Discrete Mathematics 309 (2009) 3458–3462

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Note

Average distance in graphs and eigenvalues

Sivaramakrishnan Sivasubramanian [1]

School of Technology and Computer Science, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India

ARTICLE INFO

Article history:
Received 21 December 2007
Received in revised form 7 August 2008
Accepted 26 September 2008
Available online 7 November 2008

Keywords:
Average distance
Eigenvalues
Laplacian

ABSTRACT

Brendan McKay gave the following formula relating the average distance between pairs of vertices in a tree T and the eigenvalues of its Laplacian:

$$\overline{d}_T = \frac{2}{n-1}\sum_{i=2}^{n}\frac{1}{\lambda_i}.$$

By modifying Mohar's proof of this result, we prove that for any graph G, its average distance, $\overline{d}_G$, between pairs of vertices satisfies the following inequality:

$$\overline{d}_G \geq \frac{2}{n-1}\sum_{i=2}^{n}\frac{1}{\lambda_i}.$$

This solves a conjecture of *Graffiti*. We also present a generalization of this result to the average of suitably defined distances for k subsets of a graph.

1. Introduction

We only consider undirected simple graphs in this paper. For vertices u and v of a graph G, let $d_G(u, v)$ be the distance in G from u to v. Let $\overline{d}_G$ be the average of $d_G(u, v)$, that is

$$\overline{d}_G \triangleq \binom{n}{2}^{-1} \sum_{\{u,v\}\subseteq V(G), u\neq v} d_G(u, v).$$

Brendan McKay (see Mohar and Poljak [5]) gave the following remarkable formula connecting the average distance between vertices of a tree and the eigenvalues of its Laplacian (see [3,4] for a proof). (The definition of the Laplacian of a graph is given in Section 2.)

Theorem 1. *Let T be a tree on n vertices and let $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of its Laplacian. Then,*

$$\overline{d}_T = \frac{2}{n-1}\sum_{i=2}^{n}\frac{1}{\lambda_i}. \qquad (1)$$

In this paper, we modify a proof of this theorem from [4], and derive the following theorem for general graphs.

E-mail addresses: krishnan@math.iitb.ac.in, skrishna@tcs.tifr.res.in.

[1] Currently at the Department of Mathematics, IIT Bombay, Mumbai, India.

0012-365X/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.disc.2008.09.044

Theorem 2. *Let G be a connected graph on n vertices and let $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of its Laplacian. Then,*

$$\bar{d}_G \geq \frac{2}{n-1} \sum_{i=2}^{n} \frac{1}{\lambda_i}.$$

Equality holds if and only if G is a tree.

A weaker statement, with $\frac{1}{n-1}$ instead of $\frac{2}{n-1}$ is Conjecture 167 in [7, page 68].

In the next section, we present the definitions, notation and facts that we use in our proofs. In Section 3, we recall the proof of Theorem 1 from Mohar [4]. In Section 4, we consider general graphs and prove Theorem 2, and finally, in Section 5, we obtain a generalization concerning the average over the $\binom{n}{k}$ suitably defined distances for a k subset of the graph.

2. Preliminaries

Let M be an $n \times n$ real and symmetric matrix. For $S \subseteq \{1, 2, \ldots, n\}$, let M_S be the matrix obtained by omitting from M all rows and columns whose indices appear in S.

Proposition 1. *Let M be an $n \times n$ real, symmetric matrix. The coefficient of x^k in $\mathrm{char}_M(x) \triangleq \det(xI - M)$ is $(-1)^{n-k} \sum_{S \in \binom{[n]}{k}} \det(M_S)$.*

Proof. We refer the reader to [2, page 196, Eq. 33] for a proof of this proposition. ■

The Laplacian of a graph: Let G be an undirected simple graph with n vertices. The adjacency matrix of G, denoted by $A(G)$, is the $n \times n$ symmetric matrix $[a_{uv}]_{u,v \in V(G)}$, whose entry a_{uv} is 1 if $\{u, v\} \in E(G)$ and 0 otherwise. Let $D(G) = [d_{uv}]_{u,v \in V(G)}$ be the diagonal $n \times n$ matrix with the degrees of the vertices along the diagonal, that is, $d_{vv} = \deg_G(v)$. The Laplacian of G, denoted by $L(G)$, is the symmetric matrix $D(G) - A(G)$. The characteristic polynomial of $L(G)$, denoted by $\mathrm{char}_G(x)$, is the polynomial in one variable x with real coefficients, given by $\det(xI - L(G))$. The matrix $L(G)$, is positive semidefinite and has n real eigenvalues (i.e. roots of $\mathrm{char}_G(x)$): $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. It is easy to see that $\lambda_1 = 0$ and $\lambda_2 > 0$ if and only if G is connected (see Mohar and Poljak [5, Section 2.2]). Thus, $\mathrm{char}_G(x) = x \cdot \prod_{i=2}^{n}(x - \lambda_i)$. For $S \subseteq V(G)$, let $\mathcal{F}_S$ be the set of spanning forests of G that contain exactly $|S|$ trees, where each component contains a different vertex of S. Let $\mathcal{F}$ be the set of all spanning trees of G.

Proposition 2. (a) *For $2 \leq k \leq n$, the coefficient of x^k in $\mathrm{char}_G(x)$ is*

$$(-1)^{n-k} \sum_{S \in \binom{\{2,3,\ldots,n\}}{k-1}} \prod_{j \in \{2,3,\ldots,n\} - S} \lambda_j.$$

(b) *For $S \subseteq V(G) (S \neq \emptyset)$,*

$$|\mathcal{F}_S| = \det(L_S).$$

Proof. For a proof of part (a), take the coefficient of x^k in the expansion of $x \cdot \prod_{i=2}^{n}(x - \lambda_i)$.

For part (b), we give a proof for completeness. The case when $|S| = 2$ is [1, ex 3, page 303]. It is known (see for example [6, Theorem 2.2.12]) that $L(G) = \mathrm{In} \times \mathrm{In}^t$ where In is a $|V| \times |E|$ oriented incidence matrix. Our proof of part (b) is similar to the Binet–Cauchy Theorem (see statement below) based proof of the Matrix Tree Theorem given in [6, Theorem 2.2.12].

Theorem 3 (*Binet–Cauchy Theorem (see [6, Ex 8.6.19])*). *Let $C = AB$ where A and B are $n \times m$ and $m \times n$ matrices respectively. Given $S \subseteq [m]$ of size n, let A_S be the $n \times n$ matrix whose columns are those columns of A, indexed by S and let B_S be the $n \times n$ matrix whose rows are those rows of B indexed by S. Then $\det(C) = \sum_S \det(A_S)\det(B_S)$ where the summation is over all n-sized subsets of $[m]$.*

To apply Binet–Cauchy's Theorem towards proving part (b), we note that we remove $|S| = k$ vertices. Thus we need to find all $F \subseteq E$, with $n - k$ edges such that the $(n-k) \times (n-k)$ matrix $\mathrm{In}_{V-S,F}$ (which is the submatrix of In whose rows omit those rows of In indexed by S and whose columns contain the edges in F) has a non-zero determinant.

Let $S = \{u_1, u_2, \ldots, u_k\}$. Let $V - S = \{v_1, v_2, \ldots, v_{n-k}\}$. We claim that such sets F are precisely those corresponding to elements of $\mathcal{F}_S$. Let $F = \{e_1, e_2, \ldots, e_{n-k}\}$ be a subset such that $\det(\mathrm{In}_{V-S,F}) \neq 0$. We first note that no edge $e_i \in F$ can be adjacent to two vertices in S, as then the column in $\mathrm{In}_{V-S,F}$ corresponding to e_i will only have zeroes and thus force $\det(\mathrm{In}_{V-S,F}) = 0$. It is simple to see that the graph $G_F = (V, F)$ induced on the edges F cannot have cycles if we want $\det(\mathrm{In}_{V-S,F}) \neq 0$. Thus each component of G_F is a tree and since G_F has $n - k$ edges and n vertices, it has has precisely k connected components (i.e. precisely k trees). We show that no two distinct vertices $u_i, u_j \in S$, where $i \neq j$ can be in the same component if we want $\det(\mathrm{In}_{V-S,F}) \neq 0$. This is proved will imply that only those $F \in \mathcal{F}_S$ give a non-zero determinant.

Suppose $p = \langle f_1, f_2, \ldots, f_r \rangle$ is a path from u_i to u_j contained in G_F. We can assume that all intermediate vertices (that is, vertices other than u_i and u_j) are not in S (if not, we could choose an initial prefix of p with this property). Since no edge in

G_F connects two vertices in S, it is clear that each f_i, except f_1 and f_r, has both its endpoints in $V-S$. Consider the $(n-k)\times r$ submatrix of $\mathrm{In}_{V-S,F}$ corresponding to the edges in p. By adding the column vectors of this submatrix (in the same order as the path p) with appropriate signs, we can clearly get the zero vector and thus $\det(\mathrm{In}_{V-S,F})=0$.

Thus, only the edge sets $F\in\mathcal{F}_S$ have non-zero $\det(\mathrm{In}_{V-S,F})$ value. It is easy to see by induction on the number of edges in E that this non-zero value is ± 1. Thus, by the Binet–Cauchy Theorem, $\det(L_S)=\sum_{F\in\mathcal{F}_S}(\pm 1)^2=|\mathcal{F}_S|$. The proof is complete. ■

Corollary 1 (*Matrix Tree Theorem*). *Let $S\subseteq V$ be a singleton set. Thus, $S=\{v\}$ for some $v\in V$ and for brevity, we write L_v instead of $L_{\{v\}}$. Let $\mathcal{F}$ be the set of all spanning trees of G. Applying Proposition 2*(b) *for this S, we get $|\mathcal{F}|=\det(L_v)$ for all $v\in V$.*

Just as we write L_v instead of $L_{\{v\}}$, we also write L_{uv} instead of $L_{\{u,v\}}$.

3. Trees: Proof of Theorem 1

This proof is just a restatement of the proof in Mohar [4, page 63]. It will be convenient to rewrite the right hand side of (1) as

$$\left(\frac{2}{n-1}\right)\frac{\sum_{i=2}^{n}\prod_{j=2,j\neq i}^{n}\lambda_j}{\prod_{i=2}^{n}\lambda_i}. \tag{2}$$

We will show that

$$\prod_{i=2}^{n}\lambda_i=n \tag{3}$$

$$\sum_{i=2}^{n}\prod_{j=2,j\neq i}^{n}\lambda_j=\sum_{\{u,v\}\subseteq V(T),u\neq v}d_T(u,v). \tag{4}$$

The theorem follows immediately by combining (3) and (4) with (2). To show (3) and (4), we will use Propositions 1 and 2.

Proof of (3): By Proposition 2(a), $(-1)^{n-1}\prod_{i=2}^{n}\lambda_i$ is the coefficient of x in $\mathrm{char}_T(x)$. By Proposition 1 this coefficient is equal to the $(-1)^{n-1}\sum_{v\in V(G)}\det(L_v(T))$. By Corollary 1 $\det(L_v(T))=1$ for all $v\in V(T)$. Thus,

$$\prod_{i=2}^{n}\lambda_i=\sum_{v\in V(G)}\det(L_v(T))=n.$$

Proof of (4): The coefficient of x^2 in $\mathrm{char}_T(x)$ is precisely $(-1)^{n-2}\sum_{i=2}^{n}\prod_{j=2,j\neq i}^{n}\lambda_j$. By Proposition 2(a), this is also the sum over all $\binom{n}{2}$ pairs $\{u,v\}\subseteq V(T)$ of $(-1)^{n-2}\det(L_{uv}(T))$. By Proposition 2(b), $\det(L_{uv}(T))$ is equal to the number of spanning forests that have two components, one containing u and the other containing v. In a tree the number of such spanning forests is precisely the distance $d_T(u,v)$ between u and v. Thus,

$$\sum_{i=2}^{n}\prod_{j=2,j\neq i}^{n}\lambda_j=\sum_{\{u,v\}\subseteq V(T),u\neq v}\det(L_{uv}(T))=\sum_{\{u,v\}\subseteq V(T),u\neq v}d_T(u,v).\quad\blacksquare$$

4. General graphs: Proof of Theorem 2

We again start with the expression (2). Instead of (3) and (4), we now have

$$\prod_{i=2}^{n}\lambda_i=\kappa\cdot n, \tag{5}$$

$$\sum_{i=2}^{n}\prod_{j=2,j\neq i}^{n}\lambda_j\leq\kappa\cdot\sum_{\{u,v\}\subseteq V(G),u\neq v}d_G(u,v), \tag{6}$$

where κ is the number of spanning trees in G. The equality in (5) follows from Proposition 2(a) and (b), using arguments similar to those used in the proof of (3). The proof of (6) is also similar. As before, using parts (a) and (b) of Proposition 2 we compute the coefficient of x^2 in $\mathrm{char}_G(x)$ and conclude that

$$\sum_{i=2}^{n}\prod_{j=2,j\neq i}^{n}\lambda_j=\sum_{\{u,v\}\subseteq V(G),u\neq v}\kappa_{uv}, \tag{7}$$

where κ_{uv} is the number of spanning forests of G that have two components, one containing u and the other containing v. We only need to observe that $\kappa_{uv} \leq \kappa \cdot d_G(u, v)$. Let $p_{uv} = \langle e_1, e_2, \ldots, e_d \rangle$ ($d = d_G(u, v)$) be the sequence of edges in a shortest path from u to v. With each forest $F \in \mathcal{F}_{uv}$ let us associate a pair $(T, i) \in \mathcal{F} \times \{1, 2, \ldots, d_G(u, v)\}$, as follows. Note that there is an edge in the sequence p whose addition to F makes it connected. Let e_i be the first such edge, and let T be the spanning tree obtained by adding e_i to F. It is easy to see that the map $F \mapsto (T, i)$ is one-to-one. We thus have

$$\kappa_{uv} = |\mathcal{F}_{uv}| \leq |\mathcal{F} \times \{1, 2, \ldots, d_G(u, v)\}| = \kappa \cdot d_G(u, v).$$

Only trees satisfy Theorem 2 with equality: We have seen in Section 3 that trees satisfy the above inequality with equality. We thus have to prove the *only if* part of the theorem. Suppose a connected graph satisfied Theorem 2 with equality, then by the above derivation, for this graph G, for all unordered pairs of vertices $\{u, v\}$, $\kappa_{uv} = \kappa \cdot d_G(u, v)$. This implies that the graph is a tree.

This completes the proof of Theorem 2. ■

5. A generalisation

For $S \subseteq V$ and $x \in V(G)$, let

$$d_x(S) = \sum_{u \in S} \prod_{v \in S - \{u\}} d_G(x, v).$$

Let $d(S)$ be the minimum value of $d_x(S)$ as x ranges over all vertices of G. Note that when $S = \{u, v\}$, $d_x(S) = d_G(u, v)$ where $d_G(u, v)$ is the shortest distance between u and v in the graph G. Let $\overline{d}_k(G)$ be the average of $d(S)$ as S ranges over all subsets of V of size k.

Theorem 4. *Let G be a connected graph on n vertices whose Laplacian has eigenvalues $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$. Then*

$$\overline{d}_k(G) \geq n\binom{n}{k}^{-1} \sum_{S \in \binom{\{2,3,\ldots,n\}}{k-1}} \prod_{j \in S} \frac{1}{\lambda_j}. \tag{8}$$

Proof. As in the earlier proofs, we begin by rearranging the right hand side of (8):

$$n\binom{n}{k}^{-1} \frac{\displaystyle\sum_{S \in \binom{\{2,3,\ldots,n\}}{k-1}} \prod_{j \in \{2,\ldots,n\} - S} \lambda_j}{\displaystyle\prod_{i=2}^{n} \lambda_i}.$$

We again bound the denominator and numerator separately and show

$$\prod_{i=2}^{n} \lambda_i = \kappa \cdot n \tag{9}$$

$$\sum_{S \in \binom{\{2,3,\ldots,n\}}{k-1}} \prod_{j \in \{2,3,\ldots,n\} - S} \lambda_j \leq \kappa \cdot \sum_{S \in \binom{V}{k}} d(S) \tag{10}$$

where, as before, κ is the number of spanning trees of the graph G.

Since (9) is the same as (5), we only need to prove (10).

Proof of (10): By Proposition 2(a), the expression in the left hand side of (10) is the coefficient of x^k in $\mathrm{char}_G(x)$. By Proposition 1, this is precisely $\sum_{S \subseteq \binom{V}{k}} \det(L_S)$. By Proposition 2(b), $\det(L_S)$ is the number of spanning forests of G with k components such that each element of S lies in a different component. That is,

$$\sum_{S \in \binom{\{2,3,\ldots,n\}}{k-1}} \prod_{j \in \{2,3,\ldots,n\} - S} \lambda_j = \sum_{S \subseteq \binom{V}{k}} \det(L_S) = \sum_{S \subseteq \binom{V}{k}} |\mathcal{F}_S|. \tag{11}$$

We will show a one-to-one map from the set $\mathcal{F}_S$ to the set $\mathcal{F} \times D$, where D is a set of size at most $d(S)$. Let $S = \{s_1, s_2, \ldots s_k\}$. Fix a vertex x such that $d_x(S) = d(S)$ and fix shortest (x, s_j)-paths p_j, for $j = 1, 2, \ldots, k$. Fix a forest $F \in \mathcal{F}_S$. Let x lie in the same component as s_i. For $j = 1, 2, \ldots, k$ ($j \neq i$), we will identify an edge e_j in p_j using the following sequential algorithm. Let e_j be the first edge of p_j that enters the component of s_j in the graph $F \cup \{e_\ell : 1 \leq \ell < j, \ell \neq i\}$. This way we obtain a sequence of $k - 1$ edges $\mathbf{e} = \langle e_\ell : \ell \neq i \rangle$. Note that $F \cup \{e_\ell : 1 \leq \ell \leq k, \ell \neq i\}$ is a spanning tree of G. Note that the map $F \mapsto (T, \mathbf{e})$ is a one-to-one map from the $\mathcal{F}_S$ to $\mathcal{F} \times D$, where $D = \bigcup_{i=1}^{k} \prod_{j \neq i} p_j$. Clearly, $|D| \leq d(S)$. Thus,

$$|\mathcal{F}_S| \leq |\mathcal{F} \times D| = \kappa \cdot d(S)$$

The theorem follows by combining this with (11). ■

Acknowledgements

I thank Jaikumar Radhakrishnan for several useful discussions and for enhancing the presentation of the material. I thank Siemion Fajtlowicz for very helpful emails. I also thank the anonymous referees for various suggestions which made the paper more readable.

References

[1] Chris Godsil, Gordon Royle, Algebraic Graph Theory, Springer Verlag, 2001, GTM 207.
[2] Nathan Jacobson, Basic Algebra, vol. I, second ed., W H Freeman and Company, 1985.
[3] R. Merris, An edge version of the Matrix-Tree Theorem and the Wiener index, Linear Multilinear Algebra 25 (1989) 291–296.
[4] Bojan Mohar, Eigenvalues, Diameter and Mean Distance in graphs, Graphs Combin. 7 (1991) 53–64.
[5] Bojan Mohar, Svatopluk Poljak, Combinatorial and graph-theoretical problems in linear algebra, in: Eigenvalues in Combinatorial Optimisation, in: IMA Volumes in Mathematics and Its Applications, vol. 50, 1993, pp. 107–151. Available at: http://www.fmf.uni-lj.si/~mohar/Papers/Poljak.pdf.
[6] Douglas B. West, Introduction to Graph Theory, second ed., Prentice Hall of India, 2001.
[7] Siemion Fajtlowicz, Graffiti, Written on the Wall, 2004. Available at: http://www.math.uh.edu/~clarson/wow-pre-840.ps.