Topics in Graph Theory

April 25, 2019

1 Preliminaries

A **graph** is a system $G = (V, E)$ consisting of a set V of **vertices** and a set E (disjoint from V) of **edges**, together with an **incidence function** End : $E \to M_2(V)$, where $M_2(V)$ is set of all 2-element sub-multisets of V. We usually write $V = V(G)$, $E = E(G)$, and $\mathrm{End} = \mathrm{End}_G$. For each edge $e \in E$ with $\mathrm{End}(e) = \{u, v\}$, we called u, v the **end-vertices** of e, and say that the edge e is **incident** with the vertices u, v, or the vertices u, v are **incident** with the edge e, or the vertices u, v are **adjacent** by the edge e. Sometimes it is more convenient to just write the incidence relation as $e = uv$. If $u = v$, the edge e is called a **loop**; if $u \neq v$, the edge is called a **link**. Two edges are said to be **parallel** if their end vertices are the same. Parallel edges are also referred to **multiple edges**.

A **simple graph** is a graph without loops and multiple edges. When we emphasize that a graph may have loops and multiple edges, we refer the graph as a **multigraph**. A graph is said to be (i) **finite** if it has finite number of vertices and edges; (ii) **null** if it has no vertices, and consequently has no edges; (iii) **trivial** if it has only one vertex with possible loops; (iv) **empty** if its has no edges; and (v) **nontrivial** if it is not trivial. A **complete graph** is a simple graph that every pair of vertices are adjacent. A complete graph with n vertices is denoted by K_n. A graph G is said to be **bipartite** if its vertex set $V(G)$ can be partitioned into two disjoint nonempty parts X, Y such that every edge has one end-vertex in X and the other in Y; such a partition $\{X, Y\}$ is called a **bipartition** of G, and such a bipartite graph is denoted by $G[X, Y]$. A bipartite graph $G[X, Y]$ is called a **complete bipartite graph** if each vertex in X is joined to every vertex in Y; we abbreviate $G[X, Y]$ to $K_{m,n}$ if $|X| = m$ and $|Y| = n$.

Let G be a graph. Two vertices of G are called **neighbors** each other if they are adjacent. For each vertex $v \in V(G)$, the set of neighbors of v in G is denoted by $N_v(G)$, the number of edges incident with v (loops counted twice) is called the **degree** of v in G, denoted $\deg(v)$ or $\deg_G(v)$. A vertex of degree 0 is called an **isolated vertex**; a vertex of degree 1 is called a **leaf**. A graph is said to be **regular** if its every vertex has the same degree. A graph is said to be k**-regular** if its every vertex has degree k. We always have

$$2|E(G)| = \sum_{v \in V} \deg_G(v).$$

The number of vertices of odd degree in any graph is always an even number.

A **walk** in G from a vertex u to a vertex v is a sequence $W = v_0 e_1 v_1 \cdots v_{l-1} e_l v_l$ of vertices and edges with $v_0 = u$ and $v_l = v$, whose terms are alternate between vertices and edges of G, such that the edge e_i is incident with the vertices v_{i-1} and v_i, $1 \leq i \leq l$. The vertex v_0 called the **initial vertex**, v_l the **terminal vertex** of G, and the number l the **length** of W. A walk is said to be **closed** if its initial and terminal vertices are identical. A walk is called a **trail** if its edge terms are distinct. A walk is called a **path** if its vertex terms are distinct (so are its edge terms), except possible identical initial and terminal vertices, for which it is referred to a **closed path**. A graph is said to be **connected** if there exists a path between any two distinct vertices. The maximal connected subgraph of a graph G is called a **connected component** (or just **component**) of G. The number of connected components of G is denoted by $c(G)$. For a closed path $P = v_0 e_1 v_1 \cdots v_{l-1} e_l v_\ell$ in a graph G with $v_0 = v_l$. The underlying graph of P is called a **cycle**, which is a 2-regular connected subgraph.

Proposition 1.1. *A graph is bipartite if and only if its every cycle has even length.*

2 Chinese Postman Problem and Traveling Sales Problem

An **Euler trail** of a graph G is a **walk** that uses every edge of G exactly once. A closed Euler trail is called an **Euler tour**. A graph having an Euler tour is called an **Eulerian graph**. A **Hamilton path** of a graph G is a path that contains every vertex of G and whose vertices are distinct. A closed path of G is called a **Hamilton cycle** if contains every vertex of G and all its vertices are distinct except the initial and terminal vertices. A graph is called an **even graph** if the degree of its every vertex is an even number.

Theorem 2.1. *A connected graph is Eulerian if and only if it is an even graph.*

Problem 1 (Chinese Postman Problem (CPP))**.** How to find a closed walk W of shortest length in a connected graph G so that every edge of G appears in the walk?

Problem 2 (Hamilton Cycle Problem (HCP))**.** What is a necessary and sufficient condition for a graph having a Hamilton cycle?

Problem 3 (Traveling Sales Problem (TSP))**.** How to find a closed walk of shortest length in a connected graph G so that every vertex of G is visited at least once?

3 Four Color Conjecture/Theorem

A k**-vertex-coloring** (or just k**-coloring**) of a graph $G = (V, E)$ is a function $f : V \to S$, where S is a set of k colors; it can be viewed as an assignment of k colors to the vertices of G. Usually, the color set S is taken to be $\{1, 2, \ldots, k\}$ or $\{0, 1, \ldots, k-1\}$. A coloring is said to be **proper** if adjacent vertices are assigned distinct colors. Only loopless graphs admit proper colorings. Since multiple edges do not effect proper colorings, we only consider proper colorings for simple graphs.

A graph is said to be k**-colorable** if it admits a proper k-coloring. A graphs is 1-colorable if and only if it has no edges, and 2-colorable iff it is bipartite. The **chromatic number** of a graph G is the minimum positive integer k such that G is k-colorable.

Example 3.1 (Examination Scheduling)**.** The students at a university have annual examinations in all courses they take. Naturally, exams for different courses cannot be arranged at the same time slot if the courses have students in common. How can all the exams be arranged in as few time slots as possible? To find such a schedule, consider the graph G whose vertex set is the collection of all courses, two courses being joined by an edge if they have common students. Now the problem of requiring minimum number of time slots is to find the chromatic number of the constructed graph G.

Let $\Delta(G)$ denote the maximum degree of vertices of G. It is easy to see by the *greedy algorithm* that

$$\chi(G) \le \Delta(G) + 1.$$

Proof. Trivial. □

Theorem 3.1 (Brook's Theorem)**.** *Let G be a connected graph. If G is neither an odd cycle nor a complete graph, then $\chi(G) \le \Delta(G)$.*

Proof. We proceed by induction on $\Delta = \Delta(G) \ge 3$. For each $\Delta(G)$, we apply induction on $|V(G)|$. It is obviously true when $\Delta(G) = |V(G)|$. For $|V(G)| = \Delta(G) + 1$, the theorem is true, since $G \ne K_{\Delta(G)+1}$, there is a pair of non-adjacent vertices u, v, the graph G admits a $\Delta(G)$-coloring by letting u, v receive the same color. Consider the situation $|V(G)| \ge \Delta(G) + 2$.

Case 1. There is a vertex v such that $G \smallsetminus v$ is disconnected. Let $C_1, \ldots, C_k$ $(k \ge 2)$ be components of $G \smallsetminus v$. By induction we have proper $\Delta(G)$-colorings for the induced subgraphs $G(V(C_1) \cup v), \ldots, G(V(C_k) \cup v)$. By rearranging the colors of $G(V(C_i) \cup v)$ $(i = 1, \ldots, k)$, we may assume that the colors of v are the same in all proper colorings of $G(V(C_i) \cup v)$. We thus obtain a proper $\Delta(G)$-coloring of G.

Case 2. $G \smallsetminus u$ is connected for all $u \in V(G)$, but there two non-adjacent vertices v, w such that $G \smallsetminus \{v, w\}$ is disconnected. Let $C_1, \ldots, C_k$ $(k \ge 2)$ be components of $G \smallsetminus \{v, w\}$. Suppose there is no edge from v to one component C_i; then $G \smallsetminus w$ is disconnected, which is Case 1. Thus there is at least one edge from v to each C_i, so does w. Set $G_i := G \smallsetminus C_i$ $(i = 1, \ldots, k)$. Since each G_i has fewer vertices than G, by induction each G_i admits a proper $\Delta(G)$-coloring. Note that the v, w have degree at most $\Delta(G) - 1$. It follows that $H_i := G_i \cup vw$ has fewer vertices than G and the degrees of v, w in H_i are at most $\Delta(G)$. By induction each H_i admits proper $\Delta(G)$-coloring, unless it is click on $\Delta(G) + 1$ vertices (if it is an odd cycle, we still have a proper $\Delta(G)$-coloring since $\Delta(G) \ge 3$). By rearranging the colors in the coloring of H_i, we may assume that v (w) has the same color in all colorings of $H_1, \ldots, H_k$, and v, w receive distinct colors. Combining the colorings, we obtain a proper $\Delta(G)$-coloring for G.

If one of H_i is a click on $\Delta(G) + 1$ vertices, say H_1, then the degree of v, w in G_1 are $\Delta(G) - 1$. It follows that the degree of v, w in G_j $(j \ne 1)$ must be 1, and it forces that $k = 2$. Now it is clear that G_1, G_2 admit proper $\Delta(G)$-colorings such that v, w receive the same color. Combine the two colorings to obtain a proper $\Delta(G)$-coloring for G.

Case 3. $G \smallsetminus \{v, w\}$ is connected for every pair of non-adjacent vertices v, w. Set $n := |V(G)|$. Choose a vertex u whose degree is $\Delta(G)$. Since $G \ne K_n$, the vertex u has a pair of non-adjacent neighbors v_1, v_2. The subgraph $H := G \smallsetminus \{v_1, v_2\}$ is connected. Now we can arrange a sequence of vertices $v_1 v_2 \ldots v_n$ as follows: $v_n = u$, v_i $(3 \le i < n)$ is a vertex of $H \setminus \{v_{i+1}, \ldots, v_n\}$ that is adjacent with one of $\{v_{i+1}, \ldots, v_n\}$. We thus obtain a proper $\Delta(G)$-coloring for G as follows: assign v_1, v_2 the same color; assign v_i ($i \geq 3$) a color different from the colors of those vertices of $\{v_1, \ldots, v_{i-1}\}$ adjacent to v_i (there are at most $\Delta(G) - 1$ such vertices for $i < n$ and $\Delta(G)$ vertices for $i = n$). $\square$

Let $\chi(G, k)$ denote the number of k-colorings of a graph G, where $k \geq 1$. For a graph of n vertices without edges, the number of proper colorings by k colors is just k^n. For a complete graph K_n, we have

$$\chi(K_n, k) = k(k-1)\cdots(k-n+1).$$

Given an edge $e = uv$ of G, it is easy to see that

$$\chi(G, k) = \chi(G \setminus e, k) - \chi(G/e, k).$$

By induction, one can see that $\chi(G, k)$ is a polynomial function of k with degree $|V(G)|$, called the **chromatic polynomial** of G. The chromatic number of G is just the smallest positive integer k such that $\chi(G, k) \neq 0$, i.e., the smallest positive integer k such that G has a proper k-coloring.

Proposition 3.2. *Let G be a graph with two subgraphs G_1, G_2 such that $G_1 \cup G_2 = G$ and $G_1 \cap G_2 = K_n$ for some $n \geq 1$. Then*

$$\chi(G, t) = \chi(G_1, t)\chi(G_2, t)/\chi(K_n, t).$$

Proof. Since $G_1 \cap G_2 = K_n$, we see that K_n is a subgraph of both G_1 and G_2. Clearly, $\chi(G_i) \geq n$, $i = 1, 2$. For each proper t-coloring f of K_n, let a_f denote the number of proper t-colorings for the remaining vertices of G_1. Since each f uses exactly n colors, it does not matter which n colors are used among t colors, we see that the number a_f does not depend on f, i.e., a_f is a constant. There are $\chi(K_n, t)$ ways to properly color K_n by t colors. So $\chi(G_1, t) = \chi(K_n, t)a_f$, i.e., $a_f = \chi(G_1, t)/\chi(K_n, t)$. Likewise, for each proper t-coloring f of K_n, the number ways to properly color the remaining vertices of G_2 by t colors is $\chi(G_2, t)/\chi(K_n, t)$. Thus the number of proper t-colorings of G is

$$\chi(G, t) = \chi(K_, t) \cdot \frac{\chi(G_1, t)}{\chi(K_n, t)} \cdot \frac{\chi(G_1, t)}{\chi(K_n, t)} = \frac{\chi(G_1, t)\chi(G_2, t)}{\chi(K_n, t)}.$$

$\square$

Lemma 3.3. *The chromatic polynomial of every tree T_n with n vertices is*

$$\chi(T_n, t) = t(t-1)^{n-1}.$$

Lemma 3.4. *Every 2-connected graph G has chromatic polynomial of the the form*

$$\chi(G, t) = t(t-1)f(t).$$

Let G be a graph whose chromatic number is k. The chromatic polynomial may be written as

$$\chi(G,t) = t^{c_0}(t-1)^{c_1}(t-2)^{c_2}\cdots(t-k)^{c_k} f(t) = f(t)\prod_{i=0}^{\infty}(t-i)^{c_i},$$

called the **chromatic factorization** of $\chi(G,t)$, where $c_i \geq 0$ and $c_i = 0$ for all $i > k$.

Problem What are the combinatorial interpretations for the powers c_i of the term $t-i$ in $\chi(G,t)$? It is well-known that

$$c_0 = \text{number of connected components of } G,$$

$$c_1 = \text{number of block in the block decomposition of } G.$$

What are c_i for $i \geq 2$? In particular, what is c_2?

4 Flows and Tensions

Let $D = (V, A)$ be a digraph. Let be a function.

- A real-valued function $f : A \to \mathbb{R}$ is called a **flow** (or **circulation**) of D if

$$\sum_{a\in(v^c,v)} f(a) = \sum_{a\in(v,v^c)} f(a), \quad \text{i.e.,} \quad f^-(v) = f^+(v) \quad \text{for all} \quad v \in V.$$

  The set of all flows in D is a vector space, called the **flow space** of D, denoted $F(D)$.

- A real-valued function $g : A \to \mathbb{R}$ is called a **tension** if for each directed cycle C in $(V, \vec{E})$,

$$g^+(C) - g^-(C) = 0,$$

  where

$$g^+(C) := \sum_{a\in C} g(a), \quad g^-(C) := \sum_{-a\in C} g(a),$$

  The set of all tension of D is a vector space, called the **tension space** of D, denoted $T(D)$.

- If C is a directed in $(V, \vec{E})$, then $f_C : A \to \mathbb{R}$, defined by

$$f_C(a) = \begin{cases} 1 & \text{if } a \in C \\ -1 & \text{if } -a \in C, \\ 0 & \text{otherwise} \end{cases}$$

  is a flow of D, called the flow generated by the directed cycle C.

- If U is a directed cut in $(V, \vec{E})$, i.e., either $U = (X, X^c)$ or $U = (X^c, X)$ in $(V, \vec{E})$ for a nonempty proper subset $X \subset V$, then $g_U : A \to \mathbb{R}$, defined by

$$g_U(a) := \begin{cases} 1 & \text{if } a \in U \\ -1 & \text{if } -a \in U, \\ 0 & \text{otherwise} \end{cases}$$

is a tension of D, called the flow generated by the directed cut U.

- Given a directed cycle C and a directed cut U in $(V, \vec{E})$. Then

$$\langle f_C, g_U \rangle = 0,$$

where $\langle f_C, g_U \rangle := \sum_{a \in A} f_C(a) g_U(a)$. In fact,

$$\langle f_C, g_U \rangle = |A \cap C \cap U| - |A \cap C^- \cap U| - |A \cap C \cap U^-| + |A \cap C^- \cap U^-| = 0,$$

since $|A \cap C \cap U| = |A \cap C \cap U^-|$ and $|A \cap C^- \cap U| = |A \cap C^- \cap U^-|$.

- Let $\mathbf{M}$ be incidence matrix of D, i.e., $\mathbf{M} = [m_{va}]$, where $v \in V$, $a \in A$, and

$$m_{va} = \begin{cases} 1 & \text{if } a \text{ has its tail at } v \\ -1 & \text{if } a \text{ has its head at } v. \\ 0 & \text{otherwise} \end{cases}$$

Then

$$F(D) = \ker \mathbf{M}, \quad T(D) = \operatorname{Row} \mathbf{M},$$

and

$$\dim T(D) = |V(D)| - c(D), \quad \dim F(D) = |A(D)| - |V(D)| + c(D)|,$$

where $c(D)$ is the number of connected components of D.

5 Basis Matrices

A **basis matrix** of an m-dimensional vector subspace of $\mathbb{R}^n$ is an $m \times n$ matrix whose row space is the given vector subspace. For a digraph D, we are interested in the integral basis matrix $\mathbf{B}$ of the tension space $T(D)$ and the basis matrix $\mathbf{C}$ of the flow space $F(D)$. For an edge subset $S \subseteq A(D)$, we denote by $\mathbf{B}|_S$ (or just $\mathbf{B}_S$) the submatrix of $\mathbf{B}$ consisting of the columns of $\mathbf{B}$ that are labeled by members of S.

Given a maximal spanning forest F of G. For each edge $e \in F$, $F^c \cup e$ contains a unique bond B_e, which must contain the edge e. For each edge $e' \in F^c$, $F \cup e'$ contains a unique cycle $C_{e'}$, which must contain the edge e'. It is well-know that the bond vectors g_{B_e}, $e \in F$, form an integral basis of the tension lattice of D, and the cycle vectors $f_{C_{e'}}$, $e' \in F^c$, form an integral basis of the flow lattice of G. Let the members of F be listed as $e_1, \ldots, m$, and the members of F^c as $e'_1, \ldots, e'_n$. We obtain integral basis matrices

$$\mathbf{B} = \begin{array}{c} \\ e_1 \\ e_2 \\ \vdots \\ e_m \end{array} \begin{array}{c} \begin{array}{cccccccc} e_1 & e_2 & \cdots & e_m & e'_1 & e'_2 & \cdots & e'_n \end{array} \\ \begin{pmatrix} 1 & 0 & \cdots & 0 & * & * & \cdots & * \\ 0 & 1 & \cdots & 0 & * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & * & * & \cdots & * \end{pmatrix} \end{array},$$

$$\mathbf{C} = \begin{array}{c} \\ e'_1 \\ e'_2 \\ \vdots \\ e'_n \end{array} \begin{array}{c} \begin{array}{cccccccc} e_1 & e_2 & \cdots & e_m & e'_1 & e'_2 & \cdots & e'_n \end{array} \\ \begin{pmatrix} * & * & \cdots & * & 1 & 0 & \cdots & 0 \\ * & * & \cdots & * & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * & 0 & 0 & \cdots & 1 \end{pmatrix} \end{array},$$

called **integral basis matrices** of the tension lattice and flow the lattice relative to the maximal spanning forest F respectively.

Theorem 5.1. *Let $\mathbf{B}$ be a basis matrix of the tension $T(D)$ of a digraph D, and $\mathbf{C}$ the basis matrix of the flow space $F(D)$. Given a nonempty subset $S \subset A(D)$.*

(a) The columns of $\mathbf{B}|_S$ are linearly independent iff S does not contain cycle.

(b) The columns of $\mathbf{C}|_S$ are linearly independent iff S does not contain bond.

Proof. (a) Let $\mathbf{b}(a)$ denote the column vector of $\mathbf{B}$ corresponding to the arc $a \in A(D)$. We may write $\mathbf{b}(a) = [b_1(a), \ldots, b_m(a)]^T$, where $m = \dim T(D)$. The columns $\mathbf{b}(a)$ for $a \in S$ are linearly dependent iff there exists a nonzero function $f : A \to \mathbb{R}$ such that $f|_{A \smallsetminus S} = 0$ and

$$\sum_{a \in A} f(a)\mathbf{b}(a) = 0, \quad \text{i.e.,} \quad \langle f, b_i \rangle = 0 \ \text{ for } \ 1 \le i \le m,$$

which means that f is a flow of D and its support is contained in S. Now if there is such a flow f whose support is contained in S, then the support of f contains a cycle, so does S. If S contains a cycle C, then f_C is a nonzero flow whose support is C, which is contained in S.

(b) Let $\mathbf{c}(a)$ denote the column vector of $\mathbf{C}$ corresponding to the arc $a \in A(D)$. The columns $\mathbf{c}(a)$ for $a \in S$ are linearly independent iff there exists a nonzero function g on A such that $\sum_{a \in A} f(a)\mathbf{c}(a) = 0$, i.e., there exists a nonzero tension whose support is contained in S. Now if there is such a tension g whose support is contained in S, then the support of g contains a bond, so does S. If S contains a bond B, then g_B is a nonzero tension whose support is B, which is contained in S. □

A rectangular matrix is said to be **unimodular** if its full square submatrices have determinates 1, -1, or 0; and to be **totally unimodular** if its all square submatrices have determinates 1, -1, or 0.

Lemma 5.2. *Let $\mathbf{B}$ be a basis matrix of the tension of a connected digraph D, and $\mathbf{C}$ the basis matrix of the flow space. Given a maximal spanning forest F of D.*

(a) Then $\mathbf{B}$ is uniquely determined by $\mathbf{B}|_F$, and $\mathbf{C}$ is uniquely determined by $\mathbf{C}|_{F^c}$.

(b) If $\mathbf{B}, \mathbf{C}$ are basis matrices with respect to the maximal spanning forest F, then any basis matrices $\mathbf{B}', \mathbf{C}'$ of the tension and flow spaces respectively, we have

$$\mathbf{B}' = (\mathbf{B}'|_F)\mathbf{B}, \quad \mathbf{C}' = (\mathbf{C}'|_{F^c})\mathbf{C}.$$

Proof. (a) Since F contains no cycle, we see the columns of $\mathbf{B}|_F$ are linearly independent by Lemma 5.2(a). For each arc $a \in F^c$, the set $F \cup a$ contains a cycle, it follows that the columns of $\mathbf{B}|_{F\cup a}$ are linearly dependent; so the column $\mathbf{b}(a)$ is a unique linear combination of columns of $\mathbf{B}|_F$. So $\mathbf{B}$ is uniquely determined by $\mathbf{B}|_F$.

Analogously, the set F^c contains no bond, it follows from Lemma 5.2(b) that the columns of $\mathbf{C}|_{F^c}$ are linearly independent. Since $F^c \cup a$ contains a unique bond for each $a \in F$, the columns of $\mathbf{C}|_{F^c\cup a}$ are linearly dependent, so the column $\mathbf{c}(a)$ is a unique linear combination of the columns of $\mathbf{C}|_{F^c}$. So $\mathbf{C}$ is uniquely determined by $\mathbf{C}|_{F^c}$.

(b) We order the members of A as F, F^c. Since $\mathbf{B}, \mathbf{C}$ are respective to the maximal spanning forest F, we have the form

$$\mathbf{B} = \big[\mathbf{B}|_F\ \mathbf{B}|_{F^c}\big] = \big[\mathbf{I}\ \mathbf{A}\big], \quad \mathbf{C} = \big[\mathbf{C}|_F\ \mathbf{C}|_{F^c}\big] = \big[\mathbf{G}\ \mathbf{I}\big].$$

Write the basis matrices $\mathbf{B}', \mathbf{C}'$ in the same order F, F^c as the form

$$\mathbf{B}' = \big[\mathbf{B}'|_F\ \mathbf{B}'|_{F^c}\big], \quad \mathbf{C}' = \big[\mathbf{C}'|_F\ \mathbf{C}'|_{F^c}\big].$$

It is clear that there exist square matrices $\mathbf{P}, \mathbf{Q}$ such that

$$\mathbf{B}' = \mathbf{P}\mathbf{B}, \quad \mathbf{C}' = \mathbf{Q}\mathbf{C}.$$

Then

$$\mathbf{B}' = \mathbf{P}[\mathbf{I}\ \mathbf{A}] = [\mathbf{P}, \mathbf{P}\mathbf{A}], \quad \mathbf{C}' = \mathbf{Q}[\mathbf{G}\ \mathbf{I}] = [\mathbf{Q}\mathbf{G}, \mathbf{Q}]$$

It follows that $\mathbf{P} = \mathbf{B}'|_F$ and $\mathbf{Q} = \mathbf{C}'|_F$. Hence $\mathbf{B}' = (\mathbf{B}'|_F)\mathbf{B}$ and $\mathbf{C}' = (\mathbf{C}'|_F)\mathbf{C}$. □

Theorem 5.3. *Let $\mathbf{B}$ be an integral basis matrix of the tension space, and $\mathbf{C}$ an integral basis matrix of the flow space of a graph G. Then both $\mathbf{B}$ and $\mathbf{C}$ are unimodular.*

Proof. Given a maximal spanning forest F of G. Let $\mathbf{B}', \mathbf{C}'$ be basis matrices of the tension space and the flow space of G relative to F respectively. There exists unimodular matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{B}' = \mathbf{P}\mathbf{B}$ and $\mathbf{C}' = \mathbf{Q}\mathbf{C}$. Restrict both sides to F', F'^c respectively, we obtain

$$\mathbf{B}'|_{F'} = \mathbf{P}(\mathbf{B}|_{F'}), \quad \mathbf{C}'|_{F'^c} = \mathbf{Q}(\mathbf{C}'|_{F'^c}).$$

Since $\mathbf{B}'|_{F'}, \mathbf{C}'|_{F'^c}$ are identity matrices by definition, we see that

$$\det(\mathbf{P})\det(\mathbf{B}|_{F'}) = 1, \quad \det(\mathbf{Q})\det(\mathbf{C}'|_{F'^c}) = 1.$$

It follows that $\det(\mathbf{B}|_{F'}) = \pm 1$ and $\det(\mathbf{C}|_{F'^c}) = \pm 1$.

Given edge subsets $S \subset A(D)$. If $|S| = |V(D)| - 1$ and S is not spanning tree, then S contains a cycle. Thus $\det(\mathbf{B}|_S) = 0$ by Lemma 5.2. If $|S| = |A(D)| - |V(D)| + 1$ and S^c is a not spanning tree, then S contains a bond, then $\det(\mathbf{C}|_S) = 0$ by Lemma 5.2. □

Proposition 5.4. *The incidence matrix $\mathbf{M}$ of an digraph $D = (V, A)$ is totally unimodular.*

Proof. Let $S \subseteq V$ and $F \subseteq E$ be such that $|S| = |F|$. If there exists a vertex $v \in S$ such that $v \notin V(F)$, then the v-row of $\mathbf{M}|_{S\times F}$ is a zero row; clearly, $\det(\mathbf{M}|_{S\times F}) = 0$. We may assume that $S \subseteq V(F)$. We see that $\mathbf{M}|_{S\times F}$ is the incidence matrix of the subgraph (S, F) with possible half-edges. If (S, F) contains a cycle, then the columns indexed by the edges of the cycle are linearly dependent; thus $\det(\mathbf{M}|_{S\times F}) = 0$. If (S, F) contains no cycles, we claim that S is a proper subset of $V(F)$. Otherwise, $S = V(F)$, then $(V(F), F)$ is a forest; thus $|F| = |V(F)| - c(F) = |S| - c(F) < |S|$, which is a contradiction.

Now let $e = uv \in F$ be an edge such that one of u, v is not in S, say, $v \notin S$. Then the e-column of $\mathbf{M}_{S\times F}$ has 1 or -1 at (u, e) and 0 elsewhere. Thus by the expansion along the e-column,

$$\det(\mathbf{M}|_{S\times F}) = \pm\det(\mathbf{M}|_{(S\smallsetminus u)\times(F\smallsetminus e)}) = \pm 1.$$

The second equality above follows from the fact that $(S \smallsetminus u, F \smallsetminus e)$ contains no cycles and by induction on the size of the matrix. $\square$

6 The Matrix-Tree Theorem

In many occasions one needs to compute the determinant of a product matrix AB, where A is an $m \times n$ matrix and B an $n \times m$ matrix. If $m > n$, note that

$$\operatorname{rank}(AB) \le \min\{\operatorname{rank}(A), \operatorname{rank}(B)\} \le n,$$

then $\det(AB) = 0$. If $m \le n$, we have the following Cauchy-Binet formula.

The equality $\operatorname{rank} AB \le \min\{\operatorname{rank} A, \operatorname{rank} B\}$ can be argued as follows: The rank of AB is the dimension of the column space $\operatorname{Col} AB$. Note that

$$\operatorname{Col} AB = \operatorname{Col}[Ab_1, \ldots, Ab_n] \subseteq \operatorname{Col}[Ae_1, \ldots, Ae_n] = \operatorname{Col} AI_n = \operatorname{Col} A;$$

$$\operatorname{Row} AB = \operatorname{Row}\begin{bmatrix} a_1 B \\ \vdots \\ a_m B \end{bmatrix} \subseteq \operatorname{Row}\begin{bmatrix} e_1 B \\ \vdots \\ e_m B \end{bmatrix} = \operatorname{Row} I_m B = \operatorname{Row} B.$$

Proposition 6.1 (Cauchy-Binet Formula). *Let A be an $m \times n$ matrix and B an $n \times m$ matrix. If $m \le n$, then*

$$\det(AB) = \sum_{S\subseteq[n],\,|S|=m} \det(A|_S)\det(B|_S), \tag{6.1}$$

where $A|_S$ is the $m \times m$ submatrix of A whose column index set is S, and $B|_S$ is the $m \times m$ submatrix of B whose row index set is S.

Proof. Let $A = [a_{ik}]_{m\times n}$, $B = [b_{kj}]_{n\times m}$, and $C = AB = [c_{ij}]_{m\times m}$, where $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$. Then

$$\begin{aligned}
\det(C) &= \det\begin{pmatrix} \sum_{k_1=1}^n a_{1k_1}b_{k_1 1} & \cdots & \sum_{k_m=1}^n a_{1k_m}b_{k_m m} \\ \vdots & \ddots & \vdots \\ \sum_{k_1=1}^n a_{mk_1}b_{k_1 1} & \cdots & \sum_{k_m=1}^n a_{mk_m}b_{k_m m} \end{pmatrix} \\
&= \sum_{k_1=1}^n \cdots \sum_{k_m=1}^n \det\begin{pmatrix} a_{1k_1}b_{k_1 1} & \cdots & a_{1k_m}b_{k_m m} \\ \vdots & \ddots & \vdots \\ a_{mk_1}b_{k_1 1} & \cdots & a_{mk_m}b_{k_m m} \end{pmatrix} \\
&= \sum_{k_1,\ldots,k_m=1}^n \det\begin{pmatrix} a_{1k_1} & \cdots & a_{1k_m} \\ \vdots & \ddots & \vdots \\ a_{mk_1} & \cdots & a_{mk_m} \end{pmatrix} b_{k_1 1}\cdots b_{k_m m}.
\end{aligned}$$

Rewrite the nonzero terms in the above expansion of $\det(C)$, we obtain

$$\begin{aligned}
\det(C) &= \sum_{1\le k_1,\dots,k_m\le n,\, k_i\ne k_j} \det(A|_{\{k_1,\dots,k_m\}})\, b_{k_1 1}\cdots b_{k_m m} \\
&= \sum_{1\le t_1<\cdots<t_m\le n} \sum_{\sigma\in S_m} \det(A|_{\{t_{\sigma(1)},\dots,t_{\sigma(m)}\}})\, b_{t_{\sigma(1)}1}\cdots b_{t_{\sigma(m)}m} \\
&= \sum_{1\le t_1<\cdots<t_m\le n} \det(A|_{\{t_1,\dots,t_m\}}) \sum_{\sigma\in\mathfrak{S}_m} \sin(\sigma) b_{t_{\sigma(1)}1}\cdots b_{t_{\sigma(m)}m},
\end{aligned}$$

where $\mathfrak{S}_m$ is the set of all permutations of $\{1,\dots,m\}$. Set $S=\{t_1,\dots,t_m\}$, we have $\det(C)=\sum_{S\subseteq[n],\,|S|=m}\det(A|_S)\det(B|_S)$. ☐

Theorem 6.2 (Matrix-Tree Theorem)**.** *Let* $\mathbf{B}$ *be an integral basis matrix of the tension space, and* $\mathbf{C}$ *an integral basis matrix of the flow space of a graph* G*. Then the number of maximal spanning forests of* G *is*

$$\det(\mathbf{B}\mathbf{B}^T)=\det(\mathbf{C}\mathbf{C}^T).$$

Proof. Let $t(G)$ denote the number of maximal spanning forests of G. Note that an edge subset $S\subseteq E(G)$ is a maximal forest of G iff S contains no cycles and $|S|$ is the dimension of the tension space of G. By the Cauchy-Binet formula, we have

$$\det(\mathbf{B}\mathbf{B}^T)=\sum_{\substack{S\subseteq E,\,|S|=m\\ S\text{ ayclic}}}\det(\mathbf{B}_S)\det(\mathbf{B}_S^T)=\sum_{\substack{S\subseteq E,\,|S|=m\\ S\text{ ayclic}}}(\det\mathbf{B}_S)^2=t(G),$$

$$\det(\mathbf{C}\mathbf{C}^T)=\sum_{\substack{S\subseteq E,\,|S|=m\\ S\text{ ayclic}}}\det(\mathbf{C}_{S^c})\det(\mathbf{C}_{S^c}^T)=\sum_{\substack{S\subseteq E,\,|S|=m\\ S\text{ ayclic}}}(\det\mathbf{C}_{S^c})^2=t(G).$$

The second equality follows from the fact that an set $S\subseteq E(G)$ is a maximal edge set containing no bonds of G iff S^c is a maximal edge set containing no cycles. ☐

Lemma 6.3 (Farkas' Lemma)**.** *Let* $\mathbf{A}$ *be a real* $m\times n$ *matrix and* $\mathbf{b}\in\mathbb{R}^n$*. Then exactly one of the following two statements is valid.*

(a) There exists a vector $\mathbf{x}\in\mathbb{R}^n$ *such that* $\mathbf{A}\mathbf{x}=\mathbf{b}$ *and* $\mathbf{x}\ge\mathbf{0}$*.*

(b) There exists a vector $\mathbf{y}\in\mathbb{R}^m$ *such that* $\mathbf{A}^T\mathbf{y}\ge\mathbf{0}$ *and* $\mathbf{b}^T\mathbf{y}<0$*, i.e., such that*

$$\mathbf{y}^T\mathbf{A}\ge\mathbf{0}^T,\quad \mathbf{y}^T\mathbf{b}<0.$$

Proof. Farkas's Lemma is just the geometric interpretation: Let $\mathbf{a}_1,\dots,\mathbf{a}_n$ denote the columns of A. Let $\mathrm{Cone}(\mathbf{A})$ denote the convex cone generated by $\mathbf{a}_1,\dots,\mathbf{a}_n$. Let $\mathbf{x}^T=(x_1,\dots,x_n)\ge 0$. Then $\mathbf{A}\mathbf{x}=\mathbf{b}$ means that $\mathbf{b}=x_1\mathbf{a}_1+\cdots+x_n\mathbf{a}_n$. The first statement means that $\mathbf{b}\in\mathrm{Cone}(\mathbf{A})$.

Let $\mathbf{y}^T=(y_1,\dots,y_m)$. Consider the hyperplane $H=\{\mathbf{z}\in\mathbb{R}^m:\langle\mathbf{z},\mathbf{y}\rangle=0\}$. Then $\mathbf{y}^T\mathbf{A}\ge\mathbf{0}$ means that $\langle\mathbf{a}_i,\mathbf{y}\rangle\ge 0$, $i=1,\dots,n$, i.e., $\mathrm{Cone}(\mathbf{A})$ lies in one side of H. While the strictly inequality $\mathbf{b}^T\mathbf{y}<0$ means that $\mathbf{b}$ lies in the other side of H. In other words, H separates the vector $\mathbf{b}$ and the cone $\mathrm{Cone}(\mathbf{A})$, which is equivalent to $\mathbf{b}\notin\mathrm{Cone}(\mathbf{A})$.

Assume that the first statement is true, i.e., there exists a vector $\bar{\mathbf{x}} \geq \mathbf{0}$ such that $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$. Suppose the second statement is also true, i.e., there exists a vector $\bar{\mathbf{y}}$ such that $\mathbf{A}^T\bar{\mathbf{y}} \geq \mathbf{0}$ and $\mathbf{b}^T\bar{\mathbf{y}} < 0$. Then

$$0 > \mathbf{b}^T\bar{\mathbf{y}} = (\mathbf{A}\bar{\mathbf{x}})^T\bar{\mathbf{y}} = \bar{\mathbf{x}}\mathbf{A}^T\bar{\mathbf{y}} \geq 0,$$

which is a contradiction. □

Lemma 6.4 (Farkas' Lemma – a variant version)**.** *Let* $\mathbf{A}$ *be a real* $m \times n$ *matrix and* $\mathbf{b} \in \mathbb{R}^n$*. Then exactly one of the following two statements is valid.*

(a) There exists a vector $\mathbf{x} \in \mathbb{R}^n$ *such that* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *and* $\mathbf{x} \geq \mathbf{b}$*.*

(b) There exists a vector $\mathbf{y} \in \mathbb{R}^m$ *such that* $\mathbf{y}^T\mathbf{A} \geq \mathbf{0}$ *and* $\mathbf{y}^T\mathbf{A}\mathbf{b} > 0$*.*

Proof. Let $\mathbf{b}' = -\mathbf{A}\mathbf{b}$. The second statement becomes that there exists a vector $\mathbf{y}$ such that $\mathbf{y}^T\mathbf{A} \geq \mathbf{0}^T$ and $\mathbf{y}^T\mathbf{b}' < 0$. Let $\mathbf{x} = \mathbf{x}' + \mathbf{b}$. The first statement becomes that there exists a vector $\mathbf{x}'$ such that $\mathbf{A}\mathbf{x}' = \mathbf{b}'$ and $\mathbf{x}' \geq \mathbf{0}$. □

Lemma 6.5 (Farkas' Lemma – another variant version)**.** *Let* $\mathbf{A}$ *be a real* $m \times n$ *matrix and* $\mathbf{b} \in \mathbb{R}^n$*. Then exactly one of the following two statements is valid.*

(a) There exists a vector $\mathbf{x} \in \mathbb{R}^n$ *such that* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *and* $\mathbf{x} \leq \mathbf{b}$*.*

(b) There exists a vector $\mathbf{y} \in \mathbb{R}^m$ *such that* $\mathbf{y}^T\mathbf{A} \geq \mathbf{0}^T$ *and* $\mathbf{y}^T\mathbf{A}\mathbf{b} < 0$*.*

Proof. Let $\mathbf{b}' = \mathbf{A}\mathbf{b}$. The second statement becomes that there exists a vector $\mathbf{y}$ such that $\mathbf{y}^T\mathbf{A} \geq \mathbf{0}^T$ and $\mathbf{y}^T\mathbf{b}' < 0$. Let $\mathbf{x} = -\mathbf{x}' + \mathbf{b}$. The first statement becomes that there exists a vector $\mathbf{x}'$ such that $\mathbf{A}\mathbf{x}' = \mathbf{b}'$ and $\mathbf{x}' \geq \mathbf{0}$. □

7 Graph Laplacian

Let $G = (V, E)$ be a connected graph. Given an orientation ω so that (G, ω) is a digraph. Let $\mathbf{M}$ the incidence matrix of (G, ω). For each vertex $v \in V$, let $\mathbf{M}_v$ denote the matrix obtained from $\mathbf{M}$ by deleting the row corresponding to the vertex v. A **Kirchhoff matrix** of (G, ω) is the matrix $\mathbf{K} := \mathbf{M}_v$ for a vertex v. The **Laplace matrix** of G is the matrix

$$\mathbf{L} := \mathbf{M}\mathbf{M}^T.$$

Let $\mathbf{A}$ be the adjacency matrix of G, whose (u, v)-entry is the number of edges incident with u and v, each loop is *counted twice.* Let $\mathbf{D}$ be the diagonal matrix whose diagonal (v, v)-entry is $\deg(v)$ in G, which is the number of edges at v, each loop is counted twice. Then

$$\mathbf{L} = \mathbf{D} - \mathbf{A}.$$

In fact, let $e_1, \ldots, e_r$ be the links and $e'_1, \ldots, e'_s$ the loops at v. Recall that (v, e_i)-entry in $\mathbf{M}$, denoted $m(v, e_i)$, is either 1 or -1, and (v, e'_j)-entry in $\mathbf{M}$ is always 0. So the (v, v)-entry of $\mathbf{L}$ is

$$\sum_{i=1}^{r} m(v, e_i)^2 + \sum_{j=1}^{s} m(v, e'_j)^2 = r = \text{number of links at } v.$$

The (v,v)-entry of $\mathbf{D}$ is $\deg(v)$, which is the number of edges incident with v, where each loop is counted twice. The (v,v)-entry of $\mathbf{A}$ is the twice number of loops at v. Then the (v,v)-entry of $\mathbf{D}-\mathbf{A}$ is also the number of links at v.

For distinct vertices $u,v\in V$, let $e''_1,\ldots,e''_t$ be the edges between u and v. The (u,v)-entry of $\mathbf{L}$ is

$$\sum_{i=1}^{t} m(u,e''_i)m(v,e''_i) = -t = -a_{uv} = -(\text{number of edge between } u \text{ and } v).$$

We thus have $\mathbf{L}=\mathbf{D}-\mathbf{A}$.

Recall the boundary operator ∂ and co-boundary operator δ. The **Laplace operator** is defined as $\partial\delta:\mathbb{R}^V\to\mathbb{R}^V$.

Assume that G is an Eulerian graph. Let $P=v_0e_1v_1\cdots e_nv_n$ be a directed Eulerian tour of G such that $\vec{e_i}=\overrightarrow{v_{i-1}v_i}$, and let v be appeared in P as $v_{i_1},\ldots,v_{i_k}$. Then for $f\in\mathbb{R}^V$, the difference of δf (the second order difference of f at v along the path P) is

$$\begin{aligned}(\Delta f)(v) :&= \sum_{j=1}^{k}\left[(\delta f)(e_{i_j+1})-(\delta f)(e_{i_j})\right]\\ &= \sum_{j=1}^{k}\left[\big(f(v_{i_j+1})-f(v_{i_j})\big)-\big(f(v_{i_j})-f(v_{i_j-1})\big)\right]\\ &= \sum_{j=1}^{k}\left[f(v_{i_j+1})+f(v_{i_j-1})-2f(v_{i_j})\right].\end{aligned}$$

$$\begin{aligned}(\partial\delta f)(v) &= \sum_{j=1}^{k}[(f(v_{i_j})-f(v_{i_1-1}))-(f(v_{i_j+1})-f(v_{i_j}))]\\ &= \sum_{j=1}^{k}\left[2f(v_{i_j})-f(v_{i_j-1})-f(v_{i_j+1})\right]=-(\Delta f)(v).\end{aligned}$$

Lemma 7.1. *Let G be a graph with n vertices and c connected components. Then the Laplacian $\mathbf{L}(G)$ has rank $n-c$.*

Proof. Since $\operatorname{rank}(\mathbf{M})=n-c$, it suffices to show that $\operatorname{rank}(\mathbf{L})=\operatorname{rank}(\mathbf{M})$. Given a vector $\boldsymbol{v}$. If $\mathbf{M}\mathbf{M}^T\mathbf{v}=\mathbf{0}$, then $\mathbf{v}^T\mathbf{M}\mathbf{M}^T\mathbf{v}=\mathbf{0}$, i.e., $\|\mathbf{M}^T\mathbf{v}\|=0$, thus $\mathbf{M}^T\mathbf{v}=\mathbf{0}$. Clearly, $\mathbf{M}^T\mathbf{v}=\mathbf{0}$ implies $\mathbf{M}\mathbf{M}^T\mathbf{v}=\mathbf{0}$. So $\mathbf{L}$ and $\mathbf{M}^T$ have the same kernel. Hence $\operatorname{rank}\mathbf{L}=\operatorname{rank}\mathbf{M}^T=\operatorname{rank}\mathbf{M}$. □

Since $\mathbf{L}$ is a symmetric square matrix, all eigenvalues of $\mathbf{L}$ are real. Since $\mathbf{v}^T\mathbf{L}\mathbf{v}=\|\mathbf{M}^T\mathbf{v}\|^2\ge 0$ for each vector $\mathbf{v}\in\mathbb{R}^{E(G)}$. We see that $\mathbf{L}$ is semi-positive definite. If $\mathbf{v}$ is an eigenvector for the eigenvalue λ, i.e., $\mathbf{L}\mathbf{v}=\lambda\mathbf{v}$, then $\mathbf{v}^T\mathbf{L}\mathbf{v}=\lambda\mathbf{v}^T\mathbf{v}=\lambda\|\mathbf{v}\|^2\ge 0$, thus $\lambda\ge 0$. So all eigenvalues $\mathbf{L}$ are nonnegative, and 0 is always an eigenvalue, since $\mathbf{L}$ is singular. It is easy to see that the multiplicity of the zero eigenvalue is $c(G)$, the number of components of G. Let $G_1,\ldots,G_k$ be the connected components of G. Then the eigenspace of $\mathbf{L}$ for the eigenvalue 0 is the vector space generated by $\mathbf{1}_{V(G_i)}$, $1\le i\le k$. Let $\lambda_2(G)$

denote the smallest positive eigenvalue of G, called the **second smallest eigenvalue** of $\mathbf{L}$. The eigenvalues of $\mathbf{L}(G)$ are ordered as

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n, \quad n = |V(G)|.$$

Lemma 7.2. *For two $n \times n$ matrices A and B, the determinant of $A + B$ is given by*

$$\det(A+B) = \sum_{S \subseteq [n]} \det(A_S \cup B_{S^c}),$$

where $A_S \cup B_{S^c}$ is the matrix obtained from A by replacing the rows with indices not in S with the corresponding rows of B.

Proof. Write the rows of A as $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, and the rows of B as $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. The formula follows from the following direct calculation:

$$\begin{aligned}
\det(A+B) &= \sum_{\sigma} \operatorname{sin}(\sigma) \prod_{i=1}^{n} (a_{i\sigma(i)} + b_{i\sigma(i)}) \\
&= \sum_{\sigma} \operatorname{sin}(\sigma) \sum_{c_{i\,\sigma(i)} \in \{a_{i\,\sigma(i)},\, b_{i\,\sigma(i)}\}_{i=1}^{n}} \prod_{i=1}^{n} c_{i\sigma(i)} \\
&= \sum_{\boldsymbol{c}_i \in \{\boldsymbol{a}_i, \boldsymbol{b}_i\}_{i=1}^{n}} \sum_{\sigma} \operatorname{sin}(\sigma) \prod_{i=1}^{n} c_{i\sigma(i)} \\
&= \sum_{S \subseteq [n]} \det(A_S \cup B_{S^c}).
\end{aligned}$$

□

Theorem 7.3. *The characteristic polynomial of the Laplacian $\mathbf{L}$ of a graph G with n vertices is*

$$\det(t\mathbf{I} - \mathbf{L}) = \sum_{k=0}^{n-1} (-1)^k c_k t^{n-k},$$

where c_k is the number of rooted spanning forests of G with k edges. In particular, when G is connected, c_{n-1} is the n times of the number of spanning trees of G.

Proof. Write $(t\mathbf{I} - \mathbf{L})$ as $(t\mathbf{I} + (-\mathbf{L}))$ and $\mathbf{L} = \mathbf{M}\mathbf{M}^T$, where $\mathbf{M}$ is the vertex-edge incidence matrix of G. Applying Lemma 7.2,

$$\begin{aligned}
\det(t\mathbf{I} - \mathbf{L}) &= t^n + \sum_{k=1}^{n-1} (-1)^k t^{n-k} \sum_{S \subseteq V(G),\, |S|=k} \det\left(\mathbf{M}_S \mathbf{M}_S^T\right) \\
&= \sum_{k=0}^{n-1} (-1)^k c_k t^{n-k}.
\end{aligned}$$

Since $\mathbf{M}$ is totally unimodular, applying Cauchy-Binet formula, we see that

$$\det\left(\mathbf{M}_S \mathbf{M}_S^T\right) = \#\{F \subseteq E : |F| = |S|,\, \det(\mathbf{M}_{S \times F}) \neq 0\}.$$

Note that $\det(\mathbf{M}_{S\times F}) \neq 0$ implies that $|S| = |F|$, $S \subseteq V(F)$, and the subgraph (S, F) (with possible half-edges) contains no cycle. Let (S, F) be decomposed into connected components (S_i, F_i). Then $\det(\mathbf{M}_{S\times F}) = \prod_i \det(\mathbf{M}_{S_i\times F_i})$, which implies $|S_i| = |F_i|$, $S_i \subseteq V(F_i)$, and $\det(\mathbf{M}_{S_i\times F_i}) \neq 0$ for all i. Likewise, $\det(\mathbf{M}_{S_i\times F_i}) \neq 0$ implies that (S_i, F_i) (with possible half-edges) contains no cycle. We claim that each graph $(V(F_i), F_i)$ is a tree. Suppose $(V(F_i), F_i)$ is not a tree, i.e., it contains a cycle. Then its number of independent cycles is

$$n(F_i) := |F_i| - |V(F_i)| + 1 \geq 1.$$

Consequently, $|V(F_i)| \leq |F_i| = |S_i|$. Since $S_i \subseteq V(F_i)$, we have $(S_i, F_i) = (V(F_i), F_i)$, which contains a cycle, contradictory to that (S_i, F_i) contains no cycle.

Now each $(V(F_i), F_i)$ is a tree and $|V(F_i)| - |S_i| = |V(F_i)| - |F_i| = 1$. Then $V(F_i) \smallsetminus S_i$ is a single vertex, which can be viewed as a root of the tree $(V(F_i), F_i)$. So each (S_i, F_i) may be considered as a rooted tree $(V(F_i), F_i)$ with the root v such that $\{v\} = V(F_i) \smallsetminus S_i$. Conversely, if $S_i \subseteq V(F_i)$, $|S_i| = |F_i|$, and $(V(F_i), F_i)$ is a tree, then it is clear that $\det(\mathbf{M}_{S_i\times F_i}) \neq 0$ by expansion along its v-row with v a leaf. Thus we obtain

$$c_k = \#\{\text{acyclic}\,(S, F) : F \subseteq E,\ S \subseteq V(F),\ |S| = |F| = k\},$$

where each such (S, F) is identified as a rooted spanning forest F with k edges, i.e., each component of F is specified a root. In particular, c_{n-1} is the number of rooted spanning trees, which is n times of the number of spanning trees of G. $\square$

Corollary 7.4. *Let G be a graph with n vertices. If the eigenvalues of $\mathbf{L}(G)$ are linearly ordered as $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ (multiply listed with multiplicities). Then the number of spanning trees of G equals*

$$\frac{1}{n}\prod_{i=2}^{n} \lambda_i.$$

Proof. Let $\phi(t)$ denote the characteristic polynomial of $\mathbf{L}$, i.e., $\phi(t) = \det(t\mathbf{I} - \mathbf{L})$. Then $\phi(t) = t(t-\lambda_2)\cdots(t-\lambda_n)$. The coefficient of t in $\phi(t)$ is $(-1)^{n-1}\lambda_2\cdots\lambda_n$, which is also the number of spanning trees of G times $(-1)^{n-1}n$ by Theorem 7.3. $\square$

8 Tree Encoding

Let T be a tree with n vertices labeled $1, 2, \ldots, n$, where $n \geq 2$. We can encode T as a sequence $(v_1, v_2, \ldots, v_{n-2})$ of $1, 2, \ldots, n$ of length $n-2$ as follows: Find a leaf u_1 of $T_1 := T$ with minimal label. Then u_1 is adjacent with a unique vertex v_1 in T_1. Delete u_1 and its edge from T_1 to obtain a tree $T_2 := T_1 - u_1$. In general, for the tree T_i with minimal label leaf u_i with $i \leq n-2$, find the vertex v_i adjacent with u_i in T_i. Set $T_{i+1} := T_i - u_i$. We have a sequences $(u_1, \ldots, u_{n-2})$ and $(v_1, \ldots, v_{n-2})$, where $u_i \notin \{u_1, \ldots, u_{i-1}\} \cup \{v_i, \ldots, v_{n-2}\}$. Moreover, for T_{n-1} which has exactly two vertices, we still have vertices $u_{n-1}, v_{n-1} \in T_{n-1}$ and $u_{n-1} \notin \{u_1, \ldots, u_{n-2}\} \cup \{v_{n-1}\}$.

Given a sequence $(v_1, v_2, \ldots, v_{n-2})$ of $1, 2, \ldots, n$ of length $n-2$. We may recover the tree T back from the sequence as follows: Find the minimal element u_1 from $[n] - \{v_1, \ldots, v_{n-2}\}$. Make u_1 and v_1 adjacent, and delete u_1 from $[n]$ and v_1 from $(v_1, v_2, \ldots, v_{n-2})$. In general, find the minimal element $u_i \in [n] - \{u_1, \ldots, u_{i-1}\} - \{v_i, \ldots, v_{n-2}\}$, connect u_i and v_i, $i =$

$1, \ldots, n-2$. At $i = n-2$, we must have $u_{n-2} \in [n] - \{u_1, \ldots, u_{n-3}\} - \{v_{n-2}\}$, which has exactly two vertices, and these two vertices must be adjacent.

For the complete graph K_n, each spanning tree is encoded into a sequence $(v_1, \ldots, v_{n-2})$ of $1, \ldots, n$, each v_i has n choices. So K_n has n^{n-2} spanning trees.

Let ω be an orientation ω on a graph $G = (V, E)$. If ω is acyclic, the vertices can be linearly ordered as $v_1, \ldots, v_n$ so that each directed edge is like of the form (v_i, v_j) with $i < j$. Let $U_i = [X_i, X_i^c]$, where $X_i = \{v_1, \ldots, v_i\}$, $i = 1, \ldots, n-1$. Then $g := \sum_{i=1}^{n-1} I_{U_i}$ is a tension of (G, ω) such that $g(e) > 0$ for all $e \in \omega$. If ω totally cyclic, for each edge $e \in E$ there is a directed circuit C_e of (G, ω) such that $e \in C_e$. Then $f := \sum_{e \in E} I_{C_e}$ is a flow such that $f(e) > 0$ for all $e \in \omega$.

Exercises

1. Let $G = (V, E)$ be a connected graph, and let T be a spanning tree of G. Show that the number of vertices of T is $|V| - 1$.

2. A graph G is said to be *even* if the degree of every vertex is even. Show that each even graph can be decomposed into edge-disjoint cycles.

3. Let M be the signed vertex-edge incidence matrix of a connected graph G with an orientation. (a) Given an edge subset $S \subseteq E(G)$. Show that column vectors of M indexed by members of S are linearly dependent if and only if the subgraph $G(S)$ induced by S contains a cycle. (b) Show that the rank of M is $|V| - 1$. (c) Show that $MM^T = D - A$, where D is the diagonal matrix whose vth entry is the degree of v, and A is the adjacency matrix.

4. Let M be the signed vertex-edge incidence matrix of a connected graph G with an orientation. (a) Show that the solution space of $Mx = 0$ is isomorphic to the flow space of G. (b) show that the rank of M is $|V| - 1$. (c) Show the cycle space (=flow space) of G is $|E| - |V| + 1$. (d) Show that the tension space of G is $|V| - 1$. (e) If G is not necessarily connected, say, G has $c(G)$ components, then the tension space of G has dimension $|V(G)| - c(G)$, and the cycle space of G has dimension $|E(G)| - |V(G)| + c(G)$.