vector spaces. Some good starting papers on these topics are [5] and [6]. The Laplacian matrix also helps in topological data analysis, which analyzes datasets using knowledge from topology. An excellent resource is [7].

Acknowledgments

It is a great pleasure to thank my mentor, Phillip Lo, for his help and understanding throughout the REU period. I learned to think and write more like a mathematician. I would also love to thank Professor Peter May and other faculties who helped to organize and gave talks. They really showed me how different fields of mathematics look like. Last but not least, I would thank Bradley Nelson, who kindly provided resources on topological data analysis during my early REU time. My digression from the resources eventually led to this paper.

References

[1] Sergei Treil. Linear Algebra Done Wrong. https://www.math.brown.edu/streil/papers/LADW/LADW.html
[2] Fan R.K.Chung. Spectral Graph Theory. https://mathweb.ucsd.edu/~fan/research/revised.html
[3] Daniel A. Spielman. Spectral and Algebraic Graph Theory. http://cs-www.cs.yale.edu/homes/spielman/sagt/
[4] Jiaqi Jiang. An Introduction to Spectral Graph Theory https://math.uchicago.edu/~may/REU2012/REUPapers/JiangJ.pdf
[5] Stefania Sardellitti. On the Graph Fourier Transform for Directed Graphs https://arxiv.org/pdf/1601.05972.pdf
[6] Mikhail Belkin, Partha Niyogi. Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering. https://proceedings.neurips.cc/paper/2001/file/f106b7f99d2cb30c3db1c3cc0fde9ccb-Paper.pdf

[7] Bradley Nelson. Topological Data Analysis. https://stat37411.github.io/