SPECTRAL GRAPH THEORY AND MATRIX TREE THEOREM

YI ZOU

Abstract. Spectral graph theory is powerful in studying graphs via linear algebra. The first half of this paper introduces readers to concepts and examples of graphs, reviews needed linear algebra knowledge, defines several important matrices in spectral graph theory, and explores properties of the Laplacian matrix. The second half of the paper discusses the Matrix-Tree Theorem and follows the proof in the reading materials of Professor Stanley's *Combinatorial Analysis* course on MIT Open Courseware.

Contents

1. Introduction

Graphs are structures formed by a set of vertices and a set of edges that connect between these vertices. For instance, flight information can be recorded as a graph, where each airport is a vertex and flights denote edges between vertices (airports). Another example could be relations on social media, where each person is a vertex and an edge is drawn if two people are friends. Graph theory wishes to extract information and properties from these structures.

Spectral graph theory offers an exciting way to study graphs by introducing linear algebra tools. Graphs are represented in matrices, and certain graph properties can be encapsulated in linear algebra properties such as eigenvalues and determinants. This paper will introduce spectral graph theory and then discuss and prove the Matrix-Tree Theorem, which counts the number of spanning trees in a graph by computing the determinant of a certain matrix.

2. Basic Definitions

We begin with basic definitions in graph theory.

Date: August 2024.

Definition 2.1. A *graph* G is a pair (V, E), where V is the set of vertices, and $E \subseteq 2^V$ is the set of edges, with each element being of size 2. In other words, each edge $e \in E$ is a two-element subset of V, i.e. $e = \{u, v\}$ for $u, v \in V$.

This paper uses this definition so that edges in graphs are assumed to be undirected ($\{u, v\}$ is unordered as a set) and there is no self-loop (each edge must involve two vertices). There is at most one edge between two vertices (a set has no duplicates within).

In later contexts, V denotes the vertice set and E denotes the edge set in graph G, unless otherwise noted. We assume V and E are finite sets.

Definition 2.2. Let G be a graph and $u \in V$. Then v is a *neighbor* of u if $\{u, v\} \in E$. We denote it as $u \sim v$, and sometimes say that u and v are *connected*.

Definition 2.3. Let G be a graph, and $u \in V$. The *degree* of u, denoted $\deg(u)$, is the number of neighbors of u.

Definition 2.4. Let $n \in \mathbb{Z}$ such that $n \geq 2$. A *path* p of a graph G is a tuple of vertices $(v_1, v_2, \ldots, v_n) \in V^n$ such that $\{v_i, v_{i+1}\} \in E$ for all integers $1 \leq i \leq n-1$.

Definition 2.5. A *cycle* of a graph G is a path $(v_1, v_2, \ldots, v_n)$ such that $v_1 = v_n$.

Definition 2.6. A graph is *connected* if for any pair $u, v \in V$, there exists a path between them.

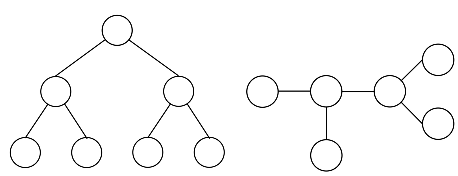
Hence, all vertices can reach others in some way, and there is no isolated vertex.

Definition 2.7. Let $G = (V_G, E_G), H = (V_H, E_H)$ be graphs. H is a *subgraph* of G if $V_H \subseteq V_G$ and $E_H \subseteq E_G$. We denote it as $H \subseteq G$.

After learning these basic concepts, we will look at two common types of graphs, trees and complete graphs.

Definition 2.8. A *tree* T is a graph that is connected and has no cycles.

The following are some common examples of trees.



The left figure is a typical binary tree, where each father node is connected to two child nodes. The tree on the right, however, has various numbers of child nodes.

Definition 2.9. Let $T = (V, E)$ be a tree. A vertex $v \in V$ is a *leaf* if it has only one neighbor.

Lemma 2.10. *There exists a leaf in any tree.*

Proof. Proceed by contradiction. Let $T = (V, E)$ be a tree of n vertices, and assume it has no leaf. Since T is connected, all $v \in V$ will have at least two neighbors. Let $v_1 \in V$, and suppose $v_2 \sim v_1$ and $v_0 \sim v_1$. Similarly, consider v_2, we will find another neighbor $v_3 \sim v_2$, and then $v_4 \sim v_3$. Following the same pattern, for each i, we can at least find $v_{i-1} \sim v_i$ and $v_{i+1} \sim v_i$ (to avoid cycles, indexes less than v_{i-1}

won't be a neighbor). At the point v_{n-1} (the n-th vertex involved), we need to find another neighbor other than $v_{n-2} \sim v_{n-1}$. While it can only be connected to one of $V \backslash \{v_{n-2}, v_{n-1}\} = \{v_0, v_1, \ldots, v_{n-3}\}$, this must form a cycle, which contradicts the definition of a tree. Thus, there must exist at least one leaf in a tree. □

Lemma 2.11. *A tree T with n vertices will have exactly $n-1$ edges.*

Proof. We will proceed by induction.
Base Case: $n = 1$

A tree with one vertex will just be a single point, without edges. This matches our claim that it has $1 - 1 = 0$ edges.
Induction Hypothesis: Let $k \in \mathbb{N}$, assume that all trees with k vertices have $k-1$ edges.
Induction Step: Consider $n = k + 1$, and let T_{k+1} be a tree with $k+1$ vertices. Then by Lemma 2.10, there exists vertex u that is a leaf of T_{k+1}.

We consider the graph $T' = T_{k+1} \backslash \{u\}$, formed by removing u and its (only) edge. We prove T' is a tree by contraposition. If T' is not a tree, then either it contains a cycle or is not connected. Then adding u back, T_{k+1} also contains a cycle or is not connected (since u, a leaf, cannot be in a cycle or the middle of a path). Thus, T_{k+1} must not be a tree, which is the opposite of our assumption.

As T' is a tree with k vertices, by induction hypothesis, T' has $k-1$ edges. Adding u and its edge back, the original tree T_{k+1} has $k - 1 + 1 = k$ edges. □

Corollary 2.12. *Let G be a graph, then*
G is connected $\iff$ G has a subgraph that is a tree and has the same vertex set. Consequently, a connected graph with n vertices has at least $n-1$ edges.

Proof. Let $G = (V, E_G)$ be a graph.
($\Leftarrow$) Assume G has a subgraph that is a tree and has the same vertex set, namely $T = (V, E_T)$. In other words, $E_T \subseteq E_G$.

As T is a tree, for any $v_1, v_2 \in V$, there exists a collection of edges in E_T that forms a path between v_1, v_2. Since this collection of edges is also in E_G, this means that G must also be connected.
($\Rightarrow$) Assume G is connected. We claim that the Breadth First Search (BFS) algorithm offers a way to create such a subgraph tree $T = (V_T, E_T)$ that covers the same vertices.

The process is the following:

We start with V_T, E_T being empty sets. Pick any vertex v, and write $S_v = \{v' \in V : v' \sim v\}$ as the set of neighbors of v.

(1) Let v be the *current vertex*. Add v to V_T.
(2) Iterate through all $v' \in S_v$. If $v' \notin V_T$, add v' to V_T and $\{v, v'\}$ to E_T (otherwise, do nothing).
(3) Pick any $v' \in S_v$ as the current vertex and repeat the same process. Stop until $V_T = V$.

The way of adding vertices ensures that the subgraph is connected, and the if-condition in step 2 ensures there is no cycle. So the subgraph is a well-defined tree.

From the forward direction, we conclude that

$$|E_G| \geq |E_T| = n - 1.$$

□

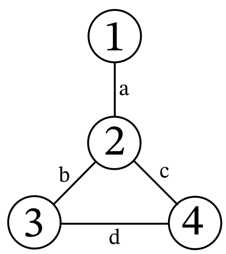
Lemma 2.13. *Let G be a graph with n vertices. If G is not a tree and has $n-1$ edges, then there must exist a cycle in G.*

Proof. By means of contradiction. Suppose there are no cycles in G. Since G is not a tree, G must not be connected, i.e. there exists a tree $T \supset G$ that uses up the same vertex set. Thus, T will have more than $n-1$ edges yet be a tree with n vertices, which contradicts Lemma 2.11. Thus, G must contain a cycle. □

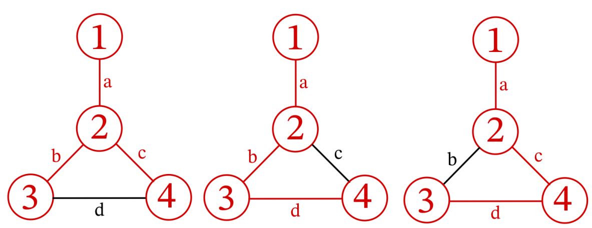
Definition 2.14. Let $G = (V_G, E_G)$ be a connected graph and $S = (V_S, E_S)$ be a subgraph. S is a *spanning tree* of G if S is a tree and $V_S = V_G$.

Basically, a spanning tree is a tree that spans all vertices in a graph. Corollary 2.12 implies that G being connected is necessary and sufficient for spanning trees to exist.

Example 2.15. We study the spanning trees in the following graph.



In this graph, since vertex 1 is only connected to 2, edge a must be included. Then to create a spanning tree that connects to vertex 3 and 4, only two edges of the triangle need to be taken. Therefore, the spanning trees are abc, abd, acd, as shown below.



Besides trees, another type of graph is the complete graph.

Definition 2.16. A graph G is a *complete graph* if any two vertices are connected by an edge, i.e. for all $u, v \in V$, $\{u, v\} \in E$. We denote a complete graph with n vertices as K_n.

Complete graphs with the same number of vertices share the same shape. The number of edges in K_n is $\binom{n}{2} = \frac{n(n-1)}{2}$, as choosing two vertices form an edge.

Observation 2.17. Let G be a graph with n vertices. Then

$$G \subseteq K_n,$$

for K_n sharing the same vertices.

3. Linear Algebra Introduction

The Spectral Theorem is an important cornerstone of linear algebra in spectral graph theory.

Theorem 3.1. *(Spectral Theorem) A real symmetric matrix has an orthonormal basis of eigenvectors.*

This version of the Spectral Theorem is modified from the version in [1].

This theorem is especially helpful when we convert a graph into a real symmetric matrix and learn about its eigenvalues. We will introduce several commonly-used matrices in spectral graph theory and then focus on the Laplacian Matrix, which is used in the proof of the Matrix-Tree Theorem.

Notation 3.2. Let $m, n, i, j \in \mathbb{Z}^+$ that $1 \leq i \leq m, 1 \leq j \leq n$.

- For a vector $v \in \mathbb{R}^m$, $v(i)$ denotes the i-th entry of v.
- For a $m \times n$ matrix M, $M(i,j)$ denotes the entry that is on the i-th row and j-th column of the matrix.

We start by defining the adjacency matrix and the degree matrix.

Definition 3.3. Let G be a graph with n vertices. Then its adjacency matrix A_G is a $n \times n$ matrix such that

$$A_G(i,j) = \begin{cases} 1, \text{ if } i \sim j \\ 0, \text{ otherwise.} \end{cases}$$

Definition 3.4. Let G be a graph with n vertices. Then its adjacency matrix D_G is a $n \times n$ matrix defined as

$$D_G(i,j) = \begin{cases} \deg(i), \text{ if } j = i \\ 0, \text{ otherwise.} \end{cases}$$

For instance, in a complete graph K_n, since all vertices have degree $n-1$, we have $D_G = (n-1) \cdot I$, where I is the $n \times n$ identity matrix.

We then define the Laplacian Matrix.

Definition 3.5. Let G be a graph with n vertices. Its Laplacian Matrix L_G is defined as $L_G = D_G - A_G$. Incorporating definitions of D_G and A_G,

$$L_G(i,j) = \begin{cases} \deg(i), \text{ if } j = i \\ -1, \text{ if } j \sim i \\ 0, \text{ otherwise.} \end{cases}$$

Lemma 3.6. *Let G be a graph of n vertices and L_G be its Laplacian matrix. For all $x \in \mathbb{R}^n$,*

$$(L_G x)(i) = \sum_{i \sim j} (x(i) - x(j)).$$

Proof.

$$\begin{aligned}(L_G x)(i) &= \sum_{j=1}^{n} L_G(i,j)x(j) \\ &= \sum_{j\sim i}((-1)\cdot x(j)) + \deg(i)x(i) \\ &= \sum_{j\sim i}((-1)\cdot x(j)) + \sum_{j\sim i} 1\cdot x(i) \\ &= \sum_{j\sim i}(x(i)-x(j))\end{aligned}$$

□

It is helpful to consider the quadratic form of the Laplacian Matrix.

Proposition 3.7. *Let G be a graph of n vertices and L_G be its Laplacian Matrix. For all $x \in \mathbb{R}^n$,*

$$x^\top L_G x = \sum_{i\sim j}(x(i)-x(j))^2.$$

Proof. By Lemma 3.6, we have

$$\begin{aligned}x^\top L_G x &= \sum_{i=1}^{n} x(i)(L_G x)(i) \\ &= \sum_{i=1}^{n}\left(x(i)\sum_{j\sim i}(x(i)-x(j))\right) \\ &= \sum_{i\sim j}[x(i)\cdot(x(i)-x(j)) + x(j)\cdot(x(j)-x(i))] \\ &= \sum_{i\sim j}(x(i)-x(j))^2.\end{aligned}$$

□

Corollary 3.8. *Let G be a graph and L_G be its Laplacian Matrix. Then L_G has an orthonormal basis of eigenvectors, and its eigenvalues, $\lambda_1, \dots, \lambda_n$, can be chosen to satisfy that $0 \le \lambda_1 \le \dots \le \lambda_n$.*

Proof. By Definition 3.5, $L_G(i,j) \in \mathbf{R}$ and $L_G(i,j) = L_G(j,i)$ for all $1 \le i,j \le n$. Thus, L_G is real and symmetric. By Theorem 3.1(Spectral Theorem), L_G has an orthonormal basis of eigenvectors. We notate its orthonormal eigenvectors as $\psi_1, \psi_2, \dots, \psi_n$, respectively corresponding to eigenvalues $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$.

For all $1 \le i \le n$, we have

$$\psi_i^\top L_G \psi_i = \psi_i^\top \lambda_i \psi_i = \lambda_i ||\psi_i||^2 = \lambda_i.$$

Meanwhile, by Proposition 3.7,

$$\lambda_i = \psi_i^\top L_G \psi_i = \sum_{a\sim b}(\psi_i(a)-\psi_i(b))^2 \ge 0.$$

Thus, all eigenvalues of L_G are non-negative. □

Corollary 3.9. *Let G be a graph and L_G be its Laplacian Matrix with eigenvalues $0 \leq \lambda_1 \leq \ldots \leq \lambda_n$. Then $\lambda_1 = 0$.*

Proof. By Lemma 3.6, for vector $\mathbf{1}$ (the all-1s vector), $L_G(i) = \sum 0 = 0$. Thus,

$$L_G \mathbf{1} = 0 \cdot \mathbf{1}.$$

This means that the smallest eigenvalue λ_1 must be zero. □

Proposition 3.10. *Let G be a graph and its Laplacian matrix L_G has eigenvalues $0 \leq \lambda_1 \leq \ldots \leq \lambda_n$. Then*

$$G \text{ is connected} \iff \lambda_2 > 0.$$

Proof. ($\Rightarrow$) Assume G is connected. Let ψ be an eigenvector such that $L_G\psi = 0$. By Proposition 3.7,

$$0 = \psi^\top L_G \psi = \sum_{i \sim j} (\psi(i) - \psi(j))^2.$$

In other words, for all $i \sim j$, $\psi(i) = \psi(j)$. Since G is connected, inductively $\psi(u) = \psi(v)$ for all $u, v \in V$. Thus, ψ must be the multiple of $\mathbf{1}$, meaning that the subspace of eigenvalue 0 must be one-dimensional. Hence, $\lambda_2 > 0$.

($\Leftarrow$) By contraposition, assume G is NOT connected. Then G could be separated into two distinct graphs G_1 and G_2, without any edges in between. We suppose G has n vertices and G_1 has m vertices ($m < n$ and G_2 has $n - m$ vertices).

By reorganizing the order of vertices, the Laplacian matrix can be expressed as

$$L_G = \begin{bmatrix} L_{G_1} & O \\ O & L_{G_2} \end{bmatrix},$$

where O denotes the all-0s matrix.

We consider the vector $v_1 = \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}$, where first m indexes are 1 and later $n - m$ indexes are 0.

Then for any integer $1 \leq i \leq m$,

$$(L_G v_1)(i) = (L_{G_1} \mathbf{1})(i) + \sum 0 \cdot 0 = 0.$$

For any integer $m < i \leq n$,

$$(L_G v_1)(i) = \sum 0 \cdot 1 + (L_{G_2} \mathbf{0})(i) = 0.$$

Thus, $L_G v_1 = 0$.

Without loss of generality, for vector $v_2 = \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}$ (where the last $n - m$ indexes are 1), $L_G v_2 = 0$. Yet v_1, v_2 are linearly independent.

Thus, there must be at least two independent eigenvectors with eigenvalues 0, so $\lambda_2 = 0$. □

We will then introduce the Incidence Matrix, closely related to the Laplacian Matrix. The Incidence Matrix is defined based on the concept of orientation.

Definition 3.11. Let G be a graph. For each edge $e_j = \{u, v\}$ (where $u, v \in V$), its *orientation*, denoted as $\delta(e_j)$, is either the ordered pair (u, v) or the ordered pair (v, u). We say that the first vertex in the ordered pair is the *starting vertex*, and the second vertex is the *ending vertex*.

In short, an orientation picks a direction for each edge, as an ordered pair.

Definition 3.12. Let G be a graph with n vertices and m edges, with $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{e_1, e_2, \ldots e_m\}$. The incidence matrix M_G is the following $n \times m$ matrix:

$$M_G(i,j) = \begin{cases} 1, & \text{if the edge } e_j \text{ has starting vertex } v_i \\ -1, & \text{if the edge } e_j \text{ has ending vertex } v_i \\ 0, & \text{otherwise.} \end{cases}$$

We observe that for each column (denoting each edge), there can only be two nonzero entries, namely 1 and -1 that represent the starting and ending vertex.

Notation 3.13. Let G be a graph and M_G be its incidence matrix.

(1) For all $v_i \in V$, $M_G(v_i)$ refers to the row vector corresponding to vertex v_i.
(2) For all $e_j \in E$, $M_G(e_j)$ refers to the column vector corresponding to edge e_j.

Similar notations apply to the reduced incidence matrix and other similar matrices defined later.

Lemma 3.14. *Let G be a graph. Despite the choice of orientation,*

$$L_G = M_G M_G^\top.$$

Proof. We check each entry i, j of matrix $M_G M_G^\top$ matches L_G.

$$M_G M_G^\top(i,j) = \sum_{k=1}^{n} M_G(i,k) \cdot M_G^\top(k,j) = \sum_{k=1}^{n} M_G(i,k) \cdot M_G(j,k).$$

Case 1: $i = j$

$$M_G M_G^\top(i,j) = \sum_{k=1}^{n} M_G(i,k)^2 = \sum_{k \sim i} (\pm 1)^2 = \deg(i).$$

Case 2: $i \neq j$

Under a given k, $M_G(i,k) \cdot M_G(j,k)$ is nonzero only if both $M_G(i,k), M_G(j,k)$ are nonzero. As a column of M_G has only two nonzero terms, this means that a nonzero value is only possible if $e_k = \{i, j\}$. As we assume no multiple edges, there is at most one such k for each i, j.

Thus, if $i \sim j$, then

$$M_G M_G^\top(i,j) = \sum_{k=1}^{n} M_G(i,k) \cdot M_G(j,k) = 1 \cdot (-1) = -1.$$

Otherwise, if i is not connected to j, then $M_G(i,k), M_G(j,k)$ cannot be simultaneously nonzero, so

$$M_G M_G^\top(i,j) = \sum_{k=1}^{n} 0 = 0.$$

All of these match the entries in a Laplacian matrix. □

Note that this result does not depend on the orientation and is always true.

Lastly, we define the determinant in a combinatorial way, which will be helpful in later proof of the Matrix-Tree Theorem.

Definition 3.15. Let $A = (a_{i,j})$ be a $n \times n$ matrix, with $1 \leq i, j \leq n$. Let S_n be the set of bijective functions (permutations) from $\{1, 2, \dots, n\}$ to itself, and let $\sigma \in S_n$. Then the determinant of A is defined as

$$\det(A) = \sum_{\sigma \in S_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \right)$$

Note that $\operatorname{sgn}(\sigma)$ is the *sign* of the permutation σ (see definition in [1]).

Proposition 3.16. *Let $A = (a_{i,j})$, an $n \times n$ matrix, and $A^\top$ be its transpose. Then*

$$\det(A^\top) = \det(A).$$

Proposition 3.17. *Let $A = (a_{i,j})$, an $n \times n$ matrix, be upper-triangular (meaning that $a_{i,j} = 0$ for all $i > j$). Then*

$$\det(A) = \prod_{i=1}^{n} a_{i,i}.$$

4. Matrix-Tree Theorem Introduction and Preparations for Proof

Then we introduce the Matrix-Tree Theorem, the main theorem that we seek to prove in this paper. The Matrix-Tree Theorem enables us to compute the number of spanning trees in a graph by converting the problem into computing the determinant of a certain matrix. We mainly explain the exposition in [4].

Definition 4.1. The *complexity* of a graph G is the number of spanning trees in G, denoted as $\kappa(G)$.

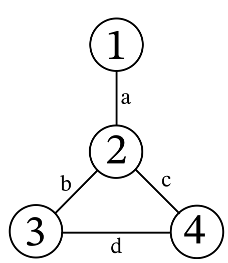
By Corollary 2.12, $\kappa(G) = 0$ if G is not connected. For the rest of the paper, we only care about G being connected, where $\kappa(G) > 0$.

Theorem 4.2. *(The Matrix-Tree Theorem) Let $G = (V, E)$ be a connected graph, and let L be its Laplacian Matrix. Let L_0 be the matrix formed by removing the last row and column (or any row and column of the same index) of L. Then*

$$\kappa(G) = \det(L_0).$$

Before laying out the ingredients for the proof, let's first see an example.

Example 4.3. We will apply this formula to Example 2.15.



In Example 2.15, we decide that there are three spanning trees here, namely abc, abd, acd. We verify this result with the Matrix-Tree Theorem.

$$L_G = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}.$$

Removing the last row and column,

$$L_0 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Then calculating by the last column,

$$\det(L_0) = -(-1) \cdot \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} + 2 \cdot \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix} = -(-1) \cdot (-1) + 2 \cdot (3 - 1) = 3.$$

This matches our computation (the same is true for removing any row and column).

The following theorems and claims are the building blocks of the proof of Theorem 4.2. One of those is the Binet-Cauchy Theorem which provides insights into determinants in matrix multiplication, which is based on the following lemma.

Lemma 4.4. *Let A and B be $n \times n$ matrices. Then*

$$\det(AB) = \det(A) \cdot \det(B).$$

A proof of this lemma can be seen in [2].

Definition 4.5. Let $A = (a_{ij})$ be an $m \times n$ matrix, and $B = (b_{ji})$ be a $n \times m$ matrix, as $1 \leq i \leq m, 1 \leq j \leq n$. Assume $m \leq n$. If $S = \{j_1, j_2, \ldots, j_m\} \subseteq \{1, 2, \ldots, n\}$ and $j_1 < j_2 < \ldots < j_m$, then $A[S]$ represents the $m \times m$ matrix such that

$$A[S](p, q) = a_{pj_q},$$

and $B[S]$ represents the $m \times m$ matrix such that

$$B[S](p, q) = b_{j_q p}.$$

Basically, only the rows/columns indexed by S are preserved, and others are discarded, to form a square matrix.

Example 4.6. Consider matrix $A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ b_1 & b_2 & b_3 & b_4 & b_5 \end{bmatrix}$ and $B = \begin{bmatrix} c_1 & d_1 \\ c_2 & d_2 \\ c_3 & d_3 \\ c_4 & d_4 \\ c_5 & d_5 \end{bmatrix}$.

Let $S = \{1, 3\} \subseteq \{1, 2, 3, 4, 5\}$.
Then by Definition 4.5,

$$A[S] = \begin{bmatrix} a_1 & a_3 \\ b_1 & b_3 \end{bmatrix}, \text{and } B[S] = \begin{bmatrix} c_1 & d_1 \\ c_3 & d_3 \end{bmatrix}.$$

The Binet-Cauchy Theorem generalizes Lemma 4.4 to multiplication of rectangular matrices.

Theorem 4.7. *(Binet-Cauchy Theorem) Let A be an $m \times n$ matrix and B be a $n \times m$ matrix.*

(1) If $m > n$, then $\det(AB) = 0$.
(2) If $m \leq n$, then

$$\det(AB) = \sum_{S \subseteq \{1,2,\ldots,n\}, |S|=m} (\det A[S])(\det B[S]).$$

Note that the matrix AB will be a $m \times m$ matrix, so it makes sense to discuss its determinant. Lemma 4.4 is the particular case of this theorem at $m = n$.

Proof. Case 1: $m > n$

Let ϕ, ψ be the linear map corresponding to matrix A, B, respectively. Then matrix AB denotes the following linear map:

$$k^m \xrightarrow{\psi} k^n \xrightarrow{\phi} k^m$$

Since $\operatorname{im}(\phi) \simeq k^n / \ker(\phi) \subset k^n$, we have $\dim(\operatorname{im}(\phi)) \leq n < m$. Therefore, $\operatorname{im}(\phi \circ \psi) \leq \dim(\operatorname{im}(\phi)) < m$. So $\phi \circ \psi$ is NOT surjective, and thus NOT invertible, meaning $\det(AB) = 0$.

Case 2: $m \leq n$

We observe that

$$\begin{bmatrix} I_{m\times m} & A \\ O_{n\times m} & I_{n\times n} \end{bmatrix} \begin{bmatrix} A & O_{m\times m} \\ -I_{n\times n} & B \end{bmatrix} = \begin{bmatrix} O_{m\times n} & AB \\ -I_{n\times n} & B \end{bmatrix} \tag{4.8}$$

Note that I represents the identity matrix and O represents the all-zero matrix. (4.8) can be verified by conducting the matrix multiplication.

For simplicity, denote

$$\begin{bmatrix} I_{m\times m} & A \\ O_{n\times m} & I_{n\times n} \end{bmatrix} = M_1, \quad \begin{bmatrix} A & O_{m\times m} \\ -I_{n\times n} & B \end{bmatrix} = M_2, \quad \begin{bmatrix} O_{m\times n} & AB \\ -I_{n\times n} & B \end{bmatrix} = M_3.$$

By Lemma 4.4, it's true that

$$\det(M_1) \cdot \det(M_2) = \det(M_3). \tag{4.9}$$

Claim: $\det(M_1) = 1$.

This is true by applying Proposition 3.17, as M_1 is an upper triangular matrix and its determinant would be products of diagonal.

Claim: $\det(M_3) = (-1)^{(m+1)n} \det(AB)$.

We need to pick $\sigma \in S_n$ such that $\prod_{i=1}^{n} a_{i,\sigma(i)} \neq 0$. To avoid picking the zeros of $O_{m\times n}$ on the upper-left, then it must be true that

$$\sigma(\{1, 2 \ldots, m\}) = \{n+1, n+2, \ldots, n+m\}.$$

This means that

$$\sigma(\{m+1, \ldots m+n\}) = \{1, 2, \ldots, n\}.$$

Looking at the lower left of M_3, the only way to pick a nonzero product is to let

$$\sigma(m+i) = i,$$

for each $1 \leq i \leq n$. Thus, the different ways to pick σ is isomorphic to S_m, depending on how σ is defined on $\{1, 2, \ldots, m\}$. We write permutation $\beta \in S_m$, and observe that for σ defined in the way above, $\text{sgn}(\sigma) = (-1)^{mn}\text{sgn}(\beta)$.

This is because for each $1 \leq i \leq m$ and $1 \leq j \leq n$, though $i < m + j$, it's true that $\sigma(i) > n \geq \sigma(m + j)$. Considering all possible i and j, $(-1)^{mn}$ should be multiplied in addition to $\text{sgn}(\beta)$. Thus,

$$\det(M_3) = (-1)^{mn} \cdot (-1)^n \sum_{\beta \in S_m} \left(\text{sgn}(\beta) \prod_{i=1}^{n} a_{i,n+\beta(i)} \right) = (-1)^{(m+1)n} \cdot \det(AB).$$

Claim: $\det(M_2) = (-1)^{(m+1)(n-m)} \det(A) \cdot \det(B)$.

Look at $\begin{bmatrix} A & O_{m\times m} \\ -I_{n\times n} & B \end{bmatrix}$.

Again, we look for permutations $\sigma : \{1, 2, \ldots, m + n\} \to \{1, 2, \ldots, m + n\}$ that forms nonzero products. To avoid $O_{m\times m}$ on the upper-right, it must be true that

$$\sigma^{-1}(\{n + 1, n + 2, \ldots, n + m\}) \subseteq \{m + 1, m + 2, \ldots, m + n\}.$$

In other words, this is similar to picking m-size subsets in $\{1, 2, \ldots, n\}$, denoted as $S \subseteq \{1, 2, \ldots, n\}$, and then consider all permutations within matrix $B[S]$.

Consequently, consider each $m+1 \leq i \leq m+n$ such that $i-m \in \{1, 2, \ldots, n\}\backslash S$, we have

$$\sigma(i) = i - m,$$

in order to match the nonzero entries in the lower-left $-I_{n\times n}$. Thus, considering the top m rows that haven't been defined, their output will exactly correspond to the columns of index in S (in order for σ to be bijective). In other words,

$$\sigma(\{1, 2, \ldots m\}) = S \subseteq \{1, 2, \ldots, n\}.$$

In this way, we explore the behavior of σ and it appears to be based on the choice of S and the permutations S_m in $A[S]$ and $B[S]$. Remind that S is m-size subset of $\{1, \ldots, n\}$, and for simplicity, we denote $S(i)$ as the i-th smallest element in S. We denote $A[S] = (a_{i,j})$ and $B[S] = (b_{i,j})$, where $1 \leq i, j \leq m$.

Let $\sigma_1, \sigma_2 \in S_m$. By a similar reasoning as the previous claim, $\text{sgn}(\sigma) = (-1)^{m(n-m)}\text{sgn}(\sigma_1)\text{sgn}(\sigma_2)$. Then

$$\begin{aligned}
\det(M_2) &= (-1)^{m(n-m)} \cdot (-1)^{n-m} \cdot \sum_{S\subseteq\{1,2,\ldots,n\},|S|=m} \left(\sum_{\sigma_1,\sigma_2\in S_m} \prod_{i,j=1}^{n} a_{i,S(\sigma_1(i))} b_{j,S(\sigma_2(j))} \right) \\
&= (-1)^{(m+1)(n-m)} \cdot \sum_{S\subseteq\{1,2,\ldots,n\},|S|=m} \left(\left(\sum_{\sigma_1\in S_m} \prod_{i=1}^{n} a_{i,S(\sigma_1(i))} \right) \cdot \left(\sum_{\sigma_2\in S_m} \prod_{j=1}^{n} b_{j,S(\sigma_2(j))} \right) \right) \\
&= (-1)^{(m+1)(n-m)} \cdot \sum_{S\subseteq\{1,2,\ldots,n\},|S|=m} \det(A[S]) \cdot \det(B[S]).
\end{aligned}$$

By (4.9),

$$\det(AB) = \sum_{S\subseteq\{1,2,\ldots,n\},|S|=m} \det(A[S]) \cdot \det(B[S]),$$

since $(-1)^{m(m+1)} = 1$ is true despite m's parity. $\square$

While this theorem is useful in computing the determinant, the following definition and lemma will relate this computation of determinant with the idea of spanning trees.

Definition 4.10. Let G be a graph, and $M[G]$ be its incidence matrix (with any orientation). The *reduced incidence matrix* of G, denoted as M_0, is the matrix with the last row of $M[G]$ removed.

Note that as G has n vertices, the matrix M_0 will have $n-1$ rows, which is exactly the number of edges in a spanning tree of G.

Lemma 4.11. *Let G be a connected graph with n vertices, and let M_0 be its reduced incidence matrix. Let $S \subseteq E$ be a set of $n-1$ edges in G.*

1. *If S forms a spanning tree in G, then* $\det(M_0[S]) = \pm 1$.
2. *If S doesn't form a spanning tree in G, then* $\det(M_0[S]) = 0$.

Note that the notation $M_0[S]$ makes sense. G is connected and M_0 has $n-1$ rows, so the number of edges in G is greater than or equal to $n-1$.

Proof. Proof of case 2:

This is equivalent to showing that the columns of $M_0[S]$ are linearly dependent. As S cannot form a tree, Lemma 2.13 implies that there must exist at least a cycle within S.

Let $(v_1, v_2, \ldots, v_n)$ be the cycle such that $v_1 = v_n$. We then define C_1, C_2, two complementary subsets of S: for each $1 \le i \le n-1$, if $\delta(\{v_i, v_{i+1}\}) = (v_i, v_{i+1})$, then $\{v_i, v_{i+1}\} \in C_1$; otherwise $\delta(\{v_i, v_{i+1}\}) = (v_{i+1}, v_i)$, so $\{v_{i+1}, v_i\} \in C_2$. Exploiting Notation 3.13, we discover that by property of cycle,

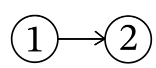
$$\sum_{e \in C_1} M_0(e) - \sum_{e' \in C_2} M_0(e') = 0.$$

These vectors are linearly dependent, hence making $\det(M_0[S]) = 0$.

Proof of case 1:

We will proceed by induction on the number of vertices in graph G (thus also S).

Base Case: $n = 2$.



Without loss of generality, we assume the orientation follows the graph above. So

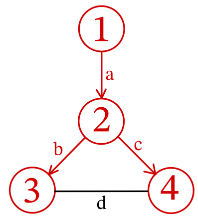
$$M[S] = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad M_0[S] = [1].$$

Thus, it's true that $\det(M_0[S]) = \pm 1$.

Induction Hypothesis: Let $k \in N$ such that $k > 2$. We assume that for all possible G with $k-1$ vertices, any spanning trees S_{k-1} in G will have $\det(M_0[S_{k-1}]) = \pm 1$.

Induction Step: We examine spanning trees S_k, in a graph G with k vertices.

(For better clarifications, I'll explain the process while simultaneously showing an example, as shown in the graph above. The red part is the spanning tree with orientation picked. Then $M[S_k] = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.)



Let v_k be the vertex represented by the last row, which is removed from $M[S_k]$ to form $M_0[S_k]$. As S_k is a spanning tree, there exists an edge $e_j \in S$ that is connected to v_k. We assume $e_j = \{v_l, v_k\}$, without loss of generality in the orientation.

(Particularly in this example, the removed vertex v_k is the 4th vertex, meaning the 4th row, and the corresponding e_j is edge c, meaning the 3rd column. Here $M_0[S_k] = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \end{bmatrix}$.)

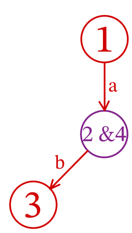
As e_j represents a column in $M_0[S_k]$, this column has only one nonzero entry, which is on the row of v_l. Write M_0', a $(k-2) \times (k-2)$ matrix by removing the row of v_l and the column of e_j in $M_0[S_k]$. Then

$$\det(M_0[S_k]) = \pm 1 \cdot \det(M_0'). \tag{4.12}$$

(Particularly, the only nonzero term in the 3rd column of $M_0[S_k]$ is the 2nd row, meaning v_l here is vertex 2. By removing the 3rd column and second row, $M_0' = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.)

We observe that $\det(M_0[S_k])$ is the reduced incidence matrix of a $k-1$ spanning tree formed by collapsing v_l and v_n as a single vertex and removing the merged vertex. This is reasonable, because the rows of both vertices are removed and the edge connecting the two vertices is removed, while all other information in the graph is preserved.

(In this example, the collapsed tree is the following graph, with the merged vertex marked purple. Note that $M = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$, and M_0' above is hence formed by removing the row of the merged vertex here.)



It needs to be clarified that only the chosen spanning tree is collapsed, and the edges not chosen in the tree won't participate in this process. Also, the collapsed graph must be a tree, because by contraposition, if it has cycles or isn't connected, the original tree cannot be a tree. Since the collapsed tree is a spanning tree with

$k-1$ vertices, by induction hypothesis and (4.12), $\det(M_0') = \pm 1$, so

$$\det(M_0[S_k]) = \pm 1 \cdot \pm 1 = \pm 1.$$

□

5. Proof of Theorem

Gathering the ingredients, we can finally prove Theorem 4.2.

Proof. Let $G = (V, E)$ be a connected graph with n vertices. Remind that L is its Laplacian Matrix and L_0 is formed by removing the last (or i-th) row and column of L. From Lemma 3.14, $L = MM^\top$ implies $L_0 = M_0 M_0^\top$.

Note that M_0 has $n-1$ rows, and as G is connected, it has more than $n-1$ edges. Thus, for any $S \subset E$ of length $n-1$, it makes sense to discuss $M_0[S]$. Further, let T be the set of spanning trees. Thus, by Theorem 4.7, Lemma 4.11 and Proposition 3.16, we have

$$\begin{aligned}
\det(L_0) &= \sum_S \det(M_0[S]) \det(M_0^\top[S]) \\
&= \sum_S \det(M_0[S]) \det(M_0[S]^\top) \\
&= \sum_S \det(M_0[S]) \det(M_0[S]) \\
&= \sum_{S \notin T} \det(M_0[S])^2 + \sum_{S \in T} \det(M_0[S])^2 \\
&= 0^2 + \sum_{S \in T} (\pm 1)^2 = \kappa(G).
\end{aligned}$$

Since the last row removed can represent any vertex, this implies that the index of the removed row/column won't affect this calculation. □

6. Refined Version and A Special Case

While the determinant of a matrix can be difficult to compute, the following corollary refines the theorem by relating it with a product of eigenvalues.

Corollary 6.1. *Let G be a connected graph with n vertices, and L is its Laplacian Matrix. Let $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_n$ be the eigenvalues of L. Then*

$$\kappa(G) = \frac{1}{n} \lambda_2 \lambda_3 \ldots \lambda_n.$$

Proof. Without loss of generality, assume L_0 is defined by removing the last row and column of L. We correlate $\det(L_0)$ with the characteristic polynomial $\det(L - xI)$.
Claim: The coefficient of x in $\det(L - xI)$ equals $-n \cdot \det(L_0)$.

We employ the fact that each row and column of L will add up to zero. Originally,

$$L - xI = \begin{bmatrix} & L_0 - xI & & \begin{matrix} a_{1n} \\ . \\ . \end{matrix} \\ a_{n1} & a_{n2} & .. & a_{nn} - x \end{bmatrix}.$$

Adding the first $n-1$ rows to the last row won't change the determinant, and we have

$$\det(L - xI) = \det\left(\begin{bmatrix} & L_0 - xI & & a_{1n} \\ & & & . \\ & & & . \\ -x & -x & .. & -x \end{bmatrix}\right) = (-x) \cdot \det\left(\begin{bmatrix} & L_0 - xI & & a_{1n} \\ & & & . \\ & & & . \\ 1 & 1 & .. & 1 \end{bmatrix}\right).$$

Thus, the coefficient of x is calculated by considering the determinant of the last matrix at $x = 0$ and adding the first $n-1$ columns to the last column.

$$\det\left(\begin{bmatrix} & L_0 & & a_{1n} \\ & & & . \\ & & & . \\ 1 & 1 & .. & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} & L_0 & & 0 \\ & & & 0 \\ & & & 0 \\ 1 & 1 & .. & n \end{bmatrix}\right) = n \cdot \det(L_0).$$

Thus, the coefficient of x here will be $-n \cdot \det(L_0)$.

Claim: The coefficient of x in the characteristic polynomial equals to $-\lambda_2\lambda_3 \ldots \lambda_n$.

By definition and property of the characteristic polynomial (seen in [2]),

$$\begin{aligned} \det(L - xI) &= (\lambda_1 - x)(\lambda_2 - x) \ldots (\lambda_n - x) \\ &= -x(\lambda_2 - x) \ldots (\lambda_n - x) \end{aligned}$$

Note that the coefficient of x here would be $-\lambda_2\lambda_3 \ldots \lambda_n$. Incorporating the two claims, we reach the conclusion that

$$\kappa(G) = \det(L_0) = \frac{1}{n}\lambda_2\lambda_3 \ldots \lambda_n.$$

□

The following result is a special case, about finding $\kappa(G)$ of complete graphs.

Example 6.2. Let K_n be the complete graph of n vertices. Then

$$\kappa(K_n) = n^{n-2}.$$

Proof. We discover that for K_n, its Laplacian Matrix

$$L_{K_n} = n \cdot I - J,$$

where I is the identity matrix and J is the all-ones matrix.

The eigenvalues of I are all 1s. For J, as its image has dimension 1, its kernel has dimension $n-1$, meaning it has eigenvalue 0 with multiplicity $n-1$. Its image is spanned by $\mathbf{1}$, corresponding to an eigenvalue n with multiplicity 1.

Thus, the eigenvalues of L_{K_n} are $\lambda_1 = 0$, and $\lambda_2 = \lambda_3 = \ldots = \lambda_n = n$. By Corollary 6.1,

$$\kappa(K_n) = \frac{1}{n} \times n^{n-1} = n^{n-2}.$$

□

Acknowledgments

It's a pleasure for me to sincerely thank my mentor, Víctor Hugo Almendra Hernández, for introducing me to the fascinating area of spectral graph theory, patiently guiding me to read, learn and discuss mathematics, and offering helpful suggestions for this paper. I'm also grateful to Professor Peter May for providing detailed feedback on my paper, and for organizing the 2024 UChicago Math REU, an amazing experience in my mathematical journey.

References

[1] Sheldon Axler. *Linear Algebra Done Right*. Springer. https://doi.org/10.1007/978-3-031-41026-0

[2] Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence. *Linear Algebra.* 223, 248.

[3] Daniel A. Spielman. *Spectral and Algebraic Graph Theory, Incomplete Draft, dated December 4, 2019.* http://cs-www.cs.yale.edu/homes/spielman/sagt

[4] Richard Stanley. "The Matrix-Tree Theorem", *Combinatorial Analysis, Fall 2014.* (MIT OpenCourseWare: Massachusetts Institute of Technology), https://shorturl.at/OaKHd (Accessed July 09, 2024). License: Creative commons BYNC-SA.