Linear Algebra and its Applications 437 (2012) 2685–2712

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

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

On the geometric interpretation of the nonnegative rank[☆]

Nicolas Gillis [a,∗,1], François Glineur [b,c]

[a] University of Waterloo, Department of Combinatorics and Optimization, Waterloo, Ontario, Canada N2L 3G1
[b] Université catholique de Louvain, CORE, Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium
[c] Université catholique de Louvain, ICTEAM Institute, Avenue Georges Lemaître 4, B-1348 Louvain-la-Neuve, Belgium

ARTICLE INFO

Article history:
Received 4 September 2010
Accepted 13 April 2012
Available online 31 July 2012

Submitted by V. Mehrmann

AMS classification:
15A23
15B48
52B05
52B11
65D99
65F30
90C27

Keywords:
Nonnegative rank
Restricted nonnegative rank
Nested polytopes problem
Computational complexity
Computational geometry
Extended formulations
Linear Euclidean distance matrices

ABSTRACT

The nonnegative rank of a nonnegative matrix is the minimum number of nonnegative rank-one factors needed to reconstruct it exactly. The problem of determining this rank and computing the corresponding nonnegative factors is difficult; however it has many potential applications, e.g., in data mining and graph theory. In particular, it can be used to characterize the minimal size of any extended reformulation of a given polytope. In this paper, we introduce and study a related quantity, called the restricted nonnegative rank. We show that computing this quantity is equivalent to a problem in computational geometry, and fully characterize its computational complexity. This in turn sheds new light on the nonnegative rank problem, and in particular allows us to provide new improved lower bounds based on its geometric interpretation. We apply these results to slack matrices and linear Euclidean distance matrices and obtain counter-examples to two conjectures of Beasley and Laffey, namely we show that the nonnegative rank of linear Euclidean distance matrices is not necessarily equal to their dimension, and that the rank of a matrix is not always greater than the nonnegative rank of its square.

☆ This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

∗ Corresponding author. Tel.: +1 519 888 4567x37529; fax: +1 519 725 5441.
E-mail addresses: ngillis@uwaterloo.ca (N. Gillis), francois.glineur@uclouvain.be (F. Glineur).
[1] This work was carried out when the author was a Research fellow of the Fonds de la Recherche Scientifique (F.R.S.-FNRS) at Université catholique de Louvain.

0024-3795/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.laa.2012.06.038

1. Introduction

The nonnegative rank of an m-by-n real nonnegative matrix $M \in \mathbb{R}^{m\times n}_+$ is the minimum number of nonnegative rank-one factors needed to reconstruct M exactly, i.e., the minimum integer k such that there exists $U \in \mathbb{R}^{m\times k}_+$ and $V \in \mathbb{R}^{k\times n}_+$ with $M = UV = \sum_{i=1}^{k} U_{:i}V_{i:}$. The pair (U, V) is called a rank-k nonnegative factorization[2] of M. The nonnegative rank of M is denoted $\text{rank}_+(M)$. Clearly,

$$\text{rank}(M) \leqslant \text{rank}_+(M) \leqslant \min(m, n).$$

Determining the nonnegative rank and computing the corresponding nonnegative factorization has been studied relatively recently in the linear algebra literature [5,15]. A lot more attention has been devoted to the approximate nonnegative factorization problem (called nonnegative matrix factorization, NMF for short [35]) consisting in finding two low-rank nonnegative factors U and V such that $M \approx UV$ or, more precisely, solving

$$\min_{U\in\mathbb{R}^{m\times k}_+, V\in\mathbb{R}^{k\times n}_+} ||M - UV||_F. \tag{NMF}$$

NMF has been widely used as a data analysis technique [7], e.g., in text mining, image processing, hyperspectral data analysis, computational biology, and clustering. Nevertheless, there are not many theoretical results about the nonnegative rank and better characterizations, in particular lower bounds, could help practitioners. For example, finding an efficient way to compute exact nonnegative factorizations could help to design new NMF algorithms using a two-step strategy [42]: first approximate M with a low-rank nonnegative matrix A (e.g., using the singular value decomposition[3]) and then compute a nonnegative factorization of A. Bounds for the nonnegative rank could also help select the factorization rank of the NMF, replacing the trial and error approach often used by practitioners. For example, in hyperspectral image analysis, the nonnegative rank corresponds to the number of materials present in the image and its computation could lead to more efficient algorithms detecting these constitutive elements, see [9,17,30,26] and the references therein.

An extended formulation (or lifting) for a polytope $P \subseteq \mathbb{R}^n$ is a polyhedron $Q \subseteq \mathbb{R}^{n+p}$ such that

$$P \;=\; \text{proj}_x(Q) := \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^p \text{ s.t. } (x, y) \in Q\}.$$

Extended formulations whose size (number of constraints plus number of variables defining Q) is polynomial in n are called *compact* and are of great importance in integer programming. They allow to reduce significantly the size of certain linear programs arising in the context of integer programming and combinatorial optimization, and therefore provide a way to solve them efficiently, i.e., in polynomial-time (see [16] for a survey). Yannakakis [45, Theorem 3] showed that the minimum size s of an extended formulation of a polytope[4]

$$P = \{x \in \mathbb{R}^n \mid Cx + d \geqslant 0, Ax = b\},$$

is of the same order[5] as the sum of its dimension n and the nonnegative rank of its slack matrix $S_M \geqslant 0$, where each column of the slack matrix is defined as

$$S_M(:, i) = Cv_i + d \geqslant 0, \quad i = 1, 2, \ldots, m, \tag{1.1}$$

and vectors v_i are the m vertices of the polytope P. Formally, we then have

$$s = \Theta(n + \text{rank}_+(S_M)).$$

[2] Note that matrices U and V in a rank-k nonnegative factorization are not required to have rank k.

[3] Even though an optimal low-rank approximation of a nonnegative matrix is not always nonnegative, it is often the case in practice [33].

[4] This can be generalized to polyhedra [16].

[5] More precisely, the minimum number of inequalities in an extended formulation of P is exactly equal to the nonnegative rank of the slack matrix of P, see, e.g., [31].

In particular, any rank-k nonnegative factorization (U, V) of $S_M = UV$ provides the following extended formulation for P with size $\Theta(n+k)$

$$Q = \{(x, y) \in \mathbb{R}^{n+k} \mid Cx + d = Uy, Ax = b, y \geqslant 0\}. \tag{1.2}$$

Indeed, $\text{proj}_x(Q) \subseteq P$ since $Uy \geqslant 0$ implies $Cx + d \geqslant 0$ for any $x \in \text{proj}_x(Q)$, and $P \subseteq \text{proj}_x(Q)$ since $Cv_i + d = S_M(:, i) = UV(:, i)$ implies that $(v_i, V(:, i)) \in Q$ for all i and therefore each vertex v_i of P belongs to $\text{proj}_x(Q)$. Intuitively, this extended formulation parametrizes the space of slacks of the original polytope with the convex cone $\{Uy \mid y \geq 0\}$.

It is therefore interesting to compute bounds for the nonnegative rank in order to estimate the size of these extended formulations. For example, Goemans [28] recently used this result to show that the size of any linear programming formulation of the permutahedron (polytope whose $n!$ vertices are permutations of $[1, 2, \ldots, n]$) is at least $\Omega(n \log(n))$ variables plus constraints (cf. Section 3).

We will see in Section 3.1 that the nonnegative rank is closely related to a problem in computational geometry that consists in finding a polytope with minimum number of vertices nested between two given polytopes. Therefore a better understanding of the properties of the nonnegative rank would presumably also allow to improve characterization of the solutions to this geometric problem.

The nonnegative rank has connections with other problems, e.g., in communication complexity theory [45,36], probability [10], and graph theory (cf. Section 3). It is also related to the cp-rank which, for a given nonnegative symmetric matrix $M \in \mathbb{R}^{m\times m}$, seeks a nonnegative factorization with factors U and V equal to each other:

$$\text{cp-rank}(M) = \min k \text{ such that } \exists U \in \mathbb{R}_+^{m\times k} \quad \text{with } M = UU^T = \sum_{i=1}^{k} U_{:i} U_{:i}^T.$$

If such a decomposition exists, M is said to be a completely positive matrix; see [6] and the references therein. Apart from the trivial relationship $\text{rank}_+(M) \leqslant \text{cp-rank}(M)$, we are not aware of any work relating these two quantities.

1.1. Summary of our results

The main contribution of this paper is to introduce improved lower bounds for the nonnegative rank. We also give some insightful results about its geometric interpretation. Below we sketch the general ideas behind our results.

In Section 2, we introduce and study a new quantity called the *restricted nonnegative rank* (RNR) of a nonnegative matrix M, denoted $\text{rank}_+^*(M)$. Its definition is similar to that of the nonnegative rank (NNR) with an additional rank condition on the first factor of the factorization.

We prove that computing the RNR is equivalent to a problem in computational geometry, referred to as the *nested polytopes problem* (NPP) (Theorem 1). Given two nested full-dimensional polytopes, say $S \subseteq P$, the NPP consists in finding a third polytope T nested between S and P with a minimum number of vertices, i.e., we want $S \subseteq T \subseteq P$. Using an appropriate reduction, we establish the equivalence between NPP and RNR computation using the following one-to-one correspondence:

RNR computation	NPP
rank of input matrix M	(dimension of input polytopes P and S)+1
columns of input matrix $M \in \mathbb{R}^{m\times n}$	n vertices of input polytope S
facets of the nonnegative orthant $\mathbb{R}_+^m$	m facets of input polytope P
columns of sought factor $U \in \mathbb{R}^{m\times k}$	k vertices of sought nested polytope T

The NPP has been widely studied in the computational geometry literature. In particular, it has been shown that

- The NPP can be solved in polynomial time when the input polytopes have dimension two [1]. This implies that the RNR can be computed efficiently when $\text{rank}(M) = 3$ (Theorem 2). Moreover, this shows that checking whether $\text{rank}_+(M) = 3$ for a rank-three matrix M can also be done in polynomial time, because we have $\text{rank}_+(M) = \text{rank}(M) \iff \text{rank}^*_+(M) = \text{rank}(M)$ (Corollary 1).
- The NPP is NP-hard in dimension three or more [20]. Therefore, computing the RNR is NP-hard as soon as $\text{rank}(M) \geqslant 4$ (Theorem 3).

We conclude Section 2 by proving some useful properties of the RNR, which sometimes differ from those of the standard and nonnegative ranks. For example, for an m-by-n matrix M, we have $\text{rank}^*_+(M) \leqslant n$ while $\text{rank}^*_+(M) \nleq m$.

In Section 3, we describe a geometric interpretation of the NNR. Similarly to RNR, computing the NNR of a nonnegative matrix M is equivalent to a geometric problem involving nested polytopes. However, this problem is more general than an NPP, in the sense that the inner polytope is no longer guaranteed to be full-dimensional. More precisely, the n points whose convex hull define the inner polytope S still correspond to the columns of M, and the m facets of the outer polytope P also correspond to the facets of the nonnegative orthant $\mathbb{R}^m_+$, but the reduction from NNR no longer produces a full-dimensional inner polytope S: the dimension of the affine space spanning S can now be strictly smaller than the dimension of the outer polytope P. We will call this type of problem a *generalized nested polytopes* problem.

Our new concept of restricted nonnegative rank has close ties with the nonnegative rank, which is also apparent from the geometric point of view: any solution to the generalized nested polytopes problem based on the NNR for matrix M can be converted to a feasible solution to the NPP, and hence to the corresponding RNR of matrix M. Indeed, given a solution polytope T, we can intersect it with the affine space spanned by S (corresponding to the column space of M) and obtain a new polytope T'. This polytope T' contains S (because T does) while living in the same affine subspace as S (by construction). Therefore T' is a solution of the NPP corresponding to the RNR of M (Theorem 3.2). Hence, we have

$$\text{rank}^*_+(M) \leqslant \#\text{vertices}(T') = \#\text{vertices}(T \cap \text{affine hull}(S)).$$

This relationship now allows us to derive new lower bounds on the nonnegative rank. Letting $\phi_r(k)$ be a standard upper bound on the number of vertices of the intersection of a polytope having k vertices (i.e., T) with a $(r-1)$-dimensional space (i.e., the affine hull of S, which has dimension $\text{rank}(M) - 1$), we finally have that

$$\text{rank}^*_+(M) \leqslant \#\text{vertices}(T \cap \text{affine hull}(S)) \leqslant \phi_{\text{rank}(M)}(\text{rank}_+(M)),$$

see Theorem 5. Inverting the non-decreasing function $\phi_r(.)$ finally gives a *lower bound for the NNR that depends on the RNR and the usual rank*:

$$\text{rank}_+(M) \geqslant \phi^{-1}_{\text{rank}(M)}(\text{rank}^*_+(M)).$$

As explained above, the RNR is, in general, NP-hard to compute. However, there are situations for which it can be computed efficiently. For example, we show in Section 4 that the RNR of slack matrices and linear Euclidean distance matrices (linear EDM's) is maximal, i.e., it is equal to the number of columns of these matrices (Theorems 7 and 8). We then apply our lower bound $\phi^{-1}_{\text{rank}(M)}(\text{rank}^*_+(M))$ on the NNR and show that it generalizes and improves the bounds of Goemans [28] for slack matrices and of Beasley and Laffey [3] for linear EDM's. We also obtain counter-examples to two conjectures of Beasley and Laffey [3], namely we show that the nonnegative rank of linear Euclidean distance matrices is not necessarily equal to their dimension, and that the rank of a matrix is not always greater than the nonnegative rank of its square.

We conclude with some open questions on the complexity of the NNR computation.

1.2. Notation

The set of real matrices of dimension m by n is denoted $\mathbb{R}^{m\times n}$; for $A \in \mathbb{R}^{m\times n}$, we denote the ith column of A by $A_{:i}$ or $A(:,i)$, the jth row of A by $A_{j:}$ or $A(j,:)$, and the entry at position (i,j) by A_{ij} or $A(i,j)$; for $b \in \mathbb{R}^{m\times 1} = \mathbb{R}^m$, we denote the ith entry of b by b_i. Notation $A(I,J)$ refers to the submatrix of A with row and column indices respectively in I and J, and $a{:}b$ is the set $\{a, a+1, \ldots, b-1, b\}$ (for a and b integers with $a \leqslant b$). The set $\mathbb{R}^{m\times n}$ with component-wise nonnegative entries is denoted $\mathbb{R}^{m\times n}_+$. Matrix A^T is the transpose of matrix A. The rank of a matrix A is denoted $\text{rank}(A)$, its column space $\text{col}(A)$. Both the convex hull of a set of points S and the convex hull of the columns of a matrix S are denoted $\text{conv}(S)$. The number of vertices of a polytope Q is denoted by $\#\,\text{vertices}(Q)$. The matrix obtained from the concatenation of the columns of two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times p}$ is denoted $[A\,B] \in \mathbb{R}^{m\times(n+p)}$. The sparsity pattern of a vector is the set of indices of its zero entries (it is the complement of its support).

2. Restricted nonnegative rank

In this section, we analyze the following quantity

Definition 1. The *restricted nonnegative rank* of a nonnegative matrix M, denoted $\text{rank}^*_+(M)$, is the minimum value of k such that there exists $U \in \mathbb{R}^{m\times k}_+$ and $V \in \mathbb{R}^{k\times n}_+$ with $M = UV$ and $\text{rank}(U) = \text{rank}(M)$.

This is the definition of the standard nonnegative rank with an additional constraint on the rank of the first factor U; note that this constraint can equivalently be formulated as $\text{col}(U) = \text{col}(M)$ (equality between column spaces of M and U). Given a nonnegative matrix M, we are interested in computing its restricted nonnegative rank $\text{rank}^*_+(M)$ and a corresponding nonnegative factorization, i.e., solve

(RNR) Given a nonnegative matrix $M \in \mathbb{R}^{m\times n}_+$, find $k = \text{rank}^*_+(M)$ and compute $U \in \mathbb{R}^{m\times k}_+$ and $V \in \mathbb{R}^{k\times n}_+$ such that $M = UV$ and $\text{rank}(U) = \text{rank}(M)$.

Motivation to study this restriction includes the following:

1. The restricted nonnegative rank provides a new *upper* bound for the nonnegative rank, since $\text{rank}_+(M) \leqslant \text{rank}^*_+(M)$.
2. The restricted nonnegative rank can be characterized more easily than the nonnegative rank. In particular, as explained in the introduction, its geometrical interpretation (Section 2.1) will lead to new improved *lower* bounds for the nonnegative rank (Sections 3 and 4).

RNR is a generalization of *exact nonnegative matrix factorization* (exact NMF) introduced by Vavasis [42], which is the problem of checking whether a matrix M with rank r satisfies $\text{rank}_+(M) = r$ and, if the answer is positive, to compute a rank-r nonnegative factorization of M. If $\text{rank}_+(M) = r$ then it is clear that $\text{rank}^*_+(M) = \text{rank}_+(M)$ since the rank of U in any rank-r nonnegative factorization (U, V) of M must be equal to r.

Vavasis studies the computational complexity of exact NMF and proves it is NP-hard by showing its equivalence with a problem called intermediate simplex. This construction requires both the dimensions of input matrix M and its rank r to increase to obtain NP-hardness. This result also implies NP-hardness of RNR when the rank of matrix M is not fixed (i.e., r is part of the input). In contrast, when the rank r of input matrix M is fixed, no results on the complexity of exact NMF are known (except in the trivial cases $r = 1, 2$ [41]). The situation for RNR is quite different: we are going to show that RNR can be solved in polynomial-time when $r = 3$ and that it is NP-hard for any fixed $r \geqslant 4$, see Theorems 2 and 3. In particular, this result implies that exact NMF can be solved in polynomial-time for rank-three nonnegative matrices, see Corollary 1.

To proceed, we first show equivalence between RNR and another problem in computational geometry, closely related to intermediate simplex (Section 2.1), and then apply results from the computational geometry literature to conclude about its computational complexity when the rank of the input matrix is fixed (Section 2.2).

2.1. Equivalence with the nested polytopes problem

Let consider the following problem called nested polytopes problem (NPP):

(NPP) Given a full-dimensional bounded polyhedron P described by m inequalities in $\mathbb{R}^{r-1}$

$$P = \{x \in \mathbb{R}^{r-1} \mid Cx + d \geq 0\},$$

(where full-dimensionality is equivalent to requiring that $(C\ d) \in \mathbb{R}^{m \times r}$ is a rank-r matrix) and a set S of n points belonging to P such that conv(S) is also full-dimensional, find the minimum number k of points belonging to P whose convex hull T contains S (and conv(S)), i.e., such that conv$(S) \subseteq T \subseteq P$.

Polytope P is referred to as the outer polytope, and conv(S) as the inner polytope; note that they are given by two distinct types of representations (facets for P, extreme points for conv(S)).

The intermediate simplex problem mentioned earlier and introduced by Vavasis [42] is a particular case of NPP in which one asks whether k is equal to r (which is the minimum possible value), i.e., if there exists a simplex T (defined by r vertices in $\mathbb{R}^{r-1}$) contained in P and containing S.

We now prove equivalence between RNR and NPP, which generalizes the result of Vavasis [42] showing equivalence between exact NMF and intermediate simplex.

Theorem 1. *There are polynomial-time reductions from RNR to NPP and from NPP to RNR.*

Proof. Let us construct a reduction from RNR to NPP. First we (1) delete the zero rows and columns of M and (2) normalize the columns of M such that it becomes column stochastic (elements in each column are nonnegative and sum to one). One can easily check that it gives a polynomially equivalent RNR instance [12]. We then decompose M as the product of two rank-r matrices (using, e.g., reduction to row-echelon form)

$$M = AB \iff M_{:i} = \sum_{l=1}^{r} A_{:l} B_{li} \ \forall i, \tag{2.1}$$

where $r = \text{rank}(M)$, $A \in \mathbb{R}^{m \times r}$ and $B \in \mathbb{R}^{r \times n}$. We first show that matrices A and B can be assumed to be column stochastic without loss of generality. Indeed, since M is column stochastic, at least one column of A does not sum to zero (otherwise all columns of $AB = M$ would sum to zero, a contradiction). One can then add that column to all columns of A that sum to zero, updating B accordingly, so that all columns of A have a positive sum. Normalizing now each column of A, updating again B accordingly, provides a column stochastic matrix A. Finally, it is easy to see that, since $M = AB$, matrices M and A being column stochastic implies that matrix B also is.

In order to find a solution of RNR, we have to find $U \in \mathbb{R}^{m \times k}_{+}$ and $V \in \mathbb{R}^{k \times n}_{+}$ such that $M = UV$ and rank$(U) = r$. For the same reasons as for A and B, U and V can be assumed to be column stochastic without loss of generality. Moreover, since

$$M = UV = AB,$$

and rank$(M) = $ rank$(A) = $ rank$(U) = r$, the column spaces of M, A and U coincide, implying that the columns of U must be a linear combination of the columns of A. The columns of U must then belong to the following set

$$Q = \left\{ u \in \mathbb{R}^m \mid u \in \text{col}(A),\ u \geqslant 0 \text{ and } \sum_{i=1}^{m} u_i = 1 \right\}. \tag{2.2}$$

One can therefore reduce the search space for those columns to the $(r-1)$-dimensional polyhedron corresponding to the coefficients of all possible linear combinations of the columns of A generating stochastic columns. Defining

$$C(:,i) = A(:,i) - A(:,r) \quad 1 \leqslant i \leqslant r-1, \quad \text{and} \quad d = A(:,r),$$

and introducing the affine function $f : \mathbb{R}^{r-1} \to \mathbb{R}^m : x \to f(x) = Cx + d$, which is injective since C is full rank (because A is full rank), this polyhedron can be defined as

$$P = \left\{ x \in \mathbb{R}^{r-1} \mid A(:,1:r-1)x + \left(1 - \sum_{i=1}^{r-1} x_i\right) A(:,r) \geqslant 0 \right\} = \{x \in \mathbb{R}^{r-1} \mid f(x) \geqslant 0\}. \quad (2.3)$$

Note that $B(1:r-1,j) \in P \; \forall j$ since $M(:,j) = AB(:,j) = f(B(1:r-1,j)) \geqslant 0 \; \forall j$.

Let us show that P is bounded: suppose P is unbounded, then

$$\exists\, x \in P, \exists\, y \neq 0 \in \mathbb{R}^{r-1}, \forall \alpha \geqslant 0 : x + \alpha y \in P,$$

$$\iff C(x + \alpha y) + d = (Cx + d) + \alpha Cy \geqslant 0.$$

Since $Cx + d \geqslant 0$, this implies that $Cy \geqslant 0$. Observe that columns of C sum to zero (since the columns of A sum to one) so that Cy sums to zero as well; moreover, C is full rank and y is nonzero implying that Cy is nonzero and therefore that Cy must contain at least one negative entry, a contradiction.

Notice that the set Q can be equivalently written as

$$Q = \{u \in \mathbb{R}^m \mid u = f(x), \; x \in P\}.$$

Noting $X = [x_1\, x_2\, \ldots x_k] \in \mathbb{R}^{r-1 \times k}$, $f(X) = [f(x_1) f(x_2) \ldots f(x_k)] = CX + [d\, d \ldots d]$, we finally have that the following statements are equivalent

(i) $\exists U \in \mathbb{R}^{m \times k}, V \in \mathbb{R}^{k \times n}$ column stochastic with $\operatorname{rank}(U) = \operatorname{rank}(M)$, and $M = UV$,
(ii) $\exists x_1, x_2, \ldots x_k \in P$ and $V \in \mathbb{R}^{k \times n}$ column stochastic such that

$$M = f(B(1:r-1,:)) = f(X)V = f(XV),$$

(iii) $\exists x_1, x_2, \ldots x_k \in P$ and $V \in \mathbb{R}^{k \times n}$ column stochastic such that

$$B(1:r-1,:) = XV.$$

The equivalence between (i) and (ii) follows from the above derivations (i.e., $U = f(X)$ for some $x_1, x_2, \ldots x_k \in P$); $f(X)V = f(XV)$ because V is column stochastic (so that $[d\, d\, \ldots\, d]V = [d\, d\, \ldots\, d]$), and the second equivalence between (ii) and (iii) is a consequence from the fact that f is an injection.

We have then reduced RNR of a rank-r matrix M to an instance of NPP with nested full-dimensional polytopes in $\mathbb{R}^{r-1}$: the inner polytope is constructed from the columns of M (more precisely it is the convex hull of the n columns in truncated matrix $B(1:r-1,:)$), the outer polytope P is defined with m inequalities $f(x) \geq 0$ (constructed from the m rows in truncated matrix $A(:,1:r)$), which correspond to the facets of $\mathbb{R}^m_+$, and the goal is to find the minimum number k of points x_i belonging to the outer polytope P whose convex hull contains the inner polytope $\operatorname{conv}(S)$.

Because all steps in the above derivation are equivalences, we have actually also defined a reduction from NPP to RNR; to map a NPP instance (based on data S, C and d) to a RNR instance, we take

$$M(:,i) = f(s_i) = Cs_i + d \geqslant 0 \text{ for each } s_i \in S \quad 1 \leqslant i \leqslant n,$$

and we have that $\operatorname{rank}(M) = r$ because the n points s_i have a full-dimensional convex hull (they affinely span P). $\square$

It is worth noting that M would be the slack matrix of polytope P if S was the set of vertices of P (cf. Section 1). This will be useful later in Section 4.1.

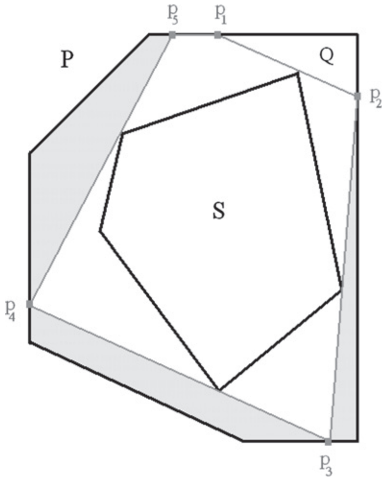

Fig. 1. Illustration of the algorithm of Aggarwal et al. [1].

2.2. Computational complexity

2.2.1. Rank-three matrices

Using Theorem 1, RNR for a rank-three matrix can be reduced to a two-dimensional nested polytopes problem.[6] Therefore, one has to find a convex polygon T with minimum number of vertices nested between two given convex polygons $\text{conv}(S) \subseteq P$. This problem has been studied by Aggarwal et al. in [1], who proposed an algorithm running in $\mathcal{O}(p\log(k))$ operations[7], where p is the total number of vertices of the given polygons $\text{conv}(S)$ and P, and k is the number of vertices of the minimal nested polygon T. If M is an m-by-n matrix then $p \leqslant m+n$ since $\text{conv}(S)$ has at most n vertices and the polygon P is defined by m inequalities (so that it has at most m vertices). Moreover we have the following upper bound on the restricted nonnegative rank: $\text{rank}^*_+(M) \leqslant \min(m, n)$, which follows from the fact that $T = \text{conv}(S)$ and $T = P$ are feasible solutions for RNR. Finally, we conclude that one can compute the restricted nonnegative rank of a rank-three m-by-n matrix in $O((m+n)\log(\min(m, n)))$ operations.

Theorem 2. *For* $\text{rank}(M) \leqslant 3$*, RNR can be solved in polynomial-time.*

Proof. Cases $r = 1, 2$ are trivial since any rank-1 (resp. 2) nonnegative matrix can always be expressed as the sum of 1 (resp. 2) nonnegative factors [41]. Case $r = 3$ follows from Theorem 1 and the polynomial-time algorithm of Aggarwal et al. [1]. □

For the sake of completeness, we sketch the main ideas behind the algorithm of Aggarwal et al. They first make the following observations: (1) any vertex of a solution T can be assumed to belong to the boundary of the polygon P (otherwise it can be projected back on P in order to generate a new solution containing the previous one), (2) any segment whose ends are on the boundary of P and tangent to $\text{conv}(S)$ (i.e., S is on one side of the segment, and the segment touches S) defines a polygon with the boundary of P which must contain a vertex of any feasible solution T (otherwise the tangent point on $\text{conv}(S)$ could not be contained in T), see, e.g., set Q on Fig. 1 delimited by the segment $[p_1, p_2]$ and the boundary of P, and such that $T \cap Q \neq \emptyset$ for any feasible solution T.

Starting from any point p_1 of the boundary of P, one can trace the tangent to $\text{conv}(S)$ and hence obtain the next intersection p_2 with P. Point p_2 is chosen as the next vertex of a solution T, and the same procedure is applied (say k times) until the algorithm can reach the initial point without going

[6] See https://sites.google.com/site/nicolasgillis/code where a MATLAB® code is provided.

[7] Wang generalized the result for non-convex polygons [43]. Bhadury and Chandrasekaran propose an algorithm to compute all possible solutions [8].

through $\text{conv}(S)$, see Fig. 1. This generates a feasible solution $T(p_1) = \text{conv}(\{p_1, p_2, \ldots, p_k\})$. Because of (1) and (2), this solution has at most one vertex more than an optimal one, i.e., $k \leqslant \text{rank}^*_+(M) + 1$ (since T determines, with the boundary of P, $k - 1$ disjoint polygons tangent to $\text{conv}(S)$). Moreover, because of (1) and (2), there must exist a vertex of an optimal solution on the boundary of P between p_1 and p_2.

The point p_1 is then replaced by the so-called 'contact change points' located on this part of the boundary of P while the corresponding solution $T(p_1)$ is updated using the procedure described above. The contact change points are: (a) the vertices of P between p_1 and p_2, and (b) the points for which one tangent point of $T(p_1)$ on $\text{conv}(S)$ is changed when p_1 is replaced by them. This (finite) set of points provides a list of candidates where the number of vertices of the solution T could potentially be reduced (i.e., where p_1 and p_k could coincide) by replacing p_1 by one of these points. It is then possible to check whether the current solution can be improved or not, and guarantee global optimality. In the example of Fig. 1, moving p_1 on the (only) vertex of P between p_1 and p_2 generates an optimal solution of this RNR instance (since it reduces the original solution from 5 to 4 vertices).

Finally, Theorem 2 also allows to shed some light on the nonnegative rank computation.

Corollary 1. *Exact NMF can be solved in polynomial-time for any nonnegative matrix M with* $\text{rank}(M) \leqslant 3$.

Proof. Given a matrix M with rank r, we observe that whenever $\text{rank}_+(M) = r$, one also has $\text{rank}^*_+(M) = r$ since in any rank-r nonnegative factorization (U, V) of M, $\text{rank}(U) = r$. Therefore $\text{rank}_+(M) = r \iff \text{rank}^*_+(M) = r$. Hence, determining whether $\text{rank}_+(M) = r$ or not and, if yes, computing a rank-r nonnegative factorization (i.e., solve exact NMF) can be done by solving RNR, which is done in polynomial-time for $r \leqslant 3$. $\square$

2.2.2. Higher-rank matrices

For a rank-four matrix, RNR reduces to the problem of finding a polytope T with the minimum number of vertices nested between two three-dimensional polytopes $S \subseteq P$. This problem has been studied by Das et al. [20,18] and has been shown to be NP-hard when minimizing the number of facets of T (the reduction is from *planar-3SAT*). From this result, we deduce using a duality argument that minimizing the number of vertices of T, hence solving RNR, is NP-hard as well [19,14].

Theorem 3. *For* $\text{rank}(M) \geqslant 4$, *RNR is NP-hard.*

Proof. The NP-hardness results of Das et al. [20,18,19] deals with the minimization of the number of facets of a polytope nested between an inner polytope described by its vertices and an outer polytope described by its facets. Taking the polar of the three nested polytopes exchanges the roles of the inner and outer polytopes, and transforms facet descriptions into vertex descriptions and vice versa, so that the descriptions of the inner and outer polytopes keep their type and one must now minimize the number of vertices of the intermediate polytope: this is the formulation of NPP, which is therefore NP-hard as well. Combining this observation with Theorem 1 concludes the proof. $\square$

Several approximation algorithms for NPP have been proposed in the literature. For example, when $\text{rank}(M) = 4$, Mitchell and Suri [38] can approximate $\text{rank}^*_+(M)$ within a $\mathcal{O}(\log(p))$ factor, where p is the total number of vertices of the given polytopes $\text{conv}(S)$ and P. Clarkson [14] proposes a randomized algorithm finding a polytope T with at most $\mathcal{O}(dr^*_+ \ln(r^*_+))$ vertices and running in $\mathcal{O}({r^*_+}^2 p^{1+\delta})$ expected time (with $r^*_+ = \text{rank}^*_+(M)$, $d = \text{rank}(M) - 1$ and δ is any fixed value > 0).

2.3. Some properties

In this section, we derive some useful properties of the restricted nonnegative rank. It is well-known that for a matrix $M \in \mathbb{R}^{m \times n}_+$, we have $\text{rank}_+(M) \leq \min(m, n)$; surprisingly, this does not hold for the restricted nonnegative rank.

Lemma 1. *For $M \in \mathbb{R}_+^{m \times n}$,*

$$\text{rank}_+^*(M) \leqslant n \quad \textit{but} \quad \text{rank}_+^*(M) \nleqslant m.$$

Proof. The first inequality is trivial since $M = MI$ (I being the identity matrix) provides a valid factorization for RNR. Example 1 below describes a situation where $\text{rank}_+^*(M) = 8$ for a 6-by-8 matrix M. $\square$

Example 1. Construct M using the following NPP instance: P is the three-dimensional cube $P = \{x \in \mathbb{R}^3 \mid 0 \leqslant x_i \leqslant 1, \ 1 \leqslant i \leqslant 3\}$, with 6 facets and S is the set of its 8 vertices $S = \{x \in \mathbb{R}^3 \mid x_i \in \{0, 1\}, \ 1 \leqslant i \leqslant 3\}$. By construction, the convex hull of S is equal to P and the unique and optimal solution to this NPP instance is $T = P = \text{conv}(S)$ with 8 vertices. By Theorem 1, the corresponding matrix M of the RNR instance

$$M = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \tag{2.4}$$

has restricted nonnegative rank equal to 8 (note that its rank is 4 and its nonnegative rank is 6, see Section 3).

Lemma 1 implies that in general $\text{rank}_+^*(M)$ and $\text{rank}_+^*(M^T)$ can differ, unlike the rank and nonnegative rank [15]. Note however that when $\text{rank}(M) \leqslant 3$, we have

$$\text{rank}_+^*(M) \leqslant \min(m, n),$$

because the number of vertices of the outer polygon P in the NPP instance is smaller or equal to its number of facets m in the two-dimensional case or lower, and that the solution $T = P$ is always feasible.

Lemma 2. *Let $A \in \mathbb{R}_+^{m \times n}$ and $B \in \mathbb{R}_+^{m \times r}$, then*

$$\text{rank}_+^*([A\,B]) \leqslant \text{rank}_+^*(A) + \text{rank}_+^*(B).$$

Proof. Let (U_a, V_a) and (U_b, V_b) be solutions of RNR for A and B respectively, then

$$[A\,B] = [U_a\,U_b] \begin{bmatrix} V_a & 0 \\ 0 & V_b \end{bmatrix},$$

and $\text{rank}([U_a\,U_b]) = \text{rank}([A\,B])$ since $\text{col}(U_a) = \text{col}(A)$ and $\text{col}(U_b) = \text{col}(B)$ by definition. $\square$

Lemma 3. *Let $M \in \mathbb{R}_+^{m \times n}$ with $\text{rank}(M) = r$ and $\text{rank}_+(M) = r_+$, $U \in \mathbb{R}_+^{m \times r_+}$ and $V \in \mathbb{R}_+^{r_+ \times n}$ with $M = UV$. Then*

$$r_+ < \text{rank}_+^*(M) \quad \Rightarrow \quad r < \text{rank}(U) \leqslant r_+ \ \textit{and} \ r \leqslant \text{rank}(V) < r_+.$$

Moreover, if M is symmetric,

$$r_+ < \text{rank}_+^*(M) \quad \Rightarrow \quad r < \text{rank}(U) < r_+ \ \textit{and} \ r < \text{rank}(V) < r_+.$$

Proof. Clearly,

$$r \leqslant \operatorname{rank}(U) \leqslant r_+ \quad \text{and} \quad r \leqslant \operatorname{rank}(V) \leqslant r_+.$$

If $\operatorname{rank}(U) = r$, we would have $\operatorname{rank}_+^*(M) = r_+$ which is a contradiction, and $\operatorname{rank}(V) = r_+$ would imply that V has a right pseudo-inverse $V^\dagger$, so that we could write $U = MV^\dagger$ and then $r \leqslant \operatorname{rank}(U) \leqslant \min(r, \operatorname{rank}(V^\dagger)) \leqslant r$, a contradiction for the same reason.

In case M is symmetric, to show that $\operatorname{rank}(U) < r_+$ and $\operatorname{rank}(V) > r$, we use symmetry and observe that $UV = M = M^T = V^T U^T$. $\square$

Corollary 2. *Given a nonnegative matrix M,*

$$\operatorname{rank}_+^*(M) \leqslant \operatorname{rank}(M) + 1 \quad \Rightarrow \quad \operatorname{rank}_+(M) = \operatorname{rank}_+^*(M).$$

If M is symmetric,

$$\operatorname{rank}_+^*(M) \leqslant \operatorname{rank}(M) + 2 \quad \Rightarrow \quad \operatorname{rank}_+(M) = \operatorname{rank}_+^*(M).$$

Proof. Let $r = \operatorname{rank}(M)$, $r_+ = \operatorname{rank}_+(M)$, and $U \in \mathbb{R}_+^{m \times r_+}$ and $V \in \mathbb{R}_+^{r_+ \times n}$ such that $M = UV$. If $r_+ < \operatorname{rank}_+^*(M)$, by Lemma 3, we have

$$r < \operatorname{rank}(U) \leqslant r_+ < \operatorname{rank}_+^*(M),$$

which is a contradiction if $\operatorname{rank}_+^*(M) \leqslant r + 1$. If M is symmetric, we have $\operatorname{rank}(U) < r_+$ and the above equation is a contradiction if $\operatorname{rank}_+^*(M) \leqslant r + 2$. $\square$

These results imply for example that to find a symmetric rank-three nonnegative matrix with $\operatorname{rank}_+(M) < \operatorname{rank}_+^*(M)$, we need $\operatorname{rank}_+^*(M) > \operatorname{rank}(M) + 2 = 5$ and therefore have to consider matrices of size at least 6-by-6 with $\operatorname{rank}_+^*(M) = 6$.

Example 2. Let us consider the following matrix M and the rank-5 nonnegative factorization (U, V),

$$M = \begin{pmatrix} 0 & 1 & 4 & 9 & 16 & 25 \\ 1 & 0 & 1 & 4 & 9 & 16 \\ 4 & 1 & 0 & 1 & 4 & 9 \\ 9 & 4 & 1 & 0 & 1 & 4 \\ 16 & 9 & 4 & 1 & 0 & 1 \\ 25 & 16 & 9 & 4 & 1 & 0 \end{pmatrix} = UV, \quad U = \begin{pmatrix} 5 & 0 & 4 & 0 & 1 \\ 3 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 4 & 1 \\ 0 & 1 & 0 & 4 & 1 \\ 0 & 3 & 1 & 1 & 0 \\ 0 & 5 & 4 & 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & 0 & 0 & 1 & 3 & 5 \\ 5 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

One can check that $\operatorname{rank}(M) = 3$, and, using the algorithm of Aggarwal et al. [1], the restricted nonnegative rank can be computed [8] and is equal to 6. Using the above decomposition, it is clear that $\operatorname{rank}_+(M) \leqslant 5 < \operatorname{rank}_+^*(M) = 6$. By Lemma 3, for any $\operatorname{rank}_+(M)$-nonnegative factorization (U, V) of M, we then must have $3 < \operatorname{rank}(U) = 4 < \operatorname{rank}_+(M)$ implying that $\operatorname{rank}_+(M) = 5$.

As we have already seen with Lemma 1, the restricted nonnegative rank does not share all the nice properties of the rank and the nonnegative rank functions [15]. The next three lemmas exploit Example 2 further to show different behavior between nonnegative rank and restricted nonnegative rank.

[8] The problem is actually trivial because each vertex of the inner polygon $\operatorname{conv}(S)$ is located on a different edge of the polygon P, so that they define with the boundary of P 6 disjoint polygons tangent to $\operatorname{conv}(S)$. This implies that $\operatorname{rank}_+^*(M) = 6$, cf. Section 2.2.1. This matrix is actually a linear Euclidean distance matrix which will be analyzed later in Section 4.2.

Lemma 4. *Let* $A \in \mathbb{R}^{m\times n}_+$ *and* $B \in \mathbb{R}^{m\times r}_+$*, then*

$$\operatorname{rank}^*_+(A+B) \nleqslant \operatorname{rank}^*_+(A) + \operatorname{rank}^*_+(B).$$

Proof. Take M, U and V from Example 2 and construct $A = U(:, 1{:}3)V(1{:}3, :)$ with $\operatorname{rank}^*_+(A) = 3$ (since $\operatorname{rank}(A) = 3$), $B = U(:, 4{:}5)V(4{:}5, :)$ with $\operatorname{rank}^*_+(B) = 2$ (trivial) and $\operatorname{rank}^*_+(A+B) = 6$ since $A + B = M$. □

Lemma 5. *Let* $A \in \mathbb{R}^{m\times n}_+$ *and* $B \in \mathbb{R}^{m\times r}_+$*, then*

$$\operatorname{rank}^*_+([A\,B]) \ngeqslant \operatorname{rank}^*_+(A).$$

where $[A\,B] \in \mathbb{R}^{m\times(n+r)}_+$ *denotes the concatenation of the columns of A and B.*

Proof. Let us take M, U and V from Example 2, and construct $A = M$ and $B = U(:, 1)$. One has $\operatorname{rank}^*_+([A\,B]) \leqslant 5$, because $\operatorname{rank}([A\,B]) = 4$ and $[A\,B] = U[V\,e_1]$ with $\operatorname{rank}(U) = 4$ (where e_i denotes the ith column of the identity matrix of appropriate dimension). □

Lemma 6. *Let* $B \in \mathbb{R}^{m\times r}_+$ *and* $C \in \mathbb{R}^{r\times n}_+$*, then*

$$\operatorname{rank}^*_+(BC) \nleqslant \min(\operatorname{rank}^*_+(B), \operatorname{rank}^*_+(C)).$$

Proof. See Example 2 in which $\operatorname{rank}^*_+(M) = 6$ and $\operatorname{rank}^*_+(U) \leqslant 5$ by Lemma 1. □

3. Lower bounds for the nonnegative rank

In this section, we provide new lower bounds for the nonnegative rank based on the restricted nonnegative rank. Recall that the restricted nonnegative rank already provides an upper bound for the nonnegative rank since for an m-by-n nonnegative matrix M,

$$0 \leqslant \operatorname{rank}(M) \leqslant \operatorname{rank}_+(M) \leqslant \operatorname{rank}^*_+(M) \leqslant n. \tag{3.1}$$

Note that this bound can in general only be computed efficiently in the case $\operatorname{rank}(M) = 3$ (see Theorems 2 and 3).

As mentioned in the introduction, it might also be interesting to compute lower bounds on the nonnegative rank. Some work has already been done in this direction, including the following

1. Let $M \in \mathbb{R}^{m\times n}_+$ be any weighted biadjacency matrix of a bipartite graph $G = (V_1 \cup V_2, E \subseteq V_1 \times V_2)$ with $M(i,j) > 0 \iff (V_1(i), V_2(j)) \in E$. A biclique of G is a complete bipartite subgraph (it corresponds to a positive rectangular submatrix of M). One can easily check that each rank-one factor $(U_{:k}, V_{k:})$ of any rank-k nonnegative factorization (U, V) of M can be interpreted as a biclique of M (i.e., as a positive rectangular submatrix) since $M = \sum_{i=1}^{k} U_{:i}V_{i:}$. Moreover, these bicliques $(U_{:k}, V_{k:})$ must cover G completely since $M = UV$. The minimum number of bicliques needed to cover G is then a lower bound for the nonnegative rank (sometimes referred to as the rectangle covering bound [31]), is called the biclique partition number and is denoted $b(G)$, see [44] and the references therein. Its computation is NP-complete [39] and is directly related to the minimum biclique cover problem (MBC)[9]. Consider for example the matrix M from Example 1. The largest biclique of the graph G generated by M has 4 edges[10]. Since G has 24 edges, we have $b(G) \geqslant \frac{24}{4} = 6$ and therefore $6 \leqslant \operatorname{rank}_+(M) \leqslant \min(m, n) = 6$.

[9] A similar result exists for the cp-rank (when M is symmetric and $U = V$, see Section 1) bounded below by the minimum number of cliques needed to cover the graph generated by M [6].

[10] This can be computed explicitly, e.g., with a brute force approach. Note however that finding the biclique with the maximum number of edges is a combinatorial NP-hard optimization problem [40]. It is closely related to a variant of the approximate nonnegative factorization problem [24].

A crown graph G is a bipartite graph with $|V_1| = |V_2| = n$ and $E = \{(V_1(i), V_2(j)) \mid i \neq j\}$ (it can be viewed as a biclique where the horizontal edges have been removed). de Caen, Gregory and Pullman [21] showed that

$$b(G) = \min_k \left\{ k \mid n \leqslant \binom{k}{\lfloor k/2 \rfloor} \right\} = \Theta(\log n).$$

Beasley and Laffey [3] studied linear Euclidean distance matrices defined as $M(i,j) = (a_i - a_j)^2$ for $1 \leqslant i, j \leqslant n$, $a_i \in \mathbb{R}$, $a_i \neq a_j$ when $i \neq j$. They proved that such matrices have rank three and that

$$\min_k \left\{ k \mid n \leqslant \binom{k}{\lfloor k/2 \rfloor} \right\} \leqslant r_+ \quad \text{which means} \quad n \leqslant \binom{r_+}{\lfloor r_+/2 \rfloor}, \tag{3.2}$$

where $r_+ = \text{rank}_+(M)$. In fact, such matrices are biadjacency matrices of crown graphs (only the diagonal entries are equal to zero).

2. Goemans makes [28] the following observation [11]: if we compute $r_+ = \text{rank}_+(M)$ and $U \in \mathbb{R}_+^{m \times r_+}$ and $V \in \mathbb{R}_+^{r_+ \times n}$ such that $M = UV$, then each column of M is equal to a linear combination of the columns of U:

$$M_{:j} = \sum_{k=1}^{r_+} U_{:k} V_{kj}.$$

Because U and V are nonnegative, the support of $M_{:j}$ is equal to the union of the supports of the columns $U_{:k}$ of U associated with a non-zero weight V_{kj}. In other terms, noting $\text{supp}(x) = \{i \mid x_i \neq 0\}$, we have

$$\text{supp}(M_{:j}) = \bigcup_{\{k \mid V_{kj} \neq 0\}} \text{supp}(U_{:k}), \quad \forall j.$$

Since U has r_+ columns, the columns of M can have at most 2^{r_+} different supports. Letting s_M be the number of columns of $M \in \mathbb{R}_+^{m \times n}$ having different supports, we have

$$s_M \leqslant 2^{r_+} \quad \iff \quad r_+ \geqslant \log_2(s_M),$$

which gives a simple lower bound for the nonnegative rank of M based on the support of its columns. Clearly, the same analysis holds for the rows of M. For example, if all the columns and rows of M have different supports, then

$$\text{rank}_+(M) \geqslant \log_2(\max(m, n)). \tag{3.3}$$

Goemans uses this bound to show that any extended formulation of the permutahedron (polytope whose $n!$ vertices are permutations of $[1, 2, \ldots, n]$) in dimension n must have $\Omega(n \log(n))$ variables and constraints. Observing that the slack matrix of the permutahedron has $n!$ columns (corresponding to each vertex of the polytope) with different supports, combined with Eq. (3.3), gives a lower bound for the minimal size s of any extended formulation of the permutahedron (cf. Section 1):

$$s = \Theta(\text{rank}_+(S_M) + n) \geqslant \Theta(\log(n!)) = \Theta(n \log(n)).$$

In this section, we provide some theoretical results linking the restricted nonnegative rank with the nonnegative rank, which allow us to improve and generalize the above results in Section 4 for both slack and linear Euclidean distance matrices.

[11] The preprint of the paper, available at http://www-math.mit.edu/~goemans/, does not explicitly mention the link with the nonnegative rank. We summarize here the argument given by Goemans during his talk at ISMP09.

3.1. Geometric interpretation of a nonnegative factorization as a nested polytopes problem

In the following, we lay the groundwork for the main results of this paper, introducing essential notations and observations that will be extensively used in this section. We rely on the geometric interpretation of the nonnegative rank, which was first introduced in [29]; see also [22,12,42] where related results are presented.

Our main observation is that any rank-k nonnegative factorization (U, V) of a nonnegative matrix M can be interpreted as the solution with k vertices of a nested polytopes problem in which the ambient space has dimension $\text{rank}(U) - 1$, the outer polytope is full-dimensional while the inner polytope spans an affine subspace of dimension $\text{rank}(M) - 1$. Since, as will be shown later, one can have $\text{rank}(M) < \text{rank}(U)$, the inner polytope is not necessarily full-dimensional and this nested polytopes problem is not necessarily an instance of NPP.

Without loss of generality, let $M \in \mathbb{R}^{m\times n}_+$, $U \in \mathbb{R}^{m\times k}_+$ and $V \in \mathbb{R}^{k\times n}_+$ be column stochastic with $M = UV$ (cf. proof of Theorem 1, the columns of M are convex combination of the columns of U). If the column space of U does not coincide with the column space of M, i.e., $r_u = \text{rank}(U) > \text{rank}(M) = r$, it means that the columns of U belong to a higher dimensional affine subspace containing the columns of M (otherwise, see Theorem 1).

Let factorize $U = AB$ where $A \in \mathbb{R}^{m\times r_u}$ and $B \in \mathbb{R}^{r_u\times r_+}$ are full rank and their columns sum to one. As in Theorem 1, we can construct the polytope of the coefficients of the linear combinations of the columns of A that generate stochastic vectors. It is defined as

$$P_u = \left\{ x \in \mathbb{R}^{r_u-1} \mid f_u(x) = A(:, 1 : r_u - 1)x + \left(1 - \sum_{i=1}^{r_u-1} x_i\right) A(:, r_u) \geqslant 0 \right\}.$$

Since $\text{col}(M) \subseteq \text{col}(U)$, there exists $B' \in \mathbb{R}^{r_u\times n}$ whose columns must sum to one such that $M = AB'$. Since $\text{rank}(M) = r$ and A is full rank, we must have $\text{rank}(B') = r$. By construction, the columns of $B_u = B(1 : r_u - 1, :)$ (corresponding to the columns of U) and $B_m = B'(1 : r_u - 1, :)$ (corresponding to the columns of M) belong to P_u. Note that since $\text{rank}(B') = r$, the columns of B_m live in a polytope P_m spanning an $(r - 1)$-dimensional affine subspace

$$P_m = \{x \in \mathbb{R}^{r_u-1} \mid f_u(x) \geqslant 0, f_u(x) \in \text{col}(M)\} \quad \subseteq \quad P_u.$$

Polytope P_m contains the points in P_u generating vectors in the column space of M. Moreover

$$M = AB' = UV = ABV,$$

implying that (since A is full rank)

$$B' = BV \quad \text{and} \quad B_m = B_u V.$$

Finally, the columns of B_m are contained in the convex hull of the columns of B_u, inside P_u, i.e.,

$$\text{conv}(B_m) \;\subseteq\; \text{conv}(B_u) \;\subseteq\; P_u.$$

Defining the polytope T as the convex hull of the columns of B_u, and the set of points S as the columns of B_m, we can then interpret the nonnegative factorization (U, V) of M as follows. The $(r_u - 1)$-dimensional polytope T with k vertices (corresponding to the columns of U) is nested between an inner $(r - 1)$-dimensional polytope $\text{conv}(S)$ (where each point in S corresponds to a column of M) and an outer $(r_u - 1)$-dimensional polytope P_u.

Let us use the matrix M and its nonnegative factorization (U, V) of Example 2 as an illustration: since $\text{rank}(U) = 4$, our ambient space and outer polytope are three-dimensional, while $\text{rank}(M) = 3$ implies that inner polytope P_m is two-dimensional. The columns of B_u then define a three-dimensional polytope T nested between $\text{conv}(B_m)$ and P_u, see Fig. 2.

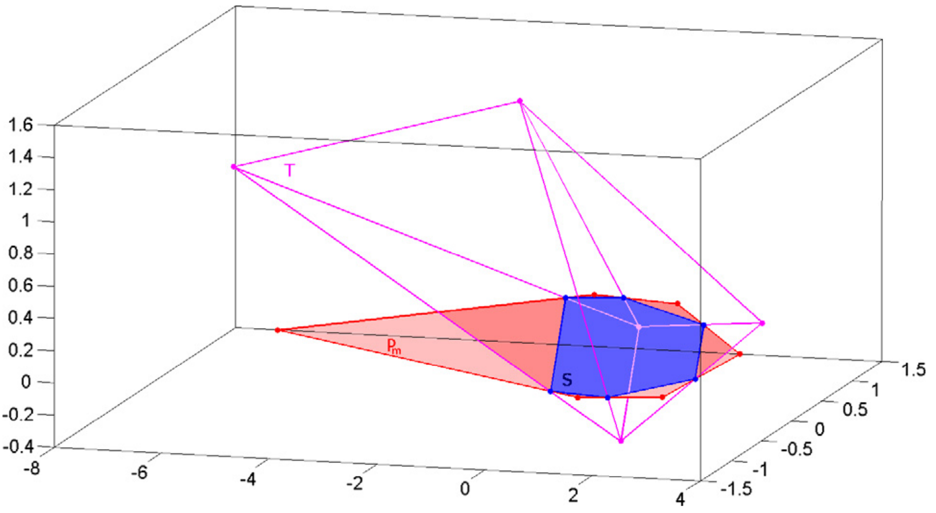

Fig. 2. Illustration of the solution from Example 2 as a nested polytopes problem, with $\text{rank}\,(M) = 3 < \text{rank}\,(U) = 4 < \text{rank}_+(M) = 5 < \text{rank}^*_+(M) = 6 = n$.

3.2. Upper bound for the restricted nonnegative rank

From the geometric interpretation introduced in the previous paragraph, we can now give the main result of this section. The idea is the following: using notations of Section 3.1, we know that (1) the polytope T (whose vertices correspond to the columns of U) contains the (lower dimensional) set of points S (corresponding to the columns of M), and (2) S is contained in P_m (which corresponds to the set of stochastic vectors in the column space of M). Therefore, the intersection between T and P_m must also contain S, i.e., the intersection $T \cap P_m$ defines a polytope which (1) is contained in the column space of M, and (2) contains S. Hence its vertices provide a feasible solution to the RNR problem, from which an upper bound for the restricted nonnegative rank can then be computed. The corresponding NPP problem is simply obtained by projection on the affine subspace spanning P_m.

In other words, any nonnegative factorization (U, V) of a nonnegative matrix M, including those where $\text{rank}(U) > \text{rank}(M)$, can be used to construct a feasible solution to the restricted nonnegative rank problem. One has simply to compute the intersection of the polytope generated by the columns of U with the column space of M (which can obviously increase the number of vertices).

Theorem 4. *Using notations of Section 3.1, we have*

$$\text{rank}^*_+(M) \leqslant \#\,\text{vertices}(T \cap P_m). \tag{3.4}$$

Proof. Let $x_1, x_2, \ldots, x_v$ be the v vertices of $T \cap P_m$ and note $X = [x_1\, x_2\, \ldots x_v]$ which has rank at most r (since it is contained in the $(r-1)$-dimensional polyhedron P_m). By construction,

$$B_m(:, j) \;\in\; T \cap P_m = \text{conv}(X) \quad 1 \leqslant j \leqslant n.$$

Therefore, there must exist a column stochastic matrix $V^* \in \mathbb{R}^{v \times n}$ such that

$$B_m = XV^*,$$

implying that

$$M = f_u(B_m) = f_u(XV^*) = f_u(X)V^* = U^*V^*,$$

where $U^* = f_u(X) \in \mathbb{R}^{m \times v}$ is nonnegative since $x_i \in P_m \subseteq P_u \; \forall i$, and U^* has rank r since $M = U^*V^*$ implies that its rank is at least r and $U^* = f_u(X)$ that it is at most r. The pair (U^*, V^*) is then a feasible solution of the corresponding RNR problem for M and therefore $\text{rank}_+^*(M) \leqslant v = \#\,\text{vertices}(T \cap P_m)$. $\square$

3.3. Lower bound for the nonnegative rank based on the restricted nonnegative rank

We can now obtain a lower bound for the nonnegative rank based on the restricted nonnegative rank. Indeed, if we consider an upper bound on the quantity $\#\,\text{vertices}(T \cap P_m)$ that increases with the nonnegative rank (i.e., the number of vertices of T), we can reinterpret Theorem 4 as providing a lower bound on the nonnegative rank. For that purpose, define the quantity $\text{faces}(n, d, k)$ to be the maximal number of k-faces of a polytope with n vertices in dimension d.

Theorem 5. *The restricted nonnegative rank of a nonnegative matrix M with $r = \text{rank}(M)$ and $r_+ = \text{rank}_+(M)$ can be bounded above by*

$$\text{rank}_+^*(M) \leqslant \max_{r \leqslant r_u \leqslant r_+} \text{faces}(r_+, r_u - 1, r_u - r). \tag{3.5}$$

Proof. Let (U, V) be a rank-r_+ nonnegative factorization of M with $\text{rank}(U) = r_u$. Using notations of Section 3.1 and the result of Theorem 4, $\text{rank}_+^*(M)$ is bounded above by the number of vertices of $T \cap P_m$. Defining $Q_m = \{x \in \mathbb{R}^{r_u - 1} \,|\, f_u(x) \in \text{col}(M)\}$, we have $P_m = Q_m \cap P_u$ and since $T \subseteq P_u$,

$$P_m \cap T = Q_m \cap P_u \cap T = Q_m \cap T.$$

Since Q_m is $(r-1)$-dimensional, the number of vertices of $T \cap Q_m$ is bounded above by the number of $(r_u - r)$-faces of T (in a $(r_u - 1)$-dimensional space, $(r_u - r)$-faces are defined by $r - 1$ equalities), we then have

$$\text{rank}_+^*(M) \leqslant \#\,\text{vertices}(T \cap P_m) = \#\,\text{vertices}(T \cap Q_m) \leqslant \text{faces}(r_+, r_u - 1, r_u - r).$$

Notice that for $r_u = r$, $\text{faces}(r_+, r - 1, 0) = r_+$ which gives $r_+ = \text{rank}_+^*(M)$ as expected. Finally, taking the maximum over all possible values of $r \leqslant r_u \leqslant r_+$ gives the above bound (3.5). $\square$

We introduce for easier reference a function ϕ_r corresponding to the upper bound in Theorem 5, i.e.,

$$\phi_r(r_+) = \max_{r \leqslant r_u \leqslant r_+} \text{faces}(r_+, r_u - 1, r_u - r).$$

Clearly, for a given fixed r, ϕ_r is an increasing function of its argument r_+, since $\text{faces}(n, d, k)$ increases with n. Therefore inequality $\text{rank}_+^*(M) \leqslant \phi_r(r_+)$ from Theorem 5 implicitly provides a lower bound on the nonnegative rank r_+ that depends on both rank r and restricted nonnegative rank $\text{rank}_+^*(M)$.

Explicit values for function ϕ_r can be computed using a tight bound for $\text{faces}(n, d, k)$ attained by cyclic polytopes [46, p.257, Corollary 8.28]

$$\text{faces}(n, d, k-1) = \sum_{i=0}^{\frac{d}{2}}{}^{*} \left(\binom{d-i}{k-i} + \binom{i}{k-d+i} \right) \binom{n-d-1+i}{i},$$

where $\sum^*$ denotes a sum where only half of the last term is taken for $i = \frac{d}{2}$ if d is even, and the whole last term is taken for $i = \lfloor \frac{d}{2} \rfloor = \frac{d-1}{2}$ if d is odd. Alternatively, simpler versions of the bound can be worked out in the following way:

Theorem 6. *The upper bound $\phi_r(r_+)$ on the restricted nonnegative rank of a nonnegative matrix M with $r = \text{rank}(M)$ and $r_+ = \text{rank}_+(M)$ satisfies*

$$\phi_r(r_+) = \max_{r \leqslant r_u \leqslant r_+} \text{faces}(r_+, r_u - 1, r_u - r)$$
$$\leq \max_{r \leqslant r_u \leqslant r_+} \binom{r_+}{r_u - r + 1} \leq \binom{r_+}{\lfloor r_+/2 \rfloor} \leq 2^{r_+}\sqrt{\frac{2}{\pi r_+}} \leq 2^{r_+} .$$

Proof. The first inequality follows from the fact that faces$(n, d, k-1) \leqslant \binom{n}{k}$, since any set of k distinct vertices defines at most one $(k-1)$-face. The second follows from the maximality of central binomial coefficients. The third is a standard upper bound on central binomial coefficients, and the fourth is an even cruder upper bound. □

We will see in Section 4 that some of these weaker bounds correspond to existing results from the literature.

When matrix M is symmetric, the bound can be slightly strengthened, leading to a different function ϕ_r':

Corollary 3. *Given a symmetric matrix M with $r_+ = \text{rank}_+(M)$, $r = \text{rank}(M)$ and $r_+ \geqslant r + 1$, we have*

$$\text{rank}_+^*(M) \leqslant \max_{r \leqslant r_u \leqslant r_+ - 1} \text{faces}(r_+, r_u - 1, r_u - r) = \phi_r'(r_+) \leqslant \phi_r(r_+).$$

Proof. We have seen in Lemma 3 that for symmetric matrices $r_u = r_+$ implies $\text{rank}_+^*(M) = r_+$. Therefore, in case $r_+ \geqslant r + 1$, one can strengthen the result of Theorem 5 and only consider the range $r \leqslant r_u \leqslant r_+ - 1$. □

3.3.1. Improvements in the rank-three case

It is possible to improve the above bound by finding better upper bounds for # vertices$(T \cap P_m)$ in Eq. (3.4). For example, since two-dimensional polytopes (i.e., polygons) have the same number of vertices (0-faces) and edges (1-faces), we have for $\text{rank}(M) = 3$ that

$$\#\,\text{vertices}(T \cap P_m) \quad = \quad \#\text{edges}(T \cap P_m).$$

Using the same argument as in Theorem 5, the number of edges of $T \cap P_m$ is bounded above by the number of $(r_u - r + 1)$-faces of T (defined by $r - 2$ equalities) leading to

Corollary 4. *The restricted nonnegative rank of a rank-three nonnegative matrix M with $r_+ = \text{rank}_+(M)$ can be bounded above with*

$$\text{rank}_+^*(M) \leqslant \max_{3 \leqslant r_u \leqslant r_+} \min_{i=0,1} \text{faces}(r_+, r_u - 1, r_u - 3 + i) \leq \phi_3(r_+). \tag{3.6}$$

The minimum taken between 0 and 1 simply accounts for the two possible cases, i.e., the bound based on # vertices$(T \cap P_m)$ with $i = 0$ as in Theorem 5, or based on #edges$(T \cap P_m)$ with $i = 1$. A similar bound holds in the symmetric case.

Remark 1. Since $\text{rank}_+(M) = \text{rank}_+(M^T)$, all the above bounds can be further improved using both $\text{rank}_+^*(M)$ and $\text{rank}_+^*(M^T)$. Notice that $\text{rank}_+^*(M^T)$ amounts to requiring the second factor V to have the same rank as M, i.e., $\text{rank}(V) = \text{rank}(M)$.

It can be shown using a polar transformation that the NPP instance corresponding to computing $\text{rank}_+^*(M^T)$ is equivalent to the NPP instance of M where one would minimize the number of facets of T instead of its number of vertices. This observation can be used to generalize the improvements described in this section to higher ranks [12], see [23, Section 3.6].

[12] The rank-three case is very special because the number of vertices of any polygon is equal to its number of facets so that $\text{rank}_+^*(M) = \text{rank}_+^*(M^T)$ for $\text{rank}(M) \leqslant 3$.

4. Applications

So far, we have not provided explicit lower bounds for the nonnegative rank. As we have seen, inequalities (3.5) and (3.6) can be interpreted as implicit lower bounds on the nonnegative rank r_+, but have the drawback of depending on the restricted nonnegative rank, which cannot be computed efficiently unless the rank of the matrix is smaller than 3 (Theorems 2 and 3).

Nevertheless, we provide in this Section explicit lower bounds for the nonnegative rank of slack matrices (Section 4.1) and linear Euclidean distance matrices (Section 4.2), cf. introduction of Section 3. These bounds are derived by showing that the restricted nonnegative rank of such matrices is maximum, i.e., it is equal to the number of columns of these matrices (cf. Lemma 1).

4.1. Slack matrices

Let start with a simple observation: it is easy to construct an m-by-n matrix of rank $r < \min(m, n)$ with maximum restricted nonnegative rank n:

1. Take any $(r-1)$-dimensional polytope P with $n \geqslant r+1$ vertices.
2. Construct a NPP instance with $S = \text{vertices}(P)$.
3. Compute the corresponding matrix M in the equivalent RNR instance.

Clearly, the unique solution for NPP is $T = P = \text{conv}(S)$ and therefore the matrix M in the corresponding RNR instance must satisfy: $\text{rank}_+^*(M) = \#\,\text{vertices}(T) = n$; see Example 1 for an illustration with the three-dimensional cube.

Remark 2. Matrices constructed from the procedure described above also satisfy

$$\text{rank}(M) < \text{rank}_+(M).$$

Otherwise $\text{rank}_+^*(M) = \text{rank}_+(M) = \text{rank}(M) < \min(m, n)$ which is a contradiction. This is interesting because it is nontrivial to construct matrices with $\text{rank}(M) < \text{rank}_+(M)$ [37]. In fact, it is easy to check that picking two random nonnegative matrices U and V of dimensions m-by-r and r-by-n respectively, and constructing $M = UV$ will generate with probability one a matrix M of rank r, hence with $\text{rank}(M) = \text{rank}_+(M)$.

In the context of compact extended formulations (cf. Section 1), given a polytope Q with a large (exponential) number of facets, the goal is to find a smaller (polynomial-size) *lifting*, i.e. to use a smaller number of constraints in a higher-dimensional space (i.e., with additional variables). A possible way to do that is to compute a nonnegative factorization of the slack matrix S_M of Q [45] (see Eq. (1.1)). The next theorem states that the restricted nonnegative rank of any slack matrix $S_M \in \mathbb{R}_+^{f \times v}$ is maximum (f is the number of facets of Q, v its number of vertices), i.e., $\text{rank}_+^*(S_M) = v$. This is directly related to the above observation: the slack matrix of a polytope Q corresponds to a NPP instance where Q is the outer polytope and its vertices are the points defining the inner polytope. Notice that the restricted nonnegative rank used as an upper bound for the nonnegative rank is useless in this case.

Theorem 7. *Let $Q = \{x \in \mathbb{R}^q \mid Fx \geqslant h, Ex = g\}$ be a p-dimensional polytope with v vertices, $v > 1$, and let $S_M(Q)$ be its slack matrix, then $\text{rank}_+^*(S_M(Q)) = v$.*

Proof. In order to prove this result, we first construct a bijective transformation L between Q and a full-dimensional polytope $P \subseteq \mathbb{R}^p$. The vertices of P can then be easily constructed from the vertices of Q, which allows to show that P and Q share the same slack matrix. Finally, using the result of Theorem 1, we show that the slack matrix of P has maximum restricted nonnegative rank.

Since Q is a p-dimensional polytope, there exists a polytope $P \subseteq \mathbb{R}^p$ and a bijective affine transformation

$$L : Q \to P : x \to L(x) = Ax + b \quad \text{and} \quad L^{-1} : P \to Q : y \to L^{-1}(y) = A^\dagger y - A^\dagger b,$$

such that $P = L(Q)$ and $Q = L^{-1}(P)$ (where $A \in \mathbb{R}^{p\times q}$ has full rank, $A^\dagger \in \mathbb{R}^{q\times p}$ is its right inverse and $b \in \mathbb{R}^p$).

By construction,

$$\begin{aligned}P &= \{y \in \mathbb{R}^p \mid y = L(x), x \in Q\} = \{y \in \mathbb{R}^p \mid L^{-1}(y) \in Q\},\\ &= \{y \in \mathbb{R}^p \mid FL^{-1}(y) \geqslant h, EL^{-1}(y) = g\},\\ &= \{y \in \mathbb{R}^p \mid FA^\dagger y \geqslant h + FA^\dagger b\},\end{aligned}$$

since the equalities $EL^{-1}(y) = g$ must be satisfied for all $y \in \mathbb{R}^p$ since P is full-dimensional.

Noting $C = FA^\dagger$ and $d = h + FA^\dagger b$, we have $P = \{y \in \mathbb{R}^q \mid Cy \geqslant d\}$. Finally, we observe that

1. Noting q_i $1 \leqslant i \leqslant v$ the vertices of Q, we have that $L(q_i)$'s define the v vertices of P. This can easily be checked since L is bijective ($\forall y \in P$, $\exists! x \in Q$ s.t. $y = L(x)$ and vice versa).
2. P can be taken as the outer polytope of a NPP instance, i.e., P is bounded and $(C\ d)$ is full rank. P is bounded since Q is. C is full rank because P has at least one vertex ($v > 1$). If $(C\ d)$ was not full rank, then $\exists z \in \mathbb{R}^p$ such that $d = Cz$, implying that $z \in P$. Since P has at least two vertices ($v > 1$), $\exists y \in P$ with $y \neq z$, and one can check that $y + \alpha(y - z) \in P\ \forall \alpha \geqslant 0$. This is a contradiction because P is bounded.
3. The slack matrix of P is equal to the slack matrix of Q:

$$\begin{aligned}S_M(P) &= CL(V) - [d\ \dots\ d] = FA^\dagger L(V) - [h + FA^\dagger b\ \dots\ h + FA^\dagger b]\\ &= F(A^\dagger L(V) - [A^\dagger b\ \dots\ A^\dagger b]) - [h\ \dots\ h]\\ &= FL^{-1}(L(V)) - [h\ \dots\ h] = FV - [h\ \dots\ h]\\ &= S_M(Q),\end{aligned}$$

where $V = [v_1\, v_2 \dots v_v]$ is the matrix whose columns are the vertices of Q, and $L(V) = [L(v_1)\, L(v_2) \dots L(v_v)]$ is the matrix whose columns are the vertices of P.
4. The NPP instance with P as the outer polytope and its v vertices $L(q_i)$'s as the set of points S defining the inner polytope has a unique and optimal solution $T = P = \mathrm{conv}(S)$ with v vertices. The matrix M in the RNR instance corresponding to this NPP instance is given by the slack matrix $S_M(P)$ of P implying that its restricted nonnegative rank is equal to v (cf. Theorem 1).

We conclude that $\mathrm{rank}^*_+(S_M(Q)) = v$. $\square$

We can now derive a lower bound on the nonnegative rank of a slack matrix and on the size of an extended formulation, by combining Theorem 5 (cf. Eq. (3.5)), Theorem 6, Theorem 7 and the result of Yannakakis [45] (see also Section 1).

Corollary 5. *Let P be a polytope with v vertices and let $S_M \in \mathbb{R}^{f\times v}_+$ be its slack matrix of rank r (i.e., P has dimension $r-1$), then*

$$v \leqslant \phi_r(r_+) \leqslant \max_{r\leqslant r_u \leqslant r_+} \binom{r_+}{r_u - r + 1} \leqslant \binom{r_+}{\lfloor r_+/2 \rfloor} \leqslant 2^{r_+}, \tag{4.1}$$

where $r_+ = \mathrm{rank}_+(S_M)$. Therefore, the minimum size s of any extended formulation of P follows

$$s = \Theta(r_+ + n) \geqslant \Theta(\phi_r^{-1}(v)) \geqslant \Theta(\log_2(v)).$$

The last bound 2^{r_+} from Eq. (4.1) is the one of Goemans [28, Theorem 1] (see introduction of Section 3), and therefore Corollary 5 provides us with an improved lower bound, even though it is still growing as $\Omega(\log_2(v))$. Actually, a lower bound with faster growth simply cannot exist for the

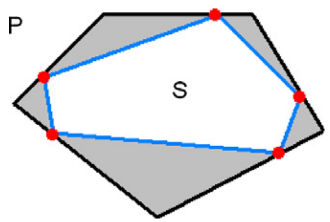

Fig. 3. Illustration of the restricted nonnegative rank of a linear EDM of dimension 5. The solution T must contain a point in each dark region, that is $\text{rank}_+^*(M) = |T| = |S| = 5$.

permutahedron because of the above-mentioned results of Goemans implying that the nonnegative rank of its slack matrix is in $\Theta(n\log(n))$.

4.2. Linear Euclidean distance matrices

Linear Euclidean distance matrices (linear EDM's) are defined by

$$M(i,j) = (a_i - a_j)^2, \quad 1 \leqslant i,j \leqslant n, \text{ for some } a \in \mathbb{R}^n. \tag{4.2}$$

In this section we assume $a_i \neq a_j$ when $i \neq j$, so that these matrices have rank three. Linear EDM's were used in [3] to show that the nonnegative rank of a matrix with fixed rank (rank 3 in this case) can be made as large as desired (while increasing the size of the matrix), implying that an upper bound for the nonnegative rank of a matrix based only on the rank cannot exist.

We refer the reader to [34] and the references therein for detailed discussions about Euclidean distance matrices, and related applications.

4.2.1. Restricted nonnegative rank of linear Euclidean distance matrices

We first show that the restricted nonnegative rank of linear EDM's is maximum, i.e., it is equal to their dimension n.

Definition 2. The columns of a matrix M have disjoint [13] sparsity patterns if and only if

$$s_i \not\subseteq s_j, \quad \forall i \neq j,$$

where $s_i = \{k | M(k,i) = 0\}$ is the sparsity pattern of the ith column of M.

Theorem 8. *Let M be a rank-three nonnegative square matrix of dimension n whose columns have disjoint sparsity patterns, then*

$$\text{rank}_+^*(M) = n.$$

In particular, linear EDM's have this property.

Proof. Let P, S and T be the polygons defined in the two-dimensional NPP instance corresponding to the RNR instance of M (cf. Theorem 1). Aggarwal et al. [1] observe that if two points in S are on different edges of P, they define a polygon with the boundary of P (see each dark regions in Fig. 3) which must contain a point of the solution T. Otherwise these two points could not be contained in T (see also Section 2.2.1). Therefore if each point of S is on a different edge of the boundary of P, any solution T to NPP must have at least $|S| = n$ vertices since S defines n disjoint polygons with the boundary of P. Finally, two points x_1 and x_2 in S are on different edges of the boundary of the polytope $P = \{x \in \mathbb{R}^2 \mid Cx + d \geqslant 0\}$ if and only if $(Cx_1 + d)$ and $(Cx_2 + d)$ have disjoint sparsity patterns or, equivalently, if and only if the two corresponding columns of M (which are precisely equal to $Cx_1 + d$

[13] This definition is slightly abusive since disjoint should refer to sets with an empty intersection.

Table 1
Comparison of the lower bounds for the nonnegative rank of linear EDM's.

Dimension n	4	5	6	7	8	9	10
Eq. (3.6)	4	5	5	6	6	6	7
Eq. (3.5)	4	5	5	5	5	5	6
Beasley and Laffey (3.2)	4	4	4	5	5	5	5
Goemans (3.3)	3	3	3	3	4	4	4

and $Cx_2 + d$) in the RNR instance have disjoint sparsity patterns. Indeed, for two vertices a and b to be located on different edges, one needs at least (1) one inequality that is active at a and inactive at b and (2) another inequality that is active at b and inactive at a. This is equivalent to requiring the sparsity patterns of the corresponding columns of the matrix M to be disjoint. □

Remark 3. This result does not hold for higher rank matrices. For example, the matrix

$$M = \begin{pmatrix} 0 & 1 & 4 & 9 & 16 & 25 \\ 2 & 0 & 1 & 4 & 9 & 16 \\ 8 & 1 & 0 & 1 & 4 & 9 \\ 13 & 4 & 1 & 0 & 1 & 4 \\ 17 & 9 & 4 & 1 & 0 & 1 \\ 25 & 16 & 9 & 4 & 1 & 0 \end{pmatrix} = UV, \quad U = \begin{pmatrix} 0 & 0 & 4 & 5 & 1 \\ 1 & 0 & 1 & 3 & 0 \\ 4 & 0 & 0 & 1 & 1 \\ 4 & 1 & 0 & 0 & 1 \\ 1 & 3 & 1 & 0 & 0 \\ 0 & 5 & 4 & 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 1 \\ 5 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 & 5 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix},$$

has $\operatorname{rank}(M) = 4$ and $\operatorname{rank}_+^*(M) \leqslant 5$ since $\operatorname{rank}(U) = 4$. Therefore we cannot conclude that higher dimensional Euclidean distance matrices have maximal restricted nonnegative rank.

4.2.2. Nonnegative rank of linear Euclidean distance matrices

Since linear EDM's are rank-three symmetric matrices, one can combine the results of Theorem 8 with Corollary 4 (cf. Eq. (3.6)) and Corollary 3 in order to obtain lower bounds for the nonnegative rank of linear EDM's.

Corollary 6. *For any linear Euclidean distance matrix M, we have*

$$\begin{aligned} \operatorname{rank}_+^*(M) = n &\leqslant \max_{3 \leqslant r_u \leqslant r_+ - 1} \min_{i=0,1} \text{faces}(r_+, r_u - 1, r_u - r + i) \\ &\leqslant \max_{3 \leqslant r_u \leqslant r_+ - 1} \text{faces}(r_+, r_u - 1, r_u - r) = \phi_r'(r_+) \\ &\leqslant \binom{r_+}{\lfloor r_+/2 \rfloor} \leqslant 2^{r_+}. \end{aligned}$$

We observe that our results (first two inequalities above, from Theorem 5 and Corollary 4) strengthen the bounds from Equations (3.2) (Beasley and Laffey [3]) and (3.3) (Goemans [28]). Fig. 4 displays the growth of the different bounds, and Table 1 compares the lower bounds on the nonnegative rank for small values of n.

For example, for a linear EDM to be guaranteed to have nonnegative rank 10, the bounds requires respectively $n = 50$ (3.6), $n = 150$ (3.5), $n = 252$ (3.2) and $n = 1024$ (3.3). This is a significant improvement, even though all the bounds are still of the same order with $r_+ = \Omega(\log(n))$.

Is it possible to further improve these bounds? Beasley and Laffey [3] conjectured that the nonnegative rank of linear EDM's is maximum, i.e., it is equal to their dimension. Lin and Chu [37, Theorem 3.1] first claimed to have proved that this equality always holds. However, Chu [11] has recently reported

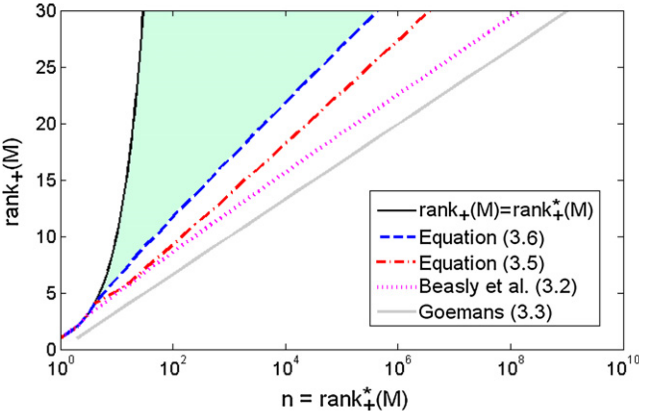

Fig. 4. Comparison of the different bounds for symmetric n-by-n matrices, with $\text{rank}^*_+(M) = n$.

an error in the proof[14]. Indeed, not all linear EDM's have maximum nonnegative rank because of the following example.

Example 3. Taking $M \in \mathbb{R}^{6\times 6}_+$ with

$$M(i,j) = (i-j)^2, \quad 1 \leqslant i,j \leqslant 6,$$

gives $\text{rank}_+(M) = 5$. In fact,

$$M = \begin{pmatrix} 0 & 1 & 4 & 9 & 16 & 25 \\ 1 & 0 & 1 & 4 & 9 & 16 \\ 4 & 1 & 0 & 1 & 4 & 9 \\ 9 & 4 & 1 & 0 & 1 & 4 \\ 16 & 9 & 4 & 1 & 0 & 1 \\ 25 & 16 & 9 & 4 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 4 & 0 & 1 \\ 3 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 4 & 1 \\ 0 & 1 & 0 & 4 & 1 \\ 0 & 3 & 1 & 1 & 0 \\ 0 & 5 & 4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & 3 & 5 \\ 5 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}, \tag{4.3}$$

$$= \begin{pmatrix} 5 & 0 & 1 & 0 & 0 \\ 3 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 3 & 0 & 1 & 0 \\ 0 & 5 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & 3 & 5 \\ 5 & 3 & 1 & 0 & 0 & 0 \\ 0 & 1 & 4 & 4 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 4 & 1 & 0 & 0 & 1 & 4 \end{pmatrix}, \tag{4.4}$$

so that $\text{rank}_+(M) \leqslant 5$, and $\text{rank}_+(M) \geqslant 5$ is guaranteed by Eq. (3.6), see Table 1 with $\text{rank}^*_+(M) = n = 6$ (or by Lemma 3, see Example 2).

[14] In their proof, they actually show that the restricted nonnegative rank is maximum (not the nonnegative rank), see Theorem 8. In fact, they only consider the case when the vertices of the solution T (corresponding to the columns of U) belong to the low-dimensional affine subspace spanned by S (corresponding to the columns of M) in the NPP instance.

Example 3 proves that linear EDM's do not necessarily have a nonnegative rank equal to their dimension. In fact, we can even show that

Theorem 9. *Linear EDM's of the following form*

$$M_n(i,j) = (i-j)^2 \quad 1 \leqslant i,j \leqslant n,$$

satisfy

$$\mathrm{rank}_+(M_n) \leqslant 2 + \left\lceil \frac{n}{2} \right\rceil,$$

where $\lceil x \rceil$ *is the smallest integer greater or equal to x.*

Proof. Let first assume that n is even and define

$$U = \left(\begin{array}{cc|c} n-1 & 0 & \\ n-3 & 0 & \\ \vdots & \vdots & I_{n/2} \\ 3 & 0 & \\ 1 & 0 & \\ \hline 0 & 1 & \\ 0 & 3 & \\ \vdots & \vdots & P_{n/2} \\ 0 & n-3 & \\ 0 & n-1 & \end{array} \right), V = \left(\begin{array}{cc|c} 0 & n-1 & \\ 0 & n-3 & \\ \vdots & \vdots & M_{n/2} \\ 0 & 3 & \\ 0 & 1 & \\ \hline 1 & 0 & \\ 3 & 0 & \\ \vdots & \vdots & P_{n/2}M_{n/2} \\ n-3 & 0 & \\ n-1 & 0 & \end{array} \right)^T,$$

where I_m is the identity matrix of dimension m and P_m is the permutation matrix with $P_m(i,j) = I_m(i, m-j+1)\ \forall i,j$; see Eq. (4.4) for an example when $n = 6$. One can check that

$$M_n = UV = \begin{pmatrix} M_{n/2} & A + P_{n/2}M_{n/2} \\ A^T + P_{n/2}M_{n/2} & M_{n/2} \end{pmatrix}, \quad \text{with } A = \begin{pmatrix} n-1 \\ n-3 \\ \vdots \\ 3 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ \vdots \\ n-3 \\ n-1 \end{pmatrix}^T.$$

If n is odd, we simply observe that $\mathrm{rank}_+(M_n) \leqslant \mathrm{rank}_+(M_{n+1}) \leqslant 2 + \frac{n+1}{2} = 2 + \lceil \frac{n}{2} \rceil$, since M_n is a submatrix of M_{n+1} [15]. □

Remark 4. In the construction of Theorem 9, one can check that $\mathrm{rank}(V) = 4$ and the factorization can then be interpreted as a nested polytopes problem (corresponding to $M^T = V^T U^T$) in which the outer polytope has dimension three only. Therefore, there is still some room for improvement and $\mathrm{rank}_+(M_n)$ is probably (much?) smaller.

This example also demonstrates that, in some cases, the structure of small size nonnegative factorizations (in this case, the one from Example 3) can be generalized to larger size nonnegative factorization problems. This might open new ways to compute large nonnegative factorizations.

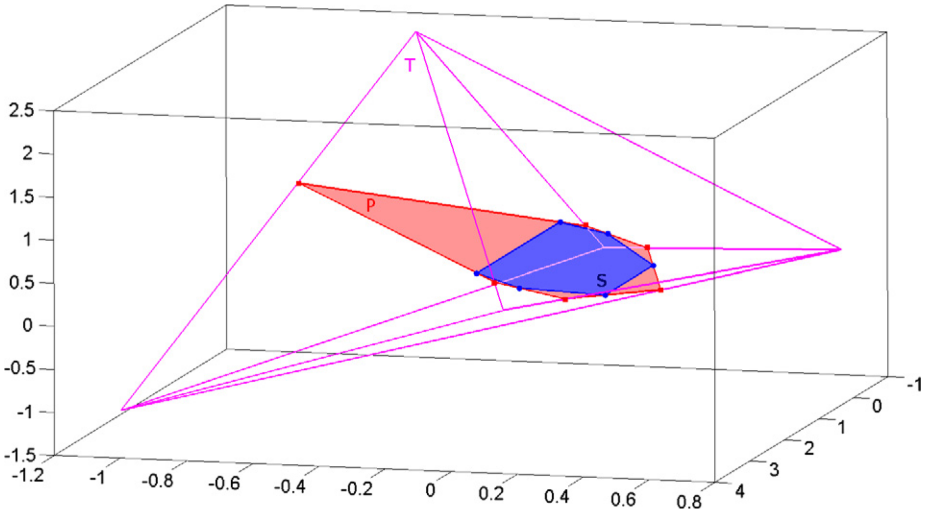

Fig. 5. Illustration of the solution from Eq. (4.4) as a nested polytopes problem, based on a linear EDM with rank $(M) = 3 <$ rank $(U) = 4 <$ $\text{rank}_+(M) = 5 < \text{rank}_+^*(M) = 6 = n$.

In Example 3, the nonnegative rank is smaller than the restricted nonnegative rank because there exists a higher dimensional polytope with only 5 vertices whose convex hull encloses the 6 vertices defined by the columns of M. Nested polytopes instances corresponding to the RNR instance with M given by Example 3 and the two above solutions are illustrated on Figs. 2 and 5 respectively (note that they are transposed to each other, but correspond to different solutions of the NPP instance), see Section 3.1. Notice that the second solution (Fig. 5) completely includes the outer polytope P; therefore, the nonnegative rank of any nonnegative matrix with the same column space as the matrix M will be at most 5.

Remark 5. The solutions of the above nonnegative rank problems have been computed with standard nonnegative matrix factorization algorithms [35,13,25] and, in general, the optimal solution is found after 10 to 100 restarts of these algorithms [15].

4.3. The nonnegative rank of a product

Beasley and Laffey [3] proved that for $A = BC$ with A, B and $C \geqslant 0$

$$\text{rank}_+(A) \leqslant \text{rank}(B)\,\text{rank}(C).$$

In particular, $\text{rank}_+(A^2) \leqslant \text{rank}(A)^2$. They also conjectured that for a nonnegative $n \times n$ matrix A,

$$\text{rank}_+(A^2) \leqslant \text{rank}(A),$$

which we prove to be false with the following counterexample (based on a circulant matrix)

[15] These algorithms are based on standard nonlinear optimization schemes (rescaled gradient descent and block-coordinate descent), and require initial matrices (U, V), which were randomly generated.

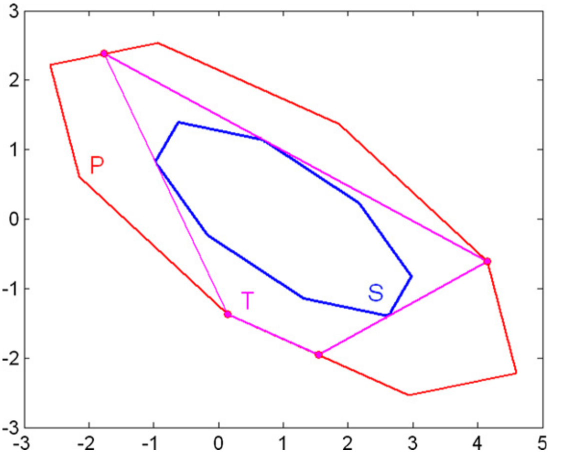

Fig. 6. Illustration of a NPP instance corresponding to A^2 and an optimal solution T, cf. Eq. (4.5).

$$A = \begin{pmatrix} 0 & 1 & a & 1+a & 1+a & a & 1 & 0 \\ 0 & 0 & 1 & a & 1+a & 1+a & a & 1 \\ 1 & 0 & 0 & 1 & a & 1+a & 1+a & a \\ a & 1 & 0 & 0 & 1 & a & 1+a & 1+a \\ 1+a & a & 1 & 0 & 0 & 1 & a & 1+a \\ 1+a & 1+a & a & 1 & 0 & 0 & 1 & a \\ a & 1+a & 1+a & a & 1 & 0 & 0 & 1 \\ 1 & a & 1+a & 1+a & a & 1 & 0 & 0 \end{pmatrix}, \tag{4.5}$$

where $a = 1+\sqrt{2}$. In fact, one can check that $\text{rank}(A) = 3$ and $\text{rank}_+(A^2) = 4$: indeed, $\text{rank}^*_+(A^2) = 4$ can be computed with the algorithm of Aggarwal et al. [1] (see Fig. 6 for an illustration [16]) and, by Corollary 2, $\text{rank}_+(A^2) = \text{rank}^*_+(A^2)$ since $\text{rank}^*_+(A^2) \leqslant \text{rank}(A^2) + 1 = 4$.

Remark 6. The matrix A from Eq. (4.5) is the slack matrix of a regular octagon with sides of length $\sqrt{2}$. By Theorem 7, we have $\text{rank}^*_+(A) = 8$. Notice also that A has rank 3 and its columns have disjoint sparsity patterns so that $\text{rank}^*_+(A) = 8$ is implied by Theorem 8 as well. What is the nonnegative rank of A? Defining

$$R = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

[16] The code used to perform the reduction is available at https://sites.google.com/site/nicolasgillis/code.

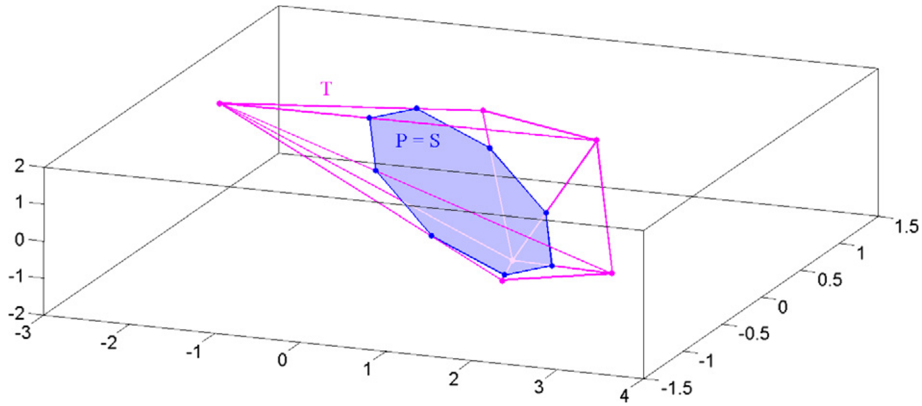

Fig. 7. Illustration of a nested polytopes instance corresponding to A and an optimal solution, cf. Eqs. (4.5) and (4.6).

we have that $B = AR$ is symmetric, has rank 3 and only has zeros on its diagonal. By Theorem 8, $\text{rank}_+^*(B) = 8$. Using Table 1, we have $\text{rank}_+(B) \geqslant 6$. Moreover

$$\text{rank}_+(AR) \leqslant \min(\text{rank}_+(A), \text{rank}_+(R)),$$

implying that $6 \leqslant \text{rank}_+(B) \leqslant \text{rank}_+(A)$. Finally, $\text{rank}_+(A) = 6$ because

$$A = UV = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & a \\ a & 0 & 0 & 0 & 1 & a+1 \\ 1 & 1 & 0 & 0 & 0 & a \\ 0 & a-1 & 1 & 0 & 0 & 1 \\ 0 & 1 & a & 0 & 1 & 0 \\ 1 & 0 & a+1 & 0 & a & 0 \\ 0 & 0 & a & 1 & 1 & 0 \\ 0 & 0 & 1 & a-1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & a & a \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & a & a & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}, \tag{4.6}$$

with $\text{rank}(U) = 4$ and $\text{rank}(V) = 5$. Fig. 7 displays the corresponding nested polytopes problem, see Section 3.1.

It is interesting to observe that, from this nonnegative factorization, one can obtain an extended formulation (lifting) Q of the regular octagon $P = \{x \in R^2 \mid Cx \leqslant d\}$, defined as $Q = \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^6 \mid Cx + Uy = d, y \geqslant 0\}$, with

$$C = \begin{pmatrix} 1 & \sqrt{2}/2 & 0 & -\sqrt{2}/2 & -1 & -\sqrt{2}/2 & 0 & \sqrt{2}/2 \\ 0 & \sqrt{2}/2 & 1 & \sqrt{2}/2 & 0 & -\sqrt{2}/2 & -1 & -\sqrt{2}/2 \end{pmatrix}^T,$$

and $d(i) = 1 + \frac{\sqrt{2}}{2}\ \forall i$, see Eq. (1.2). Since the system of equalities $Cx + Uy = d$ only defines 4 linearly independent equalities ($\text{rank}([C\ U]) = 4$), the description of Q can then be simplified and expressed with 4 variables and 6 inequality constraints.

This extended formulation is actually a particular case of a construction proposed by Ben-Tal and Nemirovski [4] (see also [27,32]) to find an extended formulation of size $\mathcal{O}(k)$ for the regular 2^k-gon in two dimensions.

Table 2
Complexity of restricted nonnegative rank and nonnegative rank computations.

$r = \text{rank}(M)$	$r_+^* = \text{rank}_+^*(M)$	$r_+ = \text{rank}_+(M)$
r not fixed	NP-hard	NP-hard [42]
$r \geqslant 4$ fixed	NP-hard (Theorem 3)	NP-hard?
$r = 3$	Polynomial (Theorem 2)	Polynomial if $r_+^* \leqslant 5$ Otherwise NP-hard?
$r \leqslant 2$	Trivial ($= r$)	Trivial ($= r$) [41]

5. Concluding remarks

In this paper, we have introduced a new quantity called the restricted nonnegative rank, whose computation amounts to solving a problem in computational geometry consisting of finding a polytope nested between two given polytopes. This allowed us to fully characterize its computational complexity (see Table 2). This geometric interpretation and the relationship between the nonnegative rank and the restricted nonnegative rank let us derive new improved lower bounds for the nonnegative rank, in particular for slack matrices and linear Euclidean distance matrices. This also allowed us to provide counterexamples to two conjectures concerning the nonnegative rank.

We conclude the paper with the following conjecture:

Conjecture 1. Computing the nonnegative rank and the corresponding nonnegative factorization of a nonnegative matrix is NP-hard when the rank of the matrix is fixed and greater or equal to 4 (or even possibly 3).

In fact, we have shown that computing a nonnegative factorization amounts to solving a nested polytopes problem in which the outer polytope might live in a higher dimensional space. Moreover, this space is not known a priori (we just know that it contains the columns of the matrix to be factorized and is contained in the unit simplex $\{x \in \mathbb{R}^m | x \geqslant 0, \sum_i x_i = 1\}$, cf. Section 3.1). Therefore, it seems plausible to assume that this problem is at least as difficult as the restricted nonnegative rank computation problem in which the outer polytope lives in the same low-dimensional space and is known. Even in the rank-three case, even though the inner polytope has dimension two, the outer polytope might have any dimension (up to the dimensions of the matrix; see, e.g., Figs. 2 and 5); therefore, it seems that the nonnegative rank computation might also be NP-hard if the rank of the matrix is equal to three. Note that, when $\text{rank}(M) = 3$ and $\text{rank}_+^*(M) \leqslant 5$, Eq. (3.5) implies $\text{rank}_+(M) = \text{rank}_+^*(M)$ so that the nonnegative rank can be computed in polynomial-time in this particular case. Moreover, Arora et al. [2] have shown very recently that checking whether the nonnegative rank of an m-by-n nonnegative matrix is equal to k can be done in polynomial time in m and n (not in k). They have also proved that if the NNR could be checked to be equal to k in time $(mn)^{o(k)}$, then 3-SAT would have a sub-exponential time algorithm, essentially ruling out polynomial time algorithms for checking whether $\text{rank}_+(M) = k$ when k is part of the input. This reinforces the conjecture above, since the nonnegative rank can be arbitrarily large even if the rank of the matrix is fixed.

Table 2 recapitulates the complexity results for the restricted nonnegative rank and the nonnegative rank of a nonnegative matrix M.

Acknowledgments

We thank Mathieu Van Vyve and Didier Henrion for helpful discussions and advice. We thank Santanu Dey for pointing our attention to the paper of Yannakakis [45]. We also would like to thank the anonymous reviewers for their insightful comments, which helped us to improve the paper.

References

[1] A. Aggarwal, H. Booth, J. O'Rourke, S. Suri, Finding minimal convex nested polygons, Inform. Comput. 83 (1) (1989) 98–110.
[2] S. Arora, R. Ge, R. Kannan, A. Moitra, Computing a nonnegative matrix factorization – provably, 2011. arXiv:1111.0952v1.

[3] L. Beasley, T. Laffey, Real rank versus nonnegative rank, Linear Algebra Appl. 431 (12) (2009) 2330–2335.
[4] A. Ben-Tal, A. Nemirovski, On polyhedral approximations of the second-order cone, Math. Oper. Res. 26 (2) (2001) 193–205.
[5] A. Berman, R. Plemmons, Rank factorization of nonnegative matrices, SIAM Rev. 15 (3) (1973) 655.
[6] A. Berman, N. Shaked-Monderer, Completely Positive Matrices, World Scientific Publishing, 2003.
[7] M. Berry, M. Browne, A. Langville, V. Pauca, R. Plemmons, Algorithms and applications for approximate nonnegative matrix factorization, Comput. Stat. Data Anal. 52 (2007) 155–173.
[8] J. Bhadury, R. Chandrasekaran, Finding the set of all minimal nested convex polygons, in: Proceedings of the 8th Canadian Conference on Computational Geometry, 1996, pp. 26–31.
[9] J. Boardman, Geometric mixture analysis of imaging spectrometry data, in: Proc. IGARSS 4, Pasadena, Calif, 1994, pp. 2369–2371.
[10] E. Carlini, F. Rapallo, Probability matrices, non-negative rank, and parameterization of mixture models, Linear Algebra Appl. 433 (2010) 424–432.
[11] M. Chu, On the nonnegative rank of Euclidean distance matrices, II. http://www4.ncsu.edu/~mtchu/Research/Papers/letter04.pdf, 2010.
[12] M. Chu, M. Lin, Low-dimensional polytope approximation and its applications to nonnegative matrix factorization, SIAM J. Sci. Comput. 30 (3) (2008) 1131–1155.
[13] A. Cichocki, R. Zdunek, S. Amari, Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization, Lecture Notes in Computer Science vol. 4666, Springer (2007) 169–176.
[14] K. Clarkson, Algorithms for polytope covering and approximation, in: Proceedings of the Third Workshop on Algorithms and Data Structures, 1993, pp. 246–252.
[15] J. Cohen, U. Rothblum, Nonnegative ranks, decompositions and factorization of nonnegative matrices, Linear Algebra Appl. 190 (1993) 149–168.
[16] M. Conforti, G. Cornuéjols, G. Zambelli, Extended formulations in combinatorial optimization, 4OR: Quart. J. Oper. Res. 10 (1) (2010) 1–48.
[17] M. Craig, Minimum-volume transforms for remotely sensed data, IEEE Trans. Geosci. Remote Sens. 32 (3) (1994) 542–552.
[18] G. Das, Approximation Schemes in Computational Geometry, Ph.D. thesis, University of Wisconsin-Madison, 1990.
[19] G. Das, M. Goodrich, On the complexity of optimization problems for three-dimensional convex polyhedra and decision trees, Comput. Geom. Theory Appl. 8 (1997) 123–137.
[20] G. Das, D. Joseph, The complexity of minimum convex nested polyhedra, in: Proc. of the 2nd Canadian Conference on Computational Geometry, 1990, pp. 296–301.
[21] D. de Caen, D. Gregory, N. Pullman, The boolean rank of zero-one matrices, in: Proc. 3rd Caribbean Conference on Combinatorics and Computing, 1981, pp. 169–173.
[22] D. Donoho, V. Stodden, When does non-negative matrix factorization give a correct decomposition into parts?, Advances in Neural Information Processing vol. 16, MIT Press (2003).
[23] N. Gillis, Nonnegative Matrix Factorization: Complexity, Algorithms and Applications, Ph.D. thesis, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 2011.
[24] N. Gillis, F. Glineur, Nonnegative Factorization and The Maximum Edge Biclique Problem, CORE Discussion paper 2008/64, 2008.
[25] N. Gillis, F. Glineur, Accelerated multiplicative updates and hierarchical ALS algorithms for nonnegative matrix factorization, Neural Comput. 24 (2012) 1085–1105.
[26] N. Gillis, R. Plemmons, Dimensionality reduction, classification, and spectral mixture analysis using nonnegative underapproximation, Opt. Eng. 50 (2011) 027001.
[27] F. Glineur, Computational experiments with a linear approximation of second order cone optimization, Image Technical Report 0001, Service de Mathématique et de Recherche Opérationnelle, Faculté Polytechnique de Mons, 2000.
[28] M. Goemans, Smallest compact formulation for the permutahedron,Talk at ISMP, Chicago, 2009.
[29] H. Hazewinkel, On Positive Vectors, Positive Matrices and the Specialization Ordering, Tech. Rep., CWI Report PM-R8407, 1984.
[30] A. Ifarraguerri, C.-I. Chang, Multispectral and hyperspectral image analysis with convex cones, IEEE Trans. Geosci.Remote Sens. 37 (2) (1999) 756–770.
[31] V. Kaibel, Extended formulations in combinatorial optimization, Optima 85 (2011) 2–7.
[32] V. Kaibel, K. Pashkovich, Constructing extended formulations from reflection relations, in: IPCO'11 Proc. of the 15th Int. Conf. on Integer Programming and Combinatorial Optimization, 2010.
[33] B. Klingenberg, J. Curry, A. Dougherty, Non-negative matrix factorization: ill-posedness and a geometric algorithm, Pattern Recognition 42 (5) (2009) 918–928.
[34] N. Krislock, H. Wolkowicz, Euclidean distance matrices and applications, in: Miguel Anjos, Jean Lasserre (Eds.), Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications, 2011.
[35] D. Lee, H. Seung, Learning the parts of objects by nonnegative matrix factorization, Nature 401 (1999) 788–791.
[36] T. Lee, A. Shraibman, Lower bounds in communication complexity, Foundations and Trends in Theoretical Computer Science, 2009.
[37] M. Lin, M. Chu, On the nonnegative rank of Euclidean distance matrices, Linear Algebra Appl. 433 (3) (2010) 681–689.
[38] J. Mitchell, S. Suri, Separation and approximation of polyhedral surfaces, Oper. Res. Lett. 11 (1992) 255–259.
[39] J. Orlin, Contentment in graph theory: covering graphs with cliques, Indag. Math. (Proceedings) 80 (5) (1977) 406–424.
[40] R. Peeters, The maximum edge biclique problem is NP-complete, Discrete Appl. Math. 131 (3) (2003) 651–654.
[41] L. Thomas, Rank factorization of nonnegative matrices, SIAM Rev. 16 (3) (1974) 393–394.
[42] S. Vavasis, On the complexity of nonnegative matrix factorization, SIAM J. Optim. 20 (2009) 1364–1377.
[43] C. Wang, Finding minimal nested polygons, BIT 31 (1991) 230–236.
[44] V. Watts, Fractional biclique covers and partitions of graphs, Electron. J. Combin. 13 (2006) #R74.
[45] M. Yannakakis, Expressing combinatorial optimization problems by linear programs, Comput. System Sci. 43 (1991) 441–466.
[46] G. Ziegler, Lectures on Polytopes, Springer-Verlag, 1995.