Journal of Machine Learning Research 13 (2012) 3349-3386 Submitted 4/12; Revised 10/12; Published 11/12

Sparse and Unique Nonnegative Matrix Factorization Through Data Preprocessing

Nicolas Gillis* NICOLAS.GILLIS@UCLOUVAIN.BE
ICTEAM Institute
Université catholique de Louvain
B-1348 Louvain-la-Neuve, Belgium

Editor: Inderjit Dhillon

Abstract

Nonnegative matrix factorization (NMF) has become a very popular technique in machine learning because it automatically extracts meaningful features through a sparse and part-based representation. However, NMF has the drawback of being highly ill-posed, that is, there typically exist many different but equivalent factorizations. In this paper, we introduce a completely new way to obtaining more well-posed NMF problems whose solutions are sparser. Our technique is based on the preprocessing of the nonnegative input data matrix, and relies on the theory of M-matrices and the geometric interpretation of NMF. This approach provably leads to optimal and sparse solutions under the separability assumption of Donoho and Stodden (2003), and, for rank-three matrices, makes the number of exact factorizations finite. We illustrate the effectiveness of our technique on several image data sets.

Keywords: nonnegative matrix factorization, data preprocessing, uniqueness, sparsity, inverse-positive matrices

1. Introduction

Given an m-by-n nonnegative matrix $M \geq 0$ and a factorization rank r, nonnegative matrix factorization (NMF) looks for two nonnegative matrices U and V of dimension m-by-r and r-by-n respectively such that $M \approx UV$. To assess the quality of an approximation, a popular choice is the Frobenius norm of the residual $||M-UV||_F$ and NMF can for example be formulated as the following optimization problem

$$\min_{U \in \mathbb{R}^{m \times r}, V \in \mathbb{R}^{r \times n}} ||M-UV||_F^2 \quad \text{such that } U \geq 0 \text{ and } V \geq 0. \tag{1}$$

Assuming that M is a matrix where each column represents an element of a data set (for example, a vectorized image of pixel intensities), NMF can be interpreted in the following way. Since $M_{:j} \approx \sum_{k=1}^{r} U_{:k} V_{kj} \ \forall j$, each column $M_{:j}$ of M is reconstructed using an additive linear combination of nonnegative basis elements (the columns of U). These basis elements can be interpreted in the same way as the columns of M (for example, as images). Moreover, they can only be summed up (since V is nonnegative) in order to approximate the original data matrix M which leads to a part-based representation: NMF will automatically extract localized and meaningful features from the data set.

*. The author is a postdoctoral researcher with the fonds de la recherche scientifique-FNRS (F.R.S.-FNRS). 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 author.

The most famous illustration of such a decomposition is when the columns of M represent facial images for which NMF is able to extract common features such as eyes, noses and lips (Lee and Seung, 1999); see Figure 8 in Section 6.

NMF has become a very popular data analysis technique and has been successfully used in many different areas such as hyperspectral imaging (Pauca et al., 2006), text mining (Xu et al., 2003), clustering (Ding et al., 2005), air emission control (Paatero and Tapper, 1994), blind source separation (Cichocki et al., 2009), and music analysis (Févotte et al., 2009).

1.1 Geometric Interpretation of NMF

A very useful tool for understanding NMF better is its geometric interpretation. In fact, NMF is closely related to a problem in computational geometry consisting in finding a polytope nested between two given polytopes. In this section, we briefly recall this connection, which will be extensively used throughout the paper.

Let (U,V) be an exact NMF of M (that is, $M=UV$, $U\geq 0$ and $V\geq 0$), and let us assume that no column of U or M is all zeros; otherwise they can be removed without loss of generality.

Definition 1 (Pullback map) *Given an m-by-n nonnegative matrix X without all-zero column, $D(X)$ is the n-by-n diagonal matrix whose diagonal elements are the inverse of the ℓ_1-norms of the columns of X:*

$$D(X)_{ii}=||X_{:i}||_1^{-1}=\left(\sum_{k=1}^{m}|X_{ki}|\right)^{-1}\quad\forall i,\quad D(X)_{ij}=0\;\forall\, i\neq j, \tag{2}$$

and $\theta(X)=XD(X)$ is the pullback map of X so that $\theta(X)$ is column stochastic, that is, $\theta(X)$ is nonnegative and its columns sum to one.

We have that (see Chu and Lin, 2008)

$$M=UV \iff \theta(M)=MD(M)=\underbrace{UD(U)}_{\theta(U)}\underbrace{D(U)^{-1}VD(M)}_{V'} \iff \theta(M)=\theta(U)V',$$

where V' must be column stochastic since $\theta(M)$ and $\theta(U)$ are both column stochastic and $\theta(M)=\theta(U)V'$. Therefore, the columns of $\theta(M)$ are convex combinations (linear combinations with nonnegative weights summing to one) of the columns of $\theta(U)$. This implies that

$$\text{conv}(\theta(M)) \quad\subseteq\quad \text{conv}(\theta(U)) \quad\subseteq\quad \Delta^m, \tag{3}$$

where $\text{conv}(X)$ denotes the convex hull of the columns of matrix X, and $\Delta^m=\{x\in\mathbb{R}^m \mid \sum_i^m x_i=1, x_i\geq 0\; 1\leq i\leq m\}$ is the unit simplex (of dimension $m-1$). An exact NMF $M=UV$ can then be geometrically interpreted as a polytope $T=\text{conv}(\theta(U))$ nested between an inner polytope $\text{conv}(\theta(M))$ and an outer polytope Δ^m.

> Hence finding the minimal number of nonnegative rank-one factors to reconstruct M exactly is equivalent to finding a polytope T with minimum number of vertices nested between two given polytopes: the inner polytope $\text{conv}(\theta(M))$ and the outer polytope Δ^m.

This problem is referred to as the nested polytopes problem (NPP), and is then equivalent to computing an exact nonnegative matrix factorization (Hazewinkel, 1984); see also Gillis and Glineur (2012a) and the references therein. In the remaining of the paper, we will denote NPP(M) the NPP instance corresponding to M with inner polytope $\text{conv}(\theta(M))$ and outer polytope Δ^m.

Remark 2 *The geometric interpretation can also be equivalently characterized in terms of cones, see Donoho and Stodden (2003), for which we have*

$$\text{cone}(M) \quad \subseteq \quad \text{cone}(U) \quad \subseteq \quad \mathbb{R}^m_+,$$

where $\text{cone}(X) = \{x | x = Xa, a \geq 0\}$. *The geometric interpretation based on convex hulls from Equation* (3) *amounts to the intersection of the cones with the hyperplane* $\{x | \sum x_i = 1\}$ *(this is the reason why zero columns of M and U need to be discarded in that case).*

1.2 Uniqueness of NMF

There are several difficulties in using NMF in practice. In particular, the optimization problem (1) is NP-hard (Vavasis, 2009), and typically only convergence to stationary points is guaranteed by standard algorithms. There does not seem to be an easy way to go around this (except if the factorization rank is very small, see Arora et al., 2012) since NMF problems typically have many local minima.

Another difficulty is the non-uniqueness: even if one is given an optimal (or good) NMF (U,V) of M, there might exist many equivalent solutions $(UQ, Q^{-1}V)$ for non-monomial[1] matrices Q with $UQ \geq 0$ and $Q^{-1}V \geq 0$, see Laurberg et al. (2008). Such transformations lead to different interpretations, especially when the supports of U and V change. For example, in document classification, each entry M_{ij} of matrix M indicates the 'importance' of word i in document j (for example, the number of appearances of word i in text j). The factors (U,V) of NMF are interpreted as follows: the columns of U represent the topics (that is, bags of words) while the columns of V link the documents to these topics. The sparsity patterns of U and V are then a crucial characteristic since they indicate which words belong to which topics and which topics is discussed by which documents.

Different approaches exist to obtain (more) well-posed NMF problems and most of them are based on the incorporation of additional constraints into the NMF model, for example,

- **Sparsity**. Require the factors in NMF to be sparse. Under some appropriate assumptions, this leads to a unique solution (Theis et al., 2005). Geometrically, requiring the matrix U to be sparse is equivalent to requiring the vertices of the nested polytope $\text{conv}(\theta(U))$ to be located on the low-dimensional faces of the outer polytope Δ^m, hence making the problem more well posed. In practice, the most popular technique to obtain sparser solutions is to add sparsity inducing penalty terms, such as a ℓ_1-norm penalty (Kim and Park, 2007) (see also Section 6). Another possibility is to use a projection onto the set of sparse matrices (Hoyer, 2004).

- **Minimum Volume**. Require the polytope $\text{conv}(\theta(U))$ to have minimum volume (Miao and Qi, 2007; Huck et al., 2010; Zhou et al., 2011) which has a long history in hyperspectral imaging (Craig, 1994). Again, this constraint is typically enforced using a proper penalty term in the objective function. Volume maximization of $\text{conv}(\theta(U))$ is also possible, leading to a sparser factor U (since the columns of U will be encouraged to be on the faces of Δ^m), see Wang et al. (2010), which is essentially equivalent to performing volume minimization for the matrix transpose. In fact, taking the polar of the three polytopes in Equation (3) interchanges the role of the inner and outer polytopes, while the polar of $\text{conv}(\theta(M))$ is given by $\text{conv}(\theta(M^T))$, see Gillis (2011, Section 3.6).

1. A monomial matrix is a permutation of a diagonal matrix with positive diagonal elements.

- **Orthogonality**. Require the columns of matrix U to be orthogonal (Ding et al., 2006). Geometrically, it amounts to position the vertices of $\text{conv}(\theta(U))$ on the low-dimensional faces of Δ^m so that if one of the columns of $\theta(U)$ is not on a facet of Δ^m (that is, $U_{ik} > 0$ for some i,k), then all the other columns of U must be on that facet (that is, $U_{ip} = 0\ \forall p \neq k$). This condition is rather restrictive, but proved successful in some situations, for example in clustering; see Ding et al. (2005) and Pompili et al. (2011).

1.3 Outline of the Paper

In this paper, we address the problem of uniqueness and introduce a completely new approach to make NMF problems more well posed, and obtain sparser solutions. Our technique is based on a preprocessing of the input matrix M to make it sparser while preserving its nonnegativity and its column space. The motivation is based on the geometric interpretation of NMF which shows that *sparser matrices will correspond to more well-posed NMF problems whose solutions are sparser*.

In Section 2, we recall how sparsity of M makes the corresponding NMF problem more well posed. In particular, we give a new result linking the support of M and the uniqueness of the corresponding NMF problem. In Section 3, we introduce a preprocessing $\mathcal{P}(M) = MQ$ of M where Q is an inverse-positive matrix, that is, Q has full rank and its inverse Q^{-1} is nonnegative. Hence, if (U,V') is an NMF of $\mathcal{P}(M)$ with $\mathcal{P}(M) \approx UV'$, then $(U,V'Q^{-1})$ is an NMF of M since $M = \mathcal{P}(M)Q^{-1} \approx UV'Q^{-1}$ and $V'Q^{-1} \geq 0$. In Section 4, we prove some important properties of the preprocessing; in particular that it is well-defined, invariant to permutation and scaling, and optimal under the separability assumption of Donoho and Stodden (2003). Moreover, in the exact case for rank-three matrices (that is, $M = UV$ and $\text{rank}(M) = 3$), we show how the preprocessing can be used to obtain an equivalent NMF problem with a finite number of solutions. In Section 5, we address some practical issues of using the preprocessing: the computational cost, the rescaling of the columns $\mathcal{P}(M)$ and the ability to dealing with sparse and noisy matrices. In Section 6, we present some very promising numerical experiments on facial and hyperspectral image data sets.

2. Non-Uniqueness, Geometry and Sparsity

Let $M \in \mathbb{R}^{m\times n}_+$ and $(U,V) \in \mathbb{R}^{m\times r}_+ \times \mathbb{R}^{r\times n}_+$ be an exact nonnegative matrix factorization of $M = UV$. The minimum r such that such a decomposition exists is the nonnegative rank of M and will be denoted $\text{rank}_+(M)$. If U is not full rank (that is, $\text{rank}(U) < r$), then the decomposition is typically not unique. In fact, the convex combinations (given by $V \geq 0$) cannot in general be uniquely determined: the polytope $T = \text{conv}(\theta(U))$ has r vertices while its dimension is strictly smaller than $r-1$ implying that any point in the interior of T can be reconstructed with infinitely many convex combinations of the r vertices of T. However, if all columns of $\text{conv}(\theta(M))$ are located on k-dimensional faces of T having exactly $k+1$ vertices, then the convex combinations given by V are unique (Sun and Xin, 2011).

In practice, it is therefore often implicitly assumed that $\text{rank}_+(M) = \text{rank}(M) = r$ hence $\text{rank}(U) = r$ (since U has r columns and spans the column space of M of dimension r); see the discussion by Arora et al. (2012) and the references therein. In this situation, the uniqueness can be characterized as follows:

Theorem 3 (Laurberg et al., 2008) *Let $(U,V) \in \mathbb{R}^{m\times r}_+ \times \mathbb{R}^{r\times n}_+$ and $M = UV$ with $\text{rank}(M) = \text{rank}(U) = r$. Then the following statments are equivalent:*

(i) The exact NMF (U,V) of M is unique (up to permutation and scaling).

(ii) There does not exist a non-monomial invertible matrix Q such that $U' = UQ \geq 0$ and $V' = Q^{-1}V \geq 0$.

(iii) The polytope $\text{conv}(\theta(U))$ *is the unique solution of NPP(M) with r vertices.*

It is interesting to notice that the columns of M containing zero entries are located on the boundary of the outer polytope Δ^m, and these points must be on the boundary of any solution T of NPP(M). Therefore, if M contains many zero entries, it is more likely that the set of exact NMF of M will be smaller, since there is less degree of freedom to fill in the space between the inner and outer polytopes. In particular, Donoho and Stodden (2003) showed that "requiring that some of the data are spread across the faces of the nonnegative orthant, there is unique simplicial cone", that is, there is a unique $\text{conv}(\theta(U))$.

In the following, based on the assumption that $\text{rank}(M) = \text{rank}_+(M)$, we provide a new uniqueness result using the geometric interpretation of NMF and the sparsity pattern of M.

Lemma 4 *Let $M \in \mathbb{R}^{m\times n}$ with $r = \text{rank}(M) = \text{rank}_+(M)$, and M have no all-zero columns. If r columns of $\theta(M)$ coincide with r different vertices of $\Delta^m \cap \text{col}(\theta(M))$, then the exact NMF of M is unique.*

Proof Let $(U,V) \in \mathbb{R}_+^{m\times r} \times \mathbb{R}_+^{r\times n}$ be such that $M = UV$. Since $r = \text{rank}(M) = \text{rank}_+(M)$, we must have $\text{rank}(U) = r$ and $\text{col}(U) = \text{col}(M)$ (where $\text{col}(X)$ denotes the column space of matrix X), hence

$$\text{conv}(\theta(M)) \subseteq \text{conv}(\theta(U)) \subseteq \Delta^m \cap \text{col}(\theta(M)).$$

Since r columns of $\theta(M)$ coincide with r vertices of $\Delta^m \cap \text{col}(\theta(M))$, we have that $\text{conv}(\theta(U)) = \text{conv}(\theta(M))$ is the unique solution of NPP(M), and Theorem 3 allows to conclude. ■

In order to identify such matrices, it would be nice to characterize the vertices of $\Delta^m \cap \text{col}(\theta(M))$ based solely on the sparsity pattern of M. By definition, the vertices of $\Delta^m \cap \text{col}(\theta(M))$ are the intersection of $r-1$ of its facets, and the facets of $\Delta^m \cap \text{col}(\theta(M))$ are given by

$$F_i = \{x \in \Delta^m \cap \text{col}(\theta(M)) \mid x_i = 0\}.$$

Therefore, a vertex of $\Delta^m \cap \text{col}(\theta(M))$ must contain at least $r-1$ zero entries. However, this is not a sufficient condition because some facets might be redundant, for example, if the ith row of M is identically equal to zero (for which $F_i = \Delta^m \cap \text{col}(\theta(M))$) or if the ith and jth row of M are equal to each other (for which $F_i = F_j$).

Lemma 5 *A column of M containing $r-1$ zeros whose corresponding rows have different sparsity patterns corresponds to a vertex of* $\text{conv}(\theta(M)) \cap \Delta^m$.

Proof Let c be one of the columns of M with at least $r-1$ zeros corresponding to rows with different sparsity patterns, that is, there exists $\mathcal{J} \subseteq \{i \mid c_i = 0\}$ with $|\mathcal{J}| = r-1$ such that the rows of $M(\mathcal{J},:)$ have different sparsity patterns. Let also $F_k = \{x \mid x_{\mathcal{J}(k)} = 0\}$ for $1 \leq k \leq r-1$ denote the $r-1$ facets with $\theta(c) \in F_k \ \forall k$. To show that $\theta(c)$ is a vertex of $\text{conv}(\theta(M)) \cap \Delta^m$, it suffices to show that the

$r-1$ facets are not redundant: for all $1 \leq k < p \leq r-1$, there exist x_k and x_p in $\text{conv}(\theta(M)) \cap \Delta^m$ such that $x_k \in F_k, x_k \notin F_p$ and $x_p \in F_p, x_p \notin F_k$. Because the rows of $M(\mathcal{J},:)$ have different sparsity patterns, for all $1 \leq k < p \leq r-1$, there must exist two indices h and l such that $M(\mathcal{J}(k),h) = 0$ and $M(\mathcal{J}(p),h) > 0$ while $M(\mathcal{J}(k),l) > 0$ and $M(\mathcal{J}(p),l) = 0$. Therefore, $\theta(M_{:h}) \in F_k, \theta(M_{:h}) \notin F_p$ and $\theta(M_{:l}) \in F_p, \theta(M_{:l}) \notin F_k$ and the proof is complete. ■

Theorem 6 *Let $M \in \mathbb{R}^{m \times n}$ with $r = \text{rank}(M) = \text{rank}_+(M)$. If M has r non-zero columns each having $r-1$ zero entries whose corresponding rows have different sparsity patterns, then the NMF of M is unique.*

Proof This follows directly from Lemma 4 and 5. ■

Here is an example,

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

with $\text{rank}(M) = \text{rank}_+(M) = 3$ whose unique NMF is $M = MI$, where I is the identity matrix of appropriate dimension. Other examples include matrices containing an r-by-r monomial submatrix; see also Kalofolias and Gallopoulos (2012) and the references therein. It is interesting to notice that this result implies that the only 3-by-3 rank-three nonnegative matrices having a unique exact NMF are the monomial matrices (permutation and scaling of the identity matrix) since all other matrices have at least two distinct exact NMF: $M = MI = IM$.

Finally, although sparsity is neither a necessary (see Remark 7 below) nor a sufficient condition for uniqueness (except in some cases, see for example Theorem 6 or Donoho and Stodden, 2003), the geometric interpretation of NMF shows that sparser matrices M lead to more well-posed NMF problems. In fact, many points of the inner polytope in NPP(M) are located on the boundary of the outer polytope Δ^m. Moreover, because the solution T must contain these points, it will have zero entries as well. In particular, assuming M does not contain a zero column, it is easy to check that for $M = UV$ we have

$$M_{ij} = 0 \quad \Rightarrow \quad \exists k \text{ such that } U_{ik} = 0.$$

Remark 7 *Having many zero entries in M is not a necessary condition for having an unique NMF. In fact, Laurberg et al. (2008) showed that there exist positive matrices with unique NMF. However, for an NMF (U,V) to be unique, the support of each column of U (resp. row of V) cannot be contained in the support of any another column (resp. row) so that each column of U (resp. row of V) must have at least one zero entry. In fact, assume the support of the kth column of U is contained in the support of lth column. Then noting $\bar{p} = \text{argmin}_{\{p|U(p,k)\neq 0\}} \frac{U(p,l)}{U(p,k)}$, $\varepsilon = \frac{U(\bar{p},l)}{U(\bar{p},k)}$, and*

$$D_{kl} = -\varepsilon, \quad D_{ii} = 1 \; \forall i, \quad D_{ij} = 0 \text{ otherwise},$$

one can check that D^{-1} is as follows

$$D^{-1}_{kl} = \varepsilon, \quad D^{-1}_{ii} = 1 \; \forall i, \quad D^{-1}_{ij} = 0 \text{ otherwise},$$

that is $D^{-1} \geq 0$. Therefore $(UD, D^{-1}V)$ is an equivalent NMF with a different sparsity pattern since $(UD)_{:l} = UD_{:l} = U_{:l} - \varepsilon U_{:k} \geq 0$, and $U_{\bar{p}l} > 0$ while $(UD)_{\bar{p}l} = 0$.

3. Preprocessing for More Well-Posed and Sparser NMF

In this section, we introduce a completely new approach to obtain more well-posed NMF problems whose solutions are sparser. As it was shown in the previous paragraph, this can be achieved by working with sparser nonnegative matrices. Hence, we look for an n-by-n matrix Q such that $MQ = M'$ is nonnegative, sparse and Q is inverse-positive. In other words, we would like to solve the following problem:

$$\min_{Q \in \mathbb{R}^{n \times n}} ||MQ||_0 \quad \text{such that} \quad MQ \geq 0 \text{ and } Q^{-1} \geq 0, \tag{4}$$

where $||X||_0$ is the ℓ_0-'norm' which counts the number of non-zero entries in X. Assuming we can solve (4) and obtain a matrix $M' = MQ$, then any NMF (U, V') of M' with $M' \approx UV'$ gives a NMF for M. In fact,

$$M \;=\; M'Q^{-1} \;\approx\; UV'Q^{-1} \;=\; UV, \;\text{ where } V = V'Q^{-1} \geq 0,$$

for which we have

$$||M - UV||_F = ||M'Q^{-1} - UV'Q^{-1}||_F = ||(M' - UV')Q^{-1}||_F \leq ||M' - UV'||_F \; ||Q^{-1}||_2.$$

In particular, if the NMF of M' is exact, then we also have an exact NMF for $M = M'Q^{-1} = UV'Q^{-1} = UV$. The converse direction, however, is not always true. We return to this point in Section 4.3.

In the remaining of this section, we propose a way to finding approximate solutions to problem (4). First, we briefly review some properties of inverse-positive matrices (Section 3.1) in order to deal with the constraint $Q^{-1} \geq 0$. Then, we replace the ℓ_0-'norm' with the ℓ_2-norm and solve the corresponding optimization problem using constrained linear least squares (Section 3.2).

3.1 Inverse-Positive Matrices

In this section, we recall the definition of three types of matrices: Z-matrices, M-matrices and inverse-positive matrices, briefly recall how they are related and provide some useful properties. We refer the reader to the book of Berman and Plemmons (1994) and the references therein for more details on the subject.

Definition 8 *An n-by-n Z-matrix is a real matrix with non-positive off-diagonal entries.*

Definition 9 *An n-by-n M-matrix is a real matrix of the following form:*

$$A = sI - B, \quad s > 0, \quad B \geq 0,$$

where the spectral radius[2] $\rho(B)$ *of B satisfies* $s \geq \rho(B)$.

It is easy to see that an M-matrix is also a Z-matrix.

Definition 10 *An n-by-n matrix Q is inverse positive if and only if* Q^{-1} *exists and* Q^{-1} *is nonnegative. We will denote this set* $I\mathcal{P}^n$*:*

$$I\mathcal{P}^n = \{Q \in \mathbb{R}^{n \times n} \mid Q \text{ is full rank and } Q^{-1} \geq 0\}.$$

2. The spectral radius $\rho(B)$ of a n-by-n matrix B is the supremum among all the absolute values of the eigenvalues of B: $\rho(B) = \max_i |\lambda_i(B)|$.

It can be shown that inverse-positive Z-matrices are M-matrices:

Theorem 11 (Berman and Plemmons 1994, Theorem 2.3) *Let A be a Z-matrix. Then the following conditions are equivalent :*

- *A is an invertible M-matrix.*
- $A = sI - B$ *with* $B \geq 0$, $s > \rho(B)$.
- $A \in I\mathcal{P}^n$.

Here is another well-known theorem in matrix theory which will be useful, see Taussky (1949) and the references therein.

Definition 12 *An n-by-n matrix A is irreducible if and only if there does not exist an n-by-n permutation matrix P such that*

$$P^T A P = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix},$$

where B and D are square matrices.

Definition 13 *An n-by-n matrix A is irreducibly diagonally dominant if A is irreducible,*

$$|A_{ii}| \geq \sum_{k \neq i} |A_{ki}|, \quad \text{for } i = 1, 2, \ldots, n,$$

and the inequality is strict for at least one i.

Theorem 14 *If A is irreducibly diagonally dominant, then A is nonsingular.*

3.2 Constrained Linear Least Squares Formulation for (4)

The ℓ_0-'norm' is of combinatorial nature and typically leads to intractable optimization problems. The standard approach is to use the ℓ_1-norm instead but we propose here to use the ℓ_2-norm. The reason is twofold:

- When looking at the structure of problem (4), we observe that any (reasonable) norm will induce solutions with zero entries. In fact, some of the constraints $MQ \geq 0$ will always be active at optimality because of the objective function $||MQ||$.
- The ℓ_2-norm is smooth hence its optimization can be performed more efficiently.[3]

We then would like to solve

$$\min_{Q \in I\mathcal{P}^n} ||MQ||_F^2 \quad \text{such that} \quad MQ \geq 0. \tag{5}$$

Optimizing over the set of inverse-positive matrices $I\mathcal{P}^n$ seems to be very difficult. At least, describing $I\mathcal{P}^n$ explicitly as a semi-algebraic set requires about n^2 polynomial inequalities of degree

3. Because of the constraint $MQ \geq 0$, the ℓ_1-norm problem can actually be decoupled into n linear programs (LP) in n variables and $m+n$ constraints, and can be solved effectively. However, in the noisy case (cf. Section 5.3), we would need to introduce mn auxiliary variables (one for each term of the objective function) which turns out to be impractical.

up to n, each with up to $n!$ terms. However, we are not aware of a rigorous analysis of the complexity of this type of problems; this is a topic for further research.

For this reason, we will restrict the search space to the subset of Z-matrices, that is, inverse-positive matrices of the form $Q = sI - B$, where s is a nonnegative scalar, I is the identity matrix of appropriate dimension and B is a nonnegative matrix such that $\rho(B) < s$, see Section 3.1. It is important to notice that

- The scalar s cannot be chosen arbitrarily. In fact, making s go to zero and $B = 0$, the objective function value goes to zero, which is optimal for (5). The same degree of freedom is in fact present in the original problem (4) since Q and αQ for any $\alpha > 0$ are equivalent solutions. Therefore, without loss of generality, we fix s to one .

- The diagonal entries of B cannot be chosen arbitrarily. In fact, taking B arbitrarily close (but smaller) to the identity matrix, the infimum of (5) will be equal to zero. We then have to set an upper bound (smaller than one) for the diagonal entries of B. It can be checked that this upper bound will always be attained (because of the minimization), and that the optimal solutions corresponding to different upper bounds will be multiples of each other. We therefore fix the bound to zero implying $B_{ii} = 0$ for all i so that $Q_{ii} = 1$ for all i.

Finally, we would like to solve

$$\min_{Q \in \mathcal{Q}^n} \quad ||MQ||_F^2 \quad \text{such that} \quad MQ \geq 0,$$

where

$$\mathcal{Q}^n = \{Q \in \mathbb{R}^{n \times n} \mid Q = I - B, B \geq 0, B_{ii} = 0 \; \forall i, \rho(B) < 1\} \subset I\mathcal{P}^n.$$

Since $MQ = M(I - B) \geq 0$, this problem is equivalent to

$$\begin{aligned} \min_{B \in \mathbb{R}^{n \times n}} \quad & \sum_{i=1}^{n} \left\| M_{:i} - \sum_{k \neq i} M_{:k} B_{ki} \right\|_2^2 \\ \text{such that} \quad & M \geq MB, \\ & \rho(B) < 1, \\ & B_{ii} = 0 \; \forall i, \; B \geq 0. \end{aligned} \tag{6}$$

Without the constraint on the spectral radius of B, this is a constrained linear least squares problem (CLLS) in $O(n^2)$ variables and $O(n^2 + mn)$ constraints. The ith column of $M' = MQ$, which is the preprocessed version of M, will then be given by the following linear combination

$$M'_{:i} = MQ_{:i} = M_{:i} - \sum_{k=1}^{n} M_{:k} B_{ki} \geq 0, \quad \text{where } B_{ki} \geq 0 \; \forall i, k \text{ and } B_{ii} = 0. \tag{7}$$

This means that we will subtract from each column of M a nonnegative linear combination of the other columns of M in order to maximize its sparsity while keeping its nonnegativity. Intuitively, this amounts to keeping only the non-redundant information from each column of M (see Section 6 for some visual examples).

3.2.1 RELAXING THE CONSTRAINT ON THE SPECTRAL RADIUS

In general, there is no easy way to deal with the non-convex constraint $\rho(B) < 1$. In particular, this constraint may lead to difficult optimization problems, for example, finding the nearest stable matrix to an unstable one:

$$\min_X ||X - A|| \quad \text{such that} \quad \rho(X) \leq 1,$$

see Polyak and Shcherbakov (2005) and the references therein. This means that even the projection on the feasible set is non-trivial.

However, we will prove in Section 4 that if the columns of M are not multiples of each other, then any optimal solution of problem (6) without the constraint on the spectral radius of B, that is, any optimal solution B^* of

$$\min_{B \in \mathbb{R}^{n \times n}_+} \quad \sum_{i=1}^{n} \left\| M_{:i} - \sum_{k \neq i} M_{:k} B_{ki} \right\|_2^2 \quad \text{such that} \quad M \geq MB,\ B_{ii} = 0\ \forall i, \tag{8}$$

automatically satisfies $\rho(B^*) < 1$ (Theorem 21). Hence, the approach may only fail when there are repetitions in the data set. The reason is that when a column is multiple of another one, say $M_{:i} = \alpha M_{:j}$ for $i \neq j$ and $\alpha > 0$, then taking $B_{ij} = \alpha$ (0 otherwise for that column) gives $MQ_{:i} = M_{:i} - \alpha M_{:j} = 0$ and similarly for $M_{:j}$. Hence we have lost a component in our data set and potentially produce a lower rank matrix MQ. In practice, it will be important to make sure that the columns of M are not multiples of each other (even though it is usually not the case for well-constructed data sets).

4. Properties of the Preprocessing

In the remainder of the paper, we denote $\mathcal{B}^*(M)$ the set of optimal solutions of problem (8) for the data matrix M, and $\mathcal{P}$ the preprocessing operator defined as

$$\mathcal{P} : \mathbb{R}^{m \times n}_+ \to \mathbb{R}^{m \times n}_+ : M \mapsto \mathcal{P}(M) = M(I - B^*), \text{ where } B^* \in \mathcal{B}^*(M).$$

In this section, we prove some important properties of $\mathcal{P}$ and $\mathcal{B}^*(M)$:

- The preprocessing operator $\mathcal{P}$ is well-defined (Theorem 15).
- The preprocessing operator $\mathcal{P}$ is invariant to permutation and scaling of the columns of M (Lemma 16).
- If the columns of $\theta(M)$ are distinct, then $\rho(B^*) < 1$ for any $B^* \in \mathcal{B}^*(M)$ (Theorem 21).
- If the vertices of $\text{conv}(\theta(M))$ are distinct then
  - There exists $B^* \in \mathcal{B}^*(M)$ such that $\rho(B^*) < 1$ (Corollary 22).
  - $\text{rank}(\mathcal{P}(M)) = \text{rank}(M)$ and $\text{rank}_+(\mathcal{P}(M)) \geq \text{rank}_+(M)$ (Corollary 19).
- If the matrix M is separable, then the preprocessing allows to recover a sparse and optimal solution of the corresponding NMF problem (Theorem 24). In particular, it is always optimal for rank-two matrices (Corollary 25).
- If the matrix has rank-three, then the preprocessing yields an instance in which the number of solutions of the exact NMF problem is finite (Theorem 29).

4.1 General Properties

A crucial property of our preprocessing is that it is well-defined.

Theorem 15 *The preprocessing $\mathcal{P}(M)$ is well-defined: for any $B_1^* \in \mathcal{B}^*(M), B_2^* \in \mathcal{B}^*(M)$, we have $M(I-B_1^*) = M(I-B_2^*) = \mathcal{P}(M)$.*

Proof Problem (8) can be decoupled into n independent CLLS (one for each column of M) of the form:

$$\min_{b \in \mathbb{R}_+^{n-1}} \|d - Cb\|^2 \text{ such that } Cb \leq d, \tag{9}$$

which is equivalent to

$$\min_{b \in \mathbb{R}_+^{n-1}, y \in \mathbb{R}^m} \quad \|d-y\|^2 \text{ such that } y \leq d, y = Cb.$$

The result follows from the fact that the ℓ_2 projection onto a polyhedral set (actually any convex set) yields a unique point. ■

Another important property of the preprocessing is its invariance to permutation and scaling of the columns of M.

Lemma 16 *Let M be a nonnegative matrix and P be a monomial matrix. Then, $\mathcal{P}(MP) = \mathcal{P}(M)P$.*

Proof We are going to show something slightly stronger; namely that B^* is an optimal solution of (8) for matrix M if and only if $P^{-1}B^*P$ is an optimal solution of (8) for matrix MP:

$$B^* \in \mathcal{B}^*(M) \quad \iff \quad P^{-1}B^*P \in \mathcal{B}^*(MP).$$

First, note that B is a feasible solution of (8) for M if and only if $P^{-1}BP$ is a feasible solution of (8) for MP. In fact, nonnegativity of B and its diagonal zero entries are clearly preserved under permutation and scaling while

$$M \geq MB \iff MP \geq MBP \iff MP \geq (MP)(P^{-1}BP).$$

Hence there is one-to-one correspondence between feasible solutions of (8) for M and (8) for MP.

Then, let B^* be an optimal solution of (8). Because (8) can be decoupled into n independent CLLS's, one for each column of B (cf. Equation (9)), we have

$$||M_{:i} - MB^*_{:i}||_2^2 \leq ||M_{:i} - MB_{:i}||_2^2, \quad \forall i,$$

for any feasible solution B of (8). Letting $p \in \mathbb{R}_+^n$ be such that p_i is equal to the non-zero entry of the ith row of P, we have

$$\begin{aligned}
\sum_i p_i^2 ||M_{:i} - MB^*_{:i}||_2^2 &= \sum_i ||M_{:i}p_i - MPP^{-1}B^*_{:i}p_i||_2^2 \\
&= ||MP - MPP^{-1}B^*P||_F^2 \\
&\leq \sum_i p_i^2 ||M_{:i} - MB_{:i}||_2^2 = ||MP - MPP^{-1}BP||_F^2,
\end{aligned}$$

for any feasible solution $B' = P^{-1}BP$ of (8) for MP. This proves $B^* \in \mathcal{B}^*(M) \Rightarrow P^{-1}B^*P \in \mathcal{B}^*(MP)$. The other direction follows directly by using the permutation P^{-1} on the matrix MP. ■

It is interesting to observe that if a column of M belongs to the convex cone generated by the other columns, then the corresponding column of $\mathcal{P}(M)$ is equal to zero.

Lemma 17 *Let $\mathcal{J} = \{1,2,\ldots,n\}\backslash\{i\}$. Then $\mathcal{P}(M)_{:i} = 0$ if and only if $M_{:i} \in \text{cone}(M(:,\mathcal{J}))$.*

Proof We have that

$$\mathcal{P}(M)_{:i} = M_{:i} - \sum_{k\neq i} B^*_{ki} M_{:k} = 0, \quad B^*_{ki} \geq 0 \iff M_{:i} = \sum_{k\neq i} B^*_{ki} M_{:k}, \quad B^*_{ki} \geq 0.$$

■

The preprocessed matrix $\mathcal{P}(M)$ may contain all-zero columns, for which the function $\theta(.)$ is not defined (cf. Definition 1). We extend the definition to matrices with zero columns as follows: $\theta(X)$ is the matrix whose columns are the normalized non-zero columns of X, that is, letting Y be the matrix X where the non-zero columns have been removed, we define $\theta(X) = \theta(Y)$. Hence $\text{conv}(\theta(X))$ denotes the convex hull of the normalized non-zero columns of X.

Another straightforward property is that the preprocessing can only inflate the convex hull defined by the columns of $\theta(M)$.

Lemma 18 *Let $M \in \mathbb{R}^{m\times n}_+$. If the vertices of $\text{conv}(\theta(M))$ are non-repeated, then*

$$\text{conv}(\theta(M)) \quad \subseteq \quad \text{conv}(\theta(\mathcal{P}(M))) \quad \subseteq \quad \Delta^m \cap \text{col}(\theta(M)).$$

Proof By construction, since $\mathcal{P}(M) = MQ$, $\text{col}(\theta(\mathcal{P}(M))) \subseteq \text{col}(\theta(M))$ and $\text{conv}(\theta(\mathcal{P}(M))) \subseteq \Delta^m \cap \text{col}(\theta(M))$. Let i be the index corresponding to a vertex of $\theta(M)$ and $\mathcal{J} = \{1,2,\ldots,n\}\backslash\{i\}$. Because vertices of $\theta(M)$ are non-repeated, we have $M_{:i} \notin \text{conv}(\theta(M(:,\mathcal{J})))$, while

$$\mathcal{P}(M)_{:i} = M_{:i} - \sum_{k\neq i} b_{ki} M_{:k} \quad \iff \quad M_{:i} = \mathcal{P}(M)_{:i} + \sum_{k\neq i} b_{ki} M_{:k}.$$

Hence $M_{:i} \in \text{conv}(\theta([\mathcal{P}(M)_{:i} M(:,\mathcal{J})]))$, which implies that

$$\text{conv}(\theta(M)) \subseteq \text{conv}(\theta([\mathcal{P}(M)_{:i} M(:,\mathcal{J})])),$$

so that replacing $M_{:i}$ by $\mathcal{P}(M)_{:i}$ extends $\text{conv}(\theta(M))$. Since this holds for all vertices, the proof is complete. ■

Corollary 19 *Let $M \in \mathbb{R}^{m\times n}_+$. If no column of M is multiple of another column, then*

$$\text{rank}(\mathcal{P}(M)) = \text{rank}(M) \quad \textit{and} \quad \text{rank}_+(\mathcal{P}(M)) \geq \text{rank}_+(M).$$

Proof Without loss of generality, we can assume that M does not have a zero column. In fact, a preprocessed zero column remains zero while it cannot influence the preprocessing of the other columns (see Equation (7)). Then, by Lemma 18, we have

$$\text{conv}(\theta(M)) \quad \subseteq \quad \text{conv}(\theta(\mathcal{P}(M))) \quad \subseteq \quad \Delta^m \cap \text{col}(\theta(M)),$$

implying $\text{rank}_+(\mathcal{P}(M)) \geq \text{rank}_+(M)$ and $\text{rank}(\mathcal{P}(M)) = \text{rank}(M)$.

Another way to prove this result is to use Corollary 22 (see below) guaranteeing the existence of an inverse-positive matrix Q such that $\mathcal{P}(M) = MQ$ which implies $\text{rank}(\mathcal{P}(M)) = \text{rank}(M)$. Moreover, any exact NMF $(U,V) \in \mathbb{R}^{m\times r} \times \mathbb{R}^{r\times n}$ of $\mathcal{P}(M)$ gives $M = UVQ^{-1}$ hence $\text{rank}_+(M) \leq \text{rank}_+(\mathcal{P}(M))$. ■

We now prove that if no column of M is multiple of another column (that is, the columns of $\theta(M)$ are distinct) then $\rho(B^*) < 1$ for any $B^* \in \mathcal{B}^*(M)$ whence $Q = I - B^*$ is an inverse positive matrix.

Lemma 20 *Let A be a column stochastic matrix and $Q = I - B$ where $B \geq 0$ and $B_{ii} = 0$ for all i be such that $AQ \geq 0$. Then,*

$$\sum_k B_{ki} \leq 1, \quad \forall i,$$

so that Q is diagonally dominant. Moreover, if $A_{:i} \notin \text{conv}(A(:,\mathcal{J}))$ where $\mathcal{J} = \{1,2,\ldots,n\}\backslash\{i\}$, then

$$\sum_k B_{ki} < 1.$$

Proof By assumption, we have for all i

$$A_{:i} \geq AB_{:i} = \sum_k A_{:k}B_{ki},$$

which implies

$$1 = ||A_{:i}||_1 \geq ||AB_{:i}||_1 = ||\sum_k A_{:k}B_{ki}||_1 = ||B_{:i}||_1 = \sum_k B_{ki},$$

because A and B are nonnegative. Moreover, if $A_{:i} \notin \text{conv}(A(:,\mathcal{J}))$, then there exists at least one index j such that $A_{ji} > A_{j:}B_{:i}$ (Lemma 17) so that the above inequality is strict. ■

Theorem 21 *If no column of M is multiple of another column, then any optimal solution B^* of* (8) *satisfies $\rho(B^*) < 1$ whence $Q = I - B^*$ is inverse positive.*

Proof By Theorem 11, $\rho(B^*) < 1$ if and only if $Q = I - B^*$ is inverse positive if and only if Q is a nonsingular M-matrix. Let us then show that Q is a nonsingular M-matrix. First, we can assume without loss of generality that

- Matrix M does not contain a column equal to zero. In fact, if M does, say the first column is equal to zero, then we must have $B_{:1} = 0$ (since $M_{:1} \geq MB_{:1}$ and there is not other zero column in M). The matrix Q is then a nonsingular M-matrix if and only if $Q(2{:}n, 2{:}n)$ is.

- The columns of M sum to one. In fact, letting $P = D(M)$ be defined as in Equation (2), by Lemma 16, B^* is an optimal solution for M if and only if $P^{-1}B^*P$ is an optimal solution for MP. Since B^* and $P^{-1}B^*P$ share the same eigenvalues, $\rho(B^*) < 1 \iff \rho(P^{-1}B^*P) < 1$.

- Let $B \in \mathcal{B}^*(M)$, $Q = I - B^*$, and P be a permutation matrix such that

$$P^T QP = \begin{pmatrix} Q^{(1)} & Q^{(12)} & Q^{(13)} & \dots & Q^{(1k)} \\ 0 & Q^{(2)} & Q^{(23)} & \dots & Q^{(2k)} \\ 0 & 0 & Q^{(3)} & \dots & Q^{(3k)} \\ \vdots & \dots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & Q^{(k)} \end{pmatrix}$$
$$= I - \begin{pmatrix} B^{(1)} & B^{(12)} & B^{(13)} & \dots & B^{(1k)} \\ 0 & B^{(2)} & B^{(23)} & \dots & B^{(2k)} \\ 0 & 0 & B^{(3)} & \dots & B^{(3k)} \\ \vdots & \dots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & B^{(k)} \end{pmatrix},$$

  where $Q^{(i)}$ are irreducible for all i. Without loss of generality, by Lemma 16, we can then assume that Q has this form.

In the following we show that $Q^{(p)}$ is nonsingular for each $1 \leq p \leq k$ hence Q is. By Theorem 14, if $Q^{(p)}$ is irreducibly diagonally dominant, then $Q^{(p)}$ is nonsingular and the proof is complete. We already have that $Q^{(p)}$ is irreducible for $1 \leq p \leq k$. Let I_p denote the index set such that $Q^{(p)} = Q(I_p, I_p)$. We have $M(I_p,:)$ is column stochastic, and

$$\mathcal{P}(M)(I_p,:) = M(I_p,:) - \sum_{l=1}^{p-1} M(I_l,:)B^{(lp)} - M(I_p,:)B^{(p)} \geq 0,$$

implying that $M(I_p,:) \geq M(I_p,:)B^{(p)}$. Moreover the columns of $M(I_p,:)$ are distinct so that there is at least one which does not belong to the convex hull of the others. Hence, by Lemma 20, $Q^{(p)}$ is irreducibly diagonally dominant. ■

Corollary 22 *Let $M \in \mathbb{R}_+^{m\times n}$. If the vertices of* $\text{conv}(\theta(M))$ *are non-repeated, then there exists an optimal solution $B^* \in \mathcal{B}^*(M)$ such that $\rho(B^*) < 1$, that is, such that $Q = I - B^*$ is an inverse-positive matrix.*

Proof Let us show that there exists an optimal solution such that Q is a nonsingular M-matrix. First, by Lemma 20, Q is diagonally dominant implying $\rho(B) \leq 1$ so that Q is an M-matrix (cf. Theorem 21). We can assume without loss of generality that the r first columns of M correspond to the vertices of $\text{conv}(\theta(M))$. This implies that there exists an optimal solution $B^* \in \mathcal{B}^*(M)$ such that

$$Q = \begin{pmatrix} Q_1 & Q_{12} \\ 0 & I \end{pmatrix} = I - \begin{pmatrix} B_1^* & B_{12}^* \\ 0 & 0 \end{pmatrix}, \text{ where } Q_1, B_1^* \in \mathbb{R}^{r\times r} \text{ and } Q_{12}, B_{12}^* \in \mathbb{R}^{r\times(n-r)}.$$

In fact, by assumption, the last columns of M belong to the convex cone of the r first ones and can then be set to zero (which is optimal) using only the first r columns (cf. Lemma 17). Lemma 20 applies on matrix Q_1 and $M(:,1{:}r)$ since

$$MQ(:,1{:}r) = M(:,1{:}r) - M(:,1{:}r)B_1^* \geq 0,$$

while by assumption no column of $M(:,1{:}r)$ belong to the convex hull of the other columns, so that Q_1 is strictly diagonally dominant hence is a nonsingular M-matrix. ■

Finally, what really matters is that the vertices of $\text{conv}(\theta(M))$ are non-repeated. In that case, the preprocessing is unique and the preprocessed matrix has the same rank as the original one. The fact that Q could be singular is not too dramatic. In fact, given an NMF (U,V') of the preprocessed matrix $\mathcal{P}(M) = MQ \approx UV'$, we can obtain the optimal factor V for matrix M by solving the nonnegative least squares problem $V = \text{argmin}_{X\geq 0} ||M - UX||_F^2$ (instead of taking $V = V'Q^{-1}$) and obtain $M \approx UV$.

4.2 Recovery Under Separability

A nonnegative matrix M is called *separable* if it can be written as $M = UV$ where $U \in \mathbb{R}_+^{m\times r}$, $V \in \mathbb{R}_+^{r\times n}$, and for each $i = 1,\ldots,r$ there is some column $f(i)$ of V that has a single nonzero entry and this entry is in the ith row, that is, V contains a monomial submatrix. In other words, each column of U appears (up to a scaling factor) as a column of M. Arora et al. (2012) showed that the NMF problem corresponding to a separable nonnegative matrix can be solved in polynomial time (while NMF is NP-hard in general; see Introduction). In this section, we show that the preprocessing is able to solve this problem while generating a sparser solution than the one obtained with the algorithm of Arora et al. (2012). We refer the reader to Gillis and Vavasis (2012) and the references therein for more details about NMF algorithms for separable matrices.

It is worth noting that the separability assumption is equivalent to the pure-pixel assumption in hyperspectral imaging (for each constitutive material present in the image, there is at least one pixel containing only that material), see Craig (1994), or, in document classification, to the assumption that, for each topic, there is at least one document corresponding only to that topic (or, considering the matrix transpose, that there is at least one word corresponding only to that topic, see Arora et al., 2012). Geometrically, separabilty means that the vertices of $\text{conv}(\theta(M))$ are given by the columns of $\theta(U)$. We have the following straightforward lemma:

Lemma 23 *$M = UV$ is separable (that is, $U \geq 0$, $V \geq 0$ and V contains a monomial submatrix) if and only if* $\text{conv}(\theta(M)) = \text{conv}(\theta(U))$.

Proof $M = UV$ where $U \geq 0, V \geq 0$ and V contains a monomial submatrix if and only if the vertices of $\theta(U)$ and $\theta(M)$ coincide if and only if $\text{conv}(\theta(M)) = \text{conv}(\theta(U))$. ■

Theorem 24 *If M is separable and the r vertices of $\theta(M)$ are non-repeated, then $\mathcal{P}(M)$ has r non-zero columns, say $S_{:1}, S_{:2}, \ldots S_{:r}$, such that* $\text{conv}(\theta(M)) \subseteq \text{conv}(\theta(S))$*, that is, there exists $R \geq 0$ such that $M = SR$.*

Proof This is a consequence of Lemmas 17, 18 and 23. ■

Theorem 24 shows that the preprocessing is able to identify the r columns of $M = UV$ corresponding to the vertices of $\theta(M)$. Moreover, it returns a sparser matrix S, namely $\mathcal{P}(U)$, whose cone contains the columns of M. Remark also that Theorem 24 does not require M to be full rank: the dimension of $\text{conv}(\theta(M))$ can be smaller than $r-1$.

Corollary 25 *For any rank-two nonnegative matrix M whose columns are not multiples of each other, $\mathcal{P}(M)$ has only two non-zero columns, say $S_{:1}$ and $S_{:2}$ such that* $\text{conv}(\theta(M)) \subseteq \text{conv}(\theta(S))$, *that is, there exists $R \geq 0$ such that $M = SR$. In other words, the preprocessing technique is optimal as it is able to identify an optimal nonnegative basis for the NMF problem corresponding to the matrix M.*

Proof A rank-two nonnegative matrix is always separable. In fact, a two-dimensional pointed cone is always spanned by two extreme vectors. In particular, $\text{rank}(M) = 2 \iff \text{rank}_+(M) = 2$ (Thomas, 1974). ■

Example 1 *Here is an example with a rank-three separable matrix*

$$M = \begin{pmatrix} 5 & 5 & 5 & 5 & 9 & 1 & 4 & 1 & 7 & 7 \\ 10 & 6 & 5 & 3 & 7 & 8 & 4 & 1 & 5 & 8 \\ 8 & 9 & 9 & 4 & 7 & 8 & 3 & 9 & 6 & 7 \end{pmatrix}^T \begin{pmatrix} 1 & 0 & 0 & 2 & 3 & 6 & 4 & 4 \\ 0 & 1 & 0 & 5 & 7 & 7 & 7 & 4 \\ 0 & 0 & 1 & 9 & 4 & 4 & 8 & 6 \end{pmatrix}. \tag{10}$$

Its (rounded) preprocessed version is given by

$$\mathcal{P}(M) = \begin{pmatrix} 3.6 & 3.9 & 3.9 & 4.3 & 7.6 & 0 & 3.3 & 0.5 & 5.9 & 5.7 \\ 6.3 & 2.5 & 1.6 & 0 & 1.5 & 6.5 & 1.5 & 0 & 0.7 & 3.4 \\ 0.8 & 2.4 & 2.7 & 0.7 & 0.7 & 1.8 & 0 & 4.2 & 0.9 & 0.6 \end{pmatrix}^T \begin{pmatrix} I_3 & 0_{3\times 5} \end{pmatrix},$$

where I_3 is the 3-by-3 identity matrix and $0_{3\times 5}$ is the 3-by-5 all-zero matrix. Figure 1 shows the geometric interpretation of the preprocessing.

4.3 Uniqueness and Robustness Through Preprocessing

A potential drawback of the preprocessing is that it might increase the nonnegative rank of M. In this section, we show how to modulate the preprocessing to prevent this behavior.

Let us define

$$\mathcal{P}^\alpha(M) = M(I - \alpha B^*) = M - \alpha MB^*,$$

where $0 \leq \alpha \leq 1$ and $B^* \in \mathcal{B}^*(M)$. Notice that $\mathcal{P}^\alpha(M)$ is well-defined because for any $B_1^*, B_2^* \in \mathcal{B}^*(M)$ we have $MB_1^* = MB_2^*$; see Theorem 15.

Lemma 26 *Let M be a nonnegative matrix such that the vertices of* $\text{conv}(\theta(M))$ *are non-repeated. Then, for any $0 \leq \alpha \leq \beta \leq 1$,*

$$\text{conv}(\theta(M)) \subseteq \text{conv}(\theta(\mathcal{P}^\alpha(M))) \subseteq \text{conv}(\theta(\mathcal{P}^\beta(M))) \subseteq \text{col}(\theta(M)) \cap \Delta^m.$$

Therefore,

$$\text{rank}_+(M) \leq \text{rank}_+(\mathcal{P}^\alpha(M)) \leq \text{rank}_+(\mathcal{P}^\beta(M)).$$

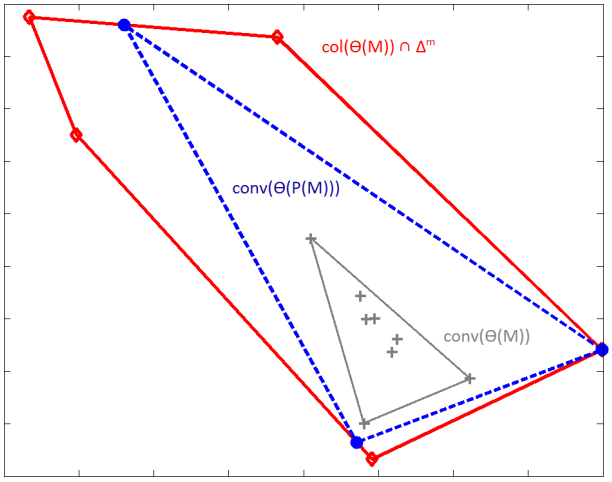

Figure 1: Geometric interpretation of the preprocessing of matrix M from Equation (10).

Proof The proof can be obtained by following exactly the same steps as the proof of Lemma 18. ■

Lemma 27 *Let M be a nonnegative matrix such that the vertices of* $\operatorname{conv}(\theta(M))$ *are non-repeated, then the supremum*

$$\bar{\alpha} = \sup_{0 \le \alpha \le 1} \alpha \quad \textit{such that} \quad \operatorname{rank}_+(\mathcal{P}^{\alpha}(M)) = \operatorname{rank}_+(M) \tag{11}$$

is attained.

Proof We can assume without loss of generality that M does not have all-zero columns. In fact, if $M_{:i} = 0$ for some i then $\mathcal{P}^{\alpha}(M)_{:i} = 0$ for all $\alpha \in [0,1]$ so that the nonnegative rank of $\mathcal{P}^{\alpha}(M)$ is not affected by the zero columns of M.

Then, if $\bar{\alpha} = 1$, the proof is complete. Otherwise, one can easily check that, for any $0 \le \bar{\alpha} < 1$, we have $\mathcal{P}^{\bar{\alpha}}(M)_{:i} \neq 0 \ \forall i$ (using a similar argument as in Lemma 17).

Finally, the result follows from the upper-semicontinuity of the nonnegative rank (Bocci et al., 2011, Theorem 3.1): 'If P is a nonnegative matrix, without zero columns and with $\operatorname{rank}_+(P) = k$, then there exists a ball $\mathcal{B}(P,\varepsilon)$ centered at P and of radius $\varepsilon > 0$ such that $\operatorname{rank}_+(N) \ge k$ for all $N \in \mathcal{B}(P,\varepsilon)$'. Therefore, if the supremum of (11) was not attained, the matrix $\mathcal{P}_{\bar{\alpha}}(M)$ would satisfy $\operatorname{rank}_+(\mathcal{P}_{\bar{\alpha}}(M)) > \operatorname{rank}_+(M)$ while for any $\alpha < \bar{\alpha}$ we would have $\operatorname{rank}_+(\mathcal{P}_{\alpha}(M)) = \operatorname{rank}_+(M)$, a contradiction. ■

Hence working with matrix $\mathcal{P}^{\bar{\alpha}}(M)$ instead of M will reduce the number of solutions of the NMF problem while preserving the nonnegative rank:

Theorem 28 *Let M be a nonnegative matrix for which the vertices of* $\operatorname{conv}(\theta(M))$ *are non-repeated, let also $\bar{\alpha}$ be defined as in Equation* (11)*. Then any exact NMF (U,V) of $\mathcal{P}^{\bar{\alpha}}(M)$ corresponds to an*

exact NMF (U, VQ^{-1}) *of* M*, while the converse is not true. In fact,*

$$\text{conv}(\theta(M)) \subseteq \text{conv}(\theta(\mathcal{P}^{\bar{\alpha}}(M))).$$

Therefore, the NMF problem for $\mathcal{P}^{\bar{\alpha}}(M)$ *is more well posed.*

Proof This follows directly from the definition of $\bar{\alpha}$, and Lemmas 26 and 27. ■

We now illustrate Corollary 28 on a simple example, which will lead to three other important results.

Example 2 (Nested Squares) *Let*

$$M = \begin{pmatrix} 5 & 3 & 3 & 5 \\ 3 & 5 & 5 & 3 \\ 5 & 5 & 3 & 3 \\ 3 & 3 & 5 & 5 \end{pmatrix}.$$

The problem NPP(M) restricted to the column space of M is made up of two nested squares, $\text{conv}(\theta(M))$ *and* $\text{col}(\theta(M)) \cap \Delta^m$*, centered at* $(0,0)$ *with side length 2 and 8 respectively, see Figure 2. The polygon corresponding to* $\mathcal{P}^{\alpha}(M)$ *is a square centered at* $(0,0)$ *with side length depending on* α*, between 2 (for* $\alpha = 0$*) and 8 (for* $\alpha = 1$*). We can show that the largest such square still included in a triangle corresponds to*

$$\mathcal{P}^{\bar{\alpha}}(M) = \mathcal{P}^{\bar{\alpha}} \begin{pmatrix} 5 & 3 & 3 & 5 \\ 3 & 5 & 5 & 3 \\ 5 & 5 & 3 & 3 \\ 3 & 3 & 5 & 5 \end{pmatrix} = \frac{1}{a} \begin{pmatrix} 1+a & 1-a & 1-a & 1+a \\ 1-a & 1+a & 1+a & 1-a \\ 1+a & 1+a & 1-a & 1-a \\ 1-a & 1-a & 1+a & 1+a \end{pmatrix}, \tag{12}$$

where $a = \sqrt{2} - 1$ *and* $\bar{\alpha} = \frac{4a-1}{3a}$ *(this follows from the proof of Theorem 29; see below). Hence, the polygon* $\text{conv}(\theta(\mathcal{P}^{\bar{\alpha}}(M)))$ *is a square centered at* $(0,0)$ *with side length* $8a$ *in between* $\text{conv}(\theta(M))$ *and* $\text{col}(\theta(M)) \cap \Delta^m$*, see Figure 2. Unfortunately, the exact NMF of* $\mathcal{P}^{\bar{\alpha}}(M)$ *is non-unique. In fact, we will see later that it has 8 solutions (the ones drawn on Figure 2 and their rotations).*

Example 2 illustrates the following three important facts:
Fact 1. Defining a well-posed NMF problem is not always possible. In other words, *there does not exist any 'reasonable' NMF formulation having always a unique solution (up to permutation and scaling).* In fact, Example 2 shows that, because of the symmetry of the problem, any solution of NPP(M) can be rotated by 90, 180 or 270 degrees to obtain a different solution with exactly the same characteristics (the rotated solutions cannot be distinguished in any reasonable way). For example, there are 4 solutions which are the sparsest, each containing one vertex of $\text{col}(\theta(M)) \cap \Delta^m$, see $\text{conv}(\theta(U_2))$ on Figure 2, including

$$U_2 = \begin{pmatrix} 1 & a & 0 \\ 0 & 1-a & 1 \\ a & 1 & 0 \\ 1-a & 0 & 1 \end{pmatrix}, \text{ and } U_2^{(180)} = \begin{pmatrix} 0 & 1-a & 1 \\ 1 & a & 0 \\ 1-a & 1 & 1 \\ a & 0 & 0 \end{pmatrix},$$

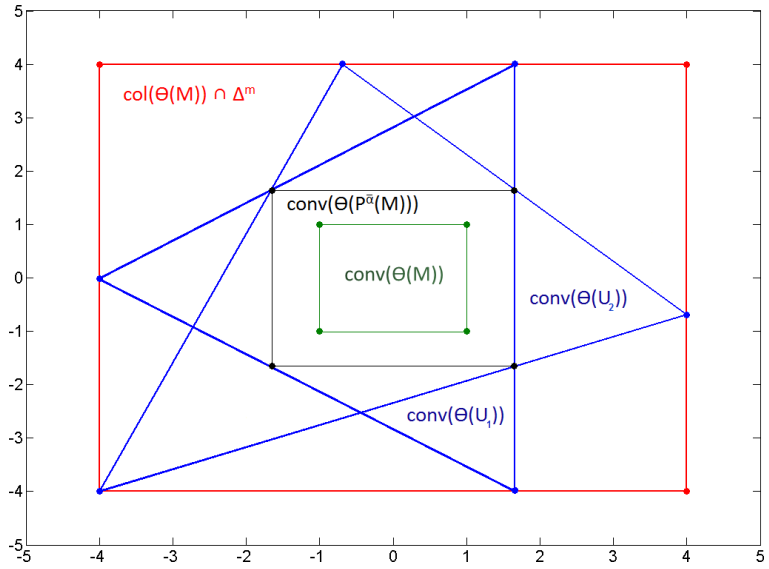

Figure 2: Geometric interpretation of the preprocessing of matrix M from Equation (12).

where $U_2^{(180)}$ is the rotation of 180 degrees of U_2.

Fact 2. The preprocessing makes NMF more robust. For any m-by-n matrix E such that $\text{col}(E) \subseteq \text{col}(M)$, $M+E \geq 0$, and

$$\text{conv}(\theta(M)) \subseteq \text{conv}(\theta(M+E)) \subseteq \text{conv}(\mathcal{P}^{\bar{\alpha}}(M)),$$

the exact NMF (U,V) of $\mathcal{P}^{\bar{\alpha}}(M)$ will still provide an optimal factor U for the perturbed matrix $M+E$. In particular, if the matrix M is positive, then one can show that[4] $\text{conv}(\theta(M))$ is strictly contained in $\text{conv}(\mathcal{P}^{\bar{\alpha}}(M))$ (given that $\bar{\alpha} > 0$) so that any sufficiently small perturbation E with $\text{col}(E) \subseteq \text{col}(M)$ will satisfy the conditions above.

In Example 2, the vertices of M can be perturbed and, as long as they remain inside the square defined by $\text{conv}(\mathcal{P}^{\bar{\alpha}}(M))$ (see Figure 2), the exact NMF of $\text{conv}(\mathcal{P}^{\bar{\alpha}}(M))$ will provide an exact NMF for the perturbed matrix M. (More precisely, any matrix E such that $\text{col}(E) \subseteq \text{col}(M)$ and $\max_{i,j}|E_{ij}| \leq \sqrt{2}-1$ will satisfy $\text{conv}(\theta(M+E)) \subseteq \text{conv}(\mathcal{P}^{\bar{\alpha}}(M))$.)

Fact 3. The preprocessing makes the NMF problem more well-posed. In Example 2, even though the NMF of $\mathcal{P}^{\bar{\alpha}}(M)$ is non-unique, the set of solutions has been drastically reduced: from a two-dimensional space to a zero-dimensional one containing eight points: $\text{conv}(\theta(U_1))$, $\text{conv}(\theta(U_2))$ and the corresponding rotated solutions, see Figure 2.

Theorem 29 *Let $M \in \mathbb{R}_+^{m\times n}$ be such that $\text{rank}(M) = \text{rank}_+(M) = 3$ and let $\bar{\alpha}$ be defined as in Equation (11). Assume also that $\text{conv}(\theta(\mathcal{P}(M)))$ has at least four vertices. Then the number of solutions of NPP($\mathcal{P}^{\bar{\alpha}}(M)$) with three vertices is smaller than $m+n$.*

Proof Let P and Q denote the outer and inner polygons of NPP($\mathcal{P}^{\bar{\alpha}}(M)$), respectively. Let us also parametrize the boundary of the outer polygon P with the parameter $t \in [0,1]$ and the function

$$x : \mathbb{R}_+ \to \mathbb{R}^2 : t \mapsto x(t) \in P,$$

4. Using the same ideas as in Lemma 18 and the fact that any preprocessed column must contain at least one zero entry.

where x is a continuous function with $x(0) = x(1)$ and $\{x(t) \mid t \in [0,1]\}$ is equal to the boundary of P. We also define the function x for values of t larger than one using $x(t) = x(t - \lfloor t \rfloor)$ where $\lfloor t \rfloor$ is the largest integer not exceeding t. Using the construction of Aggarwal et al. (1989), we define the function $f_k : \mathbb{R}_+ \to \mathbb{R}_+ : t \mapsto f_k(t)$ as follows. Let $t_1 \in [0,1)$ and $x(t_1)$ be the corresponding point on the boundary of P. From $x(t_1)$, we can trace the tangent to Q (that is, Q is on one side of the tangent, and the tangent touches Q), say in the clock-wise direction, intersect it with P and hence obtain a new point $x(t_2)$ on the boundary P (see Figure 3 for an illustration on the nested squares problem). We assume without loss of generality that $t_2 \geq t_1$ (if t_2 happens to be larger than one, we

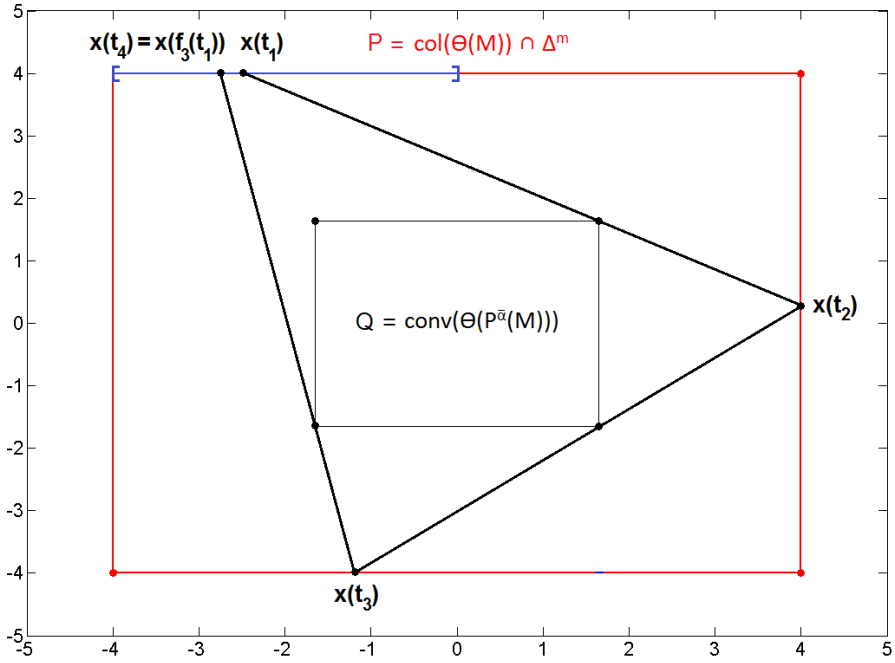

Figure 3: Mapping of the point $x(t_1)$ to $x(t_4)$ using the construction of Aggarwal et al. (1989).

do not round it down with the equivalent value $t_2 - \lfloor t_2 \rfloor$). Starting from $x(t_2)$, we can use the same procedure to obtain $x(t_3)$ and we apply this procedure k times to obtain the point $x(t_{k+1})$, where $t_{k+1} \geq \cdots \geq t_2 \geq t_1$. Finally, we define $f_k(t_1) = t_{k+1}$.

Aggarwal et al. (1989) showed that $x(t_1)$ can be taken as a vertex of a feasible solution of NPP($\mathcal{P}^{\bar{\alpha}}(M)$) with k vertices if and only if $f_k(t_1) = t_{k+1} \geq t_1 + 1$, that is, we were able to turn around Q inside P in $k+1$ steps (in fact, $x(t_1)$, $x(t_2)$, ..., and $x(t_k)$ are the vertices of a feasible solution).

Aggarwal et al. (1989) also showed that the function f_k is continuous, non-decreasing, and depends continuously on the vertices of Q (see also Appendix A). Figure 4 displays the function f_4 for the nested squares (Example 2).

If $\mathrm{col}(\theta(M)) \cap \Delta^m$ has three vertices, then $\bar{\alpha} = 1$. In fact, we have that

$$\theta(\mathcal{P}^{\alpha}(M)) \subseteq \mathrm{col}(\theta(M)) \cap \Delta^m \quad \text{for any } 0 \leq \alpha \leq 1,$$

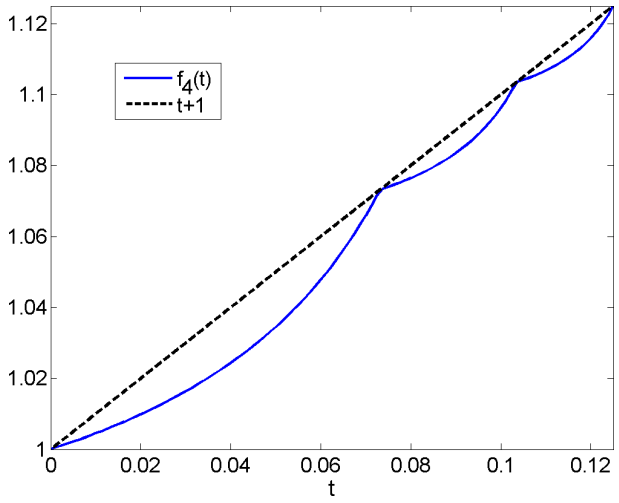

Figure 4: Function $f_4(t)$ for Example 2 using the construction of Aggarwal et al. (1989) (see also Figure 4 and Appendix A). We only plot the function f_4 in the interval $[0, \frac{1}{8}]$ because, by symmetry, $f_4(x+\frac{1}{8}) = f_4(x)+\frac{1}{8}$.

implying $\text{rank}_+(\mathcal{P}^\alpha(M)) = 3$ for all $0 \leq \alpha \leq 1$. Moreover, because $\theta(\mathcal{P}^\alpha(M))$ has at least four vertices, $\text{col}(\theta(M)) \cap \Delta^m$ is the unique solution of the corresponding NPP problem: the outer polygon is a triangle while the inner polygon has at least four vertices which are located on the edges of the outer triangle (since $\bar{\alpha} = 1$ and each column of $\mathcal{P}(M)$ contains at least one zero entry).

Let us then assume that $\text{col}(\theta(M)) \cap \Delta^m$ has at least four vertices. We show that this implies $\bar{\alpha} < 1$. Assume $\bar{\alpha} = 1$. The polygons $P = \text{col}(\theta(M)) \cap \Delta^m$ and $Q = \theta(\mathcal{P}(M))$ have at least 4 vertices. Moreover, the vertices of Q are located on the boundary of P (because $\bar{\alpha} = 1$) on at least two different sides of P (three vertices cannot be on the same side). It can be shown by inspection that the optimal solution of this NPP instance must have at least four vertices, hence $\text{rank}_+(\mathcal{P}(M)) > 3$, a contradiction.

Next, we show that $f_4(t) \leq t+1$. Assume there exists t such that $f_4(t) > t+1$. By continuity of f_4 with respect to the vertices of $Q = \text{conv}(\theta(\mathcal{P}^{\bar{\alpha}}(M)))$, there exists $\varepsilon > 0$ sufficiently small such that $\bar{\alpha}+\varepsilon < 1$ and such that the function f_4' for the NPP instance with inner polygon $Q' = \text{conv}(\theta(\mathcal{P}^{\bar{\alpha}+\varepsilon}(M)))$ and the same outer polygon P satisfies $f_4(t) > t+1$ hence $\text{rank}_+(\mathcal{P}^{\bar{\alpha}+\varepsilon}(M)) \leq 3$, a contradiction.

In Appendix A, we prove that f_k is made up of pieces which are either constant or strictly convex, with at most $m+n$ break points corresponding to different solutions to the NPP. Therefore, because f_4 is continuous and smaller than $t+1$, it can intersect the line $t+1$ only at the break points. Since there are at most $m+n$ such points corresponding to different NPP solutions, the number of solutions of NPP($\mathcal{P}^{\bar{\alpha}}(M)$) with three vertices is smaller than $m+n$. (Notice that the bound is tight for the nested squares example with 8 solutions.) ■

Remark 30 *If* $\mathrm{conv}(\theta(\mathcal{P}(M)))$ *has three vertices, they define a feasible solution for the corresponding NPP problem (that is,* $\mathcal{P}(M)$ *is separable, see Theorem 23). However, the number of solutions might be not be finite in that case. Here is an example*

$$M = \begin{pmatrix} 0 & 0.5 & 0.25 & 0 \\ 1 & 0.5 & 0.75 & 1 \\ 1 & 0 & 0.1 & 0.5 \\ 0 & 1 & 0.9 & 0.5 \end{pmatrix} \quad \textit{and} \quad \mathcal{P}(M) = \begin{pmatrix} 0 & 0.5 & 0 & 0 \\ 1 & 0.5 & 0.3 & 0.5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0.3 & 0.5 \end{pmatrix},$$

whose corresponding NPP problems are represented on Figure 5: the NPP of $\mathcal{P}(M)$ *does not have a finite number of solutions.*

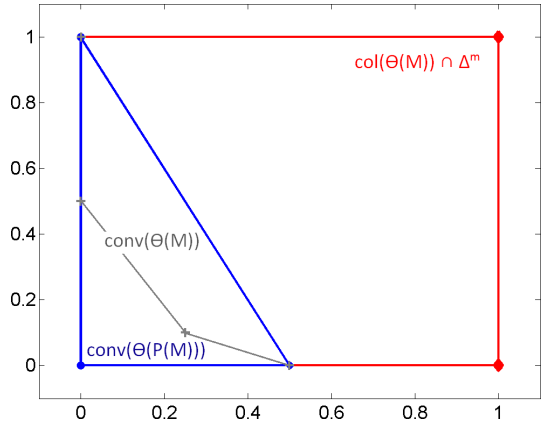

Figure 5: Counter-example for Theorem 29 when $\mathcal{P}(M)$ has three vertices.

The fact that the NPP of the matrix $\mathcal{P}^{\bar{\alpha}}(M)$ can have several different solutions is untypical and, we believe, could be due to the symmetry of the problem (as in Example 2). We conjecture that, in general, the solution to NPP($\mathcal{P}^{\bar{\alpha}}(M)$) is unique. In particular, we observed on randomly generated matrices that it was, see Example 1. In fact, as the function $f_k(.)$ defined in Theorem 29 depends continuously on the inner and outer polytopes Q and P, if these polytopes are generated randomly, there is no reason for the values of the function $f_k(.)$ at the break points to be located on the same line as on Figure 4.

We also conjecture that Theorem 29 holds true for any rank:

Conjecture 31 *Let* M *be such that* $\mathrm{rank}(M) = \mathrm{rank}_+(M) = k$ *and* $\mathrm{conv}(\theta(\mathcal{P}(M)))$ *has at least* $(k+1)$ *vertices, and* $\bar{\alpha}$ *be defined as in Equation* (11)*, then the number of solutions of NPP(*$\mathcal{P}^{\bar{\alpha}}(M)$*) is finite.*

Unfortunately, the geometric construction of Aggarwal et al. (1989) cannot be generalized to three dimensions (or higher). To prove the conjecture, we would need to show that

- Any solution of NPP($\mathcal{P}^{\bar{\alpha}}(M)$) is isolated. Intuitively, the preprocessing $\mathcal{P}^{\bar{\alpha}}(M)$ of M grows the inner polytope Q as long as the corresponding NPP instance has a solution with $\mathrm{rank}_+(M)$ vertices. If a solution was not isolated, it could be moved around while remaining feasible, which indicates that we could grow the inner polytope Q hence increase $\bar{\alpha}$.

- The number of isolated solutions is finite. We conjecture that the solutions can be characterized in terms of the faces of P and Q, which are finite (depending on m and n).

Remark 32 *Of course computing $\bar{\alpha}$ is non-trivial. However, for matrices of small rank, this could be done effectively. In fact, checking whether the nonnegative rank of an m-by-n is equal to* $\operatorname{rank}(M)$ *can be done in polynomial time in m and n provided that the rank is fixed (Arora et al., 2012). In particular, the algorithm of Aggarwal et al. (1989) does it in $O((m+n)\log(\min(m,n)))$ operations for rank-three matrices (Gillis and Glineur, 2012a). Hence, one could for example use a bisection method to find a good lower bound $\beta \lesssim \bar{\alpha}$ and use the corresponding matrix NPP($\mathcal{P}^{\beta}(M)$) to have a more well-posed NMF problem whose solutions will be solutions of the original one.*

5. Preprocessing in Practice

In this section, we address three important practical considerations of the preprocessing.

5.1 Computational Complexity of Solving (8)

It is rather straightforward to check that problem (8) can be decoupled into n independent CLLS's, each corresponding to a different column of M; for example, for the ith column of M, we have

$$\min_{b \in \mathbb{R}^n_+} \quad ||M_{:i} - Mb||_2^2 \quad \text{such that} \quad M_{:i} \geq Mb,\ b_i = 0. \tag{13}$$

We then have n CLLS's with n variables (actually $n-1$ since variable $b_i = 0$ can be removed) and $m+n$ constraints. Using interior point methods, the computational complexity for solving (13) is of the order of $O(n^{3.5})$; hence the total computational cost is of the order $O(n^{4.5})$.

Figure 6 shows the computational time needed for solving (8) with respect to m for n fixed and vice versa, for randomly generated matrices (using the *rand(.)* function of MATLAB®) on a laptop 3GHz Intel® CORE i7-2630QM CPU @2GHz 8Go RAM running MATLAB® R2011b using the function *lsqlin(.)* of MATLAB®. The computational time is linear in m while being of the order of n^3 in n, smaller than the expected $O(n^{4.5})$. Therefore, in practice, the dimension m can be rather large while, on a standard machine, n cannot be much larger than 1000. Using parallel architecture would allow to solve larger scale problems (see also Section 7).

5.2 Normalization of the Columns of $\mathcal{P}(M)$

Since the aim eventually is to provide a good approximate NMF to the original data matrix M, we observed that normalizing the columns of the preprocessed matrix $\mathcal{P}(M)$ to match the norm of the corresponding columns of M gives better results. That is, we replace $\mathcal{P}(M)$ with $D\mathcal{P}(M)$ where

$$D_{ii} = \frac{||M_{:i}||_2}{||\mathcal{P}(M)_{:i}||_2} \text{ for all } i, \quad \text{and} \quad D_{ij} = 0 \text{ for all } i \neq j.$$

This scaling does not change the nice properties of the preprocessing since D is a monomial matrix, hence QD still is an inverse-positive matrix. This scaling degree of freedom is related to the fact that we fixed the diagonal entries of Q to one, see Section 3.2.

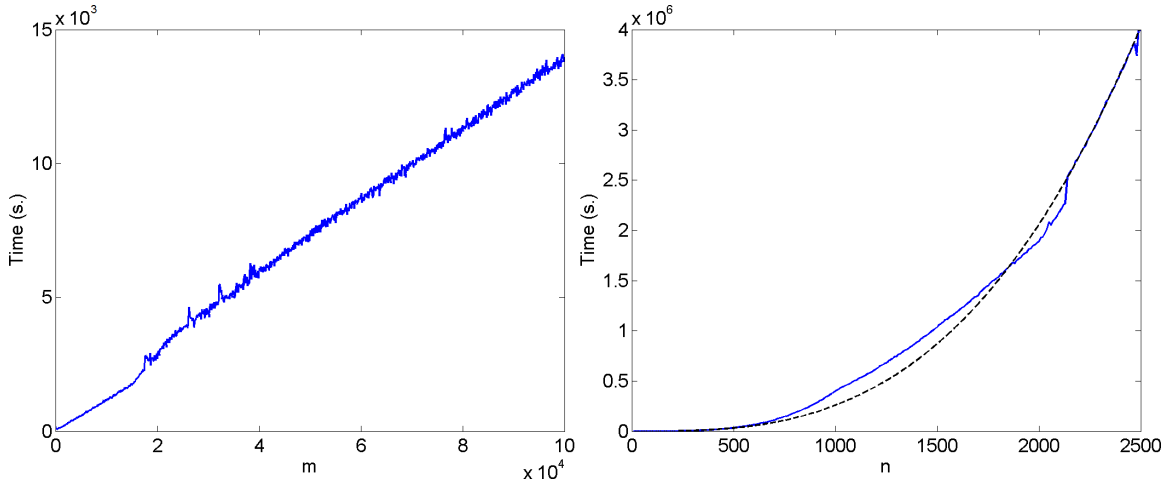

Figure 6: Computational time for solving (8). On the left, m-by-100 randomly generated matrices; on the right, 1000-by-n randomly generated matrices (plain) and the polynomial $2.6^*10^{-4}n^3$ (dashed).

The reason for this choice is that NMF algorithms are sensitive to the norm of the columns of M. In fact, when using the Frobenius norm, we have that the following two problems are equivalent

$$\min_{U\geq 0, V\geq 0} ||M-UV||_F^2 \quad \equiv \quad \min_{X\geq 0, Y\geq 0} \sum_{i=1}^{n} ||M_{:i}||_2^2 \left\| \frac{M_{:i}}{||M_{:i}||_2} - XY_{:i} \right\|_2^2.$$

Therefore, to give each column of $\mathcal{P}(M)$ the same importance in the objective function as in the original NMF problem, it makes sense to use the scaling above. This is particularly critical if there are outliers in the data set: the outliers do not look similar to the other columns of M hence their preprocessing will not reduce much their ℓ_2-norm (because they are further away from the convex cone generated by the other columns of M). Therefore, their relative importance in the objective function will increase in the NMF problem corresponding to $\mathcal{P}(M)$, which is not desirable.

5.3 Dealing with Noisy Input Matrices and/or Obtaining Sparser Preprocessing

Our technique will typically be useless when the input matrix is noisy and sparse. For example, we have

$$M = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}, \mathcal{P}(M) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{while} \quad M_\delta = \begin{pmatrix} 0 & \delta \\ 1 & 0 \\ 1 & 1 \end{pmatrix} = \mathcal{P}(M_\delta),$$

for any $\delta > 0$. This shows that the preprocessing is very sensitive to small positive entries of M. In order to deal with such noisy and sparse matrices, we propose to relax the nonnegativity constraint $MQ \geq 0$ in (8), and solve instead

$$\min_{B\in\mathbb{R}_+^{n\times n}} \quad \sum_{i=1}^{n} \left\| M_{:i} - \sum_{k\neq i} M_{:k}B_{ki} \right\|_2^2 \quad \text{such that} \quad M_{:i} + \varepsilon ||M_{:i}||_\infty e \geq \sum_{k\neq i} M_{:k}B_{ki}, \ \forall i, \qquad (14)$$

where $0 < \varepsilon \ll 1$ and e is the vector of all ones of appropriate dimension. We will denote the corresponding preprocessing $\mathcal{P}_\varepsilon(M) = M(I - B^*_\varepsilon)$ where B^*_ε is an optimal solution of (14). For the example above with $\delta = \varepsilon = 10^{-2}$, we obtain

$$\mathcal{P}_\varepsilon(M_\delta) = \begin{pmatrix} -10^{-2} & 10^{-2} \\ 1 & -10^{-2} \\ 10^{-4} & 0.99 \end{pmatrix}.$$

In practice, this technique also allows to obtain preprocessed matrices with more entries equal or smaller than zero. When choosing the parameter ε, it is very important to check whether $\rho(B^*_\varepsilon) < 1$ so that the rank of $\mathcal{P}_\varepsilon(M)$ is equal to the rank of M and no information is lost (we can recover the original matrix $M = \mathcal{P}_\varepsilon(M)(I - B^*_\varepsilon)^{-1}$ given $\mathcal{P}_\varepsilon(M)$ and B^*_ε).

6. Application to Image Processing

In this section, we apply the preprocessing technique to several image data sets. By construction, the preprocessing procedure will remove from each image a linear combination of the other images. As we will see, this will highlight certain localized parts of these images, essentially because the preprocessed matrices are sparser than the original ones. We will then show that combining the preprocessing with standard NMF algorithms naturally leads to better part-based decompositions, because sparser matrices lead to sparser NMF solutions, see Section 2.

A direct comparison between NMF applied on the original matrix and NMF applied on the preprocessed matrix is not very informative in itself: while the former will feature a lower approximation error, the latter will provide a sparser part-based representation. This does not really tell us whether the improvements in the part-based representation and sparsity are worth the increase in approximation error. For that reason, we choose to compare them with a standard sparse NMF technique, described below, in order to better assess whether the increase in sparsity achieved is worth the loss in reconstruction accuracy. Hence, we compare the following three different approaches:

- **Nonnegative matrix factorization (NMF)**. It solves the original NMF problem from Equation (1) using the accelerated HALS algorithm (A-HALS) of Gillis and Glineur (2012b) (with parameters $\alpha = 0.5$ and $\varepsilon = 0.1$ as suggested by the authors), which is a block coordinate descent method.

- **Preprocessed NMF for different values of** ε. It first computes the preprocessed matrix $\mathcal{P}_\varepsilon(M)$ (cf. Section 5.3), then solves the NMF problem for the rescaled preprocessed matrix $\mathcal{P}_\varepsilon(M)D \approx UV'$ (cf. Section 5.2) using A-HALS and finally returns (U,V) where $V = \operatorname{argmin}_{X \geq 0} ||M - UX||_F^2$. This approach will be denoted Pre-NMF(ε). (We will also indicate in brackets the error obtained when using $V = V'Q^{-1}$, which will be, by construction, always higher.) Notice that the preprocessed matrix may contain negative entries (when $\varepsilon > 0$) which is handled by A-HALS. We do not set these entries to zero for two important reasons: (i) we want to preserve the column space of M, (ii) the negative entries of M lead to sparser NMF solutions: Geometrically, a negative entry in M means that a vertex of $\text{conv}(M)$ (the inner polytope) is not contained in Δ^m (the outer polytope) making NPP(M) infeasible (as a negative entry cannot be obtained with nonnegative ones). However, the approximate solution T of NPP(M) will have to be close to the boundary of Δ^m to approximate well that vertex. In particular, Gillis and Glineur (2008) showed that if an entry of M, say at position (i, j), is

smaller than $-||\max(0,M)||_F$ then $(UV)_{ij} = 0$ for any optimal solution of NMF (1). Therefore, when indicating the sparsity of the preprocessed matrix, negative entries will be counted as zeros as they lead to even sparser NMF decompositions.

- **Sparse NMF**. The most standard technique to obtain sparse solutions for NMF problems is to use a sparsity-inducing penalty term in the objective function. In particular, it is well-known that adding an l_1-norm penalty term induces sparser solutions (Kim and Park, 2007), and we therefore solve the following problem:

$$\min_{U,V\geq 0} ||M-UV||_F^2 + \sum_{i=1}^{r} \mu_i ||U_{:i}||_1, \quad ||U_{:i}||_\infty = 1 \; \forall i,$$

where $||x||_1 = \sum_i |x_i|$, $||x||_\infty = \max_i |x_i|$ and μ_i are positive parameters controlling the sparsity of the columns of U. In order to solve sNMF, we also use A-HALS which can easily be adapted to handle this situation. The ℓ_∞-norm constraints is not restrictive because of the degree of freedom in the scaling of the columns of U and the corresponding rows of V, while it prevents matrix U to converge to zero. The theoretical motivation is that the l_1-norm is the convex envelope of the l_0-norm (that is, the largest convex function smaller than the l_0-norm) in the ℓ_∞-ball, see Recht et al. (2010) and the references therein.

In order to compare sparse NMF with Pre-NMF(ε), the parameters μ_i $1 \leq i \leq r$ are tuned in order to match the sparsity obtained by Pre-NMF(ε). The corresponding approach will be denoted sNMF(ε).

For each approach, we will keep the best solution obtained among the same ten random initializations (using the *rand(.)* function of MATLAB®) and each run was allowed 1000 (outer) iterations of the A-HALS algorithm. We will use the relative error

$$100\,\frac{||M-UV||_F}{||M||_F}$$

to asses the quality of an approximation. We will also display the error obtain by the truncated singular value decomposition (SVD) for the same factorization rank to serve as a comparison. For the sparsity, we use the percentage of non-zero entries[5]

$$s(U) = 100\,\frac{\#\text{zeros}(U)}{mr} \;\in\; [0,100], \text{ for } U \in \mathbb{R}^{m\times r}.$$

Because the solution computed with Pre-NMF does not directly aim at minimizing the error $||M-UV||_F^2$, it is not completely fair to use this measure for comparison. In fact, it would be better to compare the quality of the sparsity patterns obtained by the different techniques. For this reason, we use the same post-processing procedure as described by Gillis and Glineur (2010) which benefits all algorithms: once a solution is computed by one of the algorithms, the zero entries of U are fixed and we minimize $\min_{U\geq 0, V\geq 0} ||M-UV||_F^2$ on the remaining (nonzero) entries (again, A-HALS can easily be adapted to handle this situation and we perform 100 additional steps on each solution), and report the new relative approximation error as "Improved", while the original relative error before this postprocessing will be reported as "Plain". The code is available at `https://sites.google.com/site/nicolasgillis/code`.

5. The negative entries of the preprocessed matrix $\mathcal{P}_\varepsilon(M)$ for $\varepsilon > 0$ will be counted as zeros.

6.1 CBCL Data Set

The CBCL face data set[6] is made of 2429 gray-level images of faces with 19×19 pixels (black is one and white is zero). We look for an approximation of rank $r = 49$ as in Lee and Seung (1999). Because of the large number of images in the data set, the preprocessing is rather slow. In fact, we have seen in Section 5.1 that it is in $O(n^{4.5})$ where n is the number of images in the data set (it would take about one week on a laptop). Therefore, we only keep every third image for a total of 810 images, which takes less than three hours for the preprocessing; about 10-15 seconds per image.[7]

Table 1 reports the sparsity and the value of $\rho(B^*_\varepsilon)$ for the preprocessed matrices with different values of the parameter ε. As explained in Section 5.3, the sparsity of $\mathcal{P}_\varepsilon(M)$ increases with ε, and ε was chosen to make sure that $\rho(B^*_\varepsilon) < 1$ implying $\text{rank}(\mathcal{P}_\varepsilon(M)) = \text{rank}(M)$.

	M	$\mathcal{P}(M)$	$\mathcal{P}_{0.05}(M)$	$\mathcal{P}_{0.1}(M)$
s(.)	0	0.001	20.92	38.03
$\rho(B^*_\varepsilon)$	0	0.71	0.83	0.90

Table 1: CBCL data set: sparsity of the preprocessed matrices $\mathcal{P}_\varepsilon(M) = MQ$ and corresponding spectral radius of $B^*_\varepsilon = I - Q$.

Figure 7 displays a sample of images of the CBCL data set along with the corresponding preprocessed images for different values of ε.

Figure 7: From top to bottom: CBCL sample images, corresponding preprocessed images for $\varepsilon = 0$, $\varepsilon = 0.05$, and $\varepsilon = 0.1$.

We observe that the preprocessing is able to highlight some parts of the images: the eyes (faces 5 and 9), the eyebrows (faces 3, 4, 8, 10, 11, 13 and 16), the mustache (faces 14 and 15), the glasses (faces 6, 7 and 12), the nose (faces 1 to 4) or the mouth (faces 1 to 5). Recall that the preprocessing removes from each image of the original data set a linear combination of other images. Therefore,

6. Available at `http://cbcl.mit.edu/software-datasets/FaceData2.html`.
7. The MATLAB® function *lsqlin* for solving CLLS problems is much slower than *quadprog* with interior point (which is much faster than *quadprog* with active set).

the parts of the images which are significantly different from the other images are better preserved, hence highlighted.

We now compare the three approaches described in the introduction of this section. Table 2 gives the numerical results and shows that Pre-NMF performs competitively with sNMF in all cases (similar relative error for similar sparsity levels).

	Plain	Improved	$s(U)$	$s(V)$
SVD	7.28	7.28	0	0
NMF	7.97	7.96	53.27	11.36
Pre-NMF(0)	9.28 (9.76)	8.37	76.78	4.42
sNMF(0)	8.34	8.20	77.62	5.19
Pre-NMF(0.05)	11.12 (12.66)	9.15	90.14	2.16
sNMF(0.05)	9.24	8.90	91.12	2.22
Pre-NMF(0.1)	13.12 (23.47)	9.88	94.58	1.17
sNMF(0.1)	10.30	9.89	94.77	1.14

Table 2: Comparison of the relative approximation error and sparsity for the CBCL image data set. (In brackets, it is the error obtained when using $V = V'Q^{-1}$, instead of $V = \text{argmin}_{X \geq 0} ||M - UX||_F^2$.).

Figure 8 displays the basis elements obtained for NMF, Pre-NMF(0), Pre-NMF(0.1) and sNMF(0.1). The decomposition by parts obtained by Pre-NMF(0.1) is comparable to sNMF(0.1), reinforcing the observation above (cf. Table 2) that Pre-NMF performs competitively with sNMF.

Our technique has the advantage that only one parameter has to be chosen (namely ε) and that sparse solutions are naturally obtained. In fact, the user does not need to know in advance the desired sparsity level: one just has to try different values of $\varepsilon \in [0,1]$ (making sure $\rho(B_\varepsilon^*) < 1$) and a sparse factor U will automatically be generated (no parameters have to be tuned in the course of the optimization process). Moreover, Pre-NMF proves to be less sensitive to initialization than sNMF: we rerun both algorithms for $\varepsilon = 0.1$ with 100 different initializations (using exactly the same settings as above) and observe the following:

- Among the hundred solutions generated by sNMF(0.01), three did not achieve the required sparsity (being lower than 0.85, while all others were around 0.95 as imposed). In particular, the variance of the sparsity of the factor U for PreNMF(0.01) is $8.69\,10^{-7}$ while it is much higher $1.31\,10^{-3}$ for sNMF(0.01). (Note that after removing the three outliers, the variance of sNMF(0.01) is still higher being $3.23\,10^{-6}$.)

- The average of the relative error of Pre-NMF(0.01) is 9.94, slightly lower than sNMF(0.01) with 9.96.

- The variance of the relative error of Pre-NMF(0.01) is $5.97\,10^{-3}$, lower than sNMF(0.01) with $2.36\,10^{-2}$.

Remark 33 *We have also tested other sparse NMF techniques and they could not match the results obtained by sNMF, especially for high sparsity requirement. In particular, we tested the following*

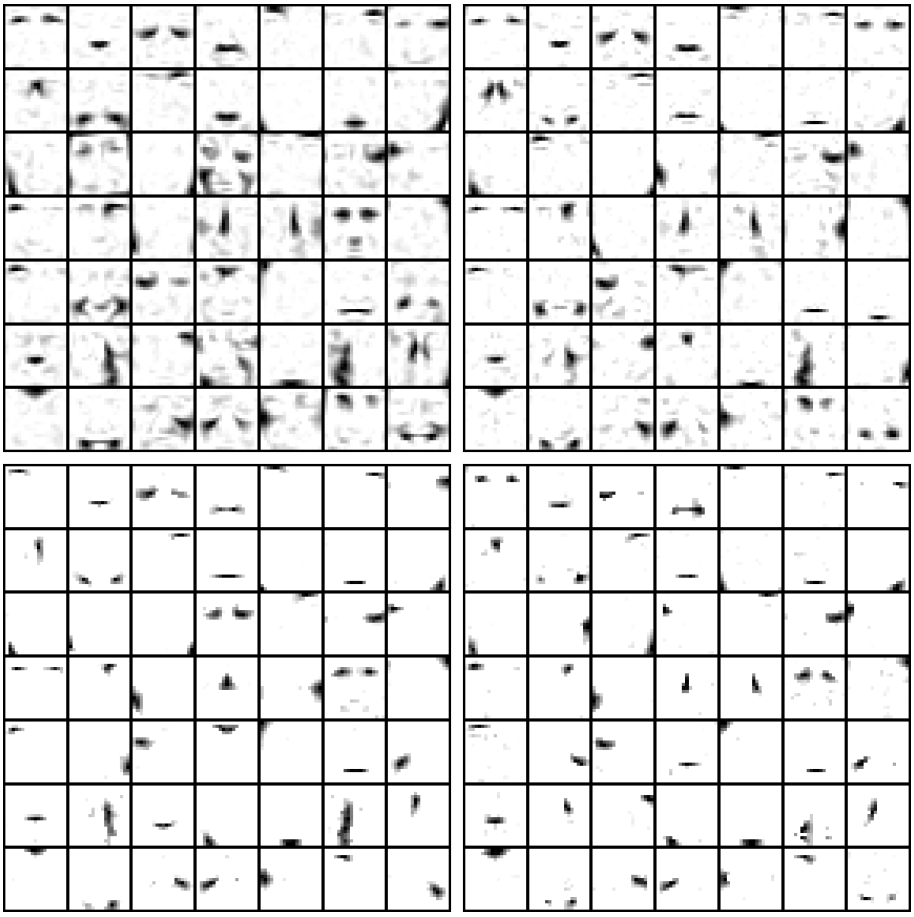

Figure 8: From left to right, top to bottom: basis elements for the CBCL data set obtained with NMF, Pre-NMF(0), Pre-NMF(0.1) and sNMF(0.1).

standard formulation using only one penalty parameter (Kim and Park, 2007)

$$\min_{U,V\geq 0} ||M-UV||_F^2 + \mu \sum_i ||U_{:i}||_1,$$

and the algorithm of Hoyer (2004).

6.2 Hubble Telescope Hyperspectral Image

The Hubble data set consists of 100 spectral images of the Hubble telescope, 128×128 pixels each (Pauca et al., 2006). It is composed of eight materials;[8] see the fourth row on Figure 10. The preprocessing took about one minute (about 0.5 second per image).[9] Figure 9 displays a sample of images of the simulated Hubble database along with the corresponding preprocessed images. The

8. These are true Hubble satellite material spectral signatures provided by the NASA Johnson Space Center.
9. The MATLAB® function *lsqlin* for solving CLLS problems was again much slower (about ten times) than *quadprog* with active set or with interior point which were comparable in this case.

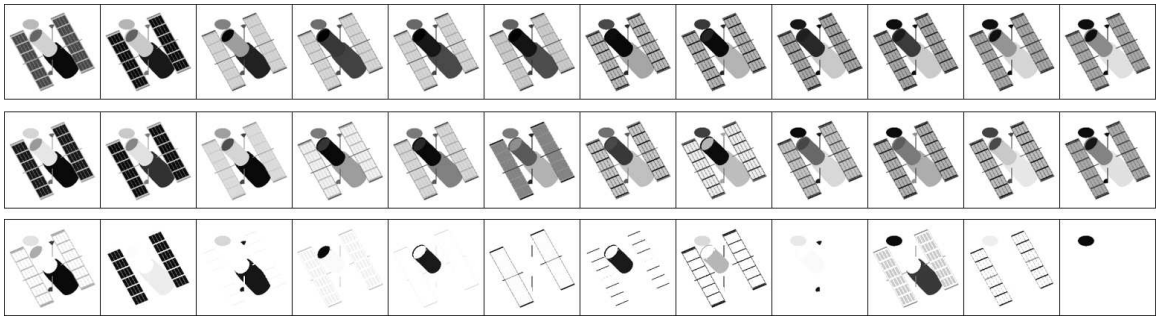

Figure 9: From top to bottom: Sample of Hubble images, corresponding preprocessed images for $\varepsilon = 0$ and $\varepsilon = 0.01$.

preprocessing for $\varepsilon = 0.01$ highlights extremely well the constitutive parts of the Hubble telescope: it is in fact able to extract some materials individually. Table 3 gives the sparsity and the value of ρ for the different preprocessed matrices.

	M	$\mathcal{P}(M)$	$\mathcal{P}_{0.01}(M)$
s(.)	57	57	80
$\rho(B_\varepsilon^*)$	0	0.9808	0.9979

Table 3: Hubble data set: sparsity of the preprocessed matrices $\mathcal{P}_\varepsilon(M) = MD$ and corresponding spectral radius of $B_\varepsilon^* = I - D$.

Table 4 reports the numerical results. Although sNMF(0.01) identifies a solution with slightly lower reconstruction error than Pre-NMF(0.01) (2.90 vs. 2.93), it is not able to identify the constitutive materials properly while Pre-NMF(0.01) *perfectly separates all eight constitutive materials*. It is also important to point out that the solutions generated by Pre-NMF(0.01) with different initializations correspond in most cases[10] to this optimal decomposition while the solutions generated by sNMF are typically very different (and with very different objective function values). This indicates that the NMF problem corresponding to the preprocessed matrix is more well posed.[11]

The comparison between sNMF(0) and Pre-NMF(0) is also interesting: the basis elements generated by Pre-NMF(0) (see second row of Figure 10) identify the constitutive materials much more effectively as six of them are almost perfectly extracted, while sNMF(0) only identifies one (while another is extracted as two separate basis elements).

10. We used 100 random initializations and obtained 61 times the optimal decomposition (in the other cases, it is always able to detect at least six of the eight materials).
11. Of course, in general, even if an NMF formulation has a unique global minimum (up to permutation and scaling), it will still have many local minima. Therefore, even in that situation, solutions generated with standard nonlinear optimization algorithms might still be rather different for different initializations.

	Plain	Improved	$s(U)$	$s(V)$
SVD	0.01	0.01	58	0
NMF	0.06	0.05	58.02	2.25
Pre-NMF(0)	0.08 (0.08)	0.07	59.16	0.13
sNMF(0)	0.37	0.36	64.14	0.63
Pre-NMF(0.01)	14.08 (75.09)	2.93	93.71	0
sNMF(0.01)	3.39	2.90	93.94	0

Table 4: Comparison of the relative approximation error and sparsity for the Hubble data set. (Notice that for $\varepsilon = 0.01$, the solution obtained using $V = V'Q^{-1}$ has a very high reconstruction error; the reason being that $Q = (I - B_\varepsilon^*)$ is close to being singular since $\rho(B_\varepsilon^*) = 0.9979$.)

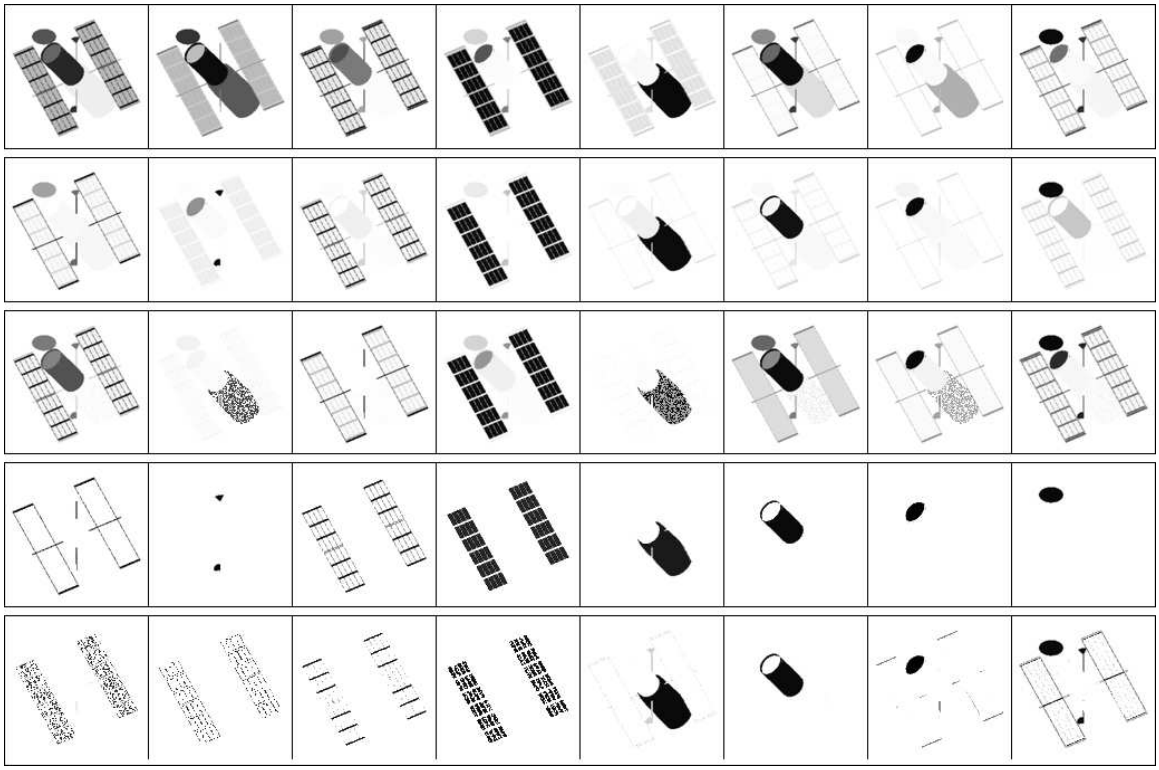

Figure 10: From top to bottom: basis elements for the Hubble data set obtained by NMF, Pre-NMF(0), sNMF(0), Pre-NMF(0.01) and sNMF(0.01).

7. Conclusion and Further Research

In this paper, we introduced a completely new approach to make NMF problems more well posed and have sparser solutions. It is based on the preprocessing of the nonnegative data matrix M:

given M, we compute an inverse positive matrix Q such that the preprocessed matrix $\mathcal{P}(M) = MQ$ remains nonnegative and is sparse. The computation of Q relies on the resolution of constrained linear least squares problems (CLLS). We proved that the preprocessing is well-defined, invariant to permutation and scaling of the columns of matrix M, and preserves the rank of M (as long as the vertices of $\text{conv}(\theta(M))$ are non repeated).

Because $\mathcal{P}(M)$ is sparser than M, the corresponding NMF problem will be more well posed and have sparser solutions. In particular, we were able to show that

- Under the separability assumption of Donoho and Stodden (2003), the preprocessing is optimal as it identifies the vertices of the convex hull of the columns of M.

- Since any rank-two matrix satisfies the separability assumption, the preprocessing is optimal for any nonnegative rank-two matrix.

- In the exact rank-three case (that is, $M = UV$, $\text{rank}(M) = \text{rank}_+(M) = 3$), the preprocessing can be used to make the set of optimal solutions of the NMF problem finite. We conjecture that, generically, it makes it unique and that this result holds for higher rank matrices.

We also proposed a more general preprocessing that relaxes the constraint that $\mathcal{P}(M)$ has to be nonnegative, which is able to deal better with noisy and sparse matrices. Moreover, it generates sparser preprocessed matrices hence sparser NMF solutions. We experimentally showed the effectiveness of this strategy on facial and hyperspectral image data sets. In particular, it performed competitively with a state-of-the-art sparse NMF technique based on ℓ_1-norm penalty functions. It is robust to high sparsity requirement and no parameters have to be tuned in the course of the optimization process. Only one parameter has to be chosen which will allow the user to generate more or less sparse preprocessed matrices.

The main drawback of the technique seems to be its computational cost: n CLLS problems in n variables and $m+n$ constraints have to be solved (where n in the number of columns of M) for a total computational cost of the order of $O(n^{4.5})$ (using MATLAB® on a standard laptop, it limits n to be smaller than 1000 for a few hours of computation). It would then be particularly interesting to investigate strategies to speed up the preprocessing. Using faster solvers is one possible approach (probably in detriment of the accuracy), for example, based on first-order methods.[12] Another possibility would be to use the following heuristic: since the preprocessing removes from each column of M a linear combinations of the other columns, one could use only a subset of k columns of M to be subtracted from the other columns of M. This amounts to fixing variables to zero in the CLLS problems and would reduce the computational complexity to $O(nk^{3.5})$. This subset of columns could for example be selected such that its convex hull has a large volume, see, for example, Klingenberg et al. (2009) for a possible heuristic; or such that they form the best possible basis for the remaining columns (that is, use a column subset selection algorithm); see Boutsidis et al. (2009) and the references therein.

Finally, a particularly challenging direction for research would be to design other data preprocessing techniques for NMF. One approach would be to characterizing the set of inverse positive matrices better: in this paper, we only worked with the subset of invertible M-matrices. For exam-

12. We have developed an alternating direction method (ADM), along with Ting Kei Pong, which allowed us to preprocess the CBCL data set in about 10 hours with 10^{-3} relative accuracy; the code is available upon request.

ple, the matrix[13]

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

would not be modified by our preprocessing (because each column contains a zero entry corresponding to positive ones in all other columns) although its NMF is not unique (cf. Section 2). In fact, we have

$$MQ = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} = 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where Q is inverse positive with $Q^{-1} = \frac{1}{2}M$, and the NMF of MQ is unique. This example shows that working with a larger set of inverse positive matrices would allow to obtain sparser preprocessed data matrices, hence more well-posed NMF problems with sparser solutions.

Acknowledgments

The author acknowledges a discussion with Mariya Ishteva about uniqueness issues of NMF which motivated the study of inverse-positive matrices in this context. The author would like to thank K.C. Sivakumar and F.-X. Orban de Xivry for helpful discussions on inverse positive matrices and on the problem of finding the closest stable matrix to a given one, respectively, and Stephen Vavasis for carefully reading and commenting a first draft of this manuscript. The authors also thanks the anonymous reviewers for their insightful comments, which helped to improve the paper.

Appendix A. Proof for Theorem 29

In this section, we prove that the function f_k defined in Theorem 29 is continuous and made up of pieces which are either constant or strictly convex (which we refer to as piecewise constant/strictly convex). The construction described below is the same as the one proposed by Aggarwal et al. (1989) and we refer the reader to that paper for more details. The novelty of our proof is to use that construction to show that f_k is piecewise constant/strictly convex (it was already shown to be continuous and nondecreasing by Aggarwal et al., 1989).

Proof Let $x(t_1)$ be on the boundary of P and define the sequence $x(t_2), \ldots, x(t_{k+1})$ as in Theorem 29 (clock-wise). As shown by Aggarwal et al. (1989), the function $f_k(t_1) = t_{k+1}$ only depends on

1. The sides of P on which the points $x(t_i)$ $1 \leq i \leq k+1$ lie ;
2. The intersections of the segments $[x(t_i), x(t_{i+1})]$ $1 \leq i \leq k$ with Q ;

and, given that these sides and intersections do not change, f_k is continuously differentiable and can be characterized in closed form (see below). These sides and intersections will change when either

- One of the points $x(t_i)$ switches from one side of the boundary of P to another. These points correspond to the vertices of P (P has at most m vertices since it is a polygon defined with m inequalities); or,

13. We thank Mariya Ishteva for providing us with this example.

- One of the intersections of the segments $[x(t_i), x(t_{i+1})]$ $1 \leq i \leq k$ with Q changes. There is a one-to-one correspondence between these points and the sides of Q (Q has at most n vertices hence at most n sides).

These points where the description of f_k changes (and where f_k is not continuously differentiable) are called the contact change points. Turning around the boundary of P, we might encounter more than $m+n$ such points. However, two contact change points corresponding to the same change are associated with the same sequence $x(t_i)$ $1 \leq i \leq k+1$ hence the same solution to the NPP. In fact, both sequences must share at least one point (either a vertex of P or the intersections of a line containing a side of Q with the boundary of P) which implies, by construction, that they are the same. Therefore, there are at most $m+n$ contact change points corresponding to different sequences $x(t_i)$ $1 \leq i \leq k+1$ on the boundary of P (Aggarwal et al., 1989).

It remains to show that the pieces of f_k between two contact change points are either constant or strictly convex.

Let us then construct the function f_k between two contact change points. Without loss of generality, we may assume that the perimeter of the outer polygon P is equal to one (otherwise scale the polygons P and Q accordingly), and that the parametrization x of the boundary of P has the following property: the distance traveled when following the boundary between $x(s)$ and $x(t)$ is equal to $|(s-\lfloor s \rfloor)-(t-\lfloor t \rfloor)|$. In particular, if $0 \leq s \leq t \leq 1$, then the distance traveled between $x(t)$ and $x(s)$ along the boundary of P is $t-s$. We may also assume without loss of generality that $x(0)=(0,0)$ is the vertex on P preceding $x(t_1)$ and that $x(t_1)=(0,t_1)$: this amounts to translating and rotating P and Q. We also define (see Figure 11 for an illustration)

- $q=(q_1,q_2)$, the tangent point on Q between $x(t_1)$ and $x(t_2)$.
- θ, the angle between the sides of P on which $x(t_1)$ and $x(t_2)$ are.
- p, the intersection between the sides on which $x(t_1)$ and $x(t_2)$ are (note that p is on the boundary of P if and only if there is one and only one vertex of P between $x(t_1)$ and $x(t_2)$).
- d, the distance between $x(0)$ and p.
- s, the distance between p and $x(t_2)$.
- a, the projection of q on the line $[x(0),p]$.
- b, the projection of $x(t_2)$ on the line $[x(0),p]$.

Case 1: The point q is on the same side as $x(t_1)$. This implies that $x(t_2)=p$ for any $t_1 < q_1$ and no other points of the sequence is changed since $x(t_2)$ remains the same. Therefore, the function $t_{k+1}=f_k(t_1)$ is constant. (Notice that $x(q_1)$ is a contact change point since $x(t_2)$ will switch side when $t_1=q_1$.)
Case 2: The point q is on the same side as $x(t_2)$. This implies that $x(t_2)=q$ for any $t_1 < d$. Therefore, the function $t_{k+1}=f_k(t_1)$ is constant. (Notice that the next contact change point will be the first vertex of P that $x(t_1)$ crosses.)
Case 3: The point q is not on the same side as $x(t_1)$ or $x(t_2)$ (it is in the interior of P). Using the similarity between the triangles $\Delta x(t_1)aq$ and $\Delta x(t_1)bx(t_2)$, we have that (Aggarwal et al., 1989, Equation (1))

$$\frac{q_2}{q_1-t_1}=\frac{s\sin(\theta)}{d-t_1+s\cos(\theta)},$$

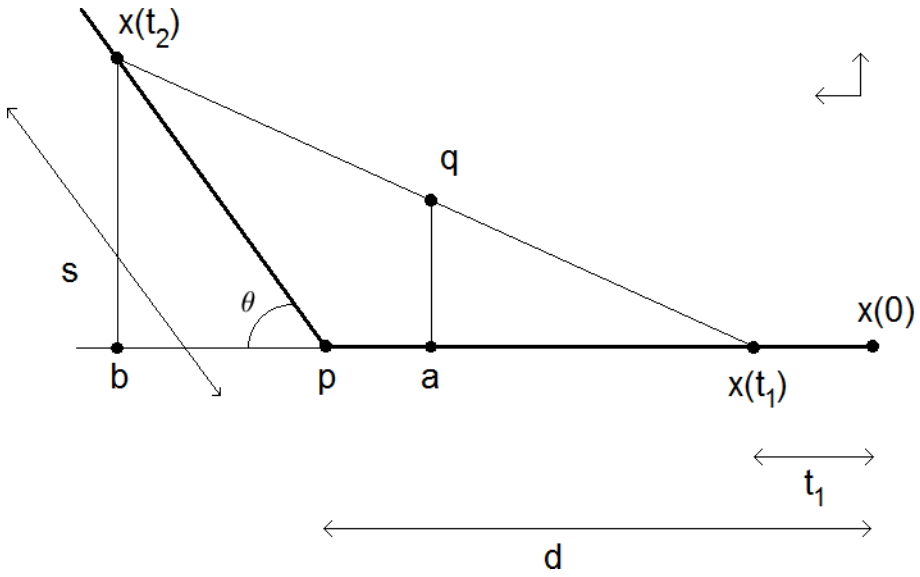

Figure 11: Construction of the function f_1 between two contact change points (see Aggarwal et al. 1989, Figure 3, for a similar illustration).

implying

$$s = \frac{q_2}{\sin(\theta)} \frac{d - t_1}{q_1 - q_2 \cot(\theta) - t_1} = g_1(t_1).$$

Let us show that $g_1(t_1)$ is strictly convex, that is, $g_1''(t_1) > 0$. Since q is not on the same side as $x(t_1)$ or $x(t_2)$, we have $q_2 > 0$ and $0 < \theta < \pi$ implying $\frac{q_2}{\sin(\theta)} > 0$. Hence it suffices to show that $h(t_1) = \frac{d-t_1}{l-t_1}$ is strictly convex, where $l = q_1 - q_2 \cot(\theta)$. Since $s > 0$ and $d > t_1$, we must have $l - t_1 > 0$. (Notice that $x(l)$ is a contact change point. In fact, for $t_1 = l$, the segments $[x(t_1), q]$ and $[p, x(t_2)]$ become parallel implying that the intersection of Q with the segment $[x(t_1), x(t_2)]$ will change.)

We then have

$$h'(t_1) = \frac{d - l}{(l - t_1)^2}.$$

Since h is a strictly increasing function of t_1 (Aggarwal et al., 1989), $h'(t_1) > 0$ hence $d > l$ and

$$h''(t_1) = 2\frac{d - l}{(l - t_1)^3} > 0,$$

so that $g_1(t)$ is strictly convex. Finally, we have

$$f_1(t_1) = t_2 = c_1 + s = c_1 + g_1(t_1),$$

where either

- $c_1 = 0$ and g_1 is a constant (cases 1. and 2.).

- c_1 is an appropriate constant and g_1 is an increasing and strictly convex function (case 3.).

By construction, the same relationship will apply between t_2 and t_3 with

$$f_2(t_1) = t_3 = c_2 + g_2(s) = c_2 + g_2(g_1(t_1)),$$

where c_2 is an appropriate constant and g_2 is either constant, or strictly convex and increasing. After $k+1$ steps, we have

$$f_k(t_1) = t_{k+1} = c_k + g_k(s) = c_k + (g_k \circ g_{k-1} \circ \cdots \circ g_1)(t_1),$$

where c_k is an appropriate constant and the functions g_i are either constant, or strictly convex and increasing. If one of the functions g_i $1 \leq i \leq k$ is constant, then f_k is constant. Otherwise the function $f_k(t_1) = c_k + (g_{k-1} \circ \cdots \circ g_1)(t_1)$ is strictly convex since it is a constant plus the composition of strictly convex and *increasing* functions. (In fact, the composition of one-dimensional increasing and strictly convex functions is increasing and strictly convex.) ■

References

A. Aggarwal, H. Booth, J. O'Rourke, and S. Suri. Finding minimal convex nested polygons. *Information and Computation*, 83(1):98–110, 1989.

S. Arora, R. Ge, R. Kannan, and A. Moitra. Computing a nonnegative matrix factorization – provably. In *Proceedings of the 44th Symposium on Theory of Computing*, STOC '12, pages 145–162, New York, NY, USA, 2012. ACM.

A. Berman and R.J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. SIAM, 1994.

C. Bocci, E. Carlini, and F. Rapallo. Perturbation of matrices and non-negative rank with a view toward statistical models. *SIAM. J. Matrix Anal. & Appl.*, 32(4):1500–1512, 2011.

C. Boutsidis, M.W. Mahoney, and P. Drineas. An improved approximation algorithm for the column subset selection problem. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, pages 968–977, Philadelphia, PA, USA, 2009. Society for Industrial and Applied Mathematics.

M.T. Chu and M.M. Lin. Low-dimensional polytope approximation and its applications to nonnegative matrix factorization. *SIAM J. Sci. Comput.*, 30(3):1131–1155, 2008.

A. Cichocki, S. Amari, R. Zdunek, and A.H. Phan. *Non-negative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. Wiley-Blackwell, 2009.

M.D. Craig. Minimum-volume tranforms for remotely sensed data. *IEEE Trans. on Geoscience and Remote Sensing*, 32(3):542–552, 1994.

C. Ding, X. He, and H.D. Simon. On the equivalence of nonnegative matrix factorization and spectral clustering. In *SIAM Int. Conf. Data Mining (SDM'05)*, pages 606–610, 2005.

C. Ding, T. Li, W. Peng, and H. Park. Orthogonal nonnegative matrix tri-factorizations for clustering. In *In Proc. of the 12th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, pages 126–135, 2006.

D. Donoho and V. Stodden. When does non-negative matrix factorization give a correct decomposition into parts? In *Advances in Neural Information Processing Systems 16*, 2003.

C. Févotte, N. Bertin, and J.L. Durrieu. Nonnegative matrix factorization with the itakura-saito divergence: With application to music analysis. *Neural Computation*, 21(3):793–830, 2009.

N. Gillis. *Nonnegative Matrix Factorization: Complexity, Algorithms and Applications*. PhD thesis, Université catholique de Louvain, 2011.

N. Gillis and F. Glineur. Nonnegative factorization and the maximum edge biclique problem. CORE discussion paper 2008/64, 2008.

N. Gillis and F. Glineur. Using underapproximations for sparse nonnegative matrix factorization. *Pattern Recognition*, 43(4):1676–1687, 2010.

N. Gillis and F. Glineur. On the geometric interpretation of the nonnegative rank. *Linear Algebra and its Applications*, 437(11):2685–2712, 2012a.

N. Gillis and F. Glineur. Accelerated multiplicative updates and hierarchical ALS algorithms for nonnegative matrix factorization. *Neural Computation*, 24(4):1085–1105, 2012b.

N. Gillis and S.A. Vavasis. Fast and robust recursive algorithms for separable nonnegative matrix factorization. arXiv:1208.1237, 2012.

H. Hazewinkel. On positive vectors, positive matrices and the specialization ordering. Technical report, CWI Report PM-R8407, 1984.

P.O. Hoyer. Nonnegative matrix factorization with sparseness constraints. *J. Machine Learning Research*, 5:1457–1469, 2004.

A. Huck, M. Guillaume, and J. Blanc-Talon. Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data. *IEEE Trans. on Geoscience and Remote Sensing*, 48 (6):2590–2602, 2010.

V. Kalofolias and E. Gallopoulos. Computing symmetric nonnegative rank factorizations. *Linear Algebra and its Applications*, 436(2):421–435, 2012.

H. Kim and H. Park. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics*, 23(12):1495–1502, 2007.

B. Klingenberg, J. Curry, and A. Dougherty. Non-negative matrix factorization: Ill-posedness and a geometric algorithm. *Pattern Recognition*, 42(5):918–928, 2009.

H. Laurberg, M.G. Christensen, M.D. Plumbley, L.K. Hansen, and S.H. Jensen. Theorems on positive data: On the uniqueness of NMF. *Computational Intelligence and Neuroscience*, 2008. Article ID 764206.

D.D. Lee and H.S. Seung. Learning the parts of objects by nonnegative matrix factorization. *Nature*, 401:788–791, 1999.

L. Miao and H. Qi. Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. *IEEE Trans. on Geoscience and Remote Sensing*, 45 (3):765–777, 2007.

P. Paatero and U. Tapper. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5:111–126, 1994.

V.P. Pauca, J. Piper, and R.J. Plemmons. Nonnegative matrix factorization for spectral data analysis. *Linear Algebra and its Applications*, 406(1):29–47, 2006.

B.T. Polyak and P.S. Shcherbakov. Hard problems in linear control theory: Possible approaches to solution. *Automation and Remote Control*, 66:681–718, 2005.

F. Pompili, N. Gillis, P.-A. Absil, and F. Glineur. Two algorithms for orthogonal nonnegative matrix factorization with application to clustering. arXiv:1201.090, 2011.

G.B. Recht, M. Fazel, and P.A. Parrilo. Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization. *SIAM Review*, 52(3):471–501, 2010.

Y. Sun and J. Xin. Underdetermined sparse blind source separation of nonnegative and partially overlapped data. *SIAM Journal on Scientific Computing*, 33(4):2063–2094, 2011.

O. Taussky. A recurring theorem on determinants. *The American Mathematical Monthly*, 56(10): 672–676, 1949.

F.J. Theis, K. Stadlthanner, and T. Tanaka. First results on uniqueness of sparse non-negative matrix factorization. In *13th European Signal Processing Conference, EUSIPCO, Antalya, Turkey*, 2005.

L.B. Thomas. Rank factorization of nonnegative matrices. *SIAM Review*, 16(3):393–394, 1974.

S.A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM J. on Optimization*, 20 (3):1364–1377, 2009.

F.-Y. Wang, C.-Y. Chi, T.-H. Chan, and Y. Wang. Nonnegative least-correlated component analysis for separation of dependent sources by volume maximization. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(5):875–888, 2010.

W. Xu, X. Liu, and Y. Gong. Document clustering based on non-negative matrix factorization. In *Proc. of the 26th Annual Int. ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '03, pages 267–273, New York, NY, USA, 2003. ACM.

G. Zhou, S. Xie, Z. Yang, J.-M. Yang, and Z. He. Minimum-volume-constrained nonnegative matrix factorization: Enhanced ability of learning parts. *IEEE Trans. on Neural Networks*, 22(10):1626–1637, 2011.