Math 184

Lecture Notes Section 5.1 *

Instructor: Swee Hong Chan

NOTE: The notes is a summary for materials discussed in the class and is not supposed to substitute the textbook. In particular, the proofs here might omit some details for brevity, and are not supposed to be how you write proofs in the exam. Please refer back to the textbook when studying for exams; materials that appear in the textbook but do not appear in the lecture notes might still be tested. Please send me an email if you find typos.

1 Graphs

We will spend the remainder of this course talking about graphs. Here is a scenario you should have in mind when talking about graphs.

Example 1. You are selected as the CEO of Friendly Instructor Farm. You are in charge of six farms in a rural area, no two of which are connected by roads. Because of this, you want to build some roads between some of the farms. Give three examples of a plan to connect these farms. △

Analysis: Here are pictures of three plans that you might want to consider.

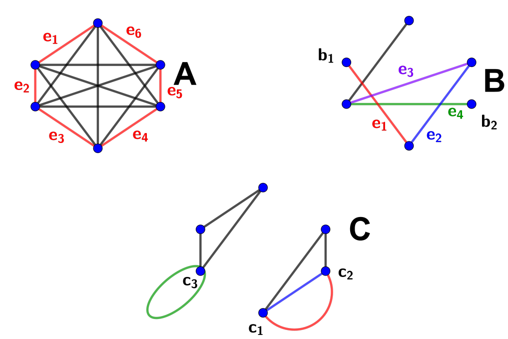

Figure 1: Three possible road systems among six farms

Let's analyze each plan separately:

- Plan A provides the following road system:

*Version date: Tuesday 10^{th} March, 2020, 01:10

- Each farm is connected to another farm by a direct road, and there are $\binom{6}{2} = 15$ roads in total. This plan might be prohibitively expensive.
- Note that some of the roads in plan A might cross other roads' path. We allow this in our road planning, as we can build elevated highways at those crossings.

- Plan B provides the following road system:
  - In this plan, every farm is connected to any other farm through some (possibly indirect) route, e.g., Farm b_1 is connected to Farm b_2 by going through the route e_1, e_2, e_3, e_4.
  - This plan is cheapest to build as it uses the least number of roads.
- Plan C provides the following road system:
  - In this plan, we have two different groups of farms; there are no roads connecting these two groups. This is not an ideal plan.
  - The red road and the blue road serve the same purpose: connecting Farm c_1 to Farm c_2, so one of them is redundant.
  - The green road starts and ends at the same farm, so this road is also redundant.

□

Plan A, B, and C are examples of *(undirected) graphs*, where the farms are the *vertices* of the graph, and the roads are the *edges* of the graph.

Definition 2 (Graph)**.** A *graph* is a pair $G = (V, E)$, where

$$V \text{ is the set of } \textit{vertices} \text{ (points in the pictue);} \tag{1}$$

$$E \text{ is the set of } \textit{edges} \text{ connecting two vertices in } G. \tag{2}$$

We will usually denote an edge e connecting vertex x and y as $\{x, y\}$. We also denote by $|V|$ the number of vertices of G, and by $|E|$ the number of edges of G. △

Here are some basic terminologies of graphs that you need to know, which we split into three groups

Definition 3.
- A *loop* is an edge that connects a vertex to itself, e.g. the green edge in Plan C in Figure 1.
- A *multiple edges* is a group of two or more edges that connect the same pair of vertices, e.g. the red and blue edge in Plan C.
- A graph is *simple* if it has neither loops nor multiple edges, e.g., plan A and B.
- A graph is *complete* if every pair of vertices is connected by an edge, e.g. plan A. △

Definition 4.
- A *walk* is a series of edges $e_1 e_2 \ldots e_k$ that lead from a vertex to another, e.g., the route $e_1 e_2 e_1 e_2 e_3 e_4$ in Plan B.
- A *path* is a walk that does not go through any vertex twice, e.g., the route $e_1 e_2 e_3 e_4$ in plan B.
- A *cycle* is a walk whose starting point is the same as its endpoints, but which otherwise has no repeated vertices, e.g., the route $e_1 e_2 e_3 e_4 e_5 e_6$ in plan A.
- A graph is *connected* if there is a walk between every pair of vertices in the graph, e.g., plan A and B are connected, plan C is NOT connected.
- A *connected component* of a graph is a maximal connected subgraph of the graph, e.g., plan A and B have one connected component, plan C have two connected components. △

Definition 5. The *degree* of a vertex v, denoted by $\deg(v)$, is the number of edges adjacent to v (i.e., the number of edges that has v as one of its endpoints). △

Remark 6. It is not hard to see that graph theory has many applications in real-life, e.g.,

- Used in engineering for planning roads between cities;
- Used in computer science to represent networks of communications, data organization, etc;
- Used in biology and conservation efforts where a vertex can represent regions where certain species exist (or inhabit) and the edges represent migration paths or movement between the regions. △

The following is the most basic result in graph theory.

Lemma 7 (Handshaking lemma)**.** *Let $d_1, d_2, \ldots, d_{|V|}$ be the degree of vertices of a graph. Then*

$$d_1 + d_2 + \ldots + d_{|V|} = 2|E|.$$

Proof. Note that both sides of the lemma count all *endpoints* of all edges of G. The left side of the equation counts them by the vertices (as vertex i is the endpoint of d_i edges), and the right side of the equation counts them by the edges (each edge has two endpoints). □

Remark 8. We adopt the convention that a loop vertex contributes +2 instead of +1 to a vertex, e.g., vertex c_3 in Plan C has degree 4 instead of 3. This convention is so that the conclusion of Handshaking lemma holds even for nonsimple graphs. △

2 Trees and forests

We return to the scenario of Example 1 where you are tasked to build roads to connect farms. As we have mentioned, plan B in Figure 1 is the cheapest to build as it connects all the farms with the least number of roads. We will formalize this property by the following definition.

Definition 9. A *tree* T is a simple connected graph with no cycles of length more than or equal to 3. △

As it turns out, there are many equivalent definitions of trees, as shown below.

Lemma 10. *Let $G = (V, E)$ be a connected simple graph. Then the following are equivalent:*

- *G has no cycles of length more than or equal to 3 (i.e., G is a tree by the definition above);*
- *G is* minimally connected*, i.e., removing any edges of G will make it no longer connected.*
- *G has exactly $|V| - 1$ edges.*

Proof. Please read the textbook for the proof of this lemma. □

A generalization of trees is, not surprisingly, called forests.

Definition 11 (Forest)**.** A *forest* is a graph whose connected components are all trees, e.g., Figure 2. △

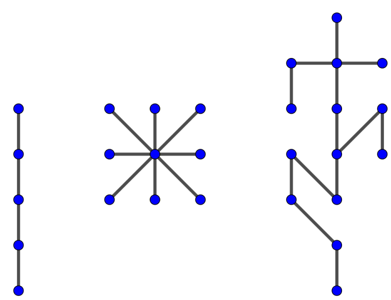

Figure 2: A forest consisting of three trees.

3 Counting trees via the matrix tree theorem

In this course we will focus on counting all kinds of graphs. For the purposes of this section, we assume these two things:

- The vertices of a graph will be labeled by the set $[n] = \{1, 2, \ldots, n\}$. These graphs are called *labeled graphs.*
- We consider two labeled graphs as *identical* if they have the same set of vertices and edges, i.e., when the number of edges between i and j is the same in both graphs for all $i, j \in V$, e.g., the two labeled graphs in Figure 3 are identical.

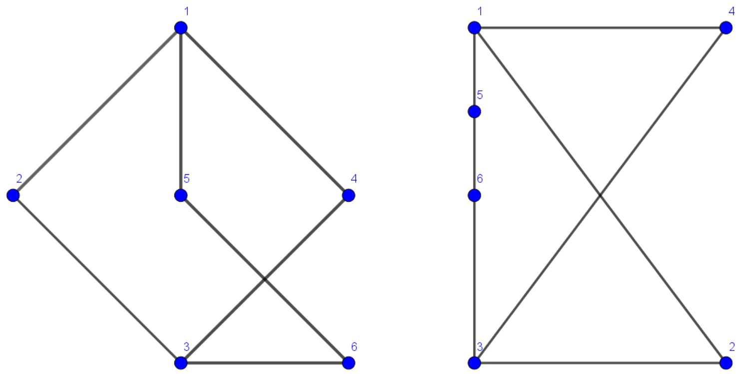

Figure 3: These two graphs are identical as labeled graphs, even if they are drawn differently.

As it turns out, there is a nice, simple, elegant formula for counting the number of labeled trees.

Theorem 12 (Cayley's formula)**.** *For all positive integers n, the number of labeled trees on vertex set $[n]$ is*

$$n^{n-2}.$$

Example 13. Let $n = 3$. There are three labeled trees on [3] (see Figure 4) as predicted by Cayley's formula. △

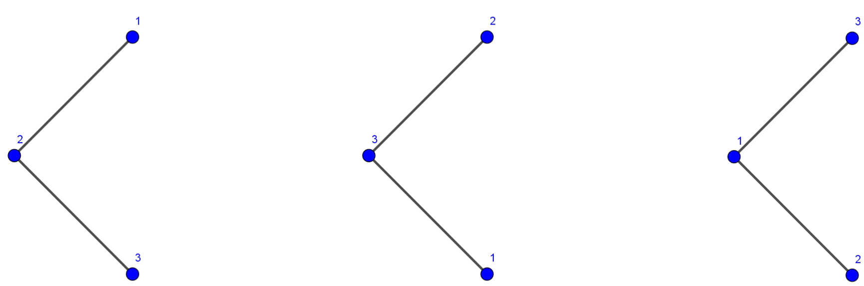

Figure 4: Three labeled trees on [3].

Many proofs of Cayley's formula are known (i.e., more then a dozen). The textbook chooses to present a bijective proof that is due to Joyal. We will choose to present a different proof using the matrix-tree theorem, which count the number of spanning trees of any given connected graph.

Definition 14 (Spanning tree)**.** A *spanning tree* of a (connected) graph G is a subgraph of G that is connected and contains all vertices of G. See Figure 5 for examples. $\triangle$

In particular, Cayley's formula counts the number of spanning trees of the simple complete graph with $[n]$ as the vertex set.

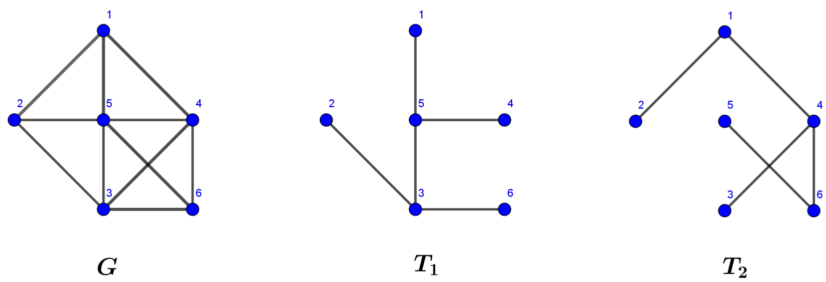

Figure 5: A graph G with two of its spanning trees T_1 and T_2.

We will need the following definition to state the matrix-tree theorem.

Definition 15 (Laplacian matrix)**.** Let $G = (V, E)$ be a simple graph with $V = [n]$. The *degree matrix* of G is the matrix

$$\begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix},$$

where d_i is the degree of the vertex i. The *adjacency matrix* of G is the matrix $A := (a_{i,j})_{1 \le i,j \le n}$ given by

$$a_{i,j} = \begin{cases} 1 & \text{if } \{i,j\} \in E; \\ 0 & \text{otherwise.} \end{cases}$$

The *Laplacian matrix* L of G is the matrix

$$L = D - A. \qquad \triangle$$

In particular, L is always a symmetric $n \times n$ matrix.

Example 16. Let G be the graph in Figure 5. The Laplacian matrix of G is

$$\begin{bmatrix} 3 & -1 & 0 & -1 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 4 & -1 & -1 & -1 \\ -1 & 0 & -1 & 4 & -1 & -1 \\ -1 & -1 & -1 & -1 & 5 & -1 \\ 0 & 0 & -1 & -1 & -1 & 3 \end{bmatrix} \qquad \triangle$$

Theorem 17 (Matrix-tree theorem). *Let G be a simple graph. Let $i \in [n]$, and let L_i be the Laplacian matrix L with row and column i removed. Then the number of spanning trees of G is equal to*

$$\det(L_i).$$

Proof. The proof of the matrix tree theorem can be found in a separate pdf file in CCLE. $\square$

Remark 18. Note that it follows from the matrix-tree theorem that $\det(L_i)$ does NOT depend on the chosen $i \in [n]$. Also note that it follows from the matrix-tree theorem that $\det(L_i) = 0$ if the graph G is not connected. $\triangle$

Example 19. Let G be the graph in Figure 5. The matrix L_1 of G is the 5×5 matrix given by

$$\begin{bmatrix} \cancel{3} & \cancel{-1} & \cancel{0} & \cancel{-1} & \cancel{-1} & \cancel{0} \\ \cancel{-1} & 3 & -1 & 0 & -1 & 0 \\ \cancel{0} & -1 & 4 & -1 & -1 & -1 \\ \cancel{-1} & 0 & -1 & 4 & -1 & -1 \\ \cancel{-1} & -1 & -1 & -1 & 5 & -1 \\ \cancel{0} & 0 & -1 & -1 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 0 & -1 & 0 \\ -1 & 4 & -1 & -1 & -1 \\ 0 & -1 & 4 & -1 & -1 \\ -1 & -1 & -1 & 5 & -1 \\ 0 & -1 & -1 & -1 & 3 \end{bmatrix}.$$

The determinant of L_1 is equal to 209 (by Zhin Pyoung's calculation), so by the matrix-tree theorem G has 209 spanning trees.

Note: The first person that find the exact value of X and sends it the instructor via an email will get 1 extra credit point. $\triangle$

We now apply the matrix-tree theorem to prove Cayley's formula.

Proof of Cayley's formula. The Laplacian matrix of the simple complete graph with n vertices is the $n \times n$ matrix

$$L = \begin{bmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{bmatrix} \qquad \text{(This is an } n \times n \text{ matrix)},$$

which is an $n \times n$ matrix with diagonal entries equal to $n-1$ and off-diagonal entries equal to -1. Deleting the first row and the first column, we get the matrix L_1

$$L_1 = \begin{bmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{bmatrix} \qquad \text{(This is an } (n-1) \times (n-1) \text{ matrix)},$$

which is the $(n-1) \times (n-1)$ matrix with diagonal entries equal to $n-1$ and off diagonal entries equal to -1 (Note that the size of L_1 is different from L!). Recall that the determinant of a matrix does not change under elementary row-and-column operations. Let's add the second, third, $\dots$, $n-1$-th row of L_1 into the first row of L_1. We then get the following matrix:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{bmatrix}.$$

which is equal to the matrix L_1 except that all the entries of the first row are replaced by 1. Now, let's add the first row of the matrix above to its second row, third row, $\dots$, $(n-1)$-th row to get:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n \end{bmatrix},$$

This matrix is a triangular matrix with the diagonal entries given by the vector $(1, n, n, \dots, n)$ of length $n-1$. It then follows the determinant of the matrix above is equal to

$$\det(L_1) = n^{n-2}.$$

It then follows from the matrix-tree theorem that the number of labeled trees on with vertices from $[n]$ is equal to n^{n-2}, which proves Cayley's formula.

□

Remark 20. Read other parts of Section 5.1 not covered in the notes! △