Spectra of Some Interesting Combinatorial Matrices Related to Oriented Spanning Trees on a Directed Graph

CHRISTOS A. ATHANASIADIS   cathan@math.mit.edu
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract. The Laplacian of a directed graph G is the matrix $L(G) = O(G) - A(G)$, where $A(G)$ is the adjacency matrix of G and $O(G)$ the diagonal matrix of vertex outdegrees. The eigenvalues of G are the eigenvalues of $A(G)$. Given a directed graph G we construct a derived directed graph $D(G)$ whose vertices are the oriented spanning trees of G. Using a counting argument, we describe the eigenvalues of $D(G)$ and their multiplicities in terms of the eigenvalues of the induced subgraphs and the Laplacian matrix of G. Finally we compute the eigenvalues of $D(G)$ for some specific directed graphs G. A recent conjecture of Propp for $D(H_n)$ follows, where H_n stands for the complete directed graph on n vertices without loops.

1. Introduction

Consider a directed graph $G = (V, E)$ on a set of n vertices V, with multiple edges and loops allowed. An *oriented rooted spanning tree* on G, or simply an *oriented spanning tree* on G, is a subgraph T of G containing all n vertices of G and having a distinguished vertex r, called the root, such that for every $v \in V$ there is a unique (directed) path in T with initial vertex v and terminal vertex r. Thus T is a rooted tree in the usual sense, if we disregard the orientation of its edges. Propp considered a directed graph $D(G)$ whose vertices are the oriented spanning trees of G. The edges of the new directed graph $D(G)$ are constructed as follows: Let T be a vertex of $D(G)$ with root r and $v \in V$. Given an edge $e \in E$ directed from r to v, let $T(e)$ be the tree obtained from T by adding the edge $e = rv$ and deleting the edge of T with initial vertex v. Note that the root of $T(e)$ is v. Then add a directed edge in $D(G)$ from T to $T(e)$.

The idea of the construction of $D(G)$ appeared for the first time implicitly in the proof of the Markov chain tree theorem by Anantharam and Tsoucas [1], where the authors needed to lift a random walk in G to a random walk in the set of arborescences of G, which coincides with the set of oriented spanning trees when G is strongly connected. On the other hand, Propp's motivation for defining $D(G)$ came from problems related to domino tilings of regions. The "re-rooting move", described above to define the edges of $D(G)$, is analogous to a certain operation on domino tilings, called an "elementary move" in [4]. In fact, under an appropriate coding, the elementary moves can be viewed as a special case of a type of move very similar to the re-rooting move. An even more general operation is described in [7]. Proposition 2.5, stated in Section 2, is the analogue of the fact that any domino tiling of a simply connected region can be obtained from any other tiling of the same region by a

sequence of elementary moves. Thus $D(G)$ encodes the ways one can reach any oriented spanning tree on G from any other, assuming that G is strongly connected, by performing re-rooting moves.

The main problem we pose here and answer in the following section is to describe the eigenvalues of the adjacency matrix of $D(G)$ in terms of information contained in our original graph G. The motivation for posing this question comes from a conjecture of Propp [6] asserting that if G is the complete directed graph H_n on n vertices without loops, then the eigenvalues of the Laplacian matrix of $D(G)$ are all integers. In fact Propp conjectured, based on computational evidence, that these eigenvalues are integers ranging from 0 to n and that the multiplicities of 0, 1, $n - 1$ and n are 1, $n^2 - 2n$, 0 and $n^{n-1} - (n - 1)^{n-1}$ respectively. In the third section we will prove Propp's conjecture and we will find the multiplicities of other eigenvalues.

In the remaining of this section we introduce some basic background and notation.

By an l-walk, or l-path, in a directed graph $G = (V, E)$ we will always mean a directed l-walk, that is an alternating sequence $(u_0, e_1, u_1, \ldots, e_l, u_l)$ of vertices and edges of G such that for each $1 \le i \le l$, edge e_i has initial vertex u_{i-1} and terminal vertex u_i. The walk is said to be closed if $u_0 = u_l$. If S is a nonempty subset of the vertex set 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 deleting the vertices not in S and all edges incident to them. A subgraph of G is obtained from G_S, for some S, by deleting some of its edges. Finally, an oriented forest on G with root set S is a collection of vertex-disjoint oriented spanning trees on subgraphs of G which use all the vertices of G and whose roots are the elements of S.

Fix an ordering of the vertex set $V = (v_1, v_2, \ldots, v_n)$ of our directed graph G. The adjacency matrix of G, denoted by $A(G)$, is the $n \times n$ matrix whose (i, j) entry is the number of edges of G directed from v_i to v_j. The eigenvalues of $A(G)$ are called the eigenvalues of G and do not depend on the ordering chosen.

The Laplacian of G is defined to be the matrix $L(G) = O(G) - A(G)$ where $O(G)$ is the diagonal matrix whose ith entry on the diagonal equals the outdegree of v_i, that is the number of directed edges emanating from v_i. Both the adjacency and Laplacian matrix we have introduced for a directed graph are natural analogues of the corresponding matrices for undirected graphs. For an exposition of many of the known results about Laplacians of undirected graphs see [5]. Note that the Laplacian matrix is independent of the number of loops of G on each vertex. The main result we will use about Laplacians is the following extended version of the matrix tree theorem. A proof and a generalization can be found in [2].

Theorem 1.1 *For $S \subseteq G$ denote by $L(G)|_S$ the submatrix of $L(G)$ obtained by deleting the rows and columns corresponding to the vertices in S. Then $\det(L(G)|_S)$ is the number of oriented forests on G with root set S.*

Finally we mention that all our definitions and theorems can easily be extended in the general setting of *weighted* directed graphs with weights from the field of complex numbers, that is directed graphs with a nonzero complex number assigned to each edge. For example in Theorem 1.1 we would have to replace the number of oriented forests with the sum of their weights, where the weight of a forest is the product of the weights of its edges, and adjust the definitions of $L(G)$ and $A(G)$ in the obvious way. We prefer here to keep the simplicity of the classical case.

2. The main theorem

It is well known that the (i, j) entry of the matrix $A(G)^l$ equals the number of l-walks in G which start at vertex v_i and end at vertex v_j. Thus the number of closed l-walks in G, which we will denote by $w(G, l)$, equals the trace of $A(G)^l$ and hence the sum of the lth powers of the eigenvalues of G. For the weighted case, of course, we should count a walk according to its weight, that is the product of the weights of its edges.

Thus a common technique to count walks in G is to compute its eigenvalues. Our approach here will be the opposite. We will count the number of closed walks in G combinatorially and then read off its eigenvalues from the answer. This will be possible thanks to the following elementary and well known fact.

Lemma 2.1 *Suppose that for some nonzero complex numbers a_i, b_j, where $1 \leq i \leq r$ and $1 \leq j \leq s$ we have*

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

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

Proof: Multiplying (1) by x^l and summing over l gives

$$\sum_{i=1}^{r} \frac{a_i x}{1 - a_i x} = \sum_{j=1}^{s} \frac{b_j x}{1 - b_j x}$$

for small x. Since clearing denominators we obtain a polynomial equation, this has to hold for all x. Thus if γ is any nonzero complex number, we can multiply the equation above by $1 - \gamma x$ and set $x = 1/\gamma$ to conclude that γ appears among the a_i as many times as among the b_j. $\square$

We are ready to state our central result:

Theorem 2.2 *For a directed graph G on the vertex set V we have*

$$w(D(G), l) = \sum_{S \subseteq V} w(G_S, l)\det(L(G)|_S - I). \tag{2}$$

Proof: An l-walk in $D(G)$ is determined by an oriented spanning tree T_0 on G with root r, which will be the initial vertex of the walk, and an l-walk $W = (u_0, e_1, u_1, \ldots, e_l, u_l)$ in G with initial vertex $u_0 = r$. Let $a = (u_0, u_1, \ldots, u_l)$. The sequence a determines the roots of the trees $T_0, T_1, \ldots, T_l$ to be visited during the walk in $D(G)$ and the edges of W determine the trees themselves. For the walk to be closed, that is $T_l = T_0$, we should also have $u_l = r$.

Now let's fix a closed l-walk W in G together with the sequence a, as above, and let's count the number of oriented spanning trees T_0 on G with root $r = u_0$ which yield a closed walk in $D(G)$ when applying W. At each step of the walk from T_{i-1} to T_i, we add a directed edge e_i from u_{i-1} to u_i and delete the edge emanating from u_i. Therefore, at the end of our walk T_l, a vertex of G appearing for the last time in a as u_{i-1}, $2 \leq i \leq l$, will be directed

to u_i with e_i and the vertices of G not appearing in a will be directed as in T_0. It follows that the number of trees T_0 which will yield a closed l-walk in $D(G)$ when we apply W is the number $\tau(G, U)$ of oriented spanning forests on G with root set $U = \{u_0, u_1, \ldots, u_l\}$. Finally, since W is a closed l-walk in the graph G_U visiting all of its vertices, we have

$$w(D(G), l) = \sum_{U \subseteq V} g(G_U, l)\tau(G, U),\tag{3}$$

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

$$g(G_U, l) = \sum_{S \subseteq U}(-1)^{|U-S|}w(G_S, l).\tag{4}$$

Hence, after using (4) to compute $g(G_U, l)$ and changing the order of sumation, (3) becomes

$$w(D(G), l) = \sum_{S \subseteq V} w(G_S, l)\sum_{S \subseteq U \subseteq V}(-1)^{|U-S|}\tau(G, U).\tag{5}$$

Using Theorem 1.1 to compute $\tau(G, U)$ together with some elementary linear algebra yields (2). □

From Theorem 2.2 and Lemma 2.1 we immediately conclude the following corollary.

Corollary 2.3 *The nonzero eigenvalues of $D(G)$ are included in the nonzero eigenvalues of the induced subgraphs of G. Moreover if $\gamma \neq 0$, then the multiplicity of γ as an eigenvalue of $D(G)$ is*

$$\sum_{S \subseteq V} m_{G_S}(\gamma)\det(L(G)|_S - I),$$

where $m_{G_S}(\gamma)$ stands for the multiplicity of γ as an eigenvalue of G_S.

Corollary 2.4 *For all complex numbers $\gamma \neq 0$ we have*

$$\sum_{S \subseteq V} m_{G_S}(\gamma)\det(L(G)|_S - I) \geq 0.$$

We digress at this point to show directly that 0 is a simple eigenvalue of the Laplacian of $D(H_n)$, meaning that its multiplicity is 1. Recall that a directed graph is said to be *strongly connected* if for any distinct vertices u and v there exists a (directed) walk in the graph from u to v. Note that the multiplicity of 0 as an eigenvalue of the Laplacian of any digraph G which is the disjoint union of strongly connected digraphs, is the number of connected components of G. Indeed, a basis of the corresponding eigenspace is the set of vectors with entry 1 on the vertices of G belonging to a given connected component and 0 on the rest. Hence, in particular, if $D(G)$ is strongly connected then $L(D(G))$ has 0 as a simple eigenvalue. In general, $D(G)$ might be disconnected even though G is connected. The same is not true, however, with strong connectedness. The following proposition, proved independently for the first time by Propp, shows that $D(H_n)$ is indeed strongly connected.

Proposition 2.5 *If $G = (V, E)$ is strongly connected then so is $D(G)$. In particular, the Laplacian of $D(G)$ has 0 as a simple eigenvalue.*

Proof: Given two oriented spanning trees T_0 and T_1 on G with roots r_0 and r_1, we want to find a walk in $D(G)$ from T_0 to T_1. Such a walk is determined by a walk in G from r_0 to r_1. Thus we start at r_0 and follow the unique path in T_1 from r_0 to r_1. Then we pick the furthest vertex v in G away from r_1 and follow the shortest walk in G from r_1 to v. This can be done by strong connectedness of G. Now we follow the unique path in T_1 from v to r_1 and continue in the same way with the second furthest vertex in G away from r_1, until only r_1 remains. At this point we stop.

This walk has the property that the last time a vertex u other than r_1 is visited by our walk, it is followed by its successor in T_1, and hence it induces a walk in $D(G)$ from T_0 to T_1.   □

3.   Applications and the proof of Propp's conjecture

Recall that the complete directed graph H_n is the graph on the vertex set $V = [n] = \{1, 2, \ldots, n\}$ with exactly one directed edge from i to j for each $i \neq j$ and $i, j \in V$. The number of vertices of $D(H_n)$ is the number of rooted spanning trees on $[n]$, which is well known to equal n^{n-1}. We now apply the method of the previous section to give an extension and proof of Propp's conjecture.

Proposition 3.1   *The adjacency matrix $A(D(H_n))$, where $n \geq 2$, has eigenvalues -1, $1, \ldots, n - 1$. The multiplicity of i is*

$$m_{D(H_n)}(i) = \begin{cases} i \dbinom{n}{i+1} (n-1)^{n-i-2} & \text{if } 1 \leq i \leq n-1 \\ n^{n-1} - (n-1)^{n-1} & \text{if } i = -1. \end{cases}$$

Proof: The induced subgraphs of H_n are isomorphic to H_m for some $1 \leq m \leq n$ and the eigenvalues of H_m can easily be shown to be $m - 1$ with multiplicity 1 and -1 with multiplicity $m - 1$. Hence, by Corollary 2.3, the nonzero eigenvalues of $D(H_n)$ are included in the set $\{-1, 1, \ldots, n-1\}$. Moreover the eigenvalue $m - 1, 2 \leq m \leq n$ has multiplicity

$$\binom{n}{m} \det\left(L(H_n)|_{[m]} - I\right) = \binom{n}{m} (m-1)(n-1)^{n-m-1},$$

while -1 has multiplicity

$$\sum_{m=2}^{n} \binom{n}{m} (m-1)\det(L(H_n)|_{[m]} - I)$$

$$= \sum_{m=1}^{n} (m-1)^2 \binom{n}{m} (n-1)^{n-m-1} = n^{n-1} - (n-1)^{n-1},$$

where we have evaluated the last sum by classical elementary methods. The multiplicities we have so far add up to n^{n-1}, so 0 is not an eigenvalue of $D(H_n)$ and the proposition follows.   □

Corollary 3.2 *The Laplacian matrix $L(D(H_n))$, where $n \geq 2$, has eigenvalues $0, 1, \ldots,$
$n - 2, n$. The multiplicity of i is*

$$\begin{cases} (n - i - 1) \dbinom{n}{i} (n - 1)^{i-1} & \text{if } 0 \leq i \leq n - 1 \\ n^{n-1} - (n - 1)^{n-1} & \text{if } i = n. \end{cases}$$

Proof: It suffices to use Proposition 3.1 and the fact that $L(D(H_n)) = (n - 1)I - A(D(H_n))$. □

As a variation of the above result we give the following proposition. Its proof consists of a similar computation and is omitted.

Proposition 3.3 *Let $M_n(r)$ denote the directed graph on the vertex set $[n]$ with r directed edges from i to j for all $i, j \in [n]$. Then the nonzero eigenvalues of $D(M_n(r))$ are $r, 2r, \ldots, nr$ with*

$$m_{D(M_n(r))}(ir) = \begin{cases} (ir - 1) \dbinom{n}{i} (nr - 1)^{n-i-1} & \text{if } 1 \leq i \leq n \\ (nr)^{n-1} - (nr - 1)^{n-1} & \text{if } i = 0. \end{cases}$$

As a final example of the applicability of Corollary 2.3 we consider the complete bipartite directed graph $H_{r,s}$, where r, s are positive integers. Its vertex set is the disjoint union of two sets $[r] = \{1, \ldots, r\}$ and $[s]' = \{1', \ldots, s'\}$ with r and s elements respectively and for each pair of vertices (a, b) one of which is in $[r]$ and the other in $[s]'$, there is one edge directed from a to b.

Proposition 3.4 *The nonzero eigenvalues of $D(H_{r,s})$ are $\sqrt{pq}$ and $-\sqrt{pq}$ for $1 \leq p \leq r$,
$1 \leq q \leq s$ and the characteristic polynomial of its adjacency matrix equals*

$$x^t \prod_{p=1}^{r} \prod_{q=1}^{s} (x^2 - pq)^{m(p,q)},$$

where t is a nonnegative integer depending on r, s and

$$m(p, q) = \binom{r}{p} \binom{s}{q} (r - 1)^{s-q-1}(s - 1)^{r-p-1}(qr + ps - pq - r - s + 1).$$

Note that $m(p, q)$ is to be interpreted as 1 if $r = p = 1, q = s, 0$ if $r = p = 1, q < s$ and similarly for the case $s = q = 1$.

Proof: There are $\binom{r}{p} \binom{s}{q}$ subgraphs of $H_{r,s}$ isomorphic to $H_{p,q}$ for $0 \leq p \leq r, 0 \leq q \leq s$, the ones with $p = 0$ or $q = 0$ having only zero eigenvalues. It is easy to see that the nonzero eigenvalues of $H_{p,q}$, where p and q are positive, are $\sqrt{pq}$ and $-\sqrt{pq}$ (for example by counting closed walks), each with multiplicity one. Therefore Corollary 2.3 gives the set of eigenvalues proposed as the nonzero eigenvalues of $D(H_{r,s})$. Moreover the multiplicity of $\sqrt{pq}$ and $-\sqrt{pq}$ contributed by the $H_{p,q}$ induced subgraphs is

$$m(p, q) = \binom{r}{p} \binom{s}{q} \det(L(H_{r,s})|_{[p] \cup [q]'} - I).$$

The above determinant equals

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where $A = (s - 1)I_{(r-p) \times (r-p)}$, $D = (r - 1)I_{(s-q) \times (s-q)}$, $B = -J_{(r-p) \times (s-q)}$ and $C = -J_{(s-q) \times (r-p)}$ and J denotes a matrix with all entries equal to 1. This determinant can easily be shown to equal

$$(r - 1)^{s-q-1}(s - 1)^{r-p-1}(qr + ps - pq - r - s + 1)$$

using elementary row and column operations. This yields the suggested value of $m(p, q)$ and completes the proof of the proposition. □

4. Some further questions

The method of counting closed walks that we have considered here was successful in determining the eigenvalues of the adjacency matrices of the digraphs $D(G)$. It does not seem to be strong enough to give other information about these matrices, such as their eigenvectors or the structure of their Jordan canonical forms. Since adjacency matrices of directed graphs are not necessarily symmetric, in general they are not diagonalizable and hence they can have nontrivial Jordan canonical forms. We would thus like to conclude by posing the problem of describing the eigenspaces of these matrices, as it is possible to do for the 0-eigenspace of a Laplacian, and their Jordan canonical forms, in terms of information about the original graph G.

The Jordan block structure of the Propp matrices $L(D(H_n))$ has been computed for $n = 2$, 3, 4 by A. Edelman [3]. For $n = 3$ the eigenvalue 3 has one 1×1 and two 2×2 Jordan blocks and for $n = 4$ the eigenvalue 4 has four 1×1, twelve 2×2 and three 3×3 Jordan blocks. The rest of the eigenvalues for these values of n were found to be semisimple.

Finally we remark that it would be interesting to find other directed graphs whose eigenvalues can be computed more easily by the method of counting closed walks than by straightforward linear algebra techniques.

Acknowledgments

References

1. V. Anantharam and P. Tsoucas, "A proof of the Markov chain tree theorem," *Stat. Probab. Letters* **8** (1989), 189–192.
2. S. Chaiken and D.J. Kleitman, "Matrix tree theorems," *J. Combin. Theory Ser. A* **24** (1978), 377–381.
3. A. Edelman, private communication.
4. N. Elkies, G. Kuperberg, M. Larsen, and J. Propp, "Alternating sign matrices and domino tilings," *J. Alg. Combin.* **1** (1992), 111–132.
5. R. Merris, "Laplacian matrices of graphs: A survey," *Linear Algebra Appl.* **197, 198** (1994), 143–176.
6. J. Propp, private communication.
7. J. Propp, "Lattice structure for orientations of graphs," unpublished manuscript dated October 1, 1993.