MINORS OF TREE DISTANCE MATRICES

HARRY RICHMAN, FARBOD SHOKRIEH, AND CHENXI WU

Abstract. We prove a formula for the determinant of the principal minors of the distance matrix of a weighted tree. This generalizes a result of Graham and Pollak.

Contents

1. Introduction

Suppose $G = (V, E)$ is a tree with n vertices. Let D denote the distance matrix of G. In [4], Graham and Pollak proved that

$$\det D = (-1)^{n-1} 2^{n-2} (n-1). \tag{1}$$

This identity is remarkable in that the result does not depend on the tree structure, beyond the number of vertices. The identity (1) was motivated by a problem in data communication, and inspired much further research on distance matrices.

The main result of this paper is to generalize (1) by replacing $\det D$ with any of its principal minors. For a subset $S \subset V(G)$, let $D[S]$ denote the submatrix consisting of the S-indexed rows and columns of D.

Date: v1, November 22, 2022 (Preliminary draft, not for circulation).

thm:main **Theorem 1.** *Suppose G is a tree with n vertices, and distance matrix D. Let $S \subset V(G)$ be a nonempty subset of vertices. Then*

eq:main
$$\det D[S] = (-1)^{|S|-1} 2^{|S|-2} \left((n-1)\,\kappa(G;S) - \sum_{\mathcal{F}_2(G;S)} (2 - \deg^o(F,*))^2 \right), \tag{2}$$

where $\kappa(G;S)$ is the number of S-rooted spanning forests of G, $\mathcal{F}_2(G;S)$ is the set of $(S,*)$-rooted spanning forests of G, and $\deg^o(F,*)$ denotes the outdegree of the $*$-component of F.

For definitions of $(S,*)$-rooted spanning forests and other terminology, see ♢ TODO: cite section ♢ . Note that the quantity $\deg^o(F,*)$ satisfies the bounds

$$1 \le \deg^o(F,*) \le |S|.$$

When $S = V$ is the full vertex set, the set of V-rooted spanning forests is a singleton, consisting of the vertex set with no edges, so $\kappa(G;V) = 1$; and moreover the set $\mathcal{F}_2(G;V)$ of $(V,*)$-rooted spanning forests is empty. Thus (2) recovers the Graham–Pollak identity (1) when $S = V$.

1.1. Weighted trees.

A weighted version of (1) was proved by Bapat–Kirkland–Neumann [1]. If $\{\alpha_e : e \in E\}$ is a collection of positive edge weights, the α-distance matrix D_α takes the distance along edge e to be α_e; then

eq:w-full-det
$$\det D_\alpha = (-1)^{n-1} 2^{n-2} \sum_{e \in E} \alpha_e \prod_{e \in E} \alpha_e.$$

The weighted identity (1.1) reduces to (1) when taking all unit weights, $\alpha_e = 1$. We also prove the following weighted version of our main theorem.

thm:w-main **Theorem 2.** *Suppose $G = (V,E)$ is a finite, weighted tree with edge weights $\{\alpha_e : e \in E\}$, and corresponding weighted distance matrix D_α. For any nonempty subset $S \subset V$, we have*

eq:w-main
$$\det D_\alpha[S] = (-1)^{|S|-1} 2^{|S|-2} \left(\sum_{E(G)} \alpha_e \sum_{\mathcal{F}_1(G;S)} w(T) - \sum_{\mathcal{F}_2(G;S)} k(F,*)^2 w(F) \right).$$

where $\mathcal{F}_1(G;S)$ is the set of S-rooted spanning forests of G, $\mathcal{F}_2(G;S)$ is the set of $(S,*)$-rooted spanning forests of G, $w(T)$ and $w(F)$ denote the α-weights of the forests T and F, and

eq:2-minus-deg
$$k(F,*) = 2 - \deg^o(F,*)$$

where $\deg^o(F,*)$ is the outdegree of the $*$-component of F, as above.

It is worth observing that depending on the chosen subset $S \subset V$, the distances appearing in the submatrix $D[S]$ may ignore a large part of the ambient tree G. We could instead replace G by the subtree consisting of paths between vertices in S, which we call $\mathrm{conv}(S,G)$, the *convex hull* of $S \subset G$. To apply formula (2) or (2) "efficiently," we should replace G with this convex hull $\mathrm{conv}(S,G)$. However, the formulas as stated are true even without this replacement due to cancellation of terms.

Corollary 3.

$$\frac{\det D[S]}{\operatorname{cof} D[S]} = \frac{1}{2}\left(\sum_{E(G)} \alpha_e - \frac{\sum_{\mathcal{F}_2(G/S)} w(F)k(F,*)^2}{\sum_{\mathcal{F}_1(G/S)} w(T)}\right).$$

♢ add remark / theorem that det/cof is achieved as result of optimization problem ♢

We remark that the calculation of $\det D[S]$ is related to the following optimization problem: for $\mathbf{m} \in \mathbb{R}^S$,

$$\begin{aligned} \text{optimize:} \quad & D[S]\mathbf{m} \\ \text{with constraints:} \quad & \mathbf{1}^\intercal \mathbf{m} = 1. \end{aligned}$$

By the theory of Lagrange multipliers, the solution $\mathbf{m}^*$ is the vector satsifying

$$D[S]\mathbf{m}^* = \lambda \mathbf{1} \qquad \text{for some } \lambda \in \mathbb{R},$$

or equivalently $\mathbf{m}^* = \lambda(D[S]^{-1}\mathbf{1})$.

Theorem 4 (Monotonicity of principal minor ratios)**.** *Suppose $G = (V, E)$ is a finite, weighted tree with distance matrix D.*

(1) If $S \subset V(G)$ is nonempty,

$$0 \le \frac{\det D[S]}{\operatorname{cof} D[S]} \le \frac{1}{2} \sum_{E(G)} \alpha_e.$$

(2) If $\operatorname{conv}(S, G)$ denotes the subtree of G consisting of all paths between points of $S \subset V(G)$,

$$\frac{\det D[S]}{\operatorname{cof} D[S]} \le \frac{1}{2} \sum_{E(\operatorname{conv}(S,G))} \alpha_e.$$

(3) If $A, B \subset V(G)$ are nonempty subsets with $A \subset B$, then

$$\frac{\det D[A]}{\operatorname{cof} D[A]} \le \frac{\det D[B]}{\operatorname{cof} D[B]}.$$

Theorem 5 (Nonsingular minors)**.** *Let G be a finite, weighted tree with distance matrix D, and let $S \subset V(G)$ be a subset of vertices. If $|S| \ge 2$ then $\det D[S] \neq 0$.*

1.2. **Previous work.** A formula for the inverse matrix D^{-1} was found by Graham and Lovász in graham-lovasz [3].

1.3. **Notation.** G a finite graph, loops and parallel edges allowed, possibly disconnected

$E(G)$ edge set of G
$V(G)$ vertex set of G
$\kappa(G)$ number of spanning trees of G
$\mathcal{F}_1(G)$ the set of spanning trees of G
$\mathcal{F}_2(G)$ the set of 2-component forests of G
$\mathcal{F}_1(G; S)$ the set of S-rooted spanning forests of G
$\mathcal{F}_2(G; S)$ the set of $(S, *)$-rooted spanning forests of G

2. Graph matrices

For background on enumeration problems for graphs and trees, see Moon moon [5].

Given a graph $G = (V, E)$, let $L \in \mathbb{R}^{V \times V}$ denote the *Laplacian matrix* of G. If G is a weighted graph with edge weights $\alpha_e \in \mathbb{R}_{>0}$ for $e \in E$, let L denote the weighted Laplacian matrix of G.

Given $S \subset V$, let $L[\overline{S}]$ denote the matrix obtained from L by removing the rows and columns indexed by S.

For any graph G, let $\kappa(G)$ denote the number of spanning trees of G. The following theorem is due to Kirchhoff. For any graph G, let $\kappa(G)$ denote the number of spanning trees of G.

thm:matrix-tree **Theorem 6** (All-minors matrix tree theorem). *Let $G = (V, E)$ be a finite graph, and let L denote the Laplacian matrix of G. Then for any nonempty vertex set $S \subset V(G)$,*

$$\det L[\overline{S}] = \kappa(G/S).$$

Note that $\kappa(G/S)$ is also the number of S-rooted spanning forests of G.

The following result is due to Bapat–Siviasubramanian.

Theorem 7 (Distance matrix cofactor sums bapat-sivasubramanian [2]). *Given a tree G, let D be the distance matrix of G, and L the Laplacian matrix. Let $S \subset V(G)$ be a nonempty subset of vertices of G. Then*

$$\operatorname{cof} D[S] = (-2)^{|S|-1} \det L[\overline{S}].$$

2.1. **Distance in trees.** Given an edge $e \in E$ and vertices $v, w \in V$, let

$$\delta(e; v, w) = \begin{cases} 1 & \text{if} \\ 0 & \text{if} . \end{cases}$$

Note that we can express the tree distance $d(v, w)$ as a sum over edges

$$d(v, w) = \sum_{e \in E(G)} \delta(e; v, w) \qquad \text{where } \delta(e; v, w) = \begin{cases} 1 & \text{if } e \text{ lies on } v \sim w \text{ path}, \\ 0 & \text{otherwise}. \end{cases}$$

- If we fix v and w, then $\delta(-; v, w) : E(G) \to \{0, 1\}$ is the indicator function for the unique v w path in G.
- On the other hand if we fix e and v, then the deletion $G \backslash e$ has two connected components, and $\delta(e; v, -) : V(G) \to \{0, 1\}$ is the indicator function for the component of $G \setminus e$ not containing v.

prop:distance-sum **Proposition 8.**

$$d(v, w) = \sum_{e \in E} \ell(e) \delta(e; v, w).$$

2.2. **Tree distance.** For $e \in E$ and $v, w \in V$, let

$$\delta(e; v, w) = \begin{cases} 1 & \text{if } e \text{ separates } v \text{ from } w, \\ 0 & \text{otherwise}. \end{cases}$$

(Note that $\delta(e; v, v) = 0$ for any e and v.)

2.3. **Principal cuts.** Given a tree $G = (V, E)$ and an edge $e \in E$, the edge deletion $G \setminus e$ contains two connected components. Using the implicit orientation on $e = (e^+, e^-)$, we let $(G \setminus e)^+$ denote the component that contains endpoint e^+, respectively $(G \setminus e)^-$ and endpoint e^-.

For any $e \in E$ and $v \in V$, we let $(G \setminus e)^v$ denote the component of $G \setminus e$ containing v, respectively $(G \setminus e)^{\overline{v}}$ ♢ or $(G \setminus e)^{-v}$ ♢ for the component not containing v.

2.4. **Outdegree of rooted forest.** Given a rooted forest F in $\mathcal{F}(G; S)$ and $s \in S$, let $F(s)$ denote the s-component of F. We define the *outdegree* $\deg^o(F, s)$ by

eq:outdeg

$$\deg^o(F, s) = \#\{e = (a, b) \in E : a \in F(s),\, b \notin F(s)\}.$$

In words, $\deg^o(F, s)$ is the number of edges which connect the s-component of F to a different component.

$$\#\{e \in E : e \text{ connects the } s\text{-component of } F \text{ to a different component}\}$$

If F is a forest in $\mathcal{F}_2(G/S)$, let $\deg^o(F, *)$ denote the outdegree of the floating component.

lem:outdeg-sum

Lemma 9. *Suppose G is a tree and $H \subset G$ is a (nonempty) connected subgraph. Then*

$$\sum_{v \in V(H)} (2 - \deg(v)) = 2 - \deg^o(H).$$

Proof. This is straightforward to check by induction on $|V(H)|$, with base case $|V(H)| = 1$: if $H = \{v\}$ consists of a single vertex, then $\deg^o(H) = \deg(v)$. □

2.5. **Spanning trees and forests.** A *spanning tree* of a graph G is a subgraph which is connected, has no cycles, and contains all vertices of G. A *spanning forest* of a graph G is a subgraph which has no cycles and contains all vertices of G.

Given a set of vertices $S = \{v_1, v_2, \ldots, v_k\}$, an *$S$-rooted spanning forest* of G is a spanning forest which has exactly one vertex v_i in each connected component. An *$(S, *)$-rooted spanning forest* of G is a spanning forest which has $|S|+1$ components, where $|S|$ components each contain one vertex of S, and the additional component is disjoint from S. We call the component disjoint from S the "floating component."

Let

$$\kappa_k(v_1|v_2|\cdots|v_k)$$

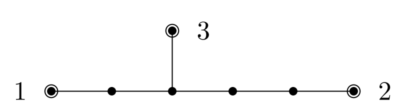
denote the number of k-component spanning trees which have a vertex v_i in each component. If $S = \{v_1, \ldots, v_k\}$, then $\kappa_k(v_1|\cdots|v_k) = \kappa(G/S)$.

If u, v, w are vertices, then let

$$\kappa_2(uv|w)$$

denote the number of two-forests which have u, v in one component and w in the other component.

Example 10. Suppose G is the tree with unit edge lengths shown below.



Let S be the set of three leaf vertices. Then $\mathcal{F}_1(G; S)$ contains 11 forests, while $\mathcal{F}_2(G; S)$ contains 19 forests. These are shown in Figures fig:1-forests ?? and fig:2-forests ??, respectively.

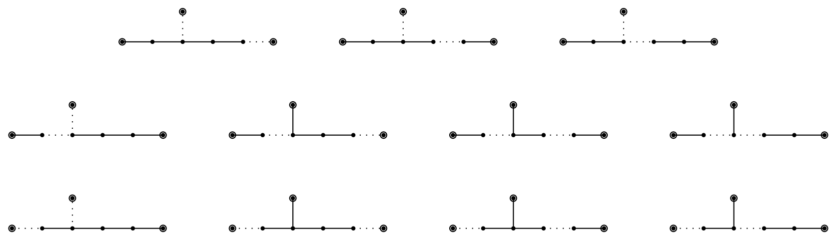


FIGURE 1. Forests in $\mathcal{F}_1(G;S)$. fig:1-forests

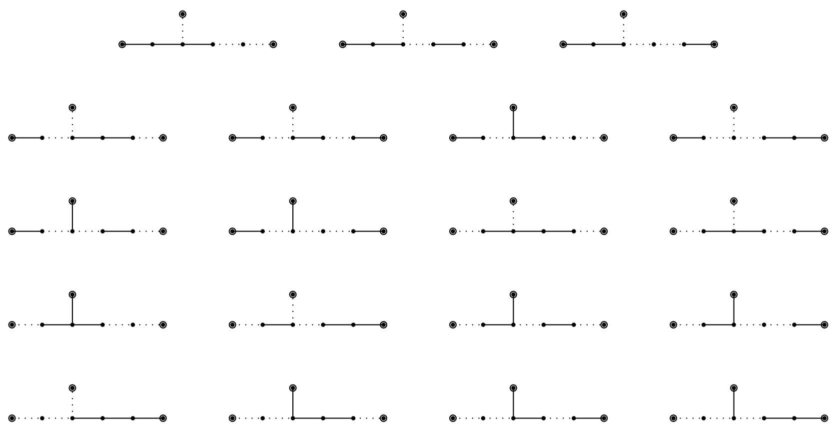


FIGURE 2. Forests in $\mathcal{F}_2(G;S)$. fig:2-forests

3. Proofs

In this section we prove Theorem thm:w-main 2.

Outline of proof: given a subset $S \subset V$ and distance submatrix $D[S]$, we will

(i) Find vector $\mathbf{m} \in \mathbb{R}^S$ such that $D[S]\mathbf{m} = \lambda \mathbf{1} \in \mathbb{R}^S$.

(ii) Compute the sum of entries of $\mathbf{m}$, i.e. $\mathbf{1}^\intercal \mathbf{m}$.

(iii) Using (i), note the identity

$$\mathbf{1}^\intercal \mathbf{m} = \lambda(\mathbf{1}^\intercal D[S]^{-1}\mathbf{1}) = \lambda \frac{\operatorname{cof} D[S]}{\det D[S]}.$$

where $\operatorname{cof} D[S]$ is the sum of cofactors of $D[S]$.

(iv) Use known expression for $\operatorname{cof} D[S]$ to compute

$$\det D[S] = \lambda(\operatorname{cof} D[S])\,(\mathbf{1}^\intercal \mathbf{m})^{-1}\,.$$

The interesting part of this expression will turn out to be in the constant λ.

Example 11. Suppose G is a tree consisting of three paths joined at a central vertex. Let S consist of the central vertex, and the three endpoints of the paths. The corresponding submatrix of the distance matrix is

$$D[S] = \begin{bmatrix} 0 & a & b & c \\ a & 0 & a+b & a+c \\ b & a+b & 0 & b+c \\ c & a+c & b+c & 0 \end{bmatrix}.$$

Following the steps outlined above:

(i) The vector $\mathbf{m} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ satisfies $D[S]\mathbf{m} = (a+b+c)\mathbf{1}$

(ii) The sum of entries of $\mathbf{m}$ is $\mathbf{1}^\intercal \mathbf{m} = 2$.

(iii) We have

$$2 = \mathbf{1}^\intercal \mathbf{m} = \lambda(\mathbf{1}^\intercal D[S]^{-1}\mathbf{1}) = \lambda \frac{\operatorname{cof} D[S]}{\det D[S]}.$$

(iv) The cofactor sum $\operatorname{cof} D[S]$ is $-8abc$, so the determinant is

$$\det D[S] = \lambda \frac{\operatorname{cof} A}{\mathbf{1}^\intercal \mathbf{m}} = (a+b+c)(-8abc)\frac{1}{2} = -4(a+b+c)abc.$$

3.1. **Warmup case: $S = V$.**

Proposition 12. *Let $G = (V, E)$ a tree, and consider the vector $\mathbf{m} \in \mathbb{R}^V$ defined by*

$$\mathbf{m}_v = 2 - \deg v \qquad \textit{for each } v \in V.$$

Then $\mathbf{1}^\intercal \mathbf{m} = \sum_{v \in V}(2 - \deg v) = 2$.

Proof. For any graph, $\sum_{v \in V} \deg v = 2|E|$. Since G is a tree, $|E| = |V| - 1$. $\square$

prop:m-distance-warmup

Proposition 13. *Let $\mathbf{m}$ be the vector defined above, and let D be the distance matrix of G. Then $D\mathbf{m} = \lambda \mathbf{1}$ for some constant λ.*

Proof. It suffices to show that for each edge e, with endpoints (e^+, e^-), we have

$$(D\mathbf{m})_{(e^+)} = (D\mathbf{m})_{(e^-)}.$$

We compute

$$\begin{aligned}
(D\mathbf{m})_{(e^+)} - (D\mathbf{m})_{(e^-)} &= \sum_{v \in V} (d(v, e^+) - d(v, e^-))(2 - \deg v) \\
&= \sum_{v \in (G \setminus e)^-} \alpha_e (2 - \deg v) - \sum_{v \in (G \setminus e)^+} \alpha_e (2 - \deg v)
\end{aligned}$$

eq:12-1

since

$$d(v, e^+) - d(v, e^-) = \begin{cases} \alpha_e & \text{if } v \text{ is closer to } e^- \text{ than } e^+, \\ -\alpha_e & \text{if } v \text{ is closer to } e^+ \text{ than } e^-. \end{cases}$$

♢ TO DO: define notation $(G \setminus e)^\pm$ ♢ For each sum in (eq:12-1 ??), we apply Proposition ♢ cite ♢ to obtain

$$\alpha_e \sum_{v \in (G \setminus e)^-} (2 - \deg v) = \alpha_e (2 - \deg^o((G \setminus e)^-) = \alpha_e.$$

The same identity applies to the sum over $(G \setminus e)^+$, so $(D\mathbf{m})_{(e^+)} = (D\mathbf{m})_{(e^-)}$ as desired. $\square$

3.2. **General case: $S \subset V$.** Fix a tree $G = (V, E)$ and a subset $S \subset V$.

dfn:m-vector **Definition 14.** Let $\mathbf{m} = \mathbf{m}(G; S) \in \mathbb{R}^S$ be defined by

eq:m-vector
$$\mathbf{m}_v = \sum_{T \in \mathcal{F}_1(G;S)} 2 - \deg^o(T, v) \qquad \text{for each } v \in S.$$

where $\deg^o(T, v)$ is the outdegree of the v-component of T, (eq:outdeg ??).

Let $\mathbf{1}$ denote the all-ones vector.

Proposition 15. *For* $\mathbf{m}$ *defined above,* $\mathbf{1}^\intercal \mathbf{m} = 2\,\kappa(G; S)$.

Proof. We have

$$\begin{aligned}
\mathbf{1}^\intercal \mathbf{m} = \sum_{s \in S} \mathbf{m}_s &= \sum_{s \in S} \left(\sum_{T \in \mathcal{F}_1(G;S)} 2 - \deg^o(T, s) \right) \\
&= \sum_{T \in \mathcal{F}_1(G;S)} \left(\sum_{s \in S} \sum_{v \in T(s)} 2 - \deg(v) \right) \\
&= \sum_{T \in \mathcal{F}_1(G;S)} \left(\sum_{v \in V} 2 - \deg(v) \right) = \sum_{T \in \mathcal{F}_1(G;S)} 2.
\end{aligned}$$

In the second line we apply Lemma lem:outdeg-sum ?? and exchange the outer summations. To obtain the third line, we observe that the vertex sets of $T(s)$ for $s \in S$ form a partition of V, since T is an S-rooted spanning forest. Finally we again apply Lemma lem:outdeg-sum ?? for the last equality, as $\deg^o(G) = 0$. $\square$

Theorem 16. *With* $\mathbf{m} = \mathbf{m}(G; S)$ *defined as in* (eq:m-vector ??), $D[S]\mathbf{m} = \lambda \mathbf{1}$ *for the constant*

$$\lambda = \sum_{\mathcal{F}_1(G;S)} w(T) \sum_{E(G)} \alpha_e - \sum_{\mathcal{F}_2(G;S)} w(F)(2 - \deg^o(F, *))^2$$

where $\deg^o(F, w)$ *is the out-degree of the* w*-component of* F *(as a spanning forest).*

Proof. For $e \in E$ and $v, w \in V$, let

$$\delta(e; v, w) = \begin{cases} 1 & \text{if } e \text{ separates } v \text{ from } w, \\ 0 & \text{otherwise.} \end{cases}$$

For any $v \in S$, we have

$$\begin{aligned}
(D[S]\mathbf{m})_v &= \sum_{s \in S} d(v, s) \mathbf{m}_s \\
&= \sum_{s \in S} \left(\sum_{e \in E(G)} \alpha_e \, \delta(e; v, s) \right) \left(\sum_{T \in \mathcal{F}_1(G;S)} w(T)(2 - \deg^o(T, s)) \right) \\
&= \sum_{e \in E} \alpha_e \sum_{T \in \mathcal{F}_1} w(T) \sum_{s \in S} (2 - \deg^o(T, s)) \delta(e; v, s) \\
&= \sum_{e} \alpha_e \sum_{T} w(T) \sum_{s \in S \cap (G \setminus e)^{\overline{v}}} (2 - \deg^o(T, s)). \qquad \text{eq:14-1}
\end{aligned}$$

where

$$S(G \setminus e)^{\overline{v}} = \{ s \in S : e \text{ separates } v \text{ from } s \}$$

$$S^*(e, v) = \{s \in S : e \text{ lies on path from } v \text{ to } s\}.$$

♢ MOVE TO REMARK? If $e \in \mathrm{conv}(G, S)$, then $S(G\backslash e)^{\overline{v}}$ is nonempty and ♢ We have

$$\sum_{s \in S(G \setminus e)^{\overline{v}}} (2 - \deg^o(T, s)) = \begin{cases} 1 & \text{if } e \notin T, \\ 1 - (2 - \deg^o(T \setminus e, *)) & \text{if } e \in T(s'),\ s' \in S(G \setminus e)^{\overline{v}}, \\ 1 + (2 - \deg^o(T \setminus e, *)) & \text{if } e \in T(s'),\ s' \in S(G \setminus e)^{v} \end{cases}$$

Here $T \setminus e$ is a forest in $\mathcal{F}_2(G; S)$, and $\deg^o(T \setminus e, *)$ refers to the out-degree of the "floating" component of $T \setminus e$. These cases correspond to

$$(G \setminus e)^{\overline{v}} = \begin{cases} \displaystyle\bigcup_{s \in S(G \setminus e)^{\overline{v}}} T_s & \text{if } e \notin T, \\ \left(\displaystyle\bigcup_{s \in S(G \setminus e)^{\overline{v}}} T_s \right) \cup (T \setminus e)_* & \text{if } e \in T_{s'},\ \delta(e; v, s') = 0 \\ \left(\displaystyle\bigcup_{s \in S(G \setminus e)^{\overline{v}}} T_s \right) \setminus (T \setminus e)_* & \text{if } e \in T_{s'},\ \delta(e; v, s') = 1. \end{cases}$$

From (??) we have

$$(D[S]\mathbf{m})_v = \sum_{e \in E} \sum_{T \in \mathcal{F}_1} \alpha_e w(T)(1 - f(v, e, T))$$

where

$$f(v, e, T) = \begin{cases} 0 & \text{if } e \notin T \\ 2 - \deg^o(T \setminus e, *) & \text{if } e \in T(s') \text{ for some } s' \in S,\ \delta(e; v, s') = 0 \\ -(2 - \deg^o(T \setminus e, *) & \text{if } e \in T(s') \text{ for some } s' \in S,\ \delta(e; v, s') = 1 \end{cases}$$

Thus

$$\begin{aligned} (D[S]\mathbf{m})(v) - \sum_{e \in E} \sum_{T \in \mathcal{F}_1} \alpha_e w(T) = &- \sum_{T \in \mathcal{F}_1} \sum_{e \in T \setminus T(S^*)} \alpha_e w(T)(2 - \deg^o(T \setminus e, *)) \\ &+ \sum_{T \in \mathcal{F}_1} \sum_{e \in T(S^*)} \alpha_e w(T)(2 - \deg^o(T \setminus e, *)) \end{aligned}$$

If $e \in T$, the deletion $T \setminus e$ is an $(S, *)$-rooted spanning forest of G, so we may rewrite the above expression in terms of $\mathcal{F}_2(G; S)$.

$$\begin{aligned} (1) = &- \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F, *)) \sum_{T \in \mathcal{F}_1} \sum_{e \in T \setminus T(S^*)} \mathbb{1}(F = T \setminus e) \\ &+ \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F, *)) \sum_{T \in \mathcal{F}_1} \sum_{e \in T(S^*)} \mathbb{1}(F = T \setminus e) \end{aligned}$$

Next, we note that $F = T \setminus e$ is equivalent to $T = F \cup e$, and in particular this only occurs when we choose $e \in \partial(F, *)$:

$$
\begin{aligned}
(1) = &- \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F, *)) \sum_{e \in \partial F(*)} \mathbb{1}(\delta(e; v, F(*)) = 1) \\
&+ \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F, *)) \sum_{e \in \partial F(*)} \mathbb{1}(\delta(e; v, F(*)) = 0)
\end{aligned}
$$

Here we let $\delta(e; v, F_*)$ denote

$$
\delta(e; v, F_*) = \begin{cases} 0 & \text{if} \\ 1 & \text{if} \end{cases}
$$

Finally, we observe that for any forest F in $\mathcal{F}_2(G; S)$, there is exactly one edge e in the boundary $\partial(F, *)$ of the floating component which satisfies $\delta(e; v, F(*)) = 1$, namely the unique boundary edge on the path from $F(*)$ to v. Thus

$$
\#\{e \in \partial F(*) : \delta(e; v, F(*)) = 1\} = 1
$$

and

$$
\#\{e \in \partial(F, *) : \delta(e; v, F(*)) = 0\} = \deg^o(F, *) - 1
$$

Thus

$$
\begin{aligned}
(1) &= - \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F, *))(1) \\
&\quad + \sum_{F \in \mathcal{F}_2} w(F)(2 - \deg^o(F, *))(\deg^o(F, *) - 1) \\
&= - \sum_{F \in \mathcal{F}_2(G/S)} w(F)(2 - \deg^o(F, *))^2.
\end{aligned}
$$

as desired. $\square$

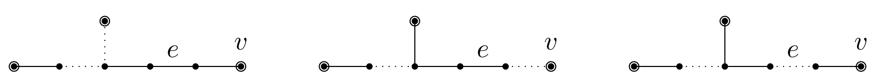


Figure 3. Components rooted in $S(G \setminus e)^{\overline{v}}$.

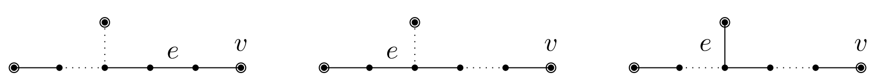


Figure 4. Components rooted in $S(G \setminus e)^{\overline{v}}$.

Remark 17. The set $\mathcal{F}_2(G; S)$ of $(S, *)$-rooted spanning forests of G can be partitioned into two types: "active" and "inactive".

$$
\mathcal{F}_2(G/S) = \mathcal{F}_2^{in}(G/S) \sqcup \mathcal{F}_2^{out}(G/S),
$$

where

$$
\begin{aligned}
\mathcal{F}_2^{in}(G/S) &= \{F \in \mathcal{F}_2(G/S) \text{ such that } \deg^o(*, F) \geq 2\}, \\
\mathcal{F}_2^{out}(G/S) &= \{F \in \mathcal{F}_2(G/S) \text{ such that } \deg^o(*, F) = 1\}.
\end{aligned}
$$

Remark 18. For a given spanning forest $F \in \mathcal{F}_2(G/S, *)$, there are exactly $\deg^o(F,*)$ choices of pairs $(T,e) \in \mathcal{F}_1(G/S) \times E(G)$ such that $F = T \setminus e$. Consider the map

$$E(G) \times \mathcal{F}_1(G/S) \to \mathcal{F}_2(G/S) \sqcup \{\text{error}\}$$

defined by ...

$$(e,T) \mapsto \begin{cases} T \setminus e & \text{if } e \in T, \\ \text{error} & \text{if } e \notin T. \end{cases}$$

For a forest F in $\mathcal{F}_2(G/S)$, the preimage under this map has $\deg^o(F,*)$ elements.

Proposition 19. *Let $G = (V,E)$ be a tree, and $S \subset V$. Suppose we label $S = \{s_1, \ldots, s_r\}$ and $V \setminus S = \{t_1, \ldots, t_{n-r}\}$. For each $t_i \in V \setminus S$, consider $\mathbf{f}_i \in \mathbb{R}^V$ defined by*

Example 20. If

$$D[S \cup t] = \begin{bmatrix} 0 & a & b & c \\ a & 0 & a+b & a+c \\ b & a+b & 0 & b+c \\ c & a+c & b+c & 0 \end{bmatrix}$$

then

$$\begin{bmatrix} 0 & a & b & c \\ a & 0 & a+b & a+c \\ b & a+b & 0 & b+c \\ c & a+c & b+c & 0 \end{bmatrix} \begin{bmatrix} ab+ac+bc \\ -bc \\ -ac \\ -ab \end{bmatrix} = \begin{bmatrix} -3abc \\ -abc \\ -abc \\ -abc \end{bmatrix}$$

4. Physical interpretation

If we consider G as a network of wires with each edge containing a unit resistor, which is grounded at all nodes in S, then $\mathbf{m}_S$ records the currents flowing to S when current is added on $V \setminus S$ in the amount $2 - \deg v$ for each $v \notin S$.

4.1. Alternate proof.

Let $\mathbf{1}$ denote the all-ones vector. When we choose a subset $S \subset V(G)$, we no longer have a single "obvious" replacement for $\mathbf{m}$ inside $\mathbb{R}^S$. Instead, we can take an average over S-rooted spanning forests.

In the outline above, our first goal is to find a "special" vector $\mathbf{m} \in \mathbb{R}^S$ satisfying $D[S]\mathbf{m} = \lambda\mathbf{1}$. We can approach this first goal as follows: consider $\mathbb{R}^S$ inside the larger vector space $\mathbb{R}^V = \mathbb{R}^S \oplus \mathbb{R}^{V \setminus S}$, and we wish to find vectors $\mathbf{n}_i \in \mathbb{R}^V$ satisfying $\pi_S(D\mathbf{n}_i) = \lambda_i\mathbf{1}$. By finding sufficiently many such vectors $\mathbf{n}_i$, we can hope to find a linear combination that lies inside $\mathbb{R}^S \oplus \{0\}$.

prop:n-vector

Proposition 21. *Suppose $v \in V \setminus S$. For each $s_j \in S$, let $\mu(v, s_j)$ = current flowing to s_j when unit current enters G at v and G is grounded at S. Explicitly,*

$$\begin{aligned} \mu(v,s) &= \frac{\#\text{ of } S\text{-rooted spanning forests of } G \text{ whose } s_j\text{-component contains } v}{\#\text{ of } S\text{-rooted spanning forests of } G} \\ &= \frac{\sum_{\mathcal{F}_1(G/S)} \mathbb{1}(v \in T(s))}{\kappa(G/S)} \\ &= \frac{\kappa_r(s_1|\cdots|s_j v|\cdots|s_r)}{\kappa_r(s_1|\cdots|s_r)} \end{aligned}$$

Consider the vector $\mathbf{n} = \mathbf{n}(G; S, v) \in \mathbb{R}^V$ defined by

$$\mathbf{n}_v = 1, \qquad \mathbf{n}_s = -\mu(v,s) \text{ if } s \in S, \qquad \mathbf{n}_w = 0 \text{ if } w \notin S \cup v$$

Then $D\mathbf{n}$ is constant on S, i.e. $\pi_S(D\mathbf{n}) = \lambda\mathbf{1}$ for some λ.

Proof sketch. For any $s, s' \in S$, consider tracking value of $D\mathbf{n}$ along path from s to s'. The value of $D\mathbf{n}$ changes according to current flow in the corresponding network, i.e. $D\mathbf{n}$ records electrical potential. By assumption S is grounded, so $D\mathbf{n}$ takes the same value at s and s'. □

Theorem 22. *Let G be a tree, S a nonempty subset of vertices, and $D[S]$ the corresponding submatrix of the distance matrix. Suppose $\mathbf{m} = \mathbf{m}(G; S) \in \mathbb{R}^S$ is defined by (eq:m-vector ??);*

$$\mathbf{m}(G;S)_v = \sum_{T \in \mathcal{F}_1(G;S)} \sum_{w \in T_v} (2 - \deg w) = \sum_{T \in \mathcal{F}_1(G;S)} 2 - \deg^o(T, v).$$

Then $D[S]\mathbf{m} = \lambda\mathbf{1}$ for some constant λ.

Proof. Note that

$$\begin{aligned}
\mathbf{m}(G;S) &= \kappa(G/S)\left(\sum_{v \in V} (2 - \deg v)\boldsymbol{\delta}(v) - \sum_{v \in V \setminus S} (2 - \deg v)\mathbf{n}(G; S, v)\right) \\
&= \kappa(G;S)\left(\mathbf{m}(G;V) - \sum_{v \in V \setminus S} (2 - \deg v)\mathbf{n}(G; S, v)\right)
\end{aligned}$$

♢ TODO: elaborate on this equation ♢ From Proposition prop:m-distance-warmup ?? we know that $D\mathbf{m}(G;V)$ is constant on V, and from Proposition prop:n-vector ?? we know that $D\mathbf{n}(G; S, v)$ is constant on S. Hence by linearity, $D\mathbf{m}(G;S)$ is constant on S. □

5. Examples

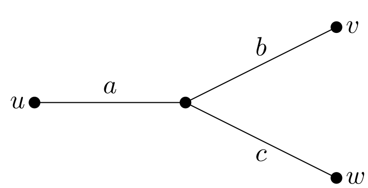
Example 23. Suppose G is a tree consisting of three paths joined at a central vertex. Let S consist of the central vertex, and the three endpoints of the paths. The corresponding minor of the distance matrix is

$$D[S] = \begin{bmatrix} 0 & a & b & c \\ a & 0 & a+b & a+c \\ b & a+b & 0 & b+c \\ c & a+c & b+c & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & a & b & c \\ a & -a & a & a \\ b & b & -b & b \\ c & c & c & -c \end{bmatrix} \sim \begin{bmatrix} 0 & a & b & c \\ a & -2a & 0 & 0 \\ b & 0 & -2b & 0 \\ c & 0 & 0 & -2c \end{bmatrix}.$$

The determinant is

$$\det D[S] = -4(a+b+c)abc.$$

Example 24. Suppose Γ is a tripod with lengths a, b, c and corresponding leaf vertices u, v, w.



Let $S = \{u, v, w\}$. Then

$$D[S] = \begin{bmatrix} 0 & a+b & a+c \\ a+b & 0 & b+c \\ a+c & b+c & 0 \end{bmatrix}.$$

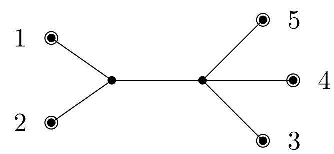
and

$$\det D[S] = 2(a+b)(a+c)(b+c) = 2\left((a+b+c)(ab+ac+bc) - abc\right).$$

The "special vector" that satisfies $D[S]\mathbf{m} = \lambda\mathbf{1}$ in this example is

$$\mathbf{m}^{\intercal} = \begin{bmatrix} a(b+c) & b(a+c) & c(a+b) \end{bmatrix}.$$

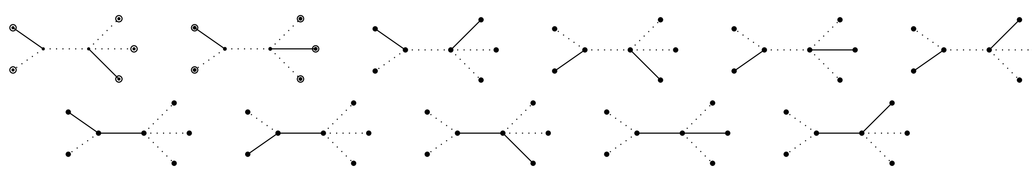
Example 25. Suppose Γ is the tree with unit edge lengths shown below, with five leaf vertices.



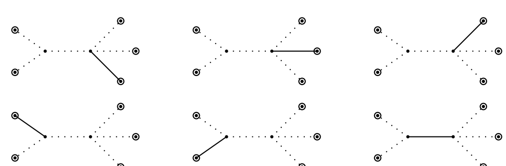
Let S denote the set of five leaf vertices. Then

$$D[S] = \begin{bmatrix} 0 & 2 & 3 & 3 & 3 \\ 2 & 0 & 3 & 3 & 3 \\ 3 & 3 & 0 & 2 & 2 \\ 3 & 3 & 2 & 0 & 2 \\ 3 & 3 & 2 & 2 & 0 \end{bmatrix}.$$

Forests in $\mathcal{F}_1(G; S)$:



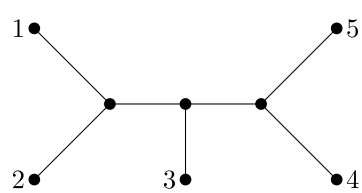
Forests in $\mathcal{F}_2(G; S)$:



and

$$\det D[S] = 368 = (-1)^4 2^3 \left(6 \cdot 11 - (3 \cdot 1^2 + 2 \cdot 2^2 + 1 \cdot 3^2)\right)$$

Example 26. Suppose Γ is the tree with unit edge lengths shown below, with five leaf vertices and three internal vertices.



Let S denote the set of five leaf vertices. Then

$$D[S] = \begin{bmatrix} 0 & 2 & 3 & 4 & 4 \\ 2 & 0 & 3 & 4 & 4 \\ 3 & 3 & 0 & 3 & 3 \\ 4 & 4 & 3 & 0 & 2 \\ 4 & 4 & 3 & 2 & 0 \end{bmatrix}$$

and

$$\det D[S] = 864 = (-1)^4 2^3 \left(7 \cdot 21 - (14 \cdot 1^2 + 4 \cdot 2^2 + 1 \cdot 3^2)\right)$$

6. Further work

richman-shokrieh-wu
See [6].

Acknowledgements

The authors would like to thank Ravindra Bapat for helpful discussion.

References

bapat-kirkland-neumann [1] R. Bapat, S. J. Kirkland, and M. Neumann. On distance matrices and Laplacians. *Linear Algebra Appl.*, 401:193–209, 2005.

bapat-sivasubramanian [2] R. B. Bapat and S. Sivasubramanian. Identities for minors of the Laplacian, resistance and distance matrices. *Linear Algebra Appl.*, 435(6):1479–1489, 2011.

graham-lovasz [3] R. L. Graham and L. Lovász. Distance matrix polynomials of trees. *Adv. in Math.*, 29(1):60–88, 1978.

graham-pollak [4] R. L. Graham and H. O. Pollak. On the addressing problem for loop switching. *Bell System Tech. J.*, 50:2495–2519, 1971.

moon [5] J. W. Moon. *Counting labelled trees*. Canadian Mathematical Monographs, No. 1. Canadian Mathematical Congress, Montreal, Que., 1970. From lectures delivered to the Twelfth Biennial Seminar of the Canadian Mathematical Congress (Vancouver, 1969).

richman-shokrieh-wu [6] D. H. Richman, F. Shokrieh, and C. Wu. Capacity on metric graphs, 2022. in preparation.