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 $\text{End} : E \to M_2(V)$, where $M_2(V)$ is set of all 2-element sub-multisets of $V$. We usually write $V = V(G)$, $E = E(G)$, and $\text{End} = \text{End}_G$. For each edge $e \in E$ with $\text{End}(e) = \{u, v\}$, 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 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 $W = v_0e_1v_1 \cdots v_{l-1}e_lv_l$ 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 \( P = v_0e_1v_1 \cdots v_{l-1}e_lv_l \) 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 affect 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 graph 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

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

Proof. Trivial. \hfill \Box

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

Proof. We proceed by induction on $\Delta = \Delta(G) \geq 3$. 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 $G \neq K_{\Delta(G)+1}$, 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 $|V(G)| \geq \Delta(G) + 2$.

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

Case 2. $G \setminus u$ is connected for all $u \in V(G)$, but there two vertices $v, w$ such that $G \setminus \{v, w\}$ is disconnected. Let $C$ be one component of $G \setminus \{v, w\}$. Suppose there is no edge from $v$ to $C$; then $G \setminus w$ is disconnected, which is Case 1, a contradiction. Thus there is at least one edge from $v$ to $C$, so does $w$. Since $G \setminus (V \cup \{v, w\})$ consists of components of $G \setminus \{v, w\}$ other than $C$, there is at least one edge from $v$ to $G \setminus (V \cup \{v, w\})$, so does $w$.

Let $G_1$ denote the subgraph obtained from $G$ by deleting $C$, and $G_2$ the subgraph obtained from $G$ by deleting $G \setminus (C \cup \{v, w\})$. By induction, $G_1$ and $G_2$ have proper $\Delta(G)$-colorings. Note that $v, w$ have degree at most $\Delta(G) - 1$ in both $G_1$ and $G_2$. It follows by induction that $G_3 := G_1 \cup vw$ and $G_4 := G_2 \cup vw$ have proper $\Delta(G)$-colorings respectively, unless one of them is complete (if either of them is an odd cycle, then we still have $\Delta(G)$-coloring since $\Delta(G) \geq 3$). By rearranging the colors we may assume that $v$ has the same color in both $\Delta(G)$-colorings of $G_3$ and $G_4$, do does $w$. We thus obtain a proper $\Delta(G)$-coloring of $G$ by combining the two colorings of $G_3$ and $G_4$.

If one of $G_3$ and $G_4$, say $G_3$, is a complete subgraph, then it is a click on $\Delta(G) + 1$ vertices, since the degree of $v, w$ in both $G_1, G_2$ are at most $\Delta(G) - 1$. Then the degree of $v, w$ must be

$$\Box$$

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 $\chi(G, k) \neq 0$.

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

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

  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 $M$ be incidence matrix of $D$, i.e., $M = [m_{va}]$, 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

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

and

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

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 $B$ of the tension space $T(D)$ and the basis matrix $C$ of the flow space $F(D)$. For an edge subset $S \subseteq A(D)$, we denote by $B|_S$ (or just $B_S$) the submatrix of $B$ consisting of the columns of $B$ 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-known 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

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

$$C = \begin{pmatrix} e'_1 & e'_2 & \cdots & e'_m & e'_1 & e'_2 & \cdots & e'_n \\ e'_1 & * & * & \cdots & * & 1 & 0 & \cdots & 0 \\ e'_2 & * & * & \cdots & * & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ e'_n & * & * & \cdots & * & 0 & 0 & \cdots & 1 \end{pmatrix},$$

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

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

(a) The columns of $B_S$ 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

$$\sum_{a \in A} f(a) \mathbf{b}(a) = 0,$$

i.e., $\langle f, b_i \rangle = 0$ for $1 \leq i \leq m$,

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 determinants 1, $-1$, or 0; and to be totally unimodular if its all square submatrices have determinants 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

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

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

$$\mathbf{B}' = [\mathbf{B}'|_{F} \mathbf{B}'|_{F^c}], \quad \mathbf{C}' = [\mathbf{C}'|_{F} \mathbf{C}'|_{F^c}].$$
It is clear that there exist square matrices $P, Q$ such that
\[ B' = PB, \quad C' = QC. \]

Then
\[ B' = P[I A] = [P, PA], \quad C' = Q[G I] = [QG, Q] \]

It follows that $P = B'|_F$ and $Q = C'|_F$. Hence $B' = (B'|_F)B$ and $C' = (C'|_F)C$. \hfill \Box

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

**Proof.** Given a maximal spanning forest $F$ of $G$. Let $B', C'$ be basis matrices of the tension space and the flow space of $G$ relative to $F$ respectively. There exists unimodular matrices $P$ and $Q$ such that $B' = PB$ and $C' = QC$. Restrict both sides to $F^i, F^{ie}$ respectively, we obtain
\[ B'|_{F^i} = P(B'|_{F^i}), \quad C'|_{F^{ie}} = Q(C'|_{F^{ie}}). \]

Since $B'|_{F^i}, C'|_{F^{ie}}$ are identity matrices by definition, we see that
\[ \det(P) \det(B'|_{F^i}) = 1, \quad \det(Q) \det(C'|_{F^{ie}}) = 1. \]

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

Given edge subsets $S \subseteq A(D)$. If $|S| = |V(D)| - 1$ and $S$ is not spanning tree, then $S$ contains a cycle. Thus $\det(B|_S) = 0$ 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 $\det(C|_S) = 0$ by Lemma 5.2. \hfill \Box

**Proposition 5.4.** The incidence matrix $M$ 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 $v \not\in V(F)$, then the $v$-row of $M|_{S \times F}$ is a zero row; clearly, $\det(M|_{S \times F}) = 0$. We may assume that $S \subseteq V(F)$. We see that $M|_{S \times F}$ 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 $\det(M|_{S \times F}) = 0$. 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, $v \not\in S$. Then the $e$-column of $M|_{S \times F}$ has 1 or $-1$ at $(u, e)$ and 0 elsewhere. Thus by the expansion along the $e$-column,
\[ \det(M|_{S \times F}) = \pm \det(M|_{(S \setminus u) \times (F \setminus e)}) = \pm 1. \]

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

**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 $m > n$, then $\det(AB) = 0$, since
\[ \text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\} \leq n. \]

If $m \leq n$, we have the following Cauchy-Binet formula.
Proposition 6.1 (Cauchy-Binet Formula). Let $A$ be an $m \times n$ matrix and $B$ an $n \times m$ matrix. If $m \leq n$, then

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

(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

$$
\det(C) = \det \left( \sum_{k_1=1}^{n} a_{1k_1} b_{k_11} \cdots \sum_{k_m=1}^{n} a_{1k_m} b_{k_m1} \begin{array}{c} \vdots \end{array} \begin{array}{c} \vdots \end{array} \right.
$$

$$
\begin{array}{ccc}
\sum_{k_1=1}^{n} a_{m_k} b_{k_11} & \cdots & \sum_{k_m=1}^{n} a_{mk_m} b_{k_m1} \\
\vdots & \ddots & \vdots \\
\sum_{k_1=1}^{n} a_{mk_1} b_{k_11} & \cdots & \sum_{k_m=1}^{n} a_{mk_m} b_{k_m1}
\end{array}
$$

$$
= \sum_{k_1=1}^{n} \cdots \sum_{k_m=1}^{n} \det \left( \begin{array}{ccc}
\sum_{k_1=1}^{n} a_{1k_1} b_{k_11} & \cdots & a_{1k_m} b_{k_m1} \\
\vdots & \ddots & \vdots \\
\sum_{k_1=1}^{n} a_{mk_1} b_{k_11} & \cdots & \sum_{k_m=1}^{n} a_{mk_m} b_{k_m1}
\end{array} \right) b_{k_11} \cdots b_{k_m1}.
$$

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

$$
\det(C) = \sum_{1 \leq k_1, \ldots, k_m \leq n, k_i \neq k_j} \det(A|_{\{k_1, \ldots, k_m\}}) b_{k_11} \cdots b_{k_m1}
$$

$$
= \sum_{1 \leq t_1 < \cdots < t_m \leq n} \sum_{\sigma \in S_m} \det(A|_{\{t_{\sigma(1)}, \ldots, t_{\sigma(m)}\}}) b_{t_{\sigma(1)}1} \cdots b_{t_{\sigma(m)}1}
$$

$$
= \sum_{1 \leq t_1 < \cdots < t_m \leq n} \det(A|_{\{t_1, \ldots, t_m\}}) \sum_{\sigma \in S_m} \sin(\sigma) b_{t_{\sigma(1)}1} \cdots b_{t_{\sigma(m)}1},
$$

where $S_m$ is the set of all permutations of $\{1, \ldots, m\}$. Set $S = \{t_1, \ldots, t_m\}$, we have

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

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

$$
\det(BB^T) = \det(CC^T).
$$

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

$$
\det(BB^T) = \sum_{S \subseteq E, |S| = m} \det(B_S) \det(B_S^T) = \sum_{S \subseteq E, |S| = m, S \text{ cyclic}} (\det B_S)^2 = t(G),
$$

where $S$ is the dimension of the flow space of $G$. Thus, we have

$$
\det(BB^T) = \sum_{S \subseteq E, |S| = m} \det(B_S) \det(B_S^T) = \sum_{S \subseteq E, |S| = m, S \text{ cyclic}} (\det B_S)^2 = t(G).
$$

9
\[ \det(\mathbf{CC}^T) = \sum_{S \in E, |S| = m} \det(\mathbf{C}_S) \det(\mathbf{C}_S^T) = \sum_{S \in E, |S| = m} (\det \mathbf{C}_S)^2 = t(G). \]

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. \( \square \)

**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 \( \mathbf{x} \geq 0 \).

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

\[
\mathbf{y}^T \mathbf{A} \geq \mathbf{0}^T, \quad \mathbf{y}^T \mathbf{b} < 0.
\]

**Proof.** Farkas’s Lemma is just the geometric interpretation: Let \( \mathbf{a}_1, \ldots, \mathbf{a}_n \) denote the columns of \( \mathbf{A} \). Let \( \text{Cone}(\mathbf{A}) \) denote the convex cone generated by \( \mathbf{a}_1, \ldots, \mathbf{a}_n \). Let \( \mathbf{x}^T = (x_1, \ldots, x_n) \geq 0 \). 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 \( \mathbf{b} \in \text{Cone}(\mathbf{A}) \).

Let \( \mathbf{y}^T = (y_1, \ldots, y_m) \). Consider the hyperplane \( H = \{ \mathbf{z} \in \mathbb{R}^m : \langle \mathbf{z}, \mathbf{y} \rangle = 0 \} \). Then \( \mathbf{y}^T \mathbf{A} \geq 0 \) means that \( \langle \mathbf{a}_i, \mathbf{y} \rangle \geq 0, \ i = 1, \ldots, n \), i.e., \( \text{Cone}(\mathbf{A}) \) 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 \( \text{Cone}(\mathbf{A}) \), which is equivalent to \( \mathbf{b} \not\in \text{Cone}(\mathbf{A}) \).

Assume that the first statement is true, i.e., there exists a vector \( \bar{\mathbf{x}} \geq 0 \) 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 \( \mathbf{A}^T \bar{\mathbf{y}} \geq 0 \) 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. \( \square \)

**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 \( \mathbf{x}' \geq 0 \). \( \square \)

**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 \( \mathbf{x}' \geq 0 \). \( \square \)
7 Graph Laplacian

Let \( D = (V, A) \) be a connected loopless digraph in which half arcs are allowed, and \( M \) the incidence matrix of \( D \). For each vertex \( v \in V \), let \( M_v \) denote the matrix obtained from \( M \) by deleting the row corresponding to the vertex \( v \). A Kirchhoff matrix of \( D \) is the matrix \( K := M_v \) for a vertex \( v \) of \( D \). The Laplace matrix of the underlying graph \( G \) of the digraph \( D \) is the matrix 
\[
L := MM^T.
\]

Let \( A \) be the adjacency matrix of \( G \), whose \((u, v)\)-entry is the number edges incident with \( u \) and \( v \), each loop is counted twice. Let \( D \) 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 
\[
L = D - A.
\]

In fact, let \( e_1, \ldots, e_k \) be the links and \( f_1, \ldots, f_l \) the loops at \( v \). Recall that \((v, e_i)\)-entry in \( M \) is either 1 or \(-1\), and \((v, f_j)\)-entry in \( M \) is always 0. So the \((v, v)\)-entry of \( L \) is 
\[
\sum_{i=1}^{k} \omega(v, e_i)^2 + \sum_{j=1}^{l} \omega(v, f_j)^2 = k = \text{number of links at } v.
\]

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

For distinct vertices \( u, v \in V \), let \( g_1, \ldots, g_m \) be the edges between \( u \) and \( v \). The \((u, v)\)-entry of \( L \) is 
\[
\sum_{i=1}^{m} \omega(u, g_i) \omega(v, g_i) = -m = -a_{uv} = \text{number of edge between } u \text{ and } v.
\]

We thus have \( L = D - A \).

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

**Proof.** Since rank \((M) = n - c \), it suffices to show that rank \((L) = \text{rank } (M) \). Given a vector \( v \). If \( MM^Tv = 0 \), then \( v^TMM^Tv = 0 \), i.e., \( \|M^Tv\| = 0 \), thus \( M^Tv = 0 \). Clearly, \( M^Tv = 0 \) implies \( MM^Tv = 0 \). So \( L \) and \( M^T \) have the same kernel. Hence rank \( L = \text{rank } M^T = \text{rank } M \). \( \square \)

Since \( L \) is a symmetric square matrix, all eigenvalues of \( L \) are real. Since \( v^TLv = \|M^Tv\|^2 \geq 0 \) for each vector \( v \in \mathbb{R}^{E(G)} \). We see that \( L \) is semi-positive definite. If \( v \) is an eigenvector for the eigenvalue \( \lambda \), i.e., \( Lv = \lambda v \), then \( v^TLv = \lambda v^Tv = \lambda \|v\|^2 \geq 0 \), thus \( \lambda \geq 0 \). So all eigenvalues \( L \) are nonnegative, and 0 is always an eigenvalue, since \( L \) 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 \( L \) for the eigenvalue 0 is the vector space generated by \( 1_{V(G_i)} \), \( 1 \leq i \leq k \). Let \( \lambda_2(G) \) denote the smallest positive eigenvalue of \( G \), called the second smallest eigenvalue of \( L \). The eigenvalues of \( L(G) \) 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 \( a_1, \ldots, a_n \), and the rows of \( B \) as \( b_1, \ldots, b_n \). The formula follows from the following direct calculation:
\[
\det(A + B) = \sum_{\sigma} \sin(\sigma) \prod_{i=1}^{n} (a_{i\sigma(i)} + b_{i\sigma(i)})
\]
\[
= \sum_{\sigma} \sin(\sigma) \sum_{c_{i\sigma(i)} \in \{a_{i\sigma(i)}, b_{i\sigma(i)}\}} \prod_{i=1}^{n} c_{i\sigma(i)}
\]
\[
= \sum_{\sigma} \sum_{c_{i} \in \{a_i, b_i\}} \sin(\sigma) \prod_{i=1}^{n} c_{i\sigma(i)}
\]
\[
= \sum_{S \subseteq [n]} \det(A_S \cup B_{S^c}).
\]
\[\square\]

Theorem 7.3. The characteristic polynomial of the Laplacian \( L \) of a graph \( G \) with \( n \) vertices is
\[
\det(tI - L) = \sum_{k=0}^{n-1} (-1)^k c_k t^{n-k},
\]
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 \((tI - L)\) as \((tI + (-L))\) and \( L = MM^T \), where \( M \) is the vertex-edge incidence matrix of \( G \). Applying Lemma 7.2,
\[
\det(tI - L) = t^n + \sum_{k=1}^{n-1} (-1)^k t^{n-k} \sum_{S \subseteq V(G), |S|=k} \det(M_S M_S^T)
\]
\[
= \sum_{k=0}^{n-1} (-1)^k c_k t^{n-k}.
\]

Since \( M \) is totally unimodular, applying Cauchy-Binet formula, we see that
\[
\det(M_S M_S^T) = \#\{F \subseteq E : |F| = |S|, \det(M_{S \times F}) \neq 0\}.
\]
Note that \( \det(M_{S \times F}) \neq 0 \) 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 \( \det(M_{S \times F}) = \prod_i \det(M_{S_i \times F_i}) \), which implies \( |S_i| = |F_i|, S_i \subseteq V(F_i) \), and
det(M_{S_i \times F_i}) \neq 0 \text{ for all } i. \text{ Likewise, } det(M_{S_i \times F_i}) \neq 0 \text{ implies that } (S_i, F_i) \ (\text{with possible half-edges}) \text{ contains no cycle. We claim that each graph } (V(F_i), F_i) \text{ is a tree. Suppose } (V(F_i), F_i) \text{ is not a tree, i.e., it contains a cycle. Then its number of independent cycles is }
\begin{equation}
n(F_i) := |F_i| - |V(F_i)| + 1 \geq 1.
\end{equation}
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, a contradictory to that \((S_i, F_i)\) contains no cycle.

Now each \((V(F_i), F_i)\) is a tree and \(V(F_i) \setminus S_i\) 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 \(\{v\} = V(F_i) \setminus S_i\). 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 \(det(M_{S_i \times F_i}) \neq 0\) 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 the number of spanning trees of \(G\).

**Corollary 7.4.** Let \(G\) be a graph with \(n\) vertices. If the eigenvalues of \(L(G)\) 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 \(L\), i.e., \(\phi(t) = det(tI - L)\). 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.

**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 \(v\)th 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 \(|E(G)| - |V(G)| + c(G)\).