The Distance Spectrum of a Tree

Russell Merris
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
CALIFORNIA STATE UNIVERSITY HAYWARD, CALIFORNIA

ABSTRACT

Let $T$ be a tree with line graph $T^*$. Define $K = 2I + A(T^*)$, where $A$ denotes the adjacency matrix. Then the eigenvalues of $-2K^{-1}$ interlace the eigenvalues of the distance matrix $D$. This permits numerous results about the spectrum of $K$ to be transcribed for the less tractable $D$.

Let $T = (V, E)$ be a tree with vertex set $V = \{1, 2, \ldots, n\}$ and edge set $E = \{e_1, e_2, \ldots, e_m\}$, $m = n - 1$. The distance matrix $D = D(T) = (d_{ij})$ is the $n$-by-$n$ matrix in which $d_{ij}$ is the number of edges in the unique path from vertex $i$ to vertex $j$. In 1971, R. L. Graham and H. O. Pollak showed that $\det(D) = (-1)^{n-1}(n - 1)2^{n-2}$, a formula depending only on $n$. It follows that $D$ is an invertible matrix with exactly one positive eigenvalue. In spite of this elegant beginning, results on the spectrum of $D$ have been few and far between.

Since any bipartite graph is 2-colorable, we may assume each vertex of $T$ has been given one of the “colors,” plus and minus, in such a way that each edge has a positive end and a negative end. The corresponding vertex-edge incidence matrix is the $n$-by-$m$ matrix $Q = Q(T) = (q_{ij})$, where $q_{ij} = 1$ if vertex $i$ is the positive end of $e_j$, $-1$ if it is the negative end, and $0$ otherwise. Define $K = K(T) = Q^TQ$. Then $K = 2I_n + A(T^*)$, where $A(T^*)$ is the 0-1 adjacency matrix of the line graph of $T$. Like $D$, the determinant of $K$ is a function only of $n$. Indeed, $\det(K) = n$. In contrast to $D$, however, all the eigenvalues of $K$ are positive. (It follows, as first observed by A. J. Hoffman, that the minimum eigenvalue of $A(T^*)$ is greater than $-2$. This has led to the notion of a “generalized line graph” and to an interesting connection with root systems [2, Section 1.1].)

A close relation of $K$ is the so-called Laplacian matrix $L(T) = QQ^T$. It turns out that $L(T) = \Delta(T) - A(T)$, where $\Delta(T)$ is the diagonal matrix of vertex degrees. The Laplacian first occurred in the Matrix-Tree Theorem of Kirchhoff. More recently, its spectrum has been the object of intense study stimulated in...
part by chemical applications [5, 9, 19] and in part by M. Fiedler’s notion of “algebraic connectivity” [10]. Of course, the m eigenvalues of $K$ are precisely the nonzero eigenvalues of $L(T)$.

**Theorem.** Let $T$ be a tree. Then the eigenvalues of $-2K^{-1}$ interlace the eigenvalues of $D$.

The key to the proof is an elementary observation implicit in [7, 8] and first proved explicitly by William Watkins [22].

**Lemma.** If $T$ is a tree on $n$ vertices and $m = n - 1$ edges, then $Q'DQ = -2I_m$.

**Proof.** $q_i d_{ij} q_j = 0$ unless $i$ is an end vertex of the $s$th edge $e_s$ and $j$ is an end vertex of $e_t$. Let $e_s = \{w, x\}$ and $e_t = \{y, z\}$, where $x$ and $z$ are the positive end vertices. Then $\sum_{i,j \in e_s} q_i d_{ij} q_j = d_{wx} - d_{wy} - d_{xy} + d_{xz}$.

If $s = t$, then $d_{wx} = d_{xz} = 0$ while $d_{wy} = d_{xy} = 1$, so the sum is $-2$. If $s \neq t$, it may still happen that $w = y$ or $x = z$. If $w = y$, then $d_{wx} = 0$, $d_{xz} = 2$, and $d_{wy} = d_{xy} = 1$, so the sum is zero. The case $x = z$ is handled similarly. If $w, x, y, z$ are four distinct vertices then either $x$ is on the (unique) path from $w$ to $e_s$ or $w$ is on the path from $x$ to $e_t$. These cases are similar. We argue the first, i.e., $d_{wx} = d_{xz} + 1$ and $d_{wy} = d_{xy} + 1$. In this case, the sum is

$$d_{xy} + 1 - (d_{xz} + 1) - d_{xy} + d_{xz} = 0.$$ 

To prove the theorem, note first that, as $K$ has rank $m$, the $m$ columns of $Q$ are linearly independent. We wish to perform a Gram-Schmidt orthonormalization process on these columns. Noting that this can be accomplished by a sequence of elementary column operations, we establish the existence of a nonsingular $m$-by-$m$ matrix $M$ (depending on $T$) such that the columns of the $n$-by-$m$ matrix $QM$ are orthonormal.

Recall that the column spaces of $Q$ and $QM$ are the same. Now, each column of $Q$ contains exactly 2 nonzero entries, one 1 and one $-1$. Denote by $F$ the $n$-by-$1$ column matrix, each of whose entries is equal to 1. Then $F$ is orthogonal to every column of $Q$, and hence to every column of $QM$. In particular, the $n$-by-$n$ partitioned matrix $U = (QM | F / \sqrt{n})$ is orthogonal.

Now,

$$U'DU = \left( \begin{array}{c} M'Q'DQM \\ R'QM/\sqrt{n} \\ M'Q'R/\sqrt{n} \\ 2W/n \end{array} \right),$$

where $R = DF$ is the column vector of row sums of $D$, and $W = F'DF/2$ is the so-called Wiener Index from chemistry [16, 20, 21]. Of course, the orthogonal similarity has preserved the spectrum of $D$. By the lemma, the leading $m$-by-$m$ principal submatrix of $U'DU$ is $-2M'M$. If we could show that $M'M$ and $K^{-1}$ have the same spectrum, we could apply Cauchy interlacing and be done. Now,
recall that $K = Q'Q$ and that $M$ was chosen so that the columns of $QM$ are orthonormal. Thus, $M'KM = M'Q'QM = I_m$. But then $M^{-1}K^{-1}(M')^{-1} = I_m$, i.e., $K^{-1} = MM'$. So, $K^{-1}$ and $M'M$ do have the same spectrum.

Before applying the results, we note some applications of the technique. Returning to (1), a new proof that $W = n(\text{trace}(K^{-1}))$ emerges from the fact that $\text{trace}(D) = 0$. (B. McKay seems to have been the first to notice this formula for the Wiener Index. Previous proofs have appeared in [16] and [18]. The first of these is based on an explicit graph-theoretic interpretation for the entries of $K^{-1}$.) Second, it follows from the lemma that $D$ has at least $m = n - 1$ negative eigenvalues. Since its Perron root is positive, we have a new proof that the inertia of $D$ is $(1, m, 0)$. (Similar arguments could be based on the observation that $Q'(xD + yI_n)Q = -2xI_m + yK$.) Finally, since $Q'DQ$ and $DQQ'$ have the same nonzero eigenvalues, the characteristic polynomial of $DL(T)$ is $x(x + 2)^{n - 1}$.

To illustrate the theorem itself, let $d_1 > 0 > d_2 > \cdots > d_n$ be the eigenvalues of $D = D(T)$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1} > 0$ be the eigenvalues of $K = K(T)$. Then $\lambda_{n-1} = a(T)$ is Fiedler's algebraic connectivity, and $1/\lambda_i, 1 \leq i \leq n$, are the eigenvalues of $K^{-1}$. Our theorem becomes

$$0 > \frac{-2}{\lambda_1} \geq \frac{-2}{\lambda_2} \geq \cdots \geq \frac{-2}{\lambda_{n-1}} \geq d_n. \quad (2)$$

A pendant vertex of $T$ is a vertex of degree 1. A pendant neighbor is a vertex adjacent to a pendant vertex. Suppose $T$ has $p$ pendant vertices and $q$ pendant neighbors.

**Corollary 1.** Let $d$ be an eigenvalue of $D(T)$ of multiplicity $k$. Then $k \leq p$.

**Proof.** By [15, Theorem 2.3], $p - 1$ is an upper bound on the multiplicity of any eigenvalue of $K(T)$. \Box

More information is available for certain specific eigenvalues. In [12], for example, the exact multiplicity of $\lambda_{n-1}$ was determined for "Type I" trees. It was shown in [15, Theorem 2.1 (ii)] that, apart from 1, $K(T)$ has no multiple integer eigenvalue. Thus, no eigenvalue of $D(T)$ of the form $-2/t$, $t = 2, 3, \ldots$, can have multiplicity greater than 2.

**Corollary 2.** Among the eigenvalues of $D(T)$, $d = -2$ occurs with multiplicity at least $p - q - 1$.

**Proof.** Isabel Faria [6] showed that the multiplicity of $\lambda = 1$ as an eigenvalue of $K(T)$ is at least $p - q$. \Box

Let $s(T)$ be the number of times $\lambda = 1$ occurs as an eigenvalue of $K(T)$, in excess of Faria's bound. Section III of [15] establishes various bounds for $s(T)$ in terms of the structure of $T$. It is proved, for example, that $s(T)$ is at most...
the covering number of the forest induced by $T$ on the vertices left after the pendants and their neighbors are removed. In [13], Faria-type bounds are obtained for other eigenvalues. Transcribing those for the distance matrix, it is clear from a glance at the tree in Figure 1 that $(x^2 - 6x + 4)^2$ exactly divides the characteristic polynomial of its distance matrix. (What is not clear is why $(x^2 - 6x + 4)^3$ should be a factor!)

**Corollary 3.** Let $\delta$ be the diameter of $T$. Then

$$d_n \leq \frac{-1}{1 - \cos(\pi/(\delta + 1))}.$$  

**Proof.** M. Doob [3] showed that the right hand side is an upper bound for $-2/a(T)$. $\blacksquare$

Many results are available in the literature concerning the algebraic connectivity $\alpha(T) = \lambda_{n-1}$. See, e.g., [1, 10-12, 14, 18, 19].

**Corollary 4.** Let $T$ be a tree with diameter $\delta$ and denote the greatest integer in $\delta/2$ by $k$. Then

(i) $d_k > -1$;
(ii) $d_q > -1$ (provided $n > 2q$);
(iii) $d_{n-q+1} < -2$;
(iv) $d_p \geq -2$; and
(v) $d_{n-p+2} \leq -2$.

**Proof.** It is proved in [15, Corollary 4.3] that $\lambda_q > 2$, in [17, Theorem 2] that $\lambda_q > 2$, and in [15, Theorem 3.11] that $\lambda_{n-q+1} < 1$. To prove (iv) and (v), note that $I_p$ is a principal submatrix of $L(T)$. By interlacing, $\lambda_p \geq 1$ and $\lambda_{n-p+1} \leq 1$. $\blacksquare$

In the exceptional case $n = 2q$, it turns out that $\lambda_q = 2$. If $n > 2q$, it may still happen that $\lambda_q = 2$ for some (at most 1) value of $t$. If so, Fiedler [11, p. 612] has shown how to determine $t$: Let $u$ be an eigenvector of $L(T)$ affording 2. Then the number of eigenvalues of $L(T)$ greater than 2 is equal to the number of edges $\{i,j\} \in E$ such that $u_i u_j > 0$. 
References


