Nonnegative Matrix Factorization

Chapter 2. Exact NMF

Taehyeong Kim

Mathematics, Pusan National University

May 7, 2021

Content

Introduction

In this section, we consider the Exact NMF problem defined as follows.

Problem 2.1 (Exact NMF)

Given a nonnegative matrix $X \in \mathbb{R}_+^{m\times n}$ and a factorization rank r, compute, if possible, two nonnegative matrices $W \in \mathbb{R}_+^{m\times r}$ and $H \in \mathbb{R}_+^{r\times n}$ such that

$$X = WH$$

We refer to WH as an Exact NMF of X of size r.

The nonnegative rank is the topic of Chapter 3. But here is notation.

Nonnegative rank

Exact NMF is closely related to the quantity referred to as the nonnegative rank of X which is the smallest r such that X admits an Exact NMF size r, and it is denoted $\mathrm{rank}_+(X)$.

2.1. Geometric interpretation

The geometric interpretation is a key aspect of NMF. (See Chapter 4, 5, 7)

In this section

- Describe the geometric interpretation of Exact NMF in term of nested convex cone.
- Describe the geometric interpretation of Exact NMF in term of nested convex hulls.
- Prove $\mathrm{rank}_+(X) = \mathrm{rank}(X)$ for any nonnegative matrix X such that $\mathrm{rank}(X) \leq 2$.
- Provide example of Exact NMF.

2.1.1. Interpretation with nested convex cone

Definition [1] (Cone)

A set C is called a *cone*, if for every $x \in C$ and $\theta \geq 0$ we have $\theta x \in C$.

2.1.1. Interpretation with nested convex cone

Definition [1] (Cone)

A set C is called a *cone*, if for every $x \in C$ and $\theta \geq 0$ we have $\theta x \in C$.

Definition [1] (Convex cone)

A set C is called a *convex cone*, if it is convex and cone, which means that for every $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$ we have $\theta_1 x_1 + \theta_2 x_2 \in C$.

2.1.1. Interpretation with nested convex cone

Definition [1] (Cone)

A set C is called a *cone*, if for every $x \in C$ and $\theta \geq 0$ we have $\theta x \in C$.

Definition [1] (Convex cone)

A set C is called a *convex cone*, if it is convex and cone, which means that for every $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$ we have $\theta_1 x_1 + \theta_2 x_2 \in C$.

Definition (Cone in this book)

Given a matrix $A \in \mathbb{R}^{m \times n}$, $\mathrm{cone}(A)$ is the convex cone generated by the columns of A, that is,

$$\mathrm{cone}(A) = \left\{ x | x = Ay \text{ for some } y \in \mathbb{R}^n, y \geq 0 \right\}$$

2.1.1. Interpretation with nested convex cone

Definition [1] (Conic hull)

The *conic hull* of a set C is the set of all conic combinations of points in C, i.e.,

$$\{\theta_1 x_1 + \cdots + \theta_k x_k | x_i \in C, \theta_i \geq 0, i = 1, \ldots, k\},$$

which is also the smallest convex cone that contains C.

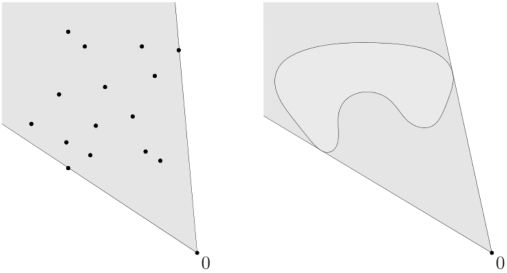


Figure 1: The conic hulls (shown shaded) of the two sets [1]

2.1.1. Interpretation with nested convex cone

Remark 2.1

dimension of $\mathrm{cone}(A)$ = *dimension of the subspace spaned by* $A = \mathrm{rank}(A)$

2.1.1. Interpretation with nested convex cone

Remark 2.1

dimension of $\mathrm{cone}(A)$ = *dimension of the subspace spaned by* $A = \mathrm{rank}(A)$

Let us consider the Exact NMF of matrix $X = WH$. Since $X(:, j) = WH(:, j), W \geq 0$, and $H \geq 0$,

$$X(:, j) \in \mathrm{cone}(W) \subseteq \mathbb{R}_+^m$$

for all j. Equivalently,

$$\mathrm{cone}(X) \subseteq \mathrm{cone}(W) \subseteq \mathbb{R}_+^m,$$

which is a nested cone problem.

2.1.1. Interpretation with nested convex cone

Remark 2.1

dimension of $\mathrm{cone}(A)$ = *dimension of the subspace spaned by* $A = \mathrm{rank}(A)$

Let us consider the Exact NMF of matrix $X = WH$. Since $X(:, j) = WH(:, j), W \geq 0$, and $H \geq 0$,

$$X(:, j) \in \mathrm{cone}(W) \subseteq \mathbb{R}^m_+$$

for all j. Equivalently,

$$\mathrm{cone}(X) \subseteq \mathrm{cone}(W) \subseteq \mathbb{R}^m_+,$$

which is a nested cone problem.

Nested cone problem

Given two cones nested in each other, namely $\mathrm{cone}(X) \subseteq \mathbb{R}^m_+$, find a cone nested between them, namely $\mathrm{cone}(W)$.

2.1.1. Interpretation with nested convex cone

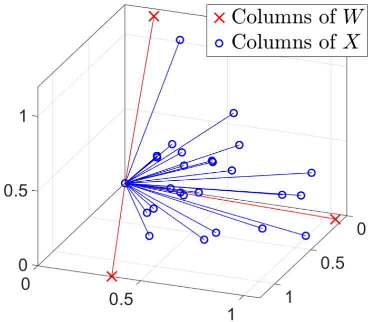


Figure 2: Geometric illustration of Exact NMF for $m = r = 3$ and $n = 25$.

Figure 2 shows that $\text{cone}(X) \subseteq \text{cone}(W) \subseteq \mathbb{R}^3_+$.

2.1.2. Interpretation with nested convex hulls

Why is nested convex hull used more than nested convex cone?

- Intuition
- Related literature
- Illustration

2.1.2. Interpretation with nested convex hulls

Before going further, let us define a few useful notations about convex hulls and polytopes.

Definition [1] (Convex)

A set C is *convex* if the line segment between any two points in C lies in C, i.e., if for any $x_1, x_2 \in C$and any θ with $0 \leq \theta \leq 1$, we have

$$\theta x_1 + (1-\theta)x_2 \in C.$$

2.1.2. Interpretation with nested convex hulls

Before going further, let us define a few useful notations about convex hulls and polytopes.

Definition [1] (Convex)

A set C is *convex* if the line segment between any two points in C lies in C, i.e., if for any $x_1, x_2 \in C$and any θ with $0 \leq \theta \leq 1$, we have

$$\theta x_1 + (1 - \theta)x_2 \in C.$$

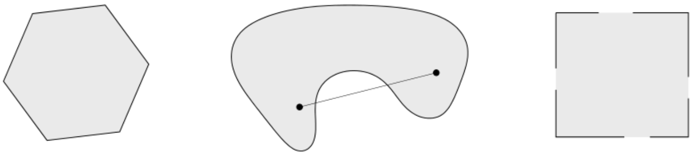


Figure 3: Some simple convex and nonconvex sets[1]

2.1.2. Interpretation with nested convex hulls

Before going further, let us define a few useful notations about convex hulls and polytopes.

Definition [1] (Convex hull)

The *convex hull* of a set C, denoted $\operatorname{conv}(C)$, is the set of all convex combinations of points in C:

$$\operatorname{conv}(C) = \{\theta_1 x_1 + \cdots + \theta_k x_k | x_i \in C, \, \theta_i \geq 0, \, i = 1, \ldots, k, \, \theta_1 + \cdots + \theta_k = 1\}$$

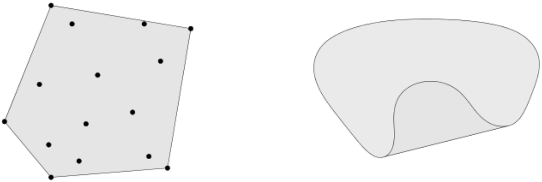


Figure 4: The convex hulls of two sets[1].

2.1.2. Interpretation with nested convex hulls

Before going further, let us define a few useful notations about convex hulls and polytopes.

Definition (Convex hull in this book)

Given a $A \in \mathbb{R}^{m \times n}$, $\mathrm{conv}(A)$ is convex hull of the columns of a A, i.e.,

$$\mathrm{conv}(A) = \left\{x | x = Ay \text{ for some } y \in \mathbb{R}^n, y \geq 0, \text{ and } e^T y = 1\right\}$$

where e is the vector of all ones of appropriate dimension, so that $e^T y = \sum_{i=1}^n y_i$.

2.1.2. Interpretation with nested convex hulls

How to select vertices of $\mathrm{conv}(A)$?

Note that all vertices of $\mathrm{conv}(A)$ are contained in the set of the columns of A.

$$
\begin{aligned}
& A(:,j) \text{ is a vertex of } \mathrm{conv}(A) \\
\Leftrightarrow & A(:,j) \notin \mathrm{conv}(A(:,\mathcal{J})) \text{ where } \mathcal{J} = \{1,2,\ldots,n\} \setminus \{j\}
\end{aligned}
$$

2.1.2. Interpretation with nested convex hulls

How to select vertices of $\mathrm{conv}(A)$?

Note that all vertices of $\mathrm{conv}(A)$ are contained in the set of the columns of A.

$$\begin{aligned} & A(:,j) \text{ is a vertex of } \mathrm{conv}(A) \\ \Leftrightarrow & A(:,j) \notin \mathrm{conv}(A(:,\mathcal{J})) \text{ where } \mathcal{J} = \{1,2,\dots,n\} \setminus \{j\} \end{aligned}$$

Definition (Simnplex in dimension r)

$$\Delta^r = \left\{ x \in \mathbb{R}^r | x \geq 0, e^T x = 1 \right\} = \mathrm{conv}(I_r) \tag{1}$$

where I_r is the identity matrix of dimension r.

In fact, the vertices of Δ^r are unit vectors, and $I_r = (e_1, e_2, \dots, e_r)$.

2.1.2. Interpretation with nested convex hulls

Definition (Column space)

$$\mathrm{col}(A) = \left\{x | x = Ay \text{ for } y \in \mathbb{R}^n\right\}$$

2.1.2. Interpretation with nested convex hulls

Definition (Column space)

$$\mathrm{col}(A) = \left\{x | x = Ay \text{ for } y \in \mathbb{R}^n\right\}$$

and its affine hull as

Definition (Affine hull of $\mathrm{col}(A)$)

$$\mathrm{aff}(A) = \left\{x | x = Ay \text{ for } y \in \mathbb{R}^n \text{ and } e^T y = 1\right\} \tag{2}$$

2.1.2. Interpretation with nested convex hulls

Note

Note that containment relationship among $\mathrm{conv}(A), \mathrm{aff}(A), \mathrm{col}(A)$

$$\mathrm{conv}(A) \subseteq \mathrm{aff}(A) \subseteq \mathrm{col}(A)$$

Note

If $\mathrm{aff}(A)$ contains the origin, then $\mathrm{aff}(A) = \mathrm{col}(A)$.

$$\mathrm{conv}(A) = \left\{x | x = Ay \text{ for some } y \in \mathbb{R}^n, y \geq 0, \text{ and } e^T y = 1\right\}$$
$$\mathrm{aff}(A) = \left\{x | x = Ay \text{ for } y \in \mathbb{R}^n \text{ and } e^T y = 1\right\}$$
$$\mathrm{col}(A) = \left\{x | x = Ay \text{ for } y \in \mathbb{R}^n\right\}$$

2.1.2. Interpretation with nested convex hulls

Let us describe the geometric interpretation of Exact NMF in term of nested convex hulls. Given an Exact NMF of $X = WH$, the following two assumptions can be made without loss of generality:

Assumtion

A1 The matrices X and W do not contain columns equal to the zero vector.

A2 $\|X(:,j)\|_1 = \|W(:,j)\|_1 = 1$ for all j,k.
Given a matrix A with nonzero columns, let D_A be the diagonal matrix as

$$D_A = \operatorname{diag}\left(\frac{1}{\|A(:,1)\|_1}, \frac{1}{\|A(:,2)\|_1}, \ldots, \frac{1}{\|A(:,n)\|_1}\right)$$

Let $\theta(A) = AD_A$, then

$$\begin{aligned} X = WH \Leftrightarrow \underbrace{XD_X}_{\theta(X)} &= \underbrace{(WD_W)}_{\theta(W)} \underbrace{\left(D_W^{-1} H D_X\right)}_{H'} \\ \Leftrightarrow \theta(X) &= \theta(W)H' \end{aligned} \tag{3}$$

2.1.2. Interpretation with nested convex hulls

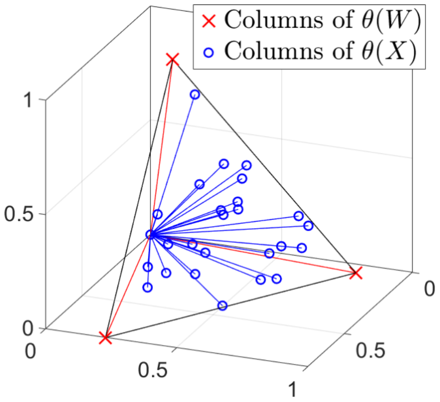


Figure 5: Geometric illustration of Exact NMF for $m = r = 3$ and $n = 25$.

2.1.2. Interpretation with nested convex hulls

Lemma 2.1. [2, Thm 3.2]

Let (W, H) be any factorization of X. If the entries in each column of X and W sum to one, then the entries in each column of H sum to one.

Proof.

Since $e^T X = e^T$ and $e^T W = e^T$, we have

$$e^T = e^T X = e^T W H = e^T H.$$

□

2.1.2. Interpretation with nested convex hulls

Theorem 2.2.

Computing an Exact NMF of X of size r is equivalent to finding r vertices within the unit simplex Δ^m whose convex hull contains the columns of $\theta(X)$.

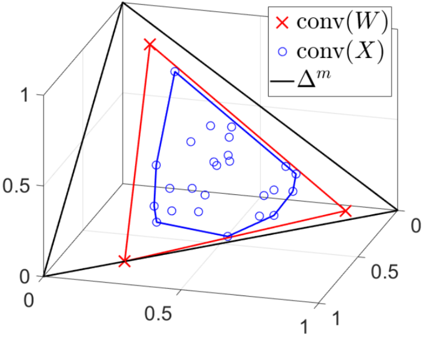


Figure 6: Geometric illustration of Theorem 2.2 with $\mathrm{conv}(X) \subseteq \mathrm{conv}(W) \subseteq \Delta^m$ where the columns of X and W have unit ℓ_1-norm.

2.1.2.1. Reducing the dimension by one

For unit simplex Δ^m, the following holds:

$$\text{for } x \in \Delta^m, x_i = 1 - \sum_{j \neq i} x_j \text{ for all } i.$$

Hence, when consider the interpretation of Exact NMF in terms of nested convex hulls, it is possible to reduce the dimension of the problem by one and represent it in a lower dimensional subspace. Let us define

$$\mathcal{S}^r = \left\{ x \in \mathbb{R}^r | x \geq 0, e^T x \leq 1 \right\},$$

which is the convex hull of the unit simplex and the origin, that is $\mathcal{S}^r = \mathrm{conv}([I_r, 0])$. For a matrix A such that $A(:, j) \in \Delta^m$, denote $\bar{A} = A(1 : m-1, :)$ such that $\bar{A}(:, j) \in \mathcal{S}^{m-1}$

2.1.2.1. Reducing the dimension by one

Lemma 2.3

For $x \in \Delta^m, W \in \mathbb{R}^{m \times r}$ such that $W(:,j) \in \Delta^m$ for all j, and $h \in \Delta^r$,

$$x = Wh \Leftrightarrow \bar{x} = \bar{W}h$$

Proof.

Let $x = \begin{pmatrix} \bar{x} \\ x_m \end{pmatrix}$ and $W = \begin{pmatrix} \bar{W} \\ W_m^T \end{pmatrix}$. Since $x \in \Delta^m$ and $W(:,j) \in \Delta^m$ for all j, $x_m = 1 - e^T \bar{x}$ and $w_m = e - \bar{W}^T e$.

$(\Rightarrow)$ Trivial.

$(\Leftarrow)$ If $\bar{x} = \bar{W}h$, since $e^T h = 1$,

$$x_m = 1 - e^T \bar{x} = 1 - e^T \bar{W} h = 1 - \left(\bar{W}^T e\right)^T h = 1 - (e - w_m)^T h = w_m^T h$$

Therefore, $x = Wh$.

$\square$

2.1.2.1. Reducing the dimension by one

Lemma 2.3 implies that for $X \geq 0$ and $W \geq 0$ whose columns have unit ℓ_1-norm, we have

$$\mathrm{conv}(X) \subseteq \mathrm{conv}(W) \subseteq \Delta^m \quad \Leftrightarrow \quad \mathrm{conv}(\bar{X}) \subseteq \mathrm{conv}(\bar{W}) \subseteq \mathcal{S}^{m-1}.$$

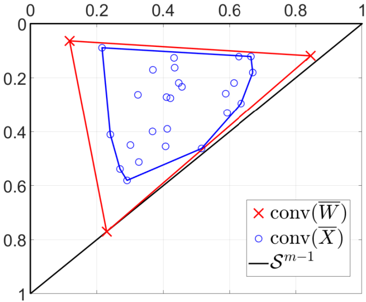


Figure 7

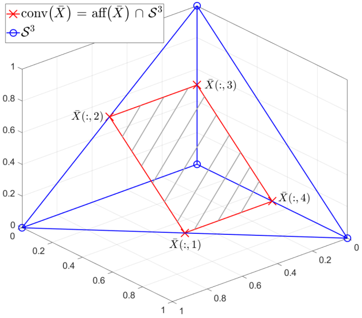


Figure 8

2.1.3. Deriving the result of Thomas

In [5], author introduce relationship between the rank and the nonnegative rank of a matrix. We need two lemmas to prove the geometric interpretation of Exact NMF in terms of convex hulls.

Lemma 2.4.

Let $X \in \mathbb{R}_+^{m \times n}$ be a matrix whose columns have unit ℓ_1-norm, that is, $X^T e = e$. We have

$$\operatorname{aff}(X) = \operatorname{col}(X) \cap \left\{ x | e^T x = 1 \right\}$$

and

$$\operatorname{conv}(X) \subseteq \operatorname{col}(X) \cap \Delta^m.$$

2.1.3. Deriving the result of Thomas

Proof.

By Definition of Affine hull of $\mathrm{col}(X)$ and columns of X have unit ℓ_1-norm, we have

$$\begin{aligned}\mathrm{aff}(X) &= \left\{x | x = X\alpha \text{ for } \alpha \in \mathbb{R}^n \text{ and } e^T\alpha = 1\right\} \\ &= \left\{x | x = X\alpha \text{ for } \alpha \in \mathbb{R}^n \text{ and } e^T x = 1\right\} \quad (\because e^T x = e^T X\alpha = e^T\alpha = 1) \\ &= \mathrm{col}(X) \cap \left\{x | e^T x = 1\right\}\end{aligned}$$

Since $X \geq 0$ and $\mathrm{conv}(X) \subseteq \mathrm{aff}(X)$,

$$\mathrm{conv}(X) \subseteq \mathrm{aff}(X) \cap \mathbb{R}^m_+ = \mathrm{col}(X) \cap \left\{x | e^T x = 1\right\} \cap \mathbb{R}^m_+ = \mathrm{col}(X) \cap \Delta^m$$

□

2.1.3. Deriving the result of Thomas

Lemma 2.5.

Let $X \in \mathbb{R}^{m \times n}$*, be a nonnegative matrix with no column equal to the zero vector. Then,* $\mathrm{conv}(\theta(X))$ *and* $\mathrm{col}(X) \cap \Delta^m$ *are polytopes of dimension* $\mathrm{rank}(X) - 1$*.*

Proof.

We have $\dim(\mathrm{conv}(\theta(X))) = \dim(\mathrm{aff}(\theta(X)))$ and $\mathrm{col}(X) = \mathrm{col}(XD_X) = \mathrm{col}(\theta(X))$. And Lemma 2.4 implies

$$\mathrm{aff}(\theta(X)) = \mathrm{col}(X) \cap \left\{x | e^T x = 1\right\}$$

and

$$\mathrm{col}(X) \cap \Delta^m = \mathrm{col}(X) \cap \left\{x | e^T x = 1\right\} \cap \mathbb{R}_+^m.$$

By definition, $\mathrm{col}(X) \cap \Delta^m$ and $\mathrm{conv}(\theta(X))$ has same dimension of the polytopes.

2.1.3. Deriving the result of Thomas

Proof (Cont.)

And $\dim(\text{aff}(\theta(X))) = \text{rank}(X) - 1$.
Let us show this rigorously using algebraic arguments.

$$\begin{aligned}
\text{aff}(\theta(X)) &= \left\{ \sum_{i=1}^{n} \alpha_i Y(:,i) | \alpha \in \mathbb{R}^n, \sum_{i=1}^{n} \alpha_i = 1 \right\} \\
&= \left\{ Y(:,n) + \sum_{i=1}^{n-1} \alpha_i (Y(:,i) - Y(:,n)) | \alpha_i \in \mathbb{R} \text{ for } i = 1, 2, \ldots n-1 \right\} \\
&= Y(:,n) + \text{col}\left(Y'\right)
\end{aligned}$$

2.1.3. Deriving the result of Thomas

The first result of Thomas [5] from Lemma 2.5

Theorem 2.6. [5]

If X is a nonnegative matrix with $\operatorname{rank}(X) \leq 2$, then $\operatorname{rank}(X) = \operatorname{rank}_+(X)$.

2.1.3. Deriving the result of Thomas

Remark 2.2 (Link with separable NMF)

When $\operatorname{rank}(X) = 2$, the two columns of W in Exact NMF of X of size $r = 2$ can be picked from the columns of X. Unfortunately, this fact does not hold for $\operatorname{rank}(X) \geq 3$.

2.1.3. Deriving the result of Thomas

Lemma 2.7.

Let $X \in \mathbb{R}_+^{m \times n}$*, and let* $X = WH$ *be an Exact NMF of* X *of size* $r = \operatorname{rank}(X)$*. Then* $\operatorname{rank}(W) = \operatorname{rank}(X)$ *and* $\operatorname{col}(W) = \operatorname{col}(X)$*.*

2.1.3. Deriving the result of Thomas

Lemma 2.8.

Let $X \in \mathbb{R}_{+}^{m\times n}$*, and* $W \in \mathbb{R}_{+}^{m\times r}$ *be matrices whose columns have unit* ℓ_1*-norm and be such that* $\mathrm{col}(W) = \mathrm{col}(X)$*. Then*

$$\mathrm{aff}(W) = \mathrm{aff}(X)$$

2.1.3. Deriving the result of Thomas

Corollary 2.9.

Let $X \in \mathbb{R}_+^{m\times n}$*, and* $W \in \mathbb{R}_+^{m\times r}$ *be matrices whose columns have unit* ℓ_1*-norm and be such that* $\operatorname{col}(W) = \operatorname{col}(X)$*. Then*

$$\operatorname{col}(\bar{W}) = \operatorname{col}(\bar{X}) \quad \textit{and} \quad \operatorname{aff}(\bar{W}) = \operatorname{aff}(\bar{X})$$

2.1.3. Deriving the result of Thomas

Theorem 2.10. [5]

The nonnegative rank of

$$X = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \tag{4}$$

is equal to 4*, while* $\operatorname{rank}(X) = 3$*.*

2.1.3. Deriving the result of Thomas

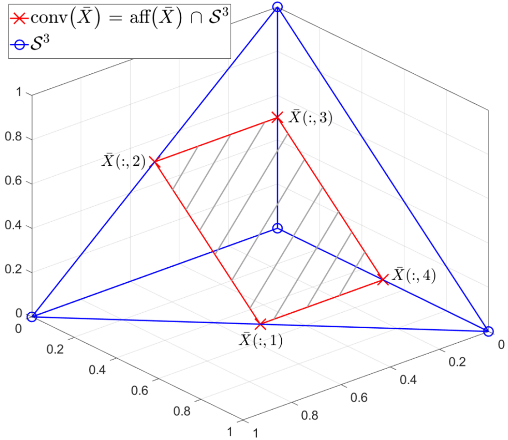


Figure 9: Geometric illustration of Exact NMF for the 4-by-4 matrix X of Thomas from (4). The matrix $\bar{X}$ is obtained by discarding the last row of matrix X. The columns of $\bar{X}$ are the vectors $(0.5, 0.5, 0), (0.5, 0, 0.5), (0, 0.5, 0)$, and $(0, 0, 0.5)$.

2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

In this section...

- Example of Exact NMF corresponding to a geometric problems in [3, 4]
- Gain more insight into the geometric interpretation of Exact NMF.

Let the nonnegative matrix $X_a \in \mathbb{R}^{n \times n}$

$$X_a = \frac{1}{6a} \begin{bmatrix} 1 & a & 2a-1 & 2a-1 & a & 1 \\ 1 & 1 & a & 2a-1 & 2a-1 & a \\ a & 1 & 1 & a & 2a-1 & 2a-1 \\ 2a-1 & a & 1 & 1 & a & 2a-1 \\ 2a-1 & 2a-1 & a & 1 & 1 & a \\ a & 2a-1 & 2a-1 & a & 1 & 1 \end{bmatrix} \tag{5}$$

then every column of X_a has unit ℓ_1-norm.
Aim : Determine the nonnegative rank of X_a depending on the parameter a.

2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

Geometry of $\mathrm{conv}(X_a)$ and $\mathrm{col}(X_a) \cap \Delta^6$

For $a > 1$, $\mathrm{rank}(X_a) = 3$ and we can factorize X_a as

$$X_a = \frac{1}{6a}\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}\begin{bmatrix} -1 & a-2 & -1 & 1-2a & 2-3a & 1-2a \\ 1 & 1 & a & 2a-1 & 2a-1 & a \\ a & 1 & 1 & a & 2a-1 & 2a-1 \end{bmatrix}. \tag{6}$$

Since the basis in (6) does not depend on the parameter a, the matrices X_a for $a > 1$ share the some column space. And $\mathrm{conv}(X_a)$ is a hexagon. This hexagon is contained in $\mathrm{aff}(X_a) \cap \Delta^6$. (For example, see Figure 10 and 11.)

2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

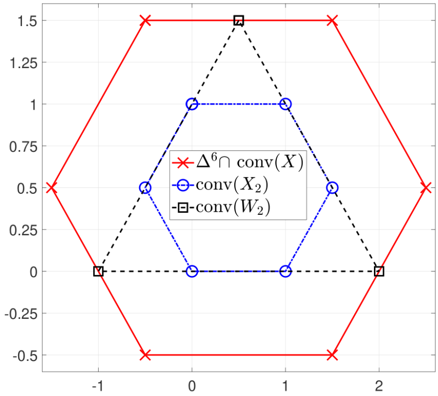


Figure 10: Geometric illustration of Exact NMF for the 6-by-6 matrix X_2 from (5).

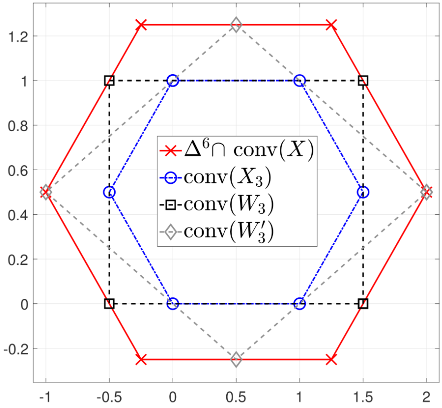


Figure 11: Geometric illustration of Exact NMF for the 6-by-6 matrix X_3 from (5).

2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

Geometry of $\mathrm{conv}(X_a)$ and $\mathrm{col}(X_a) \cap \Delta^6$

By Lemma 2.5, $\mathrm{conv}(X_a)$ is two dimensional polytope.
And by Lemma 2.4, $\mathrm{aff}(X_a) = \mathrm{col}(X_a) \cap \{x | e^T x = 1\}$, and taking the intersection with the nonnegative orthant, we obtain
$\mathrm{aff}(X_a) \cap \mathbb{R}^6_+ = \mathrm{col}(X_a) \cap \Delta^6$.
By Lemma 2.5, $\dim(\mathrm{col}(X_a) \cap \Delta^6) = 2$. So $\mathrm{col}(X_a) \cap \Delta^6$ is polygon, and hexagon.

$$X = \lim_{a \to +\infty} X_a = \frac{1}{6} \begin{bmatrix} 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 & 0 \end{bmatrix}. \tag{7}$$

Lemma 2.4.

Let $X \in \mathbb{R}^{m \times n}_+$ be a matrix whose columns have unit ℓ_1-norm, that is, $X^T e = e$. We have

$$\mathrm{aff}(X) = \mathrm{col}(X) \cap \{x | e^T x = 1\} \quad \text{and} \quad \mathrm{conv}(X) \subseteq \mathrm{col}(X) \cap \Delta^m.$$

Lemma 2.5.

Let $X \in \mathbb{R}^{m \times n}$, be a nonnegative matrix with no column equal to the zero vector. Then, $\mathrm{conv}(\theta(X))$ and $\mathrm{col}(X) \cap \Delta^m$ are polytopes of dimension $\mathrm{rank}(X) - 1$.

2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

Geometry of $\mathrm{conv}(X_a)$ and $\mathrm{col}(X_a) \cap \Delta^6$

Since $\mathrm{col}(X_a) = \mathrm{col}(X)$, we have

$$\begin{aligned} \mathrm{col}(X_a) \cap \Delta^6 &= \mathrm{col}(X) \cap \Delta^6 \\ &= \left\{ z | z = Xy, y \in \mathbb{R}^6, z \geq 0, e^T z = 1 \right\} \\ &= \mathrm{conv}(X). \end{aligned}$$

2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

What is the nonnegative rank of X_a

Let $X_a = W_a H_a$ be Exact NMF of X_a of size $\mathrm{rank}_+(X_a)$, and assume that the columns of W_a have unit ℓ_1-norm. Exact NMF is equivalent to finding W_a such that

$$\mathrm{conv}(X_a) \subseteq \mathrm{conv}(W_a) \subseteq \Delta^6$$

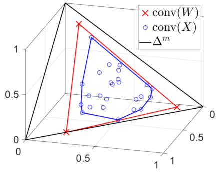
Since $\mathrm{conv}(X_a) \subseteq \mathrm{conv}(X_b)$ for $a \leq b$, $\mathrm{rank}(X_a) \leq \mathrm{rank}(X_b)$ for $a \leq b$.
We have that $\mathrm{rank}_+(X_a) \geq \mathrm{rank}(X_a) = 3$. If $\mathrm{rank}_+(X_a) = \mathrm{rank}(X_a) = 3$, Lemma 2.7 implies that $\mathrm{col}(W_a) = \mathrm{col}(X_a)$.

Theorem 2.2.

Computing an Exact NMF of X of size r is equivalent to finding r vertices with in the unit simplex Δ^m whose convex hull contains the columns of $\theta(X)$.

Lemma 2.7.

Let $X \in \mathbb{R}_+^{m\times n}$, and let $X = WH$ be an Exact NMF of X of size $r = \mathrm{rank}(X)$. Then $\mathrm{rank}(W) = \mathrm{rank}(X)$ and $\mathrm{col}(W) = \mathrm{col}(X)$.



2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

What is the nonnegative rank of X_a

And by Lemma 2.4, we have

$$\mathrm{conv}(W_a) \subseteq \mathrm{conv}(X_a) \cap \Delta^6.$$

When $\mathrm{col}(W_a) = \mathrm{col}(X_a)$.
Since W_a has three columns and $\mathrm{rank}(W_a) = 3$, $\mathrm{conv}(W_a)$ is a triangle.

Lemma 2.4.

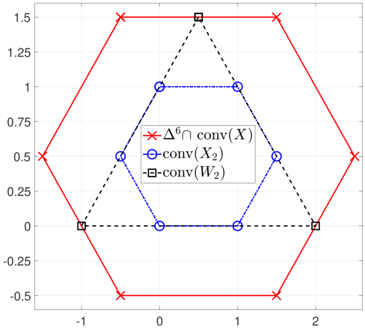
Let $X \in \mathbb{R}_+^{m \times n}$ be a matrix whose columns have unit ℓ_1-norm, that is, $X^T e = e$. We have

$$\mathrm{aff}(X) = \mathrm{col}(X) \cap \left\{ x | e^T x = 1 \right\} \quad \text{and} \quad \mathrm{conv}(X) \subseteq \mathrm{col}(X) \cap \Delta^m.$$

2.1.4 On the necessity of $\text{rank}(W) > \text{rank}(X)$

What is the nonnegative rank of X_a

Geometrically, the problem of checking whether $\text{rank}_+(X_a) = 3$ reduces to checking whether there exists a triangle nested between two hexagons, namely $\text{conv}(X_a)$ and $\text{col}(X_a) \cap \Delta^6 = \text{conv}(X)$.
For $a = 2$ shown in Figure 10, four triangles fit between the two hexagons. This implies that for any $1 < a < 2$, $\text{rank}_+(X_a) = 3$.



2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

What is the nonnegative rank of X_a

To compute an Exact NMF, the vertices of $\mathrm{conv}(W_2)$ can be obtained by averaging two consecutive vertices of $\mathrm{conv}(X)$, where X is given in (7); see Figure 10. We have

$$X_2 = \frac{1}{12}\begin{bmatrix} 1 & 2 & 3 & 3 & 2 & 1 \\ 1 & 1 & 2 & 3 & 3 & 2 \\ 2 & 1 & 1 & 2 & 3 & 3 \\ 3 & 2 & 1 & 1 & 2 & 3 \\ 3 & 3 & 2 & 1 & 1 & 2 \\ 2 & 3 & 3 & 2 & 1 & 1 \end{bmatrix}$$

$$= \underbrace{\frac{1}{12}\begin{bmatrix} 1 & 4 & 1 \\ 0 & 3 & 3 \\ 1 & 1 & 4 \\ 3 & 0 & 3 \\ 4 & 1 & 1 \\ 3 & 3 & 0 \end{bmatrix}}_{W_2} \begin{bmatrix} 2/3 & 2/3 & 1/3 & 0 & 0 & 1/3 \\ 0 & 1/3 & 2/3 & 2/3 & 1/3 & 0 \\ 1/3 & 0 & 0 & 1/3 & 2/3 & 2/3 \end{bmatrix},$$

2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

What is the nonnegative rank of X_a

where

$$W_2 = \frac{1}{12}\begin{bmatrix} 1 & 4 & 1 \\ 0 & 3 & 3 \\ 1 & 1 & 4 \\ 3 & 0 & 3 \\ 4 & 1 & 1 \\ 3 & 3 & 0 \end{bmatrix} = X\begin{bmatrix} 1/2 & 0 & 0 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \\ 0 & 0 & 1/2 \end{bmatrix}.$$

Another W of $X_2 = WH$ are

$$\frac{1}{12}\begin{bmatrix} 0 & 3 & 3 \\ 1 & 1 & 4 \\ 3 & 0 & 3 \\ 4 & 1 & 1 \\ 3 & 3 & 0 \\ 1 & 4 & 1 \end{bmatrix}, \frac{1}{6}\begin{bmatrix} 0 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \\ 2 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 0 \end{bmatrix}, \frac{1}{6}\begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

2.1.4 On the necessity of $\mathrm{rank}(W) > \mathrm{rank}(X)$

What is the nonnegative rank of X_a

Other results:

- $\mathrm{rank}_+(X_a) = 4$ for $3 \geq a > 2$.
- $\mathrm{rank}_+(X_a) \geq 5$ for $a > 3$.
- $\mathrm{rank}_+(X_a) = 5$ for $a > 3$.
- $\mathrm{rank}_+(X) = 5$

2.1.4 On the necessity of $\text{rank}(W) > \text{rank}(X)$

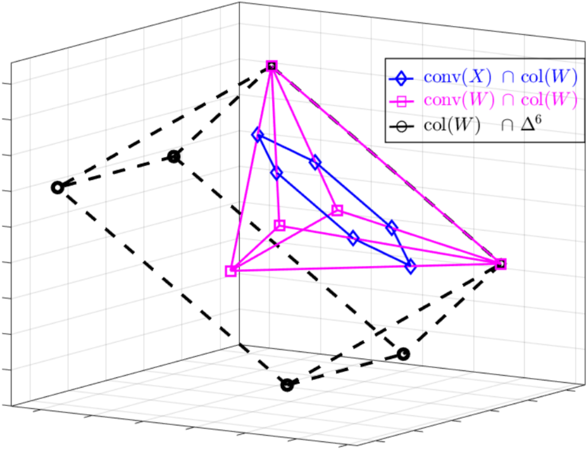


Figure 12: Geometric illustration of the Exact NMF of the 6-by-6 matrix X from (7) representing a hexagon contained within a three-dimensional polytope with five vertices within the unit simplex Δ^6. [3]

Reference

[1] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[2] Joel E Cohen and Uriel G Rothblum. “Nonnegative ranks, decompositions, and factorizations of nonnegative matrices”. In: *Linear Algebra and its Applications* 190 (1993), pp. 149–168.

[3] Nicolas Gillis. “Introduction to nonnegative matrix factorization”. In: *arXiv preprint arXiv:1703.00663* (2017).

[4] David Mond, Jim Smith, and Duco Van Straten. “Stochastic factorizations, sandwiched simplices and the topology of the space of explanations”. In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 459.2039 (2003), pp. 2821–2845.

[5] Rekha R Thomas. “Spectrahedral lifts of convex sets”. In: *Proceedings of the International Congress of Mathematicians (ICM 2018)*. World Scientific. 2018, pp. 3819–3842.

Thank you!