AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **70(2)** (2018), Pages 185–201

Distance matrices of some positive-weighted graphs

Agnese Baldisserri   Elena Rubei

Dipartimento di Matematica e Informatica "U. Dini"
viale Morgagni 67/A
50134 Firenze
Italy
agnese.baldisserri@hotmail.com   rubei@math.unifi.it

Abstract

Let $\mathcal{G} = (G, w)$ be a positive-weighted connected graph, that is, a connected graph G endowed with a function w from the edge set of G to the set of positive real numbers, and let $\{1, \ldots, n\}$ be its vertex set. For any subgraph H of G, we define the weight of H to be the sum of the weights of the edges of H and, for any $i, j \in \{1, \ldots, n\}$, we define $D_{i,j}(\mathcal{G})$ to be the minimal weight of any path in G with endpoints i and j. Obviously the $D_{i,j}(\mathcal{G})$ form a symmetric $n \times n$ matrix with zero diagonal entries and positive off-diagonal entries. It is well-known that a symmetric matrix with zero diagonal entries and positive off-diagonal entries comes from a positive-weighted connected graph if and only if its entries satisfy the triangle inequalities. In this paper we consider some particular classes of graphs: paths, caterpillars, cycles, bipartite graphs, complete graphs, planar graphs; for each of these classes, we give a criterion to establish whether, given a symmetric $n \times n$ matrix D with zero diagonal entries and positive off-diagonal entries, there exists a positive-weighted graph $\mathcal{G} = (G, w)$ in the class we have fixed, with vertex set equal to $\{1, \ldots, n\}$ and such that $D_{i,j}(\mathcal{G}) = D_{i,j}$ for every $i, j \in \{1, \ldots, n\}$.

1 Introduction

For any graph G, let $E(G)$, $V(G)$ and $L(G)$ be respectively the set of the edges, the set of the vertices and the set of the leaves of G. A *weighted graph* $\mathcal{G} = (G, w)$ is a graph G endowed with a function $w : E(G) \to \mathbb{R}$. For any edge e, the real number $w(e)$ is called the weight of the edge. We say that $\mathcal{G}$ is *positive-weighted* if all the weights of the edges are positive.

Definition 1.1. Let $\mathcal{G} = (G, w)$ be a weighted simple finite connected graph; for any subgraph H of G, we define $w(H)$ to be the sum of the weights of the edges of H and, for any $i, j \in V(G)$, we define $D_{i,j}(\mathcal{G})$ to be the minimal weight of any simple (i.e. without repeated vertices) path in G with endpoints i and j. We say that a path H with endpoints i and j *realizes* $D_{i,j}(\mathcal{G})$ if $w(H) = D_{i,j}(\mathcal{G})$. We call $D_{i,j}(\mathcal{G})$ the 2-*weight* of $\mathcal{G}$ for $\{i,j\}$ or, if $\mathcal{G}$ is positive-weighted, the *distance* between i and j in $\mathcal{G}$.

Throughout the paper the word "graph" will mean "simple finite connected graph" and the word "matrix" will mean "real matrix". Observe that in the case $\mathcal{G}$ is a weighted tree, $D_{i,j}(\mathcal{G})$ is the weight of the unique path joining i and j.

If S is a subset of $V(G)$, the 2-weights $D_{i,j}(\mathcal{G})$ with $i,j \in S$ form a *symmetric matrix*. Equivalently, we can speak of the *family of the* 2-*weights* of $(\mathcal{G}, S)$ or of the 2-*dissimilarity family* of $(\mathcal{G}, S)$.

We can wonder when a symmetric matrix is the matrix of 2-weights of some weighted graph and of some subset of the set of its vertices. We say that a symmetric $n \times n$ matrix D is *graphlike* (*p-graphlike*) if there exist a weighted graph (respectively a positive-weighted) graph $\mathcal{G} = (G, w)$ and a subset $\{1, \ldots, n\}$ of the set of its vertices such that $D_{i,j}(\mathcal{G}) = D_{i,j}$ for any $i,j \in \{1,\ldots,n\}$. In this case, we say that $\mathcal{G}$ realizes the matrix. If the graph is a weighted (positive-weighted) tree $\mathcal{T} = (T, w)$ we say that the matrix is *treelike* (respectively *p-treelike*).

The first contribution to the characterization of graphlike matrices dates back to 1965 and it is due to Hakimi and Yau; see [9]:

Theorem 1.1. (Hakimi-Yau) *A symmetric $n \times n$ matrix D with zero diagonal entries and positive off-diagonal entries is p-graphlike if and only if its entries satisfy the triangle inequalities, i.e. if and only if $D_{i,j} \leq D_{i,k} + D_{k,j}$ for any distinct $i,j,k \in \{1,\ldots,n\}$.*

In the same years, also a criterion for a metric on a finite set to be p-treelike was established; see [5, 16, 18].

Theorem 1.2. (Buneman-SimoesPereira-Zaretskii) *Let D be a symmetric $n \times n$ matrix with zero diagonal entries and positive off-diagonal entries such that its entries satisfy the triangle inequalities. It is p-treelike if and only if the $D_{i,j}$ satisfy the so-called 4-point condition, i.e., for all distinct $i,j,k,h \in \{1,\ldots,n\}$, the maximum of*

$$\{D_{i,j} + D_{k,h}, D_{i,k} + D_{j,h}, D_{i,h} + D_{k,j}\}$$

is attained at least twice.

Also the case of not necessarily positive weights has been studied. In 1972 Hakimi and Patrinos proved the following theorem (see [8]):

Theorem 1.3. (Hakimi-Patrinos) *A symmetric matrix D with zero diagonal entries is always graphlike.*

In [4], Bandelt and Steel proved a result, analogous to Theorem 1.2, for general weighted trees:

Theorem 1.4. (Bandelt-Steel) *For any symmetric matrix D with zero diagonal entries, there exists a weighted tree $\mathcal{T}$ with leaves $1,\ldots,n$ such that $D_{i,j}(\mathcal{T}) = D_{i,j}$*

for any $i, j \in \{1, \dots, n\}$ *if and only if, for any distinct* $a, b, c, d \in \{1, \dots, n\}$*, we have that at least two among*

$$D_{a,b} + D_{c,d}, \quad D_{a,c} + D_{b,d}, \quad D_{a,d} + D_{b,c}$$

are equal.

Weighted graphs have applications in several disciplines, such as biology, psychology, archeology, engineering, computer science. Phylogenetic trees are positive-weighted trees whose vertices represent species and the weight of an edge is given by how much the DNA sequences of the species represented by the vertices of the edge differ; biologists often know the "distances" (that is, how much the DNA differs) between only some of the species, and they are interested in reconstructing all the evolutionary tree from these data. There is a wide literature concerning graph-like dissimilarity families and treelike dissimilarity families, in particular concerning methods to reconstruct weighted trees from their dissimilarity families; these methods are used by biologists to reconstruct phylogenetic trees; see [13] and [17]. Also archeologists represent evolutions of manuscripts by positive-weighted trees. See for example [6], [15] for overviews on phylogenetic trees. Weighted graphs can represent hydraulic webs or railway webs where the weight of an edge is given by the length or the cost (or the difference between the earnings and the cost) of the line represented by that edge or, finally, computer or social networks. It is possible that sometimes we know the distances between some nodes of the network but not the "form" of the network and it can be interesting to have criteria to establish it from the distances.

Finally we want to mention that recently k-weights of weighted graphs for $k \geq 3$ have been introduced and studied: given a positive-weighted connected graph $\mathcal{G} = (G, w)$ and $i_1, \dots, i_k \in V(G)$, we define $D_{i_1,\dots,i_k}(\mathcal{G})$ to be the minimum of $w(R)$ where R is a connected subgraph whose vertex set contains $i_1, \dots, i_k$. In particular there are some results concerning the characterization of families of k-weights; see for instance [1, 2, 10, 11, 12, 14]. The study of k-weights for $k \geq 3$ is motivated by the fact that they are more reliable statistically than 2-weights and so the reconstruction of weighted trees from them can be more accurate than the reconstruction from 2-weights. We quote only three theorems. Let $n, k \in \mathbb{N} - \{0\}$. We say that a family of positive real numbers $\{D_I\}_{I \in \binom{\{1,\dots,n\}}{k}}$ is *ip-l*-treelike if and only if there exists a weighted tree $\mathcal{T}$ with all the weights of the edges nonnegative, all the weights of the internal edges positive and leaf set $\{1, \dots, n\}$ such that $D_I(\mathcal{T}) = D_I$ for any $I \in \binom{\{1,\dots,n\}}{k}$.

Theorem 1.5. (Herrmann, Huber, Moulton, Spillner, [10]). *Let* $n, k \in \mathbb{N} - \{0\}$*. Let* $\{D_I\}_{I \in \binom{\{1,\dots,n\}}{k}}$ *be a family of positive real numbers. If* $n \geq 2k$*, the family is ip-l-treelike if and only if its restriction to every* $2k$*-subset of* $\{1, \dots, n\}$ *is ip-l-treelike.*

Theorem 1.6. (Pachter-Speyer, [14]). *Let* $k, n \in \mathbb{N}$ *with* $3 \leq k \leq (n+1)/2$*. A positive-weighted tree* $\mathcal{T}$ *with leaves* $1, \dots, n$ *and no vertices of degree* 2 *is determined by the values* $D_I(\mathcal{T})$*, where* I *varies in* $\binom{\{1,\dots,n\}}{k}$.

Theorem 1.7. (Baldisserri-Rubei, [1]) *Let* $n \in \mathbb{N}$, $n \geq 3$. *Let* $\{D_I\}_{I \in \binom{\{1,\dots,n\}}{n-1}}$ *be a family of positive real numbers and let us denote* $D_{1,\dots,i-1,i+1,\dots,n}$ *by* $D_{\hat{i}}$.

There exists a positive weighted graph $\mathcal{G} = (G, w)$ *with exactly* n *vertices,* $1, \dots, n$, *and with* $D_{\hat{i}}(\mathcal{G}) = D_{\hat{i}}$ *for any* $i = 1, \dots, n$ *if and only if the following two conditions hold:*

(i)

$$(n-2) D_{\hat{i}} \leq \sum_{j=1,\dots,n,\ j \neq i} D_{\hat{j}} \tag{1}$$

for any $i \in \{1, \dots, n\}$,

(ii) if the maximum in $\{D_{\hat{i}}\}_{i \in \{1,\dots,n\}}$ *is achieved at least twice, the inequalities* (1) *are strict.*

In this paper we consider some particular classes of graphs: paths, caterpillars, cycles, bipartite graphs, complete graphs, planar graphs; for each of these classes, we give a criterion to establish whether, given a symmetric $n \times n$ matrix D with zero diagonal entries and positive off-diagonal entries and such that its entries satisfy the triangle inequalities, there exists a positive-weighted graph $\mathcal{G} = (G, w)$ in the class we have fixed, with $V(G) = \{1, \dots, n\}$ and such that $D_{i,j}(\mathcal{G}) = D_{i,j}$ for any $i, j \in \{1, \dots, n\}$. This kind of problem can arise, for example, in phylogenetics, because we can wonder if an evolutionary history can be represented by a caterpillar.

2 Some definitions and some remarks

We now present some notation.

Let $\mathbb{R}$ be the set of real numbers and define $\mathbb{R}_+ = \{x \in \mathbb{R} |\ x > 0\}$. Let $\mathbb{N}$ be the set of nonnegative integers. Throughout the paper, let $n \in \mathbb{N} - \{0, 1\}$ and let $[n] = \{1, \dots, n\}$.

We say that a square matrix is *a predistance matrix* if it is symmetric, all its diagonal entries are zero and its off-diagonal entries are positive. We say that a predistance matrix D is *a distance matrix* if its entries satisfy the triangle inequalities, that is $D_{i,j} \leq D_{i,k} + D_{k,j}$ for any $i, j, k \in [n]$. By Theorem 1.1, this is equivalent to be p-graphlike.

Let G be a graph. For any $v, v' \in V(G)$, we denote the edge joining v and v' by $e(v, v')$. Moreover, let $V^i(G)$ denote the set of the vertices of G of degree i and let $V^{\geq i}(G) = \cup_{j \geq i} V^j(G)$. We say that an edge of a graph is *pendant* if it is incident to a leaf. A *path* is defined to be a tree with at most 2 leaves. Finally, let T be a tree; for any $v, v' \in V(T)$, we denote the unique path in T joining v and v' by $p(v, v')$.

Definition 2.1.

- A tree C is a *caterpillar* if there is a path S in C such that $V(S) = V^{\geq 2}(C)$. We call S the *spine* of the caterpillar.
- A graph P with n vertices is a *cycle* if we can rename the vertices by $i_1, \dots, i_n$ in such a way that $E(P) = \{e(i_1, i_2), \dots, e(i_{n-1}, i_n), e(i_n, i_1)\}$.
- A graph G is *complete* if, for any $i, j \in V(G)$, we have that $e(i, j) \in E(G)$. The complete graph with n vertices is usually denoted by K_n.

- A graph B is a *bipartite graph* on two subsets X and Y of $V(B)$ if $X \cap Y = \emptyset$, $X \cup Y = V(B)$ and $E(B) \subset \{e(x,y) \,|\, x \in X,\, y \in Y\}$. A bipartite graph B on X and Y is *complete* if, for any $i \in X, j \in Y$, we have that $e(i,j) \in E(B)$. The complete bipartite graph on two sets, one of cardinality m and one of cardinality n, is usually denoted by $K_{m,n}$.

Remark 2.1. Let B be a graph and let $X, Y, P, Q \subset V(B)$ such that B is a bipartite graph on X and Y and is a bipartite graph on P and Q; then we can easily show that $X = P$ and $Y = Q$ or vice versa.

Definition 2.2. Let $\mathcal{G} = (G, w)$ a positive-weighted graph; we say that an edge e of G is *useful* if there exist $i, j \in V(G)$ such that all the paths realizing $D_{i,j}(\mathcal{G})$ contain the edge e. We say that an edge e is *useless* if it is not useful, that is, if every distance of the graph is realized by at least a path which does not contain e. Finally, we say that a graph $\mathcal{G}$ is *pruned* if all its edges are useful.

Definition 2.3. Let D be a predistance $n \times n$ matrix. We say that it is *pathlike* (*cyclelike*, *caterpillarlike*, *planar-graphlike*) if there exists a positive-weighted path (respectively, a positive-weighted cycle, a positive-weighted caterpillar, a positive-weighted planar graph) $\mathcal{G} = (G, w)$ with $V(G) = [n]$ such that $D_{i,j}(\mathcal{G}) = D_{i,j}$ for any $i, j \in [n]$. Analogously, we say that the matrix is *bigraphlike* on two subsets X and Y of $[n]$ if there exists a positive-weighted bipartite graph $\mathcal{B} = (B, w)$ with $V(B) = [n]$ on X, Y such that $D_{i,j}(\mathcal{B}) = D_{i,j}$ for any $i, j \in [n]$.

Finally we say that it is *complete-graphlike* (*complete-bigraphlike*) if there exists a pruned positive-weighted complete graph (respectively, a pruned positive-weighted complete bipartite graph) $\mathcal{G} = (G, w)$ with $V(G) = [n]$ such that $D_{i,j}(\mathcal{G}) = D_{i,j}$ for any $i, j \in [n]$.

To be precise we should say "p-pathlike, p-caterpillarlike...." to point out that we are considering positive-weighted graphs, but, since we will consider only positive-weighted graphs and so no confusion can arise, for simplicity we will omit the letter "p".

Remark 2.2. Let $\mathcal{G} = (G, w)$ be a positive-weighted graph with $V(G) = [n]$ and let $i, j \in [n]$ with $i \neq j$; if $D_{i,j}(\mathcal{G})$ is realized by a path H in G, then, for any distinct $k, t \in V(H)$, the distance $D_{k,t}(\mathcal{G})$ is realized by the path in H with endpoints k and t.

Proof. Suppose, contrary to our claim, that there exist $k, t \in V(H)$ such that any path realizing $D_{k,t}(\mathcal{G})$ is not contained in H. Call J one of the paths realizing $D_{k,t}(\mathcal{G})$ and call H' the path joining k with t contained in H. Then we would have:

$$w(H') > D_{k,t}(\mathcal{G}) = w(J); \tag{2}$$

moreover,

$$w(H) = D_{i,j}(\mathcal{G}) \leq w((H - H') \cup J) \leq w(H - H') + w(J);$$

thus $w(H') \leq w(J)$, which is absurd because it contradicts (2). □

Definition 2.4. Let D be a distance $n \times n$ matrix; we say that an entry $D_{i,j}$ with $i \neq j$ is *indecomposable* if $D_{i,j} < D_{i,z} + D_{z,j}$ for any $z \in [n] - \{i, j\}$.

Remark 2.3. Let $\mathcal{G} = (G, w)$ be a positive-weighted graph such that $V(G) = [n]$. For any $i, j \in [n]$ with $i \neq j$, the distance $D_{i,j}(\mathcal{G})$ is indecomposable if and only if $E(G)$ contains the edge $e(i, j)$ and $e(i, j)$ is useful. In this case we have that $D_{i,j}(\mathcal{G})$ is realized only by the edge $e(i, j)$ and, in particular, $D_{i,j}(\mathcal{G}) = w(e(i, j))$.

Proof. Suppose that $D_{i,j}(\mathcal{G})$ is indecomposable; if it were realized by a path joining i with j different from $e(i, j)$, it would contain another vertex $z \in [n] - \{i, j\}$, then, by Remark 2.2, we would have that $D_{i,j}(\mathcal{G}) = D_{i,z}(\mathcal{G}) + D_{z,j}(\mathcal{G})$, which is absurd. So $D_{i,j}(\mathcal{G})$ can be realized only by $e(i, j)$, thus $e(i, j) \in E(G)$ and $e(i, j)$ useful. Conversely, suppose to have a useful edge $e(i, j) \in E(G)$: by definition, there exist two vertices $a, b \in V(G)$ such that all the paths realizing $D_{a,b}(\mathcal{G})$ contain $e(i, j)$. Then $D_{i,j}(\mathcal{G})$ is realized by the edge $e(i, j)$ (by Remark 2.2) and it can be realized only by the edge $e(i, j)$, so it is indecomposable. □

3 Paths and caterpillars

In this section we give a characterization of pathlike matrices and a characterization of caterpillarlike ones. The first is rather simple and perhaps well-known to experts; we write it here because we need it for the characterization of cyclelike matrices (see Section 4).

Proposition 3.1. *Let D be a predistance $n \times n$ matrix and let $x, y \in [n]$ be such that $x \neq y$ and $D_{x,y} = max_{i,j \in [n],\, i \neq j}\{D_{i,j}\}$; the matrix is pathlike if and only if $D_{i,j} = |D_{i,x} - D_{j,x}|$ for any distinct $i, j \in [n] - \{x\}$.*

Proof. $\Longrightarrow$ Very easy to prove.

$\Longleftarrow$ First note that $D_{a,x} \neq D_{b,x}$ for any distinct $a, b \in [n] - \{x\}$; otherwise we would have:

$$D_{a,b} = |D_{a,x} - D_{b,x}| = 0,$$

which is absurd because, by assumption, the off-diagonal entries are positive. Let us denote the elements of $[n] - \{x, y\}$ by $i_1, i_2, \ldots, i_{n-2}$ in such a way that $D_{i_j,x} < D_{i_{j+1},x}$ for any $j = 1, \ldots, n-3$ and let $\mathcal{S} = (S, w)$ be the positive-weighted path defined as follows (see Figure 1): let S be the path with $V(S) = [n]$ and $E(S) = \{e(x, i_1), e(i_1, i_2), \ldots, e(i_{n-2}, y)\}$ and define the weights of the edges as in Figure 1. It is easy to check that $D_{i,j}(\mathcal{S}) = D_{i,j}$ for any distinct $i, j \in [n]$. □

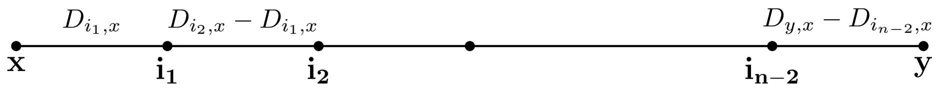

Figure 1: a positive-weighted path realizing the matrix D

Before studying caterpillarlike matrices, we introduce a definition and we state a theorem that will be useful later:

Definition 3.1. Let D be a predistance $n \times n$ matrix. We say that it is a *median matrix* if, for any distinct $a, b, c \in [n]$, there exists a unique element $m \in [n]$ such that

$$D_{i,j} = D_{i,m} + D_{j,m}$$

for any distinct $i, j \in \{a, b, c\}$.

Observe that the entries of a median matrix satisfy the triangle inequalities. The theorem below, probably well-known to experts, was suggested to us by an anonymous referee in October 2014; later we have found it also in [7]; we defer to [3] for a shorter proof. Observe that Theorem 1.2 characterizes predistance $n \times n$-matrices D such that there exists a positive-weighted tree $\mathcal{T} = (T, w)$, with $\underline{[n] \subseteq V(T)}$, such that $D_{i,j}(\mathcal{T}) = D_{i,j}$ for any $i, j \in [n]$, while the following theorem characterizes predistance $n \times n$-matrices D such that there exists a positive-weighted tree $\mathcal{T} = (T, w)$, with $\underline{[n] = V(T)}$, such that $D_{i,j}(\mathcal{T}) = D_{i,j}$ for any $i, j \in [n]$.

Theorem 3.1. *Let D be a predistance $n \times n$ matrix. There exists a positive-weighted tree $\mathcal{T} = (T, w)$, with $V(T) = [n]$, such that $D_{i,j}(\mathcal{T}) = D_{i,j}$ for any $i, j \in [n]$ if and only if the 4-point condition (see Theorem 1.2 for the definition) holds and the matrix D is median.*

Now, consider a positive-weighted caterpillar $\mathcal{C} = (C, w)$ with $V(C) = [n]$.

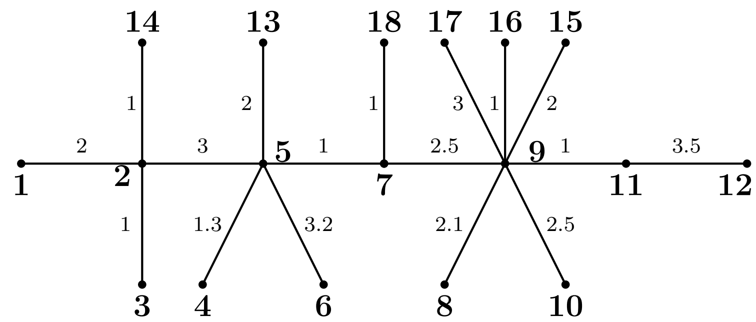

Figure 2: a positive-weighted caterpillar $\mathcal{C} = (C, w)$ with $V(C) = [18]$

Given a vertex $x \in V(C)$, we can define

$$t_x = \frac{1}{2} min_{z,y \in V(C)-\{x\}} \{D_{x,y}(\mathcal{C}) + D_{x,z}(\mathcal{C}) - D_{y,z}(\mathcal{C})\};$$

it is easy to show that, if $x \in L(C)$, then t_x is the weight of the pendant edge associated to x and that $t_x = 0$ if and only if $x \notin L(C)$, that is, x is a vertex of the spine of C.

Remark 3.1. Let $\mathcal{C} = (C, w)$ be a positive-weighted caterpillar with $V(C) = [n]$ and let $x_1, x_2 \in V(C)$ be such that $p(x_1, x_2)$ is the spine of C. Let X_1 (respectively X_2) be the set of the leaves of C adjacent to x_1 (respectively x_2) (for example, in Figure 2, we have that $\{x_1, x_2\} = \{2, 11\}$ and, if we take for instance $x_1 = 2$ and $x_2 = 11$, we have that $X_1 = \{1, 3, 14\}$ and $X_2 = \{12\}$). If we consider two distinct vertices $a, b \in V(C)$ such that

$$D_{a,b}(\mathcal{C}) - t_a - t_b = max_{i,j \in [n],\, i \neq j}\{D_{i,j}(\mathcal{C}) - t_i - t_j\},$$

we have that $a \in X_1 \cup \{x_1\}$ and $b \in X_2 \cup \{x_2\}$ or vice versa.

Proof. It is sufficient to note that, for any distinct $i, j \in V(C)$,

$$D_{i,j}(\mathcal{C}) - t_i - t_j = w(p(\overline{i}, \overline{j})),$$

where $\overline{i}$ is defined as follows: it is equal to i if $i \notin L(C)$, while it is the vertex adjacent to i if $i \in L(C)$; analogously $\overline{j}$. □

Lemma 3.1. *Let $\mathcal{C} = (C, w)$ be a positive-weighted tree with $V(C) = [n]$; call a, b two distinct vertices of C such that*

$$D_{a,b}(\mathcal{C}) - t_a - t_b = max_{i,j \in [n],\, i \neq j}\{D_{i,j}(\mathcal{C}) - t_i - t_j\}.$$

The tree C is a caterpillar if and only if for any distinct $i, j \in [n] - \{a, b\}$ we have that

$$D_{a,b}(\mathcal{C}) + D_{i,j}(\mathcal{C}) \geq max\{D_{a,i}(\mathcal{C}) + D_{b,j}(\mathcal{C}), D_{a,j}(\mathcal{C}) + D_{b,i}(\mathcal{C})\}. \quad (3)$$

Proof. If C is a caterpillar, then, using Remark 3.1 it is easy to check that (3) holds for any distinct $i, j \in [n] - \{a, b\}$. Now, suppose that C is not a caterpillar, then there must be a vertex c with degree grater than 1 which is not in $p(a, b)$.

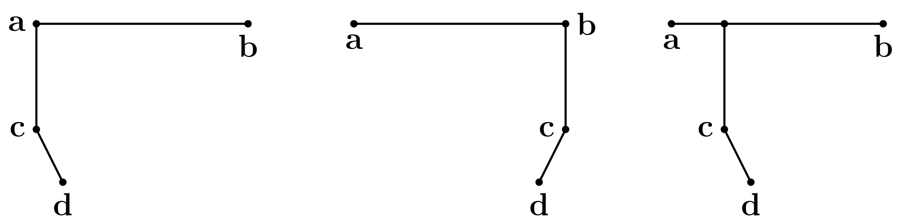

Figure 3: cases (1), (2) and (3)

We have three cases:

1. $p(a, c) \cap p(a, b) = \{a\}$;
2. $p(a, c) \cap p(a, b) = p(a, b)$;
3. $\{a\} \subsetneq p(a, c) \cap p(a, b) \subsetneq p(a, b)$.

Here we study the third case; the other ones are analogous. Call d a vertex such that $p(c,d) \cap p(c,a) = \{c\}$ (see Figure 3), we have that

$$D_{a,b}(\mathcal{C}) + D_{c,d}(\mathcal{C}) < D_{a,c}(\mathcal{C}) + D_{b,d}(\mathcal{C})$$

and

$$D_{a,b}(\mathcal{C}) + D_{c,d}(\mathcal{C}) < D_{a,d}(\mathcal{C}) + D_{b,c}(\mathcal{C}),$$

which is absurd. □

Now we are ready to give a characterization of caterpillarlike predistance matrices:

Theorem 3.2. *Let D be a predistance $n \times n$ matrix. Let a, b be two distinct elements of $[n]$ such that*

$$D_{a,b} - t_a - t_b = max_{i,j \in [n],\, i \neq j}\{D_{i,j} - t_i - t_j\};$$

the matrix is caterpillarlike if and only if the following conditions hold:

- (i) *the entries of the matrix satisfy the 4–point condition;*
- (ii) *the matrix is median;*
- (iii) $D_{a,b} + D_{i,j} \geq max\{D_{a,i} + D_{b,j}, D_{a,j} + D_{b,i}\}$ *for any distinct $i, j \in [n] - \{a, b\}$.*

Proof. It follows immediately from Theorem 3.1 and Lemma 3.1. □

4 Cycles

Let $\mathcal{P} = (P, w)$ be a positive-weighted cycle such that $V(P) = [n]$. Observe that in case $\mathcal{P}$ is not pruned, there is at most one useless edge e. So, if we delete e, we obtain a positive-weighted path $\tilde{\mathcal{P}} = (\tilde{P}, \tilde{w})$ with $V(\tilde{P}) = [n]$ and with the same distance matrix.

Suppose now that $\mathcal{P}$ is pruned: by Remark 2.3, for any $i \in V(P)$, the vertices x and y adjacent to i are exactly the ones such that $D_{i,x}(\mathcal{P})$ and $D_{i,y}(\mathcal{P})$ are indecomposable, so it is possible to recover the order of the vertices of the cycle starting from the 2-weights.

Definition 4.1. Let P be a cycle with $[n]$ as vertex set. We say that the vertex set is sequentially ordered if i and $i+1$ are adjacent for any $i \in [n-1]$ and n and 1 are adjacent.

Definition 4.2. Let D be a distance $n \times n$ matrix such that for any $i \in [n]$, there exist exactly two elements $x, y \in [n] - \{i\}$ for which $D_{i,x}$ and $D_{i,y}$ are indecomposable. We define a subset H of $[n]$ and we rename the elements of $[n]$ by the following algorithm:

rename 1 and 2 two elements of $[n]$ such that $D_{1,2} = min_{i,j \text{ with } i \neq j}\{D_{i,j}\}$ and define $H = \{1, 2\}$; observe that $D_{1,2}$ must be indecomposable; rename 3 the unique element in $[n] - \{1, 2\}$ such that $D_{2,3}$ is indecomposable and put 3 in H; recursively, call $i+1$ the unique element in $[n] - \{i-1, i\}$ such that $D_{i,i+1}$ is indecomposable; if $i+1 \in H$ stop the algorithm, otherwise put $i+1$ in H.

Observe that the algorithm above terminates because at each step we stop the algorithm if $i+1 \in H$, while we continue the algorithm only if $i+1 \notin H$ and in this case we put $i+1$ in H; so, if the algorithm does not stop, H increases, but since $[n]$ is finite, the subset H cannot increase forever.

Theorem 4.1. *Let D be a distance $n \times n$ matrix; there exists a pruned positive-weighted cycle $\mathcal{P} = (P, w)$ with $V(P) = [n]$ realizing the matrix if and only if the following conditions hold:*

(*i*) *for any $i \in [n]$ there are exactly two elements $x, y \in [n]$ such that $D_{i,x}$ and $D_{i,y}$ are indecomposable;*

(*ii*) *if H is the set described in Definition 4.2, then the cardinality of H is n;*

(*iii*) *if the elements of $[n]$ are renamed as in Definition 4.2, then, for any $a, b \in [n]$ with $a < b$, we have that*

$$D_{a,b} = min \left\{ \sum_{i=a}^{b-1} D_{i,i+1}, \sum_{i=b}^{n-1} D_{i,i+1} + D_{1,n} + \sum_{i=1}^{a-1} D_{i,i+1} \right\}. \tag{4}$$

Proof. Suppose that there exists a pruned positive-weighted cycle $\mathcal{P} = (P, w)$ with $V(P) = [n]$ such that $D_{i,j}(\mathcal{P}) = D_{i,j}$ for any $i, j \in [n]$. It is easy to check that conditions (*i*) and (*ii*) hold. Moreover, if we rename the vertices as in Definition 4.2, the vertex set is sequentially ordered. Since $\mathcal{P}$ is pruned, for any $i \in [n-1]$ the 2-weight $D_{i,i+1}(\mathcal{P})$ is realized by $e(i, i+1)$ and the 2-weight $D_{1,n}(\mathcal{P})$ is realized by $e(1, n)$ (see Remark 2.3). Obviously, for any two vertices $a, b \in [n]$, with $a < b$, a subgraph realizing the 2-weight $D_{a,b}(\mathcal{P})$ is a path with endpoints a and b and in the cycle there are exactly two different paths with endpoints a and b. Their weights are the numbers at the second member of (4), so we get condition (*iii*).

On the other hand, let D be a predistance $n \times n$ matrix satisfying conditions (*i*), (*ii*) and (*iii*). By conditions (*i*) and (*ii*) we can rename all the elements of $[n]$ as in Definition 4.2. Let $\mathcal{P} = (P, w)$ be the positive-weighted cycle with $V(P) = [n]$, with the vertex set sequentially ordered, and such that $w(e(i, i+1)) = D_{i,i+1}$ for any $i \in [n-1]$ and $w(e(1, n)) = D_{1,n}$ (see Figure 4).

We have to prove that $D_{a,b}(\mathcal{P}) = D_{a,b}$ for any distinct $a, b \in [n]$ with $a < b$; obviously a subgraph realizing $D_{a,b}(\mathcal{P})$ is a path with endpoints a and b and in the cycle there are exactly two different paths with endpoints a and b. By the definition of $\mathcal{P}$, their weights are the two numbers at the second member of (4), so we have that

$$D_{a,b}(\mathcal{P}) = min \left\{ \sum_{i=a}^{b-1} D_{i,i+1}, \sum_{i=b}^{n-1} D_{i,i+1} + D_{1,n} + \sum_{i=1}^{a-1} D_{i,i+1} \right\} = D_{a,b},$$

where the last equality holds by (4). Observe that $\mathcal{P}$ is pruned, in fact, if an edge $e(a, b)$ (with a and b adjacent vertices) were useless, then $D_{a,b}(\mathcal{P})$ would not be indecomposable, which is absurd because we have constructed $\mathcal{P}$ in such a way that two vertices are adjacent if and only if $D_{a,b}$ is indecomposable. □

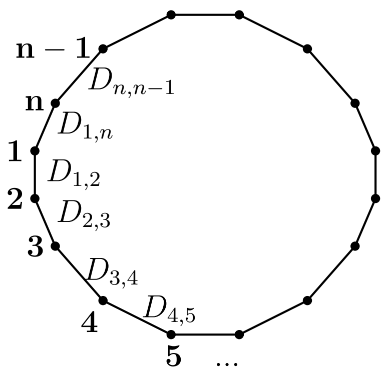

Figure 4: A cycle realizing the matrix

Now we can give a characterization of cyclelike distance matrices.

Theorem 4.2. *A distance $n \times n$ matrix is cyclelike if and only if either it is pathlike or it satisfies conditions (i), (ii) and (iii) of Theorem* 4.1.

Proof. Suppose there exists a positive-weighted cycle $\mathcal{P} = (P, w)$ with $V(P) = [n]$ realizing the matrix; if $\mathcal{P}$ is pruned, then by Theorem 4.1 the matrix must satisfy conditions (i),(ii) and (iii). If $\mathcal{P}$ is not pruned, then we can delete the unique useless edge and we obtain a positive-weighted path realizing the matrix, so the matrix is pathlike.

Conversely, suppose there exists a positive-weighted path $\mathcal{S} = (S, w)$ with $V(S) = [n]$ realizing the matrix. If i, j are the endpoints of the path, we can add to the path an edge $e(i, j)$ with weight any real number greater than or equal to $D_{i,j}$: it is easy to check that the positive-weighted cycle with n vertices we have obtained realizes the matrix D, so the matrix is also cyclelike. Finally, if the matrix satisfies conditions (i),(ii) and (iii) of Theorem 4.1, it is cyclelike by Theorem 4.1. □

5 Complete graphs and bipartite graphs

An immediate consequence of Remark 2.3 is the following characterization of the complete-graphlike distance matrices:

Lemma 5.1. *Let D be a distance $n \times n$ matrix; it is complete-graphlike if and only if $D_{i,j}$ is indecomposable for any distinct $i, j \in [n]$.*

Proof. Suppose there exists a pruned positive-weighted complete graph $\mathcal{G} = (G, w)$ with $V(G) = [n]$ realizing the matrix; thus $e(i, j)$ is useful for any distinct $i, j \in [n]$; so, by Remark 2.3, $D_{i,j}(\mathcal{G})$ is indecomposable for any distinct $i, j \in [n]$. On the other hand, suppose $D_{i,j}$ is indecomposable for any distinct $i, j \in [n]$; by Theorem 1.1, there exists a positive-weighted graph $\mathcal{G} = (G, w)$ with $V(G) = [n]$ realizing the matrix; moreover, since $D_{i,j}$ is indecomposable for any distinct $i, j \in [n]$, then, by Remark 2.3, we have that $e(i, j) \in E(G)$ and $e(i, j)$ is useful for any distinct $i, j \in [n]$. □

Now we want to characterize bigraphlike predistance matrices; first of all, given a positive-weighted bipartite graph $\mathcal{G} = (G, w)$ on $X, Y \subset V(G)$, we show that it is possible to recover X and Y from the 2-weights of $\mathcal{G}$.

Proposition 5.1. *Let $\mathcal{B} = (B, w)$ be a positive-weighted bipartite graph on X and Y with $V(B) = [n]$; let $x \in X$ and $y \in Y$; then:*

$$X = \{x\} \cup \Big\{ i \in [n] - \{x, y\} \mid \exists\, j_1, \dots, j_t \in [n], \text{ with } t \text{ odd, such that } x \neq j_1 \neq j_2 \neq \cdots \neq j_t \neq i,\ D_{x,i}(\mathcal{B}) = D_{x,j_1}(\mathcal{B}) + D_{j_1,j_2}(\mathcal{B}) + \cdots + D_{j_t,i}(\mathcal{B}) \text{ and the elements of the sum are indecomposable} \Big\} \quad (5)$$

and

$$Y = \Big\{ i \in [n] - \{x\} \mid \text{either } D_{x,i} \text{ is indecomposable or } \exists\, j_1, \dots, j_t \in [n] \text{ with } t \text{ even, such that } x \neq j_1 \neq j_2 \neq \cdots \neq j_t \neq i,\ D_{x,i}(\mathcal{B}) = D_{x,j_1}(\mathcal{B}) + D_{j_1,j_2}(\mathcal{B}) + \cdots + D_{j_t,i}(\mathcal{B}) \text{ and the elements of the sum are indecomposable} \Big\}. \quad (6)$$

Proof. Let us prove (5); the other equality can be proved analogously. Call R the second member of (5); we want to prove that $X = R$.

- $X \subset R$: let $i \in X - \{x\}$; observe that $D_{x,i}(\mathcal{B})$ is not indecomposable: otherwise by Remark 2.3, we would have $e(x, i) \in E(B)$, which is absurd; so we can write $D_{x,i}(\mathcal{B})$ as
$$D_{x,j_1}(\mathcal{B}) + D_{j_1,j_2}(\mathcal{B}) + \cdots + D_{j_t,i}(\mathcal{B})$$
for some $j_1, \dots, j_t$ with $D_{x,j_1}(\mathcal{B}), D_{j_1,j_2}(\mathcal{B}), \dots, D_{j_t,i}(\mathcal{B})$ indecomposable. By Remark 2.3, the 2-weights $D_{x,j_1}(\mathcal{B}), D_{j_1,j_2}(\mathcal{B}), \dots, D_{j_t,i}(\mathcal{B})$ are realized respectively by $e(x, j_1), e(j_1, j_2), \dots, e(j_t, i)$; thus the path given by the union of these edges realizes $D_{x,i}(\mathcal{B})$ and, since $x, i \in X$, we have that t is necessarily odd.

- $R \subset X$: if $i \in R$ then there exist $j_1, \dots, j_t \in [n]$ with t odd such that $D_{x,i}(\mathcal{B}) = D_{x,j_1}(\mathcal{B}) + D_{j_1,j_2}(\mathcal{B}) + \cdots + D_{j_t,i}(\mathcal{B})$ and the elements of the sum are indecomposable. By Remark 2.3, the 2-weights $D_{x,j_1}(\mathcal{B}), D_{j_1,j_2}(\mathcal{B}), \dots D_{j_t,i}(\mathcal{B})$ are realized respectively only by the edges $e(x, j_1)$, $e(j_1, j_2)$,...,$e(j_t, i)$, which implies that $D_{x,i}(\mathcal{B})$ is realized by the path given by these edges; so, since t is odd, $i \in X$.

□

Remark 5.1. Let $\mathcal{B} = (B, w)$ be a positive-weighted bipartite graph on X and Y with $V(B) = [n]$. Let $x, y \in [n]$, with $x \neq y$, be such that

$$D_{x,y}(\mathcal{B}) = min_{i,j \in [n],\, i \neq j} \{D_{i,j}(\mathcal{B})\};$$

hence, obviously, $D_{x,y}(\mathcal{B})$ is indecomposable, and then, by Remark 2.2, $e(x, y) \in E(B)$ and $D_{x,y}(\mathcal{B})$ is realized only by the path with unique edge $e(x, y)$.

Now we are ready to give a characterization of bigraphlike predistance matrices.

Theorem 5.1. *Let D be a distance $n \times n$ matrix and let $x, y \in [n]$, with $x \neq y$, be such that $D_{x,y} = min_{i,j\in[n],\, i\neq j}\{D_{i,j}\}$; define*

$$X = \{x\} \cup \Big\{ i \in [n] - \{x,y\} \mid \exists\, j_1, \dots, j_t \in [n],\ \text{with } t \text{ odd, such that } x \neq j_1 \neq j_2 \neq \dots \neq j_t \neq i,\ D_{x,i} = D_{x,j_1} + D_{j_1,j_2} + \dots + D_{j_t,i} \text{ and the elements of the sum are indecomposable} \Big\}$$

and

$$Y = \Big\{ i \in [n] - \{x\} \mid \text{either } D_{x,i} \text{ is indecomposable or } \exists\, j_1, \dots, j_t \in [n],\ \text{with } t \text{ even, such that } x \neq j_1 \neq j_2 \neq \dots \neq j_t \neq i,\ D_{x,i} = D_{x,j_1} + D_{j_1,j_2} + \dots + D_{j_t,i} \text{ and the elements of the sum are indecomposable} \Big\}.$$

The matrix D is bigraphlike if and only if the following conditions hold:

(1) $X \cap Y = \emptyset$

(2) for any distinct $a, b \in X$ (respectively Y), there exists $z \in Y$ (respectively X) such that:

$$D_{a,b} = D_{a,z} + D_{z,b}.$$

Proof. Suppose there exist two subsets of $[n]$, X' and Y', and a positive-weighted bipartite graph $\mathcal{B} = (B, w)$ on X' and Y' with $V(B) = [n]$ realizing the matrix. By Proposition 5.1 we have that $X = X'$ and $Y = Y'$ (or vice versa), so $X \cap Y = \emptyset$. Let $a, b \in X$; a path realizing $D_{a,b}(\mathcal{B})$ must contain a vertex $z \in Y$, so, by Remark 2.2, we have that $D_{a,b}(\mathcal{B}) = D_{a,z}(\mathcal{B}) + D_{z,b}(\mathcal{B})$. If both a and b are elements of Y, the proof is analogous.

Now, suppose that D satisfies (1) and (2). Let $\mathcal{B} = (B, w)$ be the positive-weighted bipartite graph on X and Y such that:

- $V(G) = [n]$;
- $E(G) = \{(a, b) \mid a \in X,\ b \in Y\}$;
- $w(e(a, b)) = D_{a,b}$ for any $a \in X$ and $b \in Y$.

We want to prove that $D_{a,b}(\mathcal{B}) = D_{a,b}$ for any distinct $a, b \in [n]$. Let p be a path realizing $D_{a,b}(\mathcal{B})$ and let $j_1, \dots, j_t \in [n]$ be such that p is given by $e(a, j_1), e(j_1, j_2), \dots, e(j_t, b)$; then we have that

$$D_{a,b}(\mathcal{B}) = w(e(a,j_1)) + w(e(j_1,j_2)) + \dots + w(e(j_t,b)) = D_{a,j_1} + D_{j_1,j_2} + \dots + D_{j_t,b} \geq D_{a,b}, \tag{7}$$

where the last inequality follows from the triangle inequalities.

If $a \in X$ and $b \in Y$ (or vice versa), then

$$D_{a,b}(\mathcal{B}) \leq w(e(a,b)) = D_{a,b}, \tag{8}$$

so, from (7) and (8), we get $D_{a,b}(\mathcal{B}) = D_{a,b}$.

If both a and b are in X (if they are in Y, we can argue analogously), by assumption, there exists $z \in Y$ such that $D_{a,z} + D_{b,z} = D_{a,b}$; we have that

$$D_{a,b}(\mathcal{B}) \leq D_{a,z} + D_{b,z} = D_{a,b}, \tag{9}$$

where the inequality holds because the path given by $e(a, z)$ and $e(z, b)$ contains a and b as vertices and its weight is equal to $D_{a,z} + D_{b,z}$; so, from (7) and (9), we get, also in this case, that $D_{a,b}(\mathcal{B}) = D_{a,b}$. □

Finally we also give a characterization of complete-bigraphlike matrices:

Remark 5.2. Let D be a distance matrix which is bigraphlike on $X, Y \subset [n]$. The matrix is complete-bigraphlike on X and Y if and only if $D_{i,j}$ is indecomposable for any $i \in X, j \in Y$.

Proof. Let $\mathcal{B} = (B, w)$ be a positive-weighted complete bipartite graph on X and Y, with $V(B) = [n]$, realizing the matrix. If it is pruned, then $e(i, j)$ is useful for any $i \in X, j \in Y$; so, by Remark 2.3, $D_{i,j}(\mathcal{B})$ is indecomposable for any $i \in X, j \in Y$. On the other hand, if $D_{i,j}$ is indecomposable for any $i \in X$ and $j \in Y$, then, by Remark 2.3, $e(i, j) \in E(B)$ and $e(i, j)$ is useful for any $i \in X, j \in Y$. □

6 Planar graphs

In this section we characterize planar-graphlike matrices. We start with two definitions and a famous theorem by Kuratowski.

Definition 6.1. Let G be a graph and let $e(u, v)$ be an edge of G. We say that a graph G' is obtained from G by a subdivision of the edge $e(u, v)$ if $V(G')$ is the union of $V(G)$ and a new vertex z and $E(G')$ is $E(G) - \{e(u, v)\} \cup \{e(u, z), e(z, v)\}$. We say that a graph G' is a subdivision of a graph G if it is the graph resulting from the subdivision of some edges in G.

Theorem 6.1. (Kuratowski) *A finite graph is planar if and only if it does not contain a subgraph that is a subdivision of* K_5 *or of* $K_{3,3}$.

Definition 6.2. Let G be a subdivision of K_5. We say that a vertex of G is a *true vertex* if it is a vertex of K_5. We call *verges* of G the paths that are subdivisions of the edges of K_5.

Proposition 6.1. *Let* $\mathcal{G} = (G, w)$ *be a pruned positive-weighted graph with* $V(G) = [n]$. *Let us denote* $D_{i,j}(\mathcal{G})$ *by* $D_{i,j}$ *for any* $i, j \in [n]$.

(i) G *contains a subdivision of* K_5 $\Longleftrightarrow$ *there exists* $Q \in \binom{[n]}{5}$ *such that for any distinct* $a, b \in Q$, *either* $D_{a,b}$ *is indecomposable or there exists a sequence of distinct elements* $(x_1, \ldots, x_r)$ *in* $[n] - Q$ *(depending on* $\{a, b\}$*) such that* $D_{a,x_1}, \ldots, D_{x_r,b}$ *are indecomposable and, if* $\{a, b\} \neq \{a', b'\}$, *the sequence of* $\{a, b\}$ *does not intersect the sequence of* $\{a', b'\}$.

(ii) G contains a subdivision of $K_{3,3}$ $\Longleftrightarrow$ there exist disjoint $A, B \in \binom{[n]}{3}$ such that for any $a \in A$ and $b \in B$, either $D_{a,b}$ is indecomposable or there exists a sequence of distinct elements $(x_1, \ldots, x_r)$ in $[n] - A - B$ (depending on $\{a, b\}$) such that $D_{a,x_1}, \ldots, D_{x_r,b}$ are indecomposable and, if $\{a, b\} \neq \{a', b'\}$, the sequence of $\{a, b\}$ does not intersect the sequence of $\{a', b'\}$.

Proof. Let us prove (i) (the proof of (ii) is analogous).

Sufficiency: By Remark 2.3, if $D_{i,j}$ is indecomposable, then $e(i, j) \in E(G)$. For any $a, b \in Q$, let $c_{a,b}$ be the following path: the path given only by the edge $e(a, b)$ if $D_{a,b}$ is indecomposable, the path given by the edges $e(a, x_1)$, $e(x_1, x_2)$,..., $e(x_r, b)$ if $D_{a,b}$ is not indecomposable and $(x_1, \ldots, x_r)$ is a sequence as in the statement of the proposition.

The union of the paths $c_{a,b}$ for $a, b \in Q$ gives a subgraph that is a subdivision of K_5.

Necessity: Let G' be a subdivision of K_5 in G. Let Q be the set of the true vertices of G'. Since $\mathcal{G}$ is pruned, every edge is useful, in particular, for any $x, y \in [n]$ such that $e(x, y)$ is in a verge of G', we have that $e(x, y)$ is useful, so, by Remark 2.3, the 2-weight $D_{x,y}$ is indecomposable and then we get our statement. □

Theorem 6.2. *Let D be a distance $n \times n$ matrix. It is planar-graphlike if and only if the following conditions hold:*

(a) there does not exist $Q \in \binom{[n]}{5}$ such that, for any distinct $a, b \in Q$, either $D_{a,b}$ is indecomposable or there exists a sequence of distinct elements $(x_1, \ldots, x_r)$ in $[n] - Q$ (depending on $\{a, b\}$) such that $D_{a,x_1}, \ldots, D_{x_r,b}$ are indecomposable and, if $\{a, b\} \neq \{a', b'\}$, the sequence of $\{a, b\}$ does not intersect the sequence of $\{a', b'\}$;

(b) there do not exist disjoint $A, B \in \binom{[n]}{3}$ such that, for any $a \in A$ and $b \in B$, either $D_{a,b}$ is indecomposable or there exists a sequence of distinct elements $(x_1, \ldots, x_r)$ in $[n] - A - B$ (depending on $\{a, b\}$) such that $D_{a,x_1}, \ldots, D_{x_r,b}$ are indecomposable and, if $\{a, b\} \neq \{a', b'\}$, the sequence of $\{a, b\}$ does not intersect the sequence of $\{a', b'\}$.

Proof. Necessity: Let $\mathcal{G}$ be a positive-weighted planar graph realizing the matrix. By eliminating a useless edge, then another one and so on, we get a pruned positive-weighted planar graph realizing the matrix. So we can conclude by Proposition 6.1 and Theorem 6.1.

Sufficiency: Let $\mathcal{G}$ be a positive-weighted graph realizing the matrix; by eliminating a useless edge, then another one and so on, we get a pruned positive-weighted graph realizing the matrix; by conditions (a) and (b) and using Proposition 6.1 and Theorem 6.1, we can conclude that it is planar. □

7 A final remark and open problems

It is well-known that, if a predistance $n \times n$ matrix D is the distance matrix of a positive-weighted tree with vertex set equal to $[n]$, then this tree must be unique, see for intance §2.3 of the book [6]. So, if a $n \times n$ matrix is represented by a positive-weighted path with vertex set equal to $[n]$ or a positive-weighted caterpillar with vertex set equal to $[n]$, this path (respectively caterpillar) must be unique; furthermore it is easy to prove that if a matrix is given by a pruned positive-weighted cycle or a pruned positive-weighted bipartite graph, this must be unique. It would be interesting to see, for instance, if a complete-bigraphlike matrix can be realized also by a positive-weighted graph that is not complete bipartite and analogously for the other kinds of graphs we have considered.

Moreover, in this paper we have considered only few kinds of graphs, but it would be interesting to characterize distance matrices coming from other kinds of graphs and the relative problem of unicity; for instance one could consider the following kinds of graphs:

- hypercube graphs,
- Petersen graphs,
- Kneser graphs,
- regular graphs, or k-regular graphs for some k,
- k-connected graphs for some k.

Finally it would be interesting to study the analogous problem for graphs whose vertex set contains $[n]$ but it is not necessarily equal to $[n]$ and for graphs with not necessarily positive weights, that is to study when a symmetric matrix is the matrix of 2-weights of a weighted bipartite graph or of a weighted cycle and so on.

Acknowledgments.

This work was supported by the National Group for Algebraic and Geometric Structures, and their Applications (GNSAGA-INdAM). The first author was supported by Ente Cassa di Risparmio di Firenze.

We thank the anonymous reviewers for the comments and the suggestions to improve the paper.

References

[1] A. Baldisserri and E. Rubei, On graphlike k-dissimilarity vectors, *Ann. Comb.* 18 (3) (2014), 356–381.

[2] A. Baldisserri and E. Rubei, Families of multiweights and pseudostars, *Adv. Appl. Math.* 77 (2016), 86–100.

[3] A. Baldisserri and E. Rubei, Graphlike families of multiweights, arXiv: 1606.09183.

[4] H-J Bandelt and M. A. Steel, Symmetric matrices representable by weighted trees over a cancellative abelian monoid, *SIAM J. Discrete Math.* 8 (4) (1995), 517–525.

[5] P. Buneman, A note on the metric properties of trees, *J. Combin. Theory Ser. B* 17 (1974), 48–50.

[6] A. Dress, K. T. Huber, J. Koolen, V. Moulton and A. Spillner, *Basic phylogenetic combinatorics*, Cambridge University Press, Cambridge, 2012.

[7] M. Hayamizu and K.Fukumizu, On minimum spanning tree-like metric spaces, *Discrete Appl. Math.* 226 (2017), 51–57.

[8] S. L. Hakimi and A. N. Patrinos, The distance matrix of a graph and its tree realization, *Quart. Appl. Math.* 30 (1972/73), 255–269.

[9] S. L. Hakimi and S. S. Yau, Distance matrix of a graph and its realizability, *Quart. Appl. Math.* 22 (1965), 305–317.

[10] S.Herrmann, K.Huber, V.Moulton and A.Spillner, Recognizing treelike k-dissimilarities, *J. Classification* 29 (3) (2012), 321–340.

[11] B. Iriarte Giraldo, Dissimilarity vectors of trees are contained in the tropical Grassmannian, *Electron. J. Combin.* 17 (1) N6, (2010).

[12] D. Levy, R. Yoshida and L. Pachter, Beyond pairwise distances: neighbor joining with phylogenetic diversity estimates, *Mol. Biol. Evol.* 23 (2006), 491–498.

[13] M. Nei and N. Saitou, The neighbor joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.* 4 (4) (1987), 406–425.

[14] L. Pachter and D. Speyer, Reconstructing trees from subtree weights, *Appl. Math. Lett.* 17 (6) (2004), 615–621.

[15] C. Semple and M. Steel, *Phylogenetics*, Oxford University Press, Oxford, 2003.

[16] J. M. S. Simoes Pereira, A Note on the Tree Realizability of a distance matrix, *J. Combin. Theory* 6 (1969), 303–310.

[17] J. A. Studier and K. J. Keppler, A note on the neighbor-joining algorithm of Saitou and Nei, *Mol. Biol. Evol.* 5 (6) (1988), 729–731.

[18] K. A. Zaretskii, Constructing trees from the set of distances between pendant vertices, *Uspehi Matematiceskih Nauk.* 20 (1965), 90–92.

(Received 10 Oct 2016; revised 18 Nov 2017)