MIT 18.211: COMBINATORIAL ANALYSIS

FELIX GOTTI

Lecture 27: Adjacency Matrices and The Matrix-Tree Theorem

The Adjacency Matrix. A helpful way to represent a graph G is by using a matrix that encodes the adjacency relations of G. This matrix is called the adjacency matrix of G and facilitates the use of algebraic tools to better understand graph theoretical aspects. In the first part of this lecture, we provide a couple of applications of the adjacency matrix representation.

Definition 1. Let G be a multigraph with $V(G) = [n]$. Then the *adjacency matrix* A of G is defined as follows:

- if G is undirected, then A_{jk} is the number of edges between j and k, and
- if G is directed, then A_{jk} is the number of edges from j to k.

We observe that the adjacency matrix of any undirected multigraph is symmetric. However, this is not always the case for the adjacency matrix of a directed multigraph. As we proceed to show, adjacency matrices can be used to compute number of walks in a graph.

Proposition 2. *Let G be a graph (directed or undirected) on $[n]$ with adjacency matrix A. For any $j, k \in [n]$ and $\ell \in \mathbb{N}$, there are A^{ℓ}_{jk} walks of length ℓ from j to k.*

Proof. We proceed by induction on ℓ. If $\ell = 1$, then A_{jk} is the number of edges from j to k, which is the number of walks of length 1 from j to k. Assume the statement of the proposition holds for $\ell \in \mathbb{N}$, and fix two vertices j and k of G. Set $B = A^{\ell}$. By the induction hypothesis, for any $v \in V(G)$ there are B^{ℓ}_{jv} walks of length ℓ from j to v and by the definition of A there are A_{vk} walks of length 1 (that is, edges) from v to k. As every walk of length $\ell + 1$ from j to k can be obtained by concatenating, for some $v \in [n]$, a walk of length ℓ from j to v and a walk of length 1 (an edge) from v to k, the number of walks of length $\ell + 1$ from j to k is

$$\sum_{v=1}^{n} B_{jv} A_{vk} = (BA)_{jk} = (A^{\ell} A)_{jk} = A^{\ell+1}_{jk},$$

which concludes the proof. □

In addition, we can use the adjacency matrix to check whether the corresponding graph is connected. The following proposition shows how to do this.

Proposition 3. [1] *Let G be a simple graph on $[n]$, and let A be the adjacency matrix of G. Then G is connected if and only if all the entries of $(I_n + A)^{n-1}$ are positive.*

Proof. We know that G is connected if and only if any two distinct vertices j and k of G are connected by a path (of length at most $n-1$), which happens if and only if $A^{\ell}_{j,k} > 0$ for some $\ell \in [n-1]$. Therefore the statement of the proposition follows from the following (Newton-Binomial) identity:

$$(I_n + A)^{n-1} = \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} A^{\ell}.$$

□

The Matrix-Tree Theorem. Our next goal is to introduce another important matrix related to a given directed graph G, the incidence matrix, and use it to provide a formula for the number of spanning trees of G. This formula, in turns, will allow us to prove the Matrix-Tree Theorem, which expresses the number of spanning trees of an (undirected graph) as a determinant of certain matrix.

Let G be a directed graph. In this lecture, we say that the *underlying graph* of G is the graph we obtain from G by ignoring the orientation of the edges. A *spanning tree* of a directed graph G is a subgraph T such that the underlying graph of T is a spanning tree of the underlying graph of G.

Definition 4. Let G be a directed graph with $V(G) = \{v_1, \dots, v_m\}$ and $E(G) = \{e_1, \dots, e_n\}$. The *incidence matrix* of G is the $m \times n$ matrix A with $A_{ij} = 1$ (resp., $A_{ij} = -1$) if the edge e_j starts (resp., ends) at v_i and with $A_{ij} = 0$ if e_j is not connected to v_i.

Theorem 5. *Let G be a connected directed graph (without loops), and let A be the incidence matrix of G. If A_0 is the matrix obtained from A by removing the last row, then* $\det(A_0 A_0^T)$ *is the number of spanning trees of G.*

Proof. Since G is connected, $m - 1 \le n$. Let B be an $(m-1) \times (m-1)$ submatrix of A_0, and let G' be the subgraph of G with $V(G') = V(G)$ and whose edges correspond to the columns of B. We claim that G' is a spanning tree of G if and only if $|\det B| = 1$ and that, otherwise, $\det B = 0$. We proceed by induction on $|V(G)| \in \mathbb{N}_{\ge 2}$.

If $|V(G)| = 2$, then G' consists of exactly one edge and so it is a spanning tree of G, while $B \in \{(1), (-1)\}$ and so $|\det B| = 1$. Now assume that our claim holds for any directed graph with $m-1$ vertices (with $m \ge 3$), and suppose that $V(G) = \{v_1, \dots, v_m\}$ and $E(G) = \{e_1, \dots, e_n\}$. Let B and G' be as we have described before. We split the rest of the proof into two cases.

[1] Because of time-constrains, this proposition was not covered in class, but I have included here to provide further applications of the adjacency matrix representation of a graph.

Case 1: There exists $i \in [n-1]$ such that $\operatorname{indeg}_{G'} v_i + \operatorname{outdeg}_{G'} v_i = 1$. This implies that there is exactly one nonzero entry (either 1 or -1) in the i-th row of B. Let e_j be the edge connected to v_i in G'. If we compute $\det B$ expanding along the i-th row of B, then we obtain that $|\det B| = |\det B'|$, where B' is the submatrix of B obtained by eliminating the i-th row and the column corresponding to e_j. It follows by the induction hypothesis that $\det B'$ is either 1 or -1 if and only if $G' \setminus \{v_i\}$ is a spanning tree of $G \setminus \{v\}$. As $\operatorname{indeg}_{G'} v_i + \operatorname{outdeg}_{G'} v_i = 1$, this happens if and only if G' is a spanning tree of G. Similarly, the induction hypothesis allows us to deduce that $\det B = 0$ if and only if G' is not a spanning tree of G.

Case 2: $\operatorname{indeg}_{G'} v_i + \operatorname{outdeg}_{G'} v_i \neq 1$ for any $i \in [m-1]$. Since $|V(G)| = m$ and $|E(G')| = m-1$, the graph G' must have a vertex v_j such that $\operatorname{indeg}_{G'} v_j + \operatorname{outdeg}_{G'} v_j = 0$. In particular, G' is not a spanning tree of G (as it is not even a tree itself). If $j \leq m-1$, then B has a row full of zeros and so $\det B = 0$. Otherwise, $j = n$ and, therefore, every column of B has exactly one entry -1, one entry 1, and the rest of the entries are zeros. This last statement implies that the addition of all the row vectors of B is the zero vector. In particular, the rows of B are linearly dependent, which implies that $\det B = 0$.

Now we can use our already-proved claim in tandem with the Binet-Cauchy Formula to complete the proof. It follows from the Binet-Cauchy Formula that

$$\det(A_0 A_0^T) = \sum (\det B)^2, \tag{0.1}$$

where the sum runs over all $(m-1) \times (m-1)$ submatrices B of A_0. By our claim, the spanning trees of G correspond to submatrices B with $\det B \in \{\pm 1\}$, and the determinant of the rest of the $(m-1) \times (m-1)$ submatrices of A_0 is zero. Hence the identity (0.1) implies that G contains exactly $\det(A_0 A_0^T)$ spanning trees. □

We are in a position to prove that Matrix-Tree Theorem.

Theorem 6. *Let U be a simple (undirected graph) with $V(U) = \{v_1, \dots, v_m\}$. Let L be the $(m-1) \times (m-1)$ matrix defined by*

$$L_{ij} = \begin{cases} \deg v_i & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } v_i v_j \in E(U) \\ 0 & \text{otherwise.} \end{cases}$$

for all $i, j \in [m-1]$. Then the number of spanning trees of U is $\det L$.

Proof. Let G be the directed graph that we obtain from U by replacing each edge of U by two arrows, one in each direction. Let A be the incidence matrix of G, and let A_0 be the matrix we obtain from A by removing the last row. We claim that $A_0 A_0^T = 2L$. Set $M = A_0 A_0^T$ and observe that

$$M_{ij} = A_{i1} A_{j1} + A_{i2} A_{j2} + \cdots + A_{in} A_{jn},$$

where $n = |E(G)|$. When $i = j$, the summand $A_{ik}A_{ik}$ contributes with 1 to M_{ii} if and only if the edge of G determined by the k-th column of A is incident to v_i. Therefore $M_{ii} = \text{indeg}_G v_i + \text{outdeg}_G v_i = 2\deg_U v_i = 2L_{ii}$. On the other hand, if $i \neq j$, then $A_{ik}A_{jk}$ contributes with -1 to M_{ij} if and only if the edge corresponding to the k-th column of A connects v_i and v_j, which happens exactly for two indices k. Thus, $M_{ij} = -2 = 2L_{ij}$ if $v_iv_j \in E(U)$. Otherwise, there are no edges connecting v_i and v_j in U (or in G) and so $M_{ij} = 0 = 2L_{ij}$. Hence $A_0A_0^T = M = 2L$, as claimed. Therefore

$$2^{m-1}\det L = \det(2L) = \det(A_0A_0^T),$$

which is, by virtue of Theorem 4, the number of spanning trees of G. Since each spanning tree of U gives rise to exactly 2^{m-1} spanning trees of G, we conclude that $\det L$ is the number of spanning trees of U. □

Practice Exercises

Exercise 1. *Let G be a directed graph such that $|E(G)| = |V(G)| - 1$. If $\text{indeg}\, v + \text{outdeg}\, v \neq 1$ for all $v \in V(G)$, argue that G has an isolated vertex.*

Exercise 2. *Use the Matrix-Tree Theorem to rediscover that the number of spanning trees of K_n is n^{n-2}.*

References

[1] M. Bóna: *A Walk Through Combinatorics: An Introduction to Enumeration and Graph Theory* (Fourth Edition), World Scientific, New Jersey, 2017.

Department of Mathematics, MIT, Cambridge, MA 02139
Email address: fgotti@mit.edu