On the Nonnegative Rank of Euclidean Distance Matrices

Matthew M. Lin[1]

Department of Mathematics, North Carolina State University,
Raleigh, NC 27695-8205, USA

Moody T. Chu[*,2]

Department of Mathematics, North Carolina State University,
Raleigh, NC 27695-8205, USA

Abstract

The Euclidean distance matrix for n distinct points in $\mathbb{R}^r$ is generically of rank $r+2$. It is shown in this paper via a geometric argument that its nonnegative rank for the case $r=1$ is generically n.

Key words: Euclidean distance matrix, nonnegative rank factorization, nonnegative rank

1. Introduction

Any given nonnegative matrix $A \in \mathbb{R}^{m\times n}$ can be expressed as the product $A = UV$ for some nonnegative matrices $U \in \mathbb{R}^{m\times k}$ and $V \in \mathbb{R}^{k\times n}$ with $k \le \min\{m,n\}$. The smallest k that makes this factorization possible is called the nonnegative rank of A. For convenience, we denote the nonnegative rank of A by $\operatorname{rank}_+(A)$. Trivially the nonnegative rank has bounds such as

$$\operatorname{rank}(A) \le \operatorname{rank}_+(A) \le \min\{m,n\}. \tag{1}$$

Determining the exact nonnegative rank and computing the corresponding factorization, however, are known to be NP-hard [6, 18]. If the nonnegative matrix A is such that $\operatorname{rank}_+(A) = \operatorname{rank}(A)$, then we say that A has a nonnegative rank factorization (NRF). Even in this case, there is no known effective algorithm to compute the NRF.

It is shown recently that, if $k < \min\{m,n\}$, then the probability that a matrix A with $\operatorname{rank}_+(A) = k$ should also have $\operatorname{rank}(A) = k$ is one. In other words, matrices which have an NRF are generic. To put it more plainly, if $A = UV$ where $U \in \mathbb{R}^{m\times k}$ and $V \in \mathbb{R}^{k\times n}$ are randomly generated nonnegative matrices, then with probability one we have $\operatorname{rank}(A) = k$. The converse is nevertheless not true. Indeed, the question of comoputing the probability for a 4×4 nonnegative matrix of rank 3 to have nonnegative rank 3 is not trivial at all. It is very much analogous to the Sylvester's four-point problem which, to this date, does not admit a determinate solution [14, 16]. For this reason, there has been considerable interest in the literature to identify nonnegative matrices with or without NRF.

[*]Corresponding author
Email addresses: mlin@ncsu.edu (Matthew M. Lin), chu@math.ncsu.edu (Moody T. Chu)
[1]This research was supported in part by the National Science Foundation under grants DMS-0505880 and DMS-0732299.
[2]This research was supported in part by the National Science Foundation under grants DMS-0505880 and DMS-0732299 and NIH Roadmap for Medical Research grant 1 P20 HG003900-01.

A necessary and sufficient condition qualifying whether a nonnegative matrix has an NRF can be found in [17], but that result appears too theoretical for numerical verification. A few sufficient conditions for constructing nonnegative matrices without NRF have been given in [13, 15]. The simplest example is the 4×4 matrix

$$\mathscr{C} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

with $\operatorname{rank}(\mathscr{C}) = 3$ and $\operatorname{rank}_+(\mathscr{C}) = 4$. Other known conditions for the existence of an NRF are for more restrictive subclasses of matrices such as the so called weakly monotone nonnegative matrices [12], λ-monotone [11], or matrices with nonnegative 1-inverse [4]. Still, given a nonnegative matrix, finding its (numerical) rank is computationally possible, but ensuring its nonnegative rank is an extremely hard task. Thus far, we know very little in the literature about nonnegative matrices which do not have NRF. This factorization has also be studied in the literature under the notion of prime matrices [3, 15].

The purpose of this short communication is to add the important class of Euclidean distance matrices to the list of matrices having no NRF. This note represents perhaps only a modest advance in the field, but it should be of interest to confirm the precise rank and nonnegative rank of a distance matrix.

2. Rank condition and standard form

Given n points $\mathbf{p}_1, \ldots, \mathbf{p}_n$ in the space $\mathbb{R}^r$, the corresponding *Euclidean distance matrix* (EDM) is the $n \times n$ symmetric and nonnegative matrix $Q(\mathbf{p}_1, \ldots, \mathbf{p}_n) = [q_{ij}]$ whose entry q_{ij} is defined by

$$q_{ij} = \|\mathbf{p}_i - \mathbf{p}_j\|^2, \quad i, j = 1, \ldots, n, \tag{2}$$

where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^r$. As an exhaustive record of relative spacing between any two of the n particles in $\mathbb{R}^m$, the distance matrix $Q(\mathbf{p}_1, \ldots, \mathbf{p}_n)$ has many important applications in distance geometry. See, for example, the discussions in [7, 8, 9, 10]. Our attention here is solely on the rank condition of $Q(\mathbf{p}_1, \ldots, \mathbf{p}_n)$.

Theorem 2.1. *For any $n \geq r + 2$, the rank of $Q(\mathbf{p}_1, \ldots, \mathbf{p}_n)$ is no greater than $r + 2$ and is generically $r + 2$.*

Proof. (This is a classical and well known fact. There are many elegant ways to verify this result, but for the sake of comparing the associated factorizations we find the following equality representation is most constructive and straightforward.) Regarding each $\mathbf{p}_\ell$ as a column vector and $q_{ij} = \langle \mathbf{p}_i - \mathbf{p}_j, \mathbf{p}_i - \mathbf{p}_j \rangle$ with $\langle \cdot, \cdot \rangle$ denoting the Euclidean inner product, we can write [1]

$$Q(\mathbf{p}_1, \ldots, \mathbf{p}_n) = \underbrace{\begin{bmatrix} \langle \mathbf{p}_1, \mathbf{p}_1 \rangle & 1 & -2\mathbf{p}_1^\top \\ \vdots & \vdots & \vdots \\ \langle \mathbf{p}_i, \mathbf{p}_i \rangle & 1 & -2\mathbf{p}_i^\top \\ \vdots & \vdots & \vdots \\ \langle \mathbf{p}_n, \mathbf{p}_n \rangle & 1 & -2\mathbf{p}_n^\top \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 1 & \ldots & 1 & \ldots & 1 \\ \langle \mathbf{p}_1, \mathbf{p}_1 \rangle & \ldots & \langle \mathbf{p}_j, \mathbf{p}_j \rangle & \ldots & \langle \mathbf{p}_n, \mathbf{p}_n \rangle \\ \mathbf{p}_1 & & \mathbf{p}_j & & \mathbf{p}_n \end{bmatrix}}_{V}. \tag{3}$$

Note that $U \in \mathbb{R}^{n \times (r+2)}$ and $V \in \mathbb{R}^{(r+2) \times n}$. Unless the points $\mathbf{p}_1, \ldots, \mathbf{p}_n$ satisfy some specific algebraic equations, such as $\|\mathbf{p}_\ell\| = 1$ for all $\ell = 1, \ldots n$, the matrices U and V are generically of rank $r + 2$.

The fact that the rank of an EDM depends on r, but is independent of the size n, is very interesting. The rank deficiency indicates that many entries in the matrix provide redundant information. It is curious to know whether $\operatorname{rank}_+(Q(\mathbf{p}_1,\ldots,\mathbf{p}_n))$ has similar property. Note that the two factors U and V in (3) cannot be both nonnegative, so the (minimum) nonnegative factorization of $Q(\mathbf{p}_1,\ldots,\mathbf{p}_n)$ is yet to be determined.

In a recent paper [1], it is estimated via an intriguing algebraic argument that for a nonnegative matrix of rank 3 to have nonnegative rank 10, we would need a matrix of order at least 252. The discussion in the sequel clearly indicates that the actual order can be much lower.

Suppose that a nonnegative matrix A has two factorizations, $A = BC$ and $A = FG$. We say that these two factorizations are equivalent if there exist a permutation matrix P and a diagonal matrix D with positive diagonal elements such that $BDP = F$ and $P^\top D^{-1}C = G$ [1]. With this notion in mind, it suffices to consider the nonnegative factorization for an EDM in a special form.

Lemma 2.1. *Suppose $n \geq r+2 \geq 3$. Then any nonnegative factorization of $Q(\mathbf{p}_1,\ldots,\mathbf{p}_n)$ is equivalent to the form*

$$Q(\mathbf{p}_1,\ldots,\mathbf{p}_n) = \begin{bmatrix} 1 & 0 & * & * & \ldots \\ * & 1 & 0 & * & \ldots \\ 0 & * & 1 & * & \ldots \\ * & * & * & * & \\ \vdots & & & & \end{bmatrix} \begin{bmatrix} 0 & * & + & * & \ldots \\ + & 0 & * & * & \\ * & + & 0 & * & \\ * & * & * & * & \\ \vdots & & & & \end{bmatrix} \tag{4}$$

where $*$ stands for some undetermined nonnegative numbers and $+$ stands for three undetermined positive numbers.

Proof. Suppose $Q(\mathbf{p}_1,\ldots,\mathbf{p}_n) = UV$ is a nonnegative factorization. Then there must exist an index $1 \leq k_1 \leq n$ such that $u_{1k_1}v_{k_13} > 0$. Permuting both the first and the k_1th columns of U and the first and the k_1th rows of V simultaneously will not affect the product and will place u_{1k_1} at the $(1,1)$ position and v_{k_13} at the $(1,3)$ position. After scaling u_{1k_1} to unit, rename without causing ambiguity the permuted matrices as U and V, respectively. The corresponding v_{11} in the new V must be zero. Consequently, there must exist an index $2 \leq k_2 \leq n$ such that $u_{2k_2}v_{k_21} > 0$. Permuting the second and the k_2th columns of U and the second and the k_2th rows of V simultaneously will not affect the product, will not alter the first column of U or the first row of V, and will place u_{2k_2} at the $(2,2)$ position and v_{k_21} at the $(2,1)$ position. Again, after scaling u_{2k_2} to unit and renaming the permuted matrices as U and V, it must be $u_{31} = v_{22} = 0$. It follows that there exist an index $3 \leq k_3 \leq n$ such that $u_{3k_3}v_{k_32} > 0$. Permuting the third and the k_3th columns and rows and scaling u_{3k_3} to unit will give rise to the structure specified in the lemma.

It is important to note that the procedure described in the above proof cannot be continued to the fourth or other rows or columns. For this reason, we refer to (4) as the *standard* nonnegative factorization of $Q(\mathbf{p}_1,\ldots,\mathbf{p}_n)$.

When reference to the points $\mathbf{p}_1,\ldots,\mathbf{p}_n$ is not critical, we abbreviate a generic $Q(\mathbf{p}_1,\ldots,\mathbf{p}_n)$ as Q_n. The notion of nonnegative rank has an interesting geometric meaning [5] which will be our main toll for verifying the nonnegative rank of Q_n. Let the columns of a general nonnegative matrix $A \in \mathbb{R}_+^{m\times n}$ be denoted by $A = [\mathbf{a}_1,\ldots,\mathbf{a}_n]$. Define the *scaling factor* $\sigma(A)$ by

$$\sigma(A) := \operatorname{diag}\{\|\mathbf{a}_1\|_1,\ldots,\|\mathbf{a}_n\|_1\}, \tag{5}$$

where $\|\cdot\|_1$ stands for the 1-norm in $\mathbb{R}^m$, and the *pullback map* $\vartheta(A)$ by

$$\vartheta(A) := A\sigma(A)^{-1}. \tag{6}$$

Each column of $\vartheta(A)$ can be regarded as a point on the $(m-1)$-dimensional *probability simplex* $\mathcal{D}_m$ defined by

$$\mathcal{D}_m := \left\{ \mathbf{x} \in \mathbb{R}_+^m \,|\, x_i \geq 0,\ \sum_{i=1}^{m} x_i = 1 \right\}. \tag{7}$$

Suppose a given nonnegative matrix A can be factorized as $A = UV$, where $U \in \mathbb{R}_+^{m\times p}$ and $V \in \mathbb{R}_+^{p\times n}$. Because $UV = (UD)(D^{-1}V)$ for any invertible nonnegative matrix $D \in \mathbb{R}^{p\times p}$, we may assume without loss of generality that U is already a pullback so that $\sigma(U) = I_n$. We can write

$$A = \vartheta(A)\sigma(A) = UV = \vartheta(U)\vartheta(V)\sigma(V). \tag{8}$$

Note that the product $\vartheta(U)\vartheta(V)$ itself is on the simplex $\mathcal{D}_m$. It follows that

$$\vartheta(A) = \vartheta(U)\vartheta(V), \tag{9}$$
$$\sigma(A) = \sigma(V). \tag{10}$$

In particular, if $p = \text{rank}_+(A)$, then we see that $\text{rank}_+(\vartheta(A)) = p$, and vice versa. The expression (9) means that the columns in the pullback $\vartheta(A)$ are convex combinations of columns of $\vartheta(U)$. The integer $\text{rank}_+(A)$ stands for the minimal number of vertices on $\mathcal{D}_m$ so that the resulting convex polytope encloses all columns of the pullback $\vartheta(A)$.

3. Nonnegative rank and factorization for linear EDM

Give a permutation σ of the set $\{1, 2, \ldots, n\}$, define the permutation matrix $P_\sigma := [\delta_{i\sigma(j)}]$ where δ_{st} denotes the Kronecker delta function. Then it is easy to see that

$$P_\sigma^\top Q(\mathbf{p}_1, \ldots, \mathbf{p}_n) P_\sigma = Q(\mathbf{p}_{\sigma(1)}, \ldots \mathbf{p}_{\sigma(n)}). \tag{11}$$

In other words, the conjugation of an EDM by any permutation matrix remains to be an EDM. In the one dimensional case, i.e., $r = 1$, we may assume without loss of generality that the point are arranged is ascending order, $\mathbf{p}_1 < \ldots < \mathbf{p}_n$. Define $s_i := \mathbf{p}_{i+1} - \mathbf{p}_i$, $i = 1, \ldots, n-1$. Entries in the linear EDM has a special ordering pattern that radiates away from the diagonal per column and row, i.e.

$$Q(\mathbf{p}_1, \ldots, \mathbf{p}_n) = \begin{bmatrix} 0 & s_1^2 & (s_1+s_2)^2 & (s_1+s_2+s_3)^2 & \ldots \\ s_1^2 & 0 & s_2^2 & (s_2+s_3)^2 & \ldots \\ (s_1+s_2)^2 & s_2^2 & 0 & s_3^2 & \ldots \\ (s_1+s_2+s_3)^2 & (s_2+s_3)^2 & s_3^2 & 0 & \\ \vdots & & & & \end{bmatrix} \tag{12}$$

We shall exploit this particular ordering to help to obtain some initial insight into the nonnegative rank of the EDM. Unless mentioned otherwise, the subsequent discussion is for the case $r = 1$.

It is illuminating to begin the analysis with the case $n = 4$. For convenience, we adopt the colon notation as in Matlab to pick out selected rows, columns or elements of vectors. Denote the columns of Q_4 by $Q_4 = [\mathbf{q}_1, \ldots, \mathbf{q}_4]$. The probability simplex $\mathcal{D}_4$ can easily be visualized via the unit tetrahedron S_3 in the first octant of $\mathbb{R}^3$ if we identify the 4-dimensional vector $\mathbf{x}$ by the vector $[x_1, x_2, x_3]^\top$ of its first three entries. In this way, columns of $\vartheta(Q_4)$ can be interpreted as points $\vartheta(\mathbf{q}_1), \vartheta(\mathbf{q}_2), \vartheta(\mathbf{q}_3), \vartheta(\mathbf{q}_4)$ depicted in Figure 1. Note that the four points $\vartheta(\mathbf{q}_1), \vartheta(\mathbf{q}_2), \vartheta(\mathbf{q}_3), \vartheta(\mathbf{q}_4)$ are coplanar because $\text{rank}(Q_4) = 3$. The y-intercept of this common plane is $\frac{s_1 s_2}{s_1 s_2 - s_3(s_1+s_2+s_3)}$ which is either negative or positive with value greater than 1. In either case, the plane intersects the tetrahedron as a quadrilateral. The first three points sit on three separate "ridges" of the quadrilateral and hence cannot be enclosed by any triangle within the quadrilateral except the one

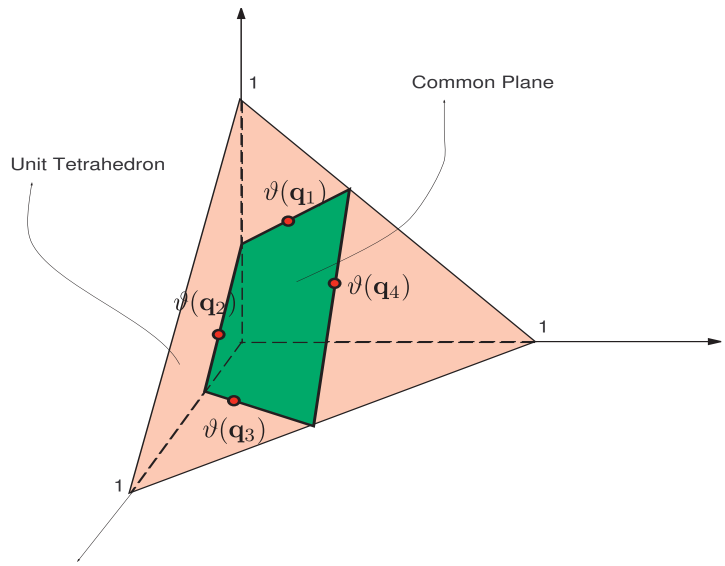

Figure 1: A geometric representation of the matrix $\vartheta(Q_4)$ when $r = 1$.

with vertices at these three points. If $\text{rank}_+(Q_4) < 4$, then $\vartheta(\mathbf{q}_4)$ must be inside this triangle and hence be a convex combination of $\vartheta(\mathbf{q}_1), \vartheta(\mathbf{q}_2), \vartheta(\mathbf{q}_3)$, which translates to that the vector $\mathbf{q}_4$ must be a nonnegative combination of $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$, but this impossible because $q_{44} = 0$. Thus $\text{rank}_+(Q_4) = 4$.

The expression $Q_4 = UV$ in the form (4) represents a polynomial system of 22 equations in 23 unknowns whereas one of the nonzero unknowns can be normalized to unit. This nonlinear system is solvable. Other than the trivial factorization $Q_4 = I_4 Q_4$ where I_4 stands for the identity matrix, we find that there are only three nontrivial nonnegative factorizations which we list in Table 1. While the first set of factorization in the table is equivalent to $Q_4 I_4$, it is important to note that the last two sets of factorizations correspond to the four vertices of the quadrilateral shown in figure 1. This observation also shows that Q_4 is not prime [2, 15].

When $n > 4$, such a visualization in geometry is not possible, but the idea remains justifiable via an algebraic argument with which we precede as follows.

Theorem 3.1. *Suppose that the linear EDM Q_n is of rank 3. Then* $\text{rank}_+(Q_n) = n$.

Proof. Because $\text{rank}(Q_n) = 3$, its columns reside on a 3-dimensional subspace of $\mathbb{R}^n$. The pull-back map ϑ can be considered as the intersection of this subspace and the hyperplane defined by $\sum_{i=1}^n x_i = 1$. Columns of $\vartheta(Q_n)$ therefore are "coplanar" whereas by their common plane we refer to a 2-dimensional affine subspace in $\mathbb{R}^n$. Identifying any n-dimensional vector $\mathbf{x} \in \mathcal{D}_n$ by its first $n-1$ entries $[x_1, \ldots, x_{n-1}]^\top$, we thus are able to "see" columns $\vartheta(\mathbf{q}_1), \ldots, \vartheta(\mathbf{q}_n)$ as n points residing within the unit polyhedron $\mathcal{S}_{n-1}$ in the first orthotant of $\mathbb{R}^{n-1}$. These points remain to be coplanar. (Indeed, the 2-dimensional affine subspace can be identified by a fixed point, say, $\vartheta(\mathbf{q}_1)$, and two coordinate axes, say, $\mathbf{v}_1 := \vartheta(\mathbf{q}_2) - \vartheta(\mathbf{q}_1)$ and $\mathbf{v}_2 := \vartheta(\mathbf{q}_3) - \vartheta(\mathbf{q}_n)$, where all points in the 2-dimensional affine subspace can be represented as $\vartheta(\mathbf{q}_1) + \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$ with scalars α_1 and α_2. The drawing in Figure 1, therefore, is still relatively instructive.)

For $1 \leq i \leq n-1$, it is clear that $\vartheta(\mathbf{q}_i)$ cannot possibly be a convex combination of any other $\vartheta(\mathbf{q}_j)$ because of the unique zero at its ith entry. We claim further that $\vartheta(\mathbf{q}_n)$ cannot possibly be in the convex hull spanned by $\vartheta(\mathbf{q}_1), \ldots, \vartheta(\mathbf{q}_{n-1})$. Assume otherwise, then we would have

$$\vartheta(\mathbf{q}_n) = \sum_{i=1}^{n-1} c_i \vartheta(\mathbf{q}_i)$$

Table 1: Standard nonnegative factorizations of Q_4.

U	V
$\begin{bmatrix} 1 & 0 & \frac{{s_1}^2}{s_2^2} & (s_1+s_2+s_3)^2 \\ \frac{{s_2}^2}{(s_1+s_2)^2} & 1 & 0 & (s_2+s_3)^2 \\ 0 & \frac{(s_1+s_2)^2}{{s_1}^2} & 1 & {s_3}^2 \\ \frac{{s_3}^2}{(s_1+s_2)^2} & \frac{(s_1+s_2+s_3)^2}{{s_1}^2} & \frac{(s_2+s_3)^2}{{s_2}^2} & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & (s_1+s_2)^2 & 0 \\ {s_1}^2 & 0 & 0 & 0 \\ 0 & {s_2}^2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0 & \frac{(s_1+s_2)s_2(s_1+s_2+s_3)}{s_2+s_3} \\ 0 & 1 & 0 & {s_2}^2 \\ 0 & \frac{s_3(s_1+s_2)}{(s_2+s_3)s_1} & 1 & 0 \\ \frac{s_3(s_2+s_3)}{(s_1+s_2)s_1} & 0 & \frac{(s_2+s_3)(s_1+s_2+s_3)}{(s_1+s_2)s_2} & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & {s_1}^2 & \frac{s_3s_1(s_1+s_2)}{s_2+s_3} & 0 \\ {s_1}^2 & 0 & 0 & \frac{s_3(s_2+s_3)s_1}{s_1+s_2} \\ \frac{(s_1+s_2)s_2(s_1+s_2+s_3)}{s_2+s_3} & {s_2}^2 & 0 & 0 \\ 0 & 0 & 1 & \frac{(s_2+s_3)(s_1+s_2+s_3)}{(s_1+s_2)s_2} \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0 & {s_1}^2 \\ \frac{s_2(s_2+s_3)}{(s_1+s_2)(s_1+s_2+s_3)} & 1 & 0 & 0 \\ 0 & \frac{s_3(s_1+s_2)}{(s_2+s_3)s_1} & 1 & 0 \\ 0 & 0 & \frac{(s_2+s_3)(s_1+s_2+s_3)}{(s_1+s_2)s_2} & \frac{s_3(s_2+s_3)s_1}{s_1+s_2} \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & \frac{s_2(s_1+s_2)(s_1+s_2+s_3)}{s_2+s_3} & (s_1+s_2+s_3)^2 \\ {s_1}^2 & 0 & 0 & \frac{s_3(s_2+s_3)s_1}{s_1+s_2} \\ \frac{s_2(s_1+s_2)(s_1+s_2+s_3)}{s_2+s_3} & {s_2}^2 & 0 & 0 \\ 0 & 1 & \frac{s_3(s_1+s_2)}{(s_2+s_3)s_1} & 0 \end{bmatrix}$

for some $c_i \geq 0$ with $\sum_{i=1}^{n-1} c_i = 1$. Note that $\|\vartheta(\mathbf{q}_n)\|_1 = 1$. However, $\|\sum_{i=1}^{n-1} c_i\vartheta(\mathbf{q}_i)\|_1 < 1$ because $\|\vartheta(\mathbf{q}_i)\|_1 < 1$ after chopping away the last row of $\vartheta(Q_n)$. This is a contradiction. The smallest number of vertices for a convex hull to enclose $\vartheta(\mathbf{q}_1), \ldots, \vartheta(\mathbf{q}_n)$, therefore, has to be n, implying that $\text{rank}_+(Q_n) = n$.

There is a subtle difference between the standard nonnegative factorization of Q_4 and that of Q_n when $n \geq 5$. Except for the trivial factorization, both factors U and V in Table 1 for Q_4 are of rank 3. This is not the case in general.

Lemma 3.1. *Suppose $n \geq 5$ and $Q_n = UV$ is a standard nonnegative factorization for the matrix Q_n. Then it cannot be such that both U and V are of rank 3 simultaneously.*

Proof. Observe first that for $n \geq 3$, assuming the generic condition $\text{rank}(Q_n) = 3$, we can partition Q_n as

$$Q_n = \left[\begin{array}{c|c} Q_3 & Q_3\Phi \\ \hline \Phi^\top Q_3 & \Phi^\top Q_3\Phi \end{array}\right] \tag{13}$$

where $\Phi \in \mathbb{R}^{3\times(n-3)}$ is uniquely determined. Indeed, if we write $\Phi = [\boldsymbol{\phi}_4, \ldots \boldsymbol{\phi}_n]$, then it can be shown that

$$\boldsymbol{\phi}_j = \begin{bmatrix} \frac{(\sum_{\ell=2}^{j-1} s_\ell)(\sum_{\ell=3}^{j-1} s_\ell)}{s_1(s_1+s_2)} \\ -\frac{(\sum_{\ell=1}^{j-1} s_\ell)(\sum_{\ell=3}^{j-1} s_\ell)}{s_1 s_2} \\ \frac{(\sum_{\ell=1}^{j-1} s_\ell)(\sum_{\ell=2}^{j-1} s_\ell)}{s_2(s_1+s_2)} \end{bmatrix}, \quad j = 4, \ldots, n. \tag{14}$$

Note that the second entry in $\boldsymbol{\phi}_j$ is always negative.

Assume by contradiction that both U and V of Q_n are of rank 3. As U and V appear in the standard form (4), their 3×3 leading principal submatrices U_{11} and V_{11} are nonsingular. Thus similar to (13), we can partition the nonnegative factors into blocks

$$Q_n = \left[\begin{array}{c|c} U_{11} & U_{11}\Theta \\ \hline \Lambda^\top U_{11} & \Lambda^\top U_{11}\Theta \end{array}\right] \left[\begin{array}{c|c} V_{11} & V_{11}\Gamma \\ \hline \Delta^\top V_{11} & \Delta^\top V_{11}\Gamma \end{array}\right], \tag{15}$$

where Θ, Λ, Γ and Δ are real matrices of compatible sizes. Upon comparison with (13), we see that $\Lambda = \Phi = \Gamma$. Taking a closer look at the product $\Lambda^\top U_{11}$, we find that the signs of its entries are given by

$$\Lambda^\top U_{11} = \begin{bmatrix} + & - & + \\ \vdots & \vdots & \vdots \\ & & \\ + & - & + \end{bmatrix} \begin{bmatrix} 1 & 0 & * \\ * & 1 & 0 \\ 0 & * & 1 \end{bmatrix} = \begin{bmatrix} * & \boxdot & + \\ \vdots & \vdots & \vdots \\ & & \\ * & \boxdot & + \end{bmatrix},$$

where, again, $*$ indicates some undetermined nonnegative numbers, $+$ some undetermined positive numbers, and $\boxdot$ some nonnegative numbers which can further be determined. Similarly, the signs for entries of $V_{11}\Gamma$ are given by

$$V_{11}\Gamma = \begin{bmatrix} 0 & * & + \\ + & 0 & * \\ * & + & 0 \end{bmatrix} \begin{bmatrix} + & \ldots & + \\ - & \ldots & - \\ + & \ldots & + \end{bmatrix} = \begin{bmatrix} * & \ldots & * \\ + & \ldots & + \\ \boxdot & \ldots & \boxdot \end{bmatrix}.$$

Being nonnegative, U and V are complementary to each other in the sense that $u_{ij}v_{ji} = 0$ for all indices i and j. It follows that the $+$'s in the middle row of $V_{11}\Gamma$ must cause the $\boxdot$'s in the middle column of $\Lambda^\top U_{11}$ to become zeros. This implies that the very same u_{32} would have to satisfy the equalities

$$-\frac{\left(\sum_{\ell=1}^{j-1} s_\ell\right)\left(\sum_{\ell=3}^{j-1} s_\ell\right)}{s_1 s_2} + \frac{\left(\sum_{\ell=1}^{j-1} s_\ell\right)\left(\sum_{\ell=2}^{j-1} s_\ell\right)}{s_2(s_1+s_2)} u_{32} = 0,$$

for all $j = 4, \dots n$, which is not possible if $n \geq 5$.

To compute the nonnegative factorization of Q_n for $n \geq 5$ is considerably harder. The case $n = 5$, for example, involves a polynomial system of 39 nonlinear equations in 41 unknowns two of which can be normalized. A short cut from a geometric point of view might be worth mentioning. Write

$$Q_5 = \left[\begin{array}{c|c} Q_4 & \mathbf{q}_5 \\ \hline \mathbf{q}_5^\top & 0 \end{array}\right], \tag{16}$$

with $\mathbf{q}_5 \in \mathbb{R}^{4\times 1}$. Consider the submatrix $[Q_4, \mathbf{q}_5]$ only. Clearly, its columns are coplanar and, hence, $\vartheta(\mathbf{q}_5)$ is a point in the interior of the quadrilateral drawn in Figure 1. In particular, if $Q_4 = UV$ is one of the two nontrivial standard nonnegative factorizations of Q_4, i.e., columns of $\vartheta(U)$ (or $\vartheta(V^\top)$) are the four vertices of the quadrilateral, then $\mathbf{q}_5$ is a nonnegative combination of columns of U (or $V^\top$). In this way, two of the nontrivial standard nonnegative factorizations of Q_5 are given by

$$Q_5 = \left[\begin{array}{c|c} U & \mathbf{0} \\ \hline \mathbf{0}^\top & 1 \end{array}\right]\left[\begin{array}{c|c} V & \mathbf{w}_5 \\ \hline \mathbf{q}_5^\top & 0 \end{array}\right] = \left[\begin{array}{c|c} U & \mathbf{q}_5 \\ \hline \mathbf{z}^\top & 0 \end{array}\right]\left[\begin{array}{c|c} V & \mathbf{0} \\ \hline \mathbf{0}^\top & 1 \end{array}\right], \tag{17}$$

respectively, where $\mathbf{w}_5$ and $\mathbf{z}_5$ are some nonnegative vectors satisfying $U\mathbf{w}_5 = V^\top \mathbf{z}_5 = \mathbf{q}_5$. This procedure can be generalized to higher n, but there might be other nonnegative factorizations which are not of this particular form specified in (17).

4. A conjecture for higher dimensional EDM

In higher dimensional vector spaces, points $\mathbf{p}_1, \dots, \mathbf{p}_n$ cannot be totally ordered. Thus, for $r > 1$ and $n \geq r+2$, the EDM will not enjoy the inherent structure indicated in (12). Nevertheless, if we denote $\mathbf{p}_j = [p_{ij}]$, then we can write

$$Q(\mathbf{p}_1, \dots, \mathbf{p}_n) = \sum_{i=1}^{r} Q(p_{i1}, \dots, p_{in}).$$

We have shown earlier that generically $\operatorname{rank}_+(Q(p_{i1}, \dots, p_{in})) = n$ for each $1 \leq i \leq r$. Representing the distance matrices for respective components, these r linear EDMs in general are not related to each other. For their summation (of nonnegative entries) to cause a reduction of rank, they must satisfy some delicate algebraic constraints. We thus conjecture that $\operatorname{rank}_+(Q(\mathbf{p}_1, \dots, \mathbf{p}_n)) = n$ generically for all r.

It might be informative to reexamine the geometric representation of the matrix Q_4 when $r > 1$. In contrast to the setting in Figure 1, columns of Q_4 are not coplanar. Their representation becomes that depicted in Figure 2. The vertex $\vartheta(\mathbf{q}_4)$ resides on the simplex $\mathcal{D}_3$. How the base plane determined by vertices $\vartheta(\mathbf{q}_1)$, $\vartheta(\mathbf{q}_2)$, and $\vartheta(\mathbf{q}_3)$ intersects the axes characterizes the zero structure of nonnegative factors. Different from the case $r = 1$, there are several possibilities and there is simply no general rules here. The one shown in Figure 2 implies that the corresponding Q_4 is prime, which is another interesting contrast to the case when $r = 1$.

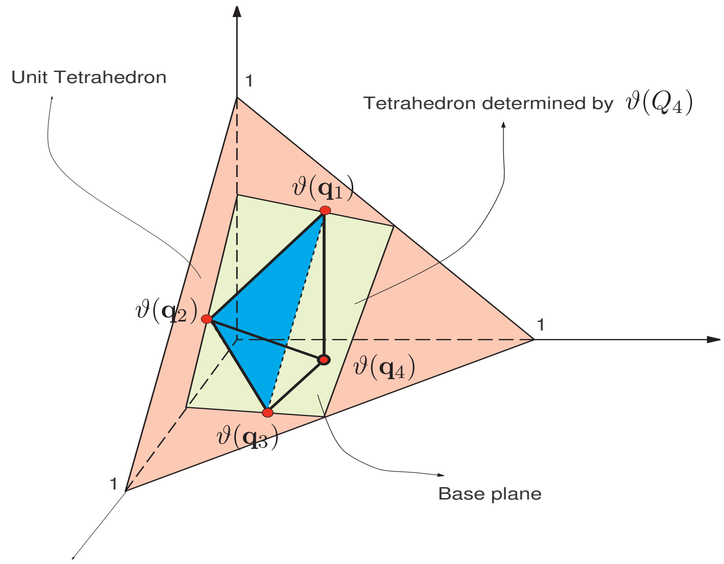

Figure 2: A geometric representation of the matrix $\vartheta(Q_4)$ when $r > 1$.

References

[1] L. B. Beasley and T. J. Laffey, *Real rank versus nonnegative rank*, Linear Algebra Appl., (2009). doi:10.1016/j.laa.2009.02.034.

[2] A. Berman and R. J. Plemmons, *Matrix group monotonicity*, Proc. Amer. Math. Soc., 46 (1974), pp. 355–359.

[3] ———, *Nonnegative matrices in the mathematical sciences*, vol. 9 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Revised reprint of the 1979 original.

[4] S. L. Campbell and G. D. Poole, *Computing nonnegative rank factorizations*, Linear Algebra Appl., 35 (1981), pp. 175–182.

[5] M. T. Chu and M. M. Lin, *Low-dimensional polytope approximation and its applications to nonnegative matrix factorization*, SIAM J. Sci. Comput., 30 (2008), pp. 1131–1155.

[6] J. E. Cohen and U. G. Rothblum, *Nonnegative ranks, decompositions, and factorizations of nonnegative matrices*, Linear Algebra Appl., 190 (1993), pp. 149–168.

[7] G. M. Crippen and T. F. Havel, *Distance geometry and molecular conformation*, vol. 15 of Chemometrics Series, Research Studies Press Ltd., Chichester, 1988.

[8] J. Dattorro, *Euclidean distance matrix*, thesis, Stanford University, 2004. `http://www.stanford.edu/~dattorro/EDM.pdf`.

[9] W. Glunt, T. L. Hayden, and M. Raydan, *Molecular conformations from distance matrices*, J. Comput. Chemistry, 14 (1993), pp. 114–120.

[10] J. C. Gower, *Euclidean distance geometry*, Math. Sci., 7 (1982), pp. 1–14.

[11] S. K. Jain and J. Tynan, *Nonnegative rank factorization of a nonnegative matrix A with $A^{\dagger}A \geq 0$*, Linear Multilinear Algebra, 51 (2003), pp. 83–95.

[12] M. W. Jeter and W. C. Pye, *A note on nonnegative rank factorizations*, Linear Algebra Appl., 38 (1981), pp. 171–173.

[13] ———, *Some nonnegative matrices without nonnegative rank factorizations*, Indust. Math., 32 (1982), pp. 37–41.

[14] D. A. Klain and G.-C. Rota, *Introduction to geometric probability*, Lezioni Lincee. [Lincei Lectures], Cambridge University Press, Cambridge, 1997.

[15] D. J. Richman and H. Schneider, *Primes in the semigroup of non-negative matrices*, Linear and Multilinear Algebra, 2 (1974), pp. 135–140.

[16] J. J. Sylvester, *On a special class of questions on the theory of probabilities*, Birmingham British Assoc. Rept., (1865), pp. 8–9.

[17] L. B. Thomas, *Solution to problem 73-14: Rank factorization of nonnegative matrices by a. berman and r. j. plemmons*, SIAM Review, 16 (1974), pp. 393–394.

[18] S. A. Vavasis, *On the complexity of nonnegative matrix factorization*, 2007. Available at `http://www.citebase.org/abstract?id=oai:arXiv.org:0708.4149`.