Applied Mathematics and Computation 216 (2010) 2016–2023

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

On properties of cell matrices

Gašper Jaklič [a,*], Jolanda Modic [b]

[a] FMF and IMFM, University of Ljubljana and PINT, University of Primorska, Jadranska 21, 1000 Ljubljana, Slovenia
[b] FMF, University of Ljubljana, Jadranska 21, 1000 Ljubljana, Slovenia

ARTICLE INFO

Keywords:
Cell matrix
Star graph
Determinant
Eigenvalues
Euclidean distance matrix
Circum-Euclidean distance matrix

ABSTRACT

In this paper properties of cell matrices are studied. A determinant of such a matrix is given in a closed form. In the proof a general method for determining a determinant of a symbolic matrix with polynomial entries, based on multivariate polynomial Lagrange interpolation, is outlined. It is shown that a cell matrix of size $n > 1$ has exactly one positive eigenvalue. Using this result it is proven that cell matrices are (Circum-)Euclidean Distance Matrices ((C)EDM), and their generalization, k-cell matrices, are CEDM under certain natural restrictions. A characterization of k-cell matrices is outlined.

1. Introduction

A matrix $M \in \mathbb{R}^{n\times n}$ is an Euclidean Distance Matrix (EDM), if there exist points $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n \in \mathbb{R}^r$ $(r \leqslant n)$, such that $m_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2$ for all $i, j = 1, 2, \dots, n$ [1,2]. These matrices were introduced by Schoenberg in [2,3] and have received a considerable attention. They are used in applications in geodesy, economics, genetics, psychology, biochemistry, engineering, etc., where frequently a question arises, what facts can be deduced given only distance information. Some examples can be found in [4]: kissing-number of sphere packing, trilateration in wireless sensor or cellular telephone network, molecular conformation, convex polyhedron construction, etc.. In bioinformatics, distance matrices are used to represent protein structures in a coordinate-independent manner, and in DNA/RNA sequential alignment, determination of the conformation of biological molecules from information given by nuclear magnetic resonance data, etc.

An EDM matrix M is circum-Euclidean (CEDM) (also spherical) if the points which generate it lie on the surface of some hypersphere [5]. Circum-Euclidean distance matrices are important because every EDM is a limit of CEDMs.

Let $\boldsymbol{a} = (a_i),\ i = 1, 2, \dots, n$, be given numbers and $\boldsymbol{a} > \boldsymbol{0}$, where the inequality is considered componentwise. A cell matrix $D \in \mathbb{R}^{n\times n}$, associated with $\boldsymbol{a}$, is defined as

$$d_{ij} := \begin{cases} a_i + a_j, & i \neq j, \\ 0, & i = j. \end{cases} \tag{1}$$

A cell matrix is a particular distance matrix of a star graph (also called a claw graph), i.e., a graph with outer vertices (leaves), connected only to one inner vertex [6], where only Euclidean distances between leaves are considered, measured through the inner vertex of the graph. Their name is derived from the approximation theory, where a triangulation (or, more generally, a simplicial partition) with only one inner vertex is called a cell (also star). Such matrices are used in graph theory and in biochemistry [7,6,8,9]. Later on the assumption on positivity of $\boldsymbol{a}$ will be loosened with $a_i \geqslant 0$, and thus the inner vertex can be included in the distance information.

If we take k star graphs and connect their inner vertices, we obtain so-called k-star graph. The matrix, whose elements are distances between leaves of a k-star graph, is a k-cell matrix.

* Corresponding author.
E-mail addresses: gasper.jaklic@fmf.uni-lj.si (G. Jaklič), jolanda.modic@gmail.com (J. Modic).

0096-3003/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.amc.2010.03.032

More precisely, let $C \in \mathbb{R}^{n\times n}$ be a k-cell matrix, where $k \leqslant n/2$. Let G be the associated connected k-star graph, consisting of star graphs $S_1, S_2, \ldots, S_k$ and let $d((u,v))$ be the distance of the edge (u,v) in G. Let v_i be a leaf of the graph G and u_ℓ the attached inner vertex. Let us denote $a_i := d(v_i, u_\ell)$ and let $h_{\ell,m} := d(u_\ell, u_m)$ be the distance between inner vertices u_ℓ and u_m of star graphs S_ℓ and S_m, respectively. Then

$$c_{ij} := \begin{cases} 0, & i = j, \\ a_i + a_j, & i \neq j,\ v_i, v_j \text{ belong to the same star graph}, \\ a_i + a_j + h_{\ell,m}, & i \neq j,\ v_i, v_j \text{ belong to distinct star graphs } S_\ell, S_m. \end{cases} \tag{2}$$

In this paper we study properties of cell matrices. Firstly, in Section 2, we establish determinants of principal submatrices of a cell matrix. In the proof a general method for confirming a determinant formula of a symbolic matrix with polynomial entries, based on multivariate polynomial Lagrange interpolation, is outlined. Using this, in Section 3, we show that such a matrix has only one positive eigenvalue, and establish that cell matrices belong to a class of well-known Euclidean distance matrices. Furthermore, they are circum-Euclidean. Also their generalization, k-cell matrices, are CEDMs under some natural assumptions. In Section 4, a characterization of k-cell matrices is presented. The paper is concluded by an example in the last section.

2. Determinant and spectrum of cell matrices

First let us determine determinants of principal submatrices of a cell matrix.

Lemma 1. *Let $D \in \mathbb{R}^{n\times n}$ be a cell matrix, associated with a vector $\boldsymbol{a} > \mathbf{0}$ and let $D^{(i)} := D(1:i, 1:i)$, $i = 1, 2, \ldots, n$, be its principal submatrices. Then*

$$\det D^{(i)} = (-1)^{i-1} 2^{i-2} \left(4(i-1) + \sum_{j=1}^{i} \sum_{\ell=1}^{j-1} \frac{(a_j - a_\ell)^2}{a_j a_\ell} \right) \prod_{k=1}^{i} a_k. \tag{3}$$

From Lemma 1 it can easily be seen that if one of the parameters a_i is zero, the determinant formula simplifies. If at least two of the parameters a_i are zero, a cell matrix is singular.

Corollary 1. *If $a_m = 0$ for some $m \in \{1, 2, \ldots, n\}$ and $a_i > 0, \forall i \neq m$, then*

$$\det D = (-1)^{n-1} 2^{n-2} \sum_{\substack{\ell=1 \\ \ell \neq m}}^{n} a_\ell^2 \prod_{\substack{k=1 \\ k \neq \ell, m}}^{n} a_k.$$

If $a_m = a_j = 0, j \neq m$, then $\det D = 0$.

Proof. Let $\mathbb{P}_i(\mathbb{R}^d)$ denote the space of polynomials of total degree $\leqslant i$ in d variables.

Elements of the matrix $D^{(i)} := D^{(i)}(a_1, a_2, \ldots, a_i)$ are linear polynomials, therefore $\det D^{(i)}$ is a polynomial in i variables $a_1, a_2, \ldots, a_i$ of total degree $\leqslant i$,

$$p_i := p_i(a_1, a_2, \ldots, a_i) := \det D^{(i)} \in \mathbb{P}_i(\mathbb{R}^i).$$

Let us choose $k = \dim \mathbb{P}_i(\mathbb{R}^i) = \binom{2i}{i}$ pairwise distinct points

$$\boldsymbol{a}^{(j)} := \left(a_1^{(j)}, a_2^{(j)}, \ldots, a_i^{(j)}\right) \in \mathbb{Z}^i, \quad 1 \leqslant j \leqslant k,$$

in such a way, that they do not lie on an algebraic hypersurface of degree $\leqslant i$. Thus the multivariate Lagrange polynomial interpolation problem is unisolvent [10]. Integer components are needed only to ensure exact computation later on. Now let us evaluate determinants of matrices $D^{(i)}$ at chosen points, $z_j = \det D^{(i)}(a_1^{(j)}, \ldots, a_i^{(j)})$ for $1 \leqslant j \leqslant k$. Let $q := q(a_1, a_2, \ldots, a_i) \in \mathbb{P}_i(\mathbb{R}^i)$ denote the polynomial on the right-hand side of (3). Now compute the values $w_j := q(a_1^{(j)}, a_2^{(j)}, \ldots, a_i^{(j)})$. If $w_j = z_j$ for all $j = 1, 2, \ldots, k$, the polynomials p_i and q have the same values at $k = \dim \mathbb{P}_i(\mathbb{R}^i)$ points and as there is precisely one Lagrange interpolation polynomial in $\mathbb{P}_i(\mathbb{R}^i)$ through prescribed data $(\boldsymbol{a}^{(j)}, z_j)$, $p_i \equiv q$. This concludes the proof. □

Note that in [11] a similar result has been proven, but is unfortunately inappropriate for our purposes. The presented approach can be efficiently applied in general for proving that a given polynomial expression is the determinant of a symbolic matrix with polynomial entries. Using this approach, the hard part is to somehow obtain the closed form of the determinant, later the proof is quite straightforward, since a powerful tool of approximation theory is applied. Further examples can be found in [12] and [13], where the dimension of a bivariate spline space was studied. An excellent overview of similar methods for determinant calculation is [14].

Since cell matrices $D \in \mathbb{R}^{n\times n}$ are symmetric, their eigenvalues λ_i are real. They have a zero diagonal, hence the sum of their eigenvalues is zero. The following theorem shows that they have exactly one positive eigenvalue, the rest of eigenvalues are negative. Thus cell matrices are nonsingular if $\boldsymbol{a} > \mathbf{0}$.