Eigenvalues of Tree Graph

Yixiao Feng
Zhiyuan College, Shanghai Jiao Tong University, Shanghai, P.R.China

January 11, 2024

1 Introduction

Tree Graph of a directed graph G emerges from proof of Markov chain tree theorem. The construction of the tree graph $\mathcal{T}G$ is recalled in section 2. Properties of $\mathcal{T}G$ have drawn many interests. In particular, eigenvalues of $\mathcal{T}G$ is a focal point. In this essay, we rearrange the proof of eigenvalues of $\mathcal{T}G$ and make conjectures of the Jordan canonical form of $\mathcal{T}K_n$, in which K_n is the directed complete graph on n vertices.

The Laplacian of G is defined as $Q = D - M$, where D is the diagonal matrix of outdegrees of vertices and M is the adjacency matrix of G.

An l-walk, or l-path, in a directed graph $G = (V, E)$ means a directed l-walk, that is a sequence $(u_0, e_1, u_1, \cdots, e_l, u_l)$ of vertices and edges of G such that for each $1 \leqslant i \leqslant l$, edge e_i has initial vertex u_{i-1} and terminal vertex u_i. The walk is closed if $u_0 = u_l$.

If S is a nonempty subset of V, we denote by G_S the induced subgraph of G on the vertex set S, that is the directed graph obtained from G by only keeping vertices in S and edges with both endpoints in S.

2 tree graph

In this section we recall some basic concepts and introduce a few previous results.

2.1 Spanning trees and forests

Let $G = (E, V)$ be a finite directed graph. For each edge we denote $s(e)$ its source and $t(e)$ its target. A (rooted) spanning tree of G is a subgraph of G containing all vertices with no cycle. There is one vertex r, called the root, having outdegree 0 and the other vertices have outdegree 1. By the nature of tree, the spanning tree of G has $|V| - 1$ edges.

Generally, if $W \subset V$ is a nonempty subset, a forest of G rooted in W is defined as the following subgraph: it contains all vertices, with no cycle, and the vertices in W have outdegree 0 while the other vertices have outdegree 1.

2.2 Tree graph $\mathcal{T}G$

The tree graph of G, denoted $\mathcal{T}G$, is the directed graph whose vertices are the spanning trees of G and whose edges are obtained by the following construction:

Let $\mathbf{a}$ be a spanning tree of G with root r. For an edge $e \in E$ with $s(e) = r$, let $\mathbf{b}$ be the subgraph of G obtained by adding edge e to $\mathbf{a}$ then deleting the edge coming out of $t(e)$ in $\mathbf{a}$.

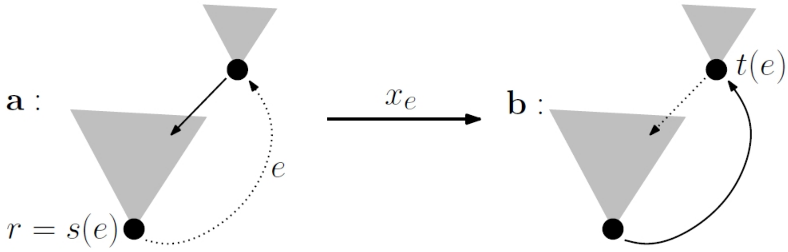

Figure 1: definition of $\mathcal{T}G$

It's easy to check that **b** is a spanning tree of G, with root $t(e)$.

In the following we will denote $\mathcal{T}V$ the set of vertices of $\mathcal{T}G$, i.e., the set of spanning trees of G.

There is a natural map p from $\mathcal{T}G$ to G. p maps each vertex of $\mathcal{T}G$ to its root, and maps each edge of $\mathcal{T}G$ to the edge e of G for its construction.

Now we give here some elementary properties of the tree graph. See previous papers [2] for proofs.

- $\mathcal{T}G$ is simple and has no loop.
- The graph $\mathcal{T}G$ is strongly connected if G is.
- The graph $\mathcal{T}G$ is a covering graph [1] of G.
- The set of edges of $\mathcal{T}G$ can be partitioned into edge-disjoint simple cycles, which project onto simple cycles of G. If C is a simple cycle of G, with vertex set W, then the number of simple cycles of $\mathcal{T}G$ lying above C is equal to the number of forests rooted in W.
- The graph $\mathcal{T}G$ is Eulerian: The number of outgoing or incoming edges of a vertex **a** are both equal to the number of outgoing edges of the root of **a** in G.

If $W \subset V$, The matrix-tree theorem gives the generating function of spanning forests of G rooted in W . Moreover, it's easy to know the total number of those spanning forests: Denote Q^W as the matrix obtained from the Laplacian matrix Q of G by deleting rows and columns indexed by elements of W, then the number of spanning forests of G rooted in W is given by $\det(Q^W)$.

In particular, let $W = \{r\}$, then the number of spanning trees of G rooted in r is given by $\det(Q^r)$. For convenience in the following we use the notation $Q_W = Q^{V \setminus W}$ to denote the matrix extracted from the Laplacian matrix Q of G by keeping only rows and columns indexed by elements of W.

It is important to notice that the matrix-tree theorem is valid for both directed and undirected graphs.

By the matrix-tree theorem, we can easily derive all spanning trees of G, then construct the tree graph $\mathcal{T}G$.

[1] https://en.wikipedia.org/wiki/Covering_graph

3 Eigenvalues of Tree Graph

Though astonishing, eigenvalues of the adjacency matrix of the tree graph $\mathcal{T}G$ can be computed explicitly. This can be expected from the following fact: the (i, j) entry of the matrix M, where M is the adjacency matrix of G, equals the number of l-walks in G which start at v_i and end at v_j. Thus, the number of closed l-walks in $\mathcal{T}G$, which we will denote by $w(\mathcal{T}G, l)$, equals the trace of M^l and hence the sum of the lth powers of the eigenvalues of M.

C. Athanasiadis proved the following theorem about eigenvalues of $\mathcal{M}$[1], i.e., the adjacency matrix of $\mathcal{T}G$.[2]

Theorem 3.1. *The nonzero eigenvalues of $\mathcal{M}$ are eigenvalues of $M_X; X \subseteq V$ where M is the adjacency matrix of G. For an eigenvalue γ, let $m_X(\gamma)$ denotes its multiplicity in M_X, then its multiplicity in $\mathcal{M}$ is*

$$\sum_{X\subseteq V} m_X(\gamma)\det(Q_{V\setminus X} - I)$$

Here we make the convention that $\det(Q_\emptyset - I) = 1$.

The original proof from Athanasiadis is a little confusing, so we make it more clear here. First we need to calculate $w(\mathcal{T}G, l)$.

Theorem 3.2. *For a directed graph G on the vertex set V we have*

$$w(\mathcal{T}G, l) = \sum_{S\subseteq V} w(G_S, l)\det(Q_{V\setminus S} - I).$$

Proof. A closed l-walk in $\mathcal{T}G$ is determined by a spanning tree T_0 on G with root r, and a closed l-walk $W = (u_0, e_1, u_1, \cdots, e_l, u_l)$ in G with $u_0 = u_l = r$. Let $a = (u_0, u_1, \cdots, u_l)$. The sequence a determines the roots of the trees $T_0, T_1, \cdots, T_l = T_0$ to be visited during the walk in $\mathcal{T}G$.

Fix a closed l-walk W in G together with the sequence a. Following Athanasiadis's argument (Theorem 2.2) the number of trees T_0 which will yield a closed l-walk in $\mathcal{T}G$ is the number $\tau(G, U)$ of spanning forests on G with root set $U = \{u_0, u_1, \cdots, u_l\}$. Finally, the number of closed l-walks in $\mathcal{T}G$ is

$$w(\mathcal{T}G, l) = \sum_{U\subseteq V} g(G_U, l)\tau(G, U), \tag{1}$$

where $g(G, l)$ stands for the number of closed l-walks in a graph G visiting all of its vertices.

We may explain (1) more explicitly. Notice that a closed l-walk in G can have repeated vertices, so if G_U has a closed l-walk visiting all of its vertices, then $|U|$ can be divided by l. Also, G_U can have more than one closed l-walk. The inclusion-exclusion principle gives

$$g(G_U, l) = \sum_{S\subseteq U} (-1)^{|U-S|} w(G_S, l). \tag{2}$$

Hence, after using (2) to compute $g(G_U, l)$ and changing the order of summation, (1) becomes

$$w(\mathcal{T}G, l) = \sum_{S\subseteq V} w(G_S, l) \sum_{S\subseteq U\subseteq V} (-1)^{|U-S|}\tau(G, U). \tag{3}$$

[2]In Biane and Chapuy's paper[2], they wrote the theorem wrongly (Proposition 3.1)

Recall that by the matrix-tree theorem, $\tau(G,U) = \det(Q_{V\setminus U})$, thus

$$w(\mathcal{T}G,l) = \sum_{S\subseteq V} w(G_S,l) \sum_{S\subseteq U\subseteq V} (-1)^{|U-S|}\det(Q_{V\setminus U}) = \sum_{S\subseteq V} w(G_S,l)\det(Q_{V\setminus S}-I).$$

□

Now that we have the number of closed l-walks in $\mathcal{T}G$, we can compute the trace of $\mathcal{M}^l$ and M^l. We need the following lemma to complete the proof.

Lemma 3.3. *Suppose that for some nonzero complex numbers a_i, b_j, where $1 \leqslant i \leqslant r$ and $1 \leqslant j \leqslant s$ we have*

$$\sum_{i=1}^{r} a_i^l = \sum_{j=1}^{s} b_j^l \tag{4}$$

for all positive integers l. Then $r = s$ and the a_i are a permutation of the b_j.

The proof can be found in Athanasiadis's paper[1].

From theorem 3.2 and Lemma 3.3 we immediately get theorem 3.1.

4 Example of Complete Graph

The complete directed graph K_n is the graph on the vertex set $V = [n] = \{1, 2, \cdots, n\}$ with exactly one directed edge from i to j for each $i \neq j$.

4.1 Eigenvalues of $\mathcal{T}K_n$

Proposition 4.1. *With $G = K_n$, $n \geqslant 2$, $\mathcal{M}$ has eigenvalues $-1, 1, \cdots, n-1$. The multiplicity of i is*

$$m_{\mathcal{M}}(i) = \begin{cases} i\binom{n}{i+1}(n-1)^{n-i-2}, & 1 \leqslant i \leqslant n-1 \\ n^{n-1} - (n-1)^{n-1}, & i = -1. \end{cases}$$

Proof. First we calculate the eigenvalues of K_n. Note that $(1, \cdots, 1)^\mathsf{T}$ is an eigenvector of M with eigenvalue $n-1$. Besides, J_n, the $n \times n$ matrix with all entries equal to 1, has a kernel of dimension $n-1$. Thus the eigenvalues of K_n are $n-1$ with multiplicity 1 and -1 with multiplicity $n-1$. Hence, by theorem 3.1, the eigenvalues of $\mathcal{M}$ are $-1, 1, \cdots, n-1$. Moreover the eigenvalue $m-1$, $2 \leqslant m \leqslant n$ has multiplicity

$$\binom{n}{m}\det(Q_{V\setminus[m]} - I) = \binom{n}{m}(m-1)(n-1)^{n-m-1}.$$

which is easy to obtain by consider eigenvalues of $Q_{V\setminus[m]} - I$.

-1 has multiplicity

$$\begin{aligned}&\sum_{m=2}^{n}\binom{n}{m}(m-1)\det(Q_{V\setminus[m]} - I)\\ =& \sum_{m=1}^{n}(m-1)^2\binom{n}{m}(n-1)^{n-m-1} = n^{n-1} - (n-1)^{n-1}.\end{aligned}$$

The multiplicities we have so far add up to n^{n-1}, so 0 is not an eigenvalue of $\mathcal{M}$ and the proposition follows. □

4.2 Properties of $\mathcal{T}K_n$

Properties of K_n have always been interesting. We have a lot more to say about $\mathcal{T}K_n$. First, since $\mathcal{T}K_n$ is a covering graph of K_n, we have the following proposition.

Proposition 4.2. *$\mathcal{T}K_n$ is a regular graph where each vertex has outdegree and indegree $n-1$.*

However, we remind readers that while every vertex in $\mathcal{T}K_n$ has outdegree $n-1$ is obvious from the construction of edges in $\mathcal{T}K_n$, the constant indegree of each vertex is not trivial at all. Given a spanning tree T_0 of K_n, writing down all $T_1, \cdots, T_{n-1}$ that $T_i \to T_0$ in $\mathcal{T}K_n$ is not that easy. The difficulty lies in irreversibility of the construction of edges in $\mathcal{T}K_n$. Actually, we can get the following proposition.

Proposition 4.3. *Let T_1 is a spanning tree of K_n with root r_1, T_2 is a spanning tree of K_n with root r_2. There is an edge $T_1 \to T_2$ in $\mathcal{T}K_n$. Then there is an edge $T_2 \to T_1$ in $\mathcal{T}K_n$ if and only if $r_2 \to r_1$ in T_1. Also, by definition, we have $r_1 \to r_2$ in T_2. In the remaining we will say that there is a bi-edge between T_1 and T_2.*

Proof. Suppose $r_2 \to r_1$ in T_1. Then in T_2, we have $r_1 \to r_2$, and r_1 has no other outgoing edges. Thus, if we add $r_2 \to r_1$ in T_2 and delete $r_1 \to r_2$ in T_2, we get exactly T_1. Hence, $T_2 \to T_1$ in $\mathcal{T}K_n$. Conversely, suppose $T_2 \to T_1$ in $\mathcal{T}K_n$. By definition, $r_2 \to r_1$ in T_1. $\square$

From the above proposition, we know that given T_0, the number of T such that T_0 and T has a bi-edge equals the indegree of r_0 in T_0, where r_0 is the root of T_0.

For any m, we now consider number of rooted trees with m bi-edges. There is a helpful simple decomposition of rooted trees: Let T be a rooted tree. If we remove all edges incident with the root, we get the root together with a set of trees. Number of trees equals indegree of the root.

Proposition 4.4. *With $G = K_n$, number of trees with m bi-edges is*

$$nm\binom{n-1}{m}(n-1)^{n-2-m}$$

Proof. First, there are n choices of root, denoted as r_0. Second, by the above discussion, there are $\binom{n-1}{m}$ choices of roots after the decomposition, denoted as $r_1, \cdots, r_m$. Finally, excluding r_0, number of forests rooted in $r_1, \cdots, r_m$ equals to $\det(n-1)I_{n-1-m} - J_{n-1-m}$, which is $m(n-1)^{n-2-m}$. By multiplicity principle, we get the proposition. $\square$

Proposition 4.5. *With $G = K_n$, if $T_0 \to T_1$, $T_1 \to T_2$ in $\mathcal{T}G$, then $T_0 \nrightarrow T_2$.*

Proof. Denote roots of T_0, T_1, T_2 as r_0, r_1, r_2 and they are distinct. From Figure 2 we see that if $T_0 \to T_1$ and $T_1 \to T_2$, then $r_0 \to r_1 \to r_2$. Thus, if we add an edge from r_0 to r_2, the spanning tree we get is different from T_2. $\square$

4.3 Three Conjectures

Properties of $\mathcal{T}K_n$ are much more than we have discussed. In particular, while eigenvalues of $\mathcal{T}K_n$ have been computed, the Jordan canonical form of $\mathcal{T}K_n$ has not been obtained. By calculating $n = 3, 4, 5, 6$ cases, we make three conjectures closely related to Jordan canonical form of $\mathcal{T}K_n$ here.

Conjecture 4.6. *For eigenvalues $\lambda = 1, 2, \cdots, n-1$ of $\mathcal{T}K_n$, the Jordan blocks are all of size 1.*

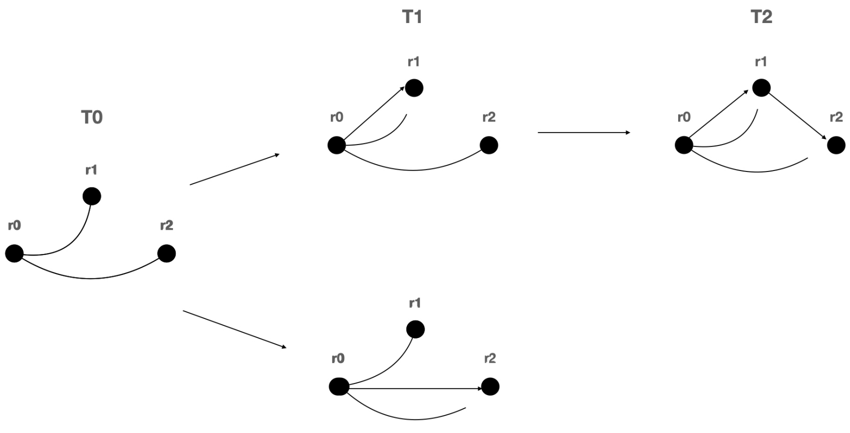

Figure 2: path in $\mathcal{T}K_n$

Conjecture 4.7. *For eigenvalue* $\lambda = -1$*, the Jordan blocks are of size* $1, 2, \cdots, n-1$*. Moreover, there are exactly* $n-1$ *Jordan blocks of size* $n-1$*.*

Conjecture 4.8. *For eigenvalue* $\lambda = -1$, $\operatorname{rank}((\mathcal{M}+I)^{n-1}) = \operatorname{rank}((\mathcal{M}+I)^{n}) = \operatorname{rank}((\mathcal{M}+I)^{n+1}) = \cdots = (n-1)^{n-1}$

Table 1: Jordan blocks of $\mathcal{T}K_n, n = 3, 4, 5, 6$

	$\lambda = -1$	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$	$\lambda = 5$
$\mathcal{T}K_3$	1×1 2×2	3×1	1×1	$\backslash$	$\backslash$	$\backslash$
$\mathcal{T}K_4$	4×1 12×2 3×3	18×1	8×1	1×1	$\backslash$	$\backslash$
$\mathcal{T}K_5$	30×1 100×2 41×3 4×4	160×1	80×1	15×1	1×1	$\backslash$
$\mathcal{T}K_6$	336×1 1096×2 578×3 91×4 5×5	1875×1	1000×1	225×1	24×1	1×1

We hope to prove or falsify these three conjectures later.

4.4 classification of vertices of $\mathcal{T}K_n$

In Biane and Chapuy's paper[2], they obtain the number of spanning trees of $\mathcal{T}K_n$, namely the number of vertices of graph $\mathcal{T}^2K_n$:

$$n^{n-2}\prod_{k=1}^{n-1}((n-k)n^{k-1})^{(k-1)(n-1)^{n-k-1}\binom{n}{k}}$$

If we want to calculate eigenvalues of $\mathcal{T}^2K_n$, we need first classify vertices of $\mathcal{T}K_n$. By relabeling this reduces to the problem of classifying unlabeled rooted trees with n vertices. Namely, we only consider number of equivalence classes under graph isomorphism.

There is no simple formula for counting unlabeled trees. [3] Denote $r(x)$ be the generating function for number of unlabeled rooted trees, i.e., $r(x) = \sum_{n=1}^{\infty} r_n x^n$ where r_n is the number of unlabeled rooted trees with n vertices. Then we have the following theorem.

Theorem 4.9. *The generating function $r(x)$ for unlabeled rooted trees satisfies*

$$r(x) = xh[r(x)] = x\exp\left(\sum_{k=1}^{\infty}\frac{r(x^k)}{k}\right). \tag{5}$$

Please refer to Gessel's paper[3] for definition for $h(A(x))$ and proof of this theorem. We find that

$$\begin{aligned} r(x) = x &+ x^2 + 2x^3 + 4x^4 + 9x^5 + 20x^6 + 48x^7 + 115x^8 + 286x^9 \\ &+ 719x^{10} + 1842x^{11} + 4766x^{12} + 12486x^{13} + 32973x^{14} \\ &+ 87811x^{15} + 235381x^{16} + 634847x^{17} + 1721159x^{18} \\ &+ 4688676x^{19} + 1282622x^{20} + 35221832x^{21} + \cdots \end{aligned}$$

These numbers of sequence A000081 in the Online Encyclopedia of Integer Sequences (OEIS)[4].

References

[1] C. A. Athanasiadis, Spectral of Some Interesting Combinatorial Matrices Related to Oriented Spanning Trees on a Directed Graph, J. Algebr. Comb 5 (1996), no.1, p. 5-11.

[2] P. Biane and G. Chapuy, Laplacian matrices and spanning trees of tree graphs, Annales de la faculté des sciences de Toulouse XXVI (2017), p. 235-261.

[3] I. M. Gessel, Good Will Hunting's Problem: Counting Homeomorphically Irreducible Trees, arXiv: 2305.03157v1 (2023).

[4] OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences, http://oeis.org (2021).

[3] There is a complicated explicit formula, as a sum over partitions, for counting unlabeled rooted trees