On Identifiable Polytope Characterization for Polytopic Matrix Factorization

ICASSP 2022

Bariscan Bozkurt and Alper T. Erdogan

KUIS AI-Center and EE Engineering

Koc University

Outline

- Matrix Factorization Problem

Outline

- Matrix Factorization Problem
- Polytopic Matrix Factorization

Outline

- Matrix Factorization Problem
- Polytopic Matrix Factorization
- Identifiable Polytopes

Outline

- Matrix Factorization Problem
- Polytopic Matrix Factorization
- Identifiable Polytopes
- Brute-force Approach

Outline

- Matrix Factorization Problem
- Polytopic Matrix Factorization
- Identifiable Polytopes
- Brute-force Approach
- Group Structure of Symmetries

Outline

- Matrix Factorization Problem
- Polytopic Matrix Factorization
- Identifiable Polytopes
- Brute-force Approach
- Group Structure of Symmetries
- Exploiting Graph Automorphism

Outline

- Matrix Factorization Problem
- Polytopic Matrix Factorization
- Identifiable Polytopes
- Brute-force Approach
- Group Structure of Symmetries
- Exploiting Graph Automorphism
- Proposed Approach

Outline

- Matrix Factorization Problem
- Polytopic Matrix Factorization
- Identifiable Polytopes
- Brute-force Approach
- Group Structure of Symmetries
- Exploiting Graph Automorphism
- Proposed Approach
- Numerical Example

Matrix Factorization Problem

Generative Model

$$\underbrace{\begin{bmatrix} | & \cdots & | \\ \mathbf{y}_g[1] & \cdots & \mathbf{y}_g[N] \\ | & \cdots & | \end{bmatrix}}_{\mathbf{Y}_g\text{(observations)}} = \underbrace{\begin{bmatrix} & & \\ & \mathbf{H}_g & \\ & & \end{bmatrix}}_{\mathbf{H}_g\text{(linear transf.)}} \underbrace{\begin{bmatrix} | & \cdots & | \\ \mathbf{s}_g[1] & \cdots & \mathbf{s}_g[N] \\ | & \cdots & | \end{bmatrix}}_{\mathbf{S}_g\text{(latent vectors)}}$$

Matrix Factorization Problem

Generative Model

$$\underbrace{\begin{bmatrix} | & \cdots & | \\ \mathbf{y}_g[1] & \cdots & \mathbf{y}_g[N] \\ | & \cdots & | \end{bmatrix}}_{\mathbf{Y}_g(\text{observations})} = \underbrace{\begin{bmatrix} & & \\ & \mathbf{H}_g & \\ & & \end{bmatrix}}_{\mathbf{H}_g(\text{linear transf.})} \underbrace{\begin{bmatrix} | & \cdots & | \\ \mathbf{s}_g[1] & \cdots & \mathbf{s}_g[N] \\ | & \cdots & | \end{bmatrix}}_{\mathbf{S}_g(\text{ latent vectors})}$$

Given $\mathbf{Y}_g$, find $\mathbf{S}_g$ and $\mathbf{H}_g$

Matrix Factorization Problem

Generative Model

$$\underbrace{\begin{bmatrix} | & \cdots & | \\ \mathbf{y}_g[1] & \dots & \mathbf{y}_g[N] \\ | & \cdots & | \end{bmatrix}}_{\mathbf{Y}_g(\text{observations})} = \underbrace{\begin{bmatrix} \\ \mathbf{H}_g \\ \\ \end{bmatrix}}_{\mathbf{H}_g(\text{linear transf.})} \underbrace{\begin{bmatrix} | & \cdots & | \\ \mathbf{s}_g[1] & \dots & \mathbf{s}_g[N] \\ | & \cdots & | \end{bmatrix}}_{\mathbf{S}_g(\text{ latent vectors})}$$

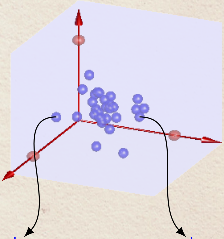
Given $\mathbf{Y}_g$, find $\mathbf{S}_g$ and $\mathbf{H}_g$

We need to make assumptions on $\mathbf{S}_g$ and $\mathbf{H}_g$

Matrix Factorization Problem

Nonnegative Matrix Factorization

The columns of $\mathbf{S}_g$ are drawn from the nonnegative orthant ($\mathbb{R}^r_+$).

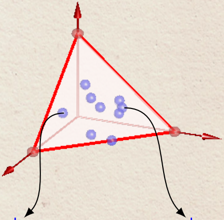


$$\begin{bmatrix} | & \cdots & | \\ \mathbf{y}_g[1] & \cdots & \mathbf{y}_g[N] \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} & \mathbf{H}_g & \end{bmatrix} \begin{bmatrix} | & \cdots & | \\ \mathbf{s}_g[1] & \cdots & \mathbf{s}_g[N] \\ | & \cdots & | \end{bmatrix}$$

Matrix Factorization Problem

Simplex-Structured Matrix Factorization

The columns of $\mathbf{S}_g$ are drawn from the unit simplex (Δ).

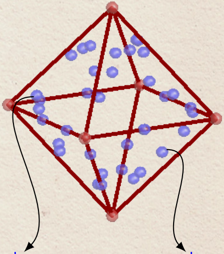


$$\begin{bmatrix} | & \cdots & | \\ \mathbf{y}_g[1] & \cdots & \mathbf{y}_g[N] \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} & \mathbf{H}_g & \end{bmatrix} \begin{bmatrix} | & \cdots & | \\ \mathbf{s}_g[1] & \cdots & \mathbf{s}_g[N] \\ | & \cdots & | \end{bmatrix}$$

Matrix Factorization Problem

Sparse Component Analysis

The columns of $\mathbf{S}_g$ are drawn from the ℓ_1-norm ball $(\mathcal{B}_1)$.

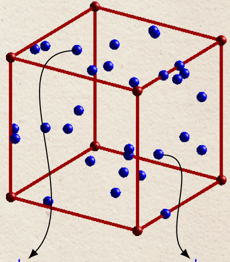


$$\begin{bmatrix} | & \cdots & | \\ \mathbf{y}_g[1] & \cdots & \mathbf{y}_g[N] \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} & & \\ & \mathbf{H}_g & \\ & & \end{bmatrix} \begin{bmatrix} | & \cdots & | \\ \mathbf{s}_g[1] & \cdots & \mathbf{s}_g[N] \\ | & \cdots & | \end{bmatrix}$$

Matrix Factorization Problem

Bounded Component Analysis

The columns of $\mathbf{S}_g$ are drawn from the ℓ_∞-norm ball ($\mathcal{B}_\infty$).

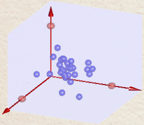


$$\begin{bmatrix} | & \cdots & | \\ \mathbf{y}_g[1] & \cdots & \mathbf{y}_g[N] \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} \\ \mathbf{H}_g \\ \\ \end{bmatrix} \begin{bmatrix} | & \cdots & | \\ \mathbf{s}_g[1] & \cdots & \mathbf{s}_g[N] \\ | & \cdots & | \end{bmatrix}$$

Matrix Factorization Problem

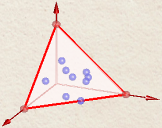
Summary: Domains Enabling Identifiability

$\mathbb{R}^r_+$



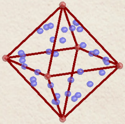
NMF

Δ



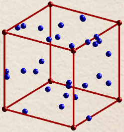
SSMF

$\mathcal{B}_1$



SCA

$\mathcal{B}_\infty$

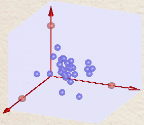


BCA

Matrix Factorization Problem

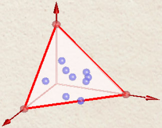
Summary: Domains Enabling Identifiability

$\mathbb{R}^r_+$



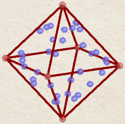
NMF

Δ



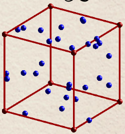
SSMF

$\mathcal{B}_1$



SCA

$\mathcal{B}_\infty$



BCA

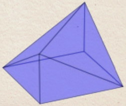
Can we find other domains enabling identifiablity?

Polytopic Matrix Factorization

PMF Generative Model

input matrix

- $\mathbf{Y} = \mathbf{H}_g \mathbf{S}_g$
- $\mathbf{H}_g$ is a full-column-rank matrix
- $\mathbf{S}_{g:,i} \in \mathcal{P}, i = 1, \ldots, N.$

where $\mathcal{P} =$ 

polytope

Polytopic Matrix Factorization

Det-Max Criterion

$$\underset{\mathbf{H} \in \mathbb{R}^{m\times n}, \mathbf{S} \in \mathbb{R}^{n\times N}}{\text{maximize}} \quad \log\det(\hat{\mathbf{R}}_S)$$

subject to

$$\mathbf{Y} = \mathbf{HS}$$

$$\hat{\mathbf{R}}_S = \frac{1}{N}\sum_{k=1}^{N} \mathbf{S}_{:,k}\mathbf{S}_{:,k}^T$$

$$\mathbf{S}_{:,k} \in \mathcal{P}, \quad k = 1, \ldots, N.$$

Spread the columns of S

inside the polytope P

Identifiable Polytopes

Theorem about Identifiable Polytopes[1]:

Fundamental Theorem of PMF

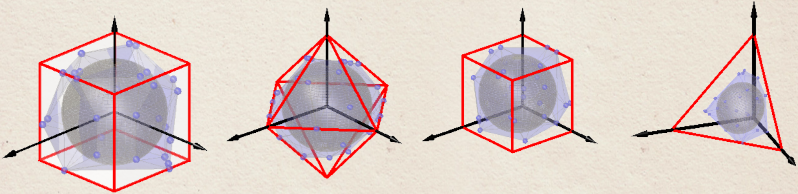
A polytope $\mathcal{P}$ is identifiable if its symmetry group is restricted to component permutations and sign alterations.

---

[1] Gokcan Tatli and Alper T. Erdogan. "Polytopic Matrix Factorization: Determinant Maximization Based Criterion and Identifiability". In: *IEEE Transactions on Signal Processing* 69 (2021), pp. 5431–5447. DOI: 10.1109/TSP.2021.3112918.

Identifiable Polytopes

Example: Four Special Polytopes

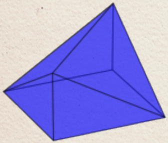


$\mathcal{B}_\infty$	$\mathcal{B}_1$	$\mathcal{B}_\infty \cap \Re^r_+$	$\mathcal{B}_1 \cap \Re^r_+$
Antisparse	Sparse	Nonnegative Antisparse	Nonnegative Sparse

Identifiable Polytopes

Example: A Polytope with Mixed Features

$$\mathcal{P} = \left\{ \mathbf{s} \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} s_1, s_2 \in [-1, 1], s_3 \in [0, 1], \\ \left\| \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \right\|_1 \leq 1, \left\| \begin{bmatrix} s_2 \\ s_3 \end{bmatrix} \right\|_1 \leq 1 \end{array} \right\}$$

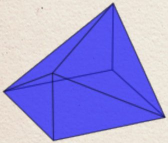


- s_1, s_2: signed, s_3: nonnegative
- s_1, s_2: mutually sparse
- s_2, s_3: mutually sparse

Identifiable Polytopes

Example: A Polytope with Mixed Features

$$\mathcal{P} = \left\{ \mathbf{s} \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} s_1, s_2 \in [-1, 1], s_3 \in [0, 1], \\ \left\| \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \right\|_1 \leq 1, \left\| \begin{bmatrix} s_2 \\ s_3 \end{bmatrix} \right\|_1 \leq 1 \end{array} \right\}$$



- s_1, s_2: signed, s_3: nonnegative
- s_1, s_2: mutually sparse
- s_2, s_3: mutually sparse

How to determine a polytope's identifiability?

Brute-force Approach

- Let $\mathbf{V}_{\mathcal{P}} \in \mathbb{R}^{n\times m}$ be the vertex matrix of $\mathcal{P} \in \mathbb{R}^{n}$ containing m vertices of $\mathcal{P}$ in its columns.

Brute-force Approach

- Let $\mathbf{V}_{\mathcal{P}} \in \mathbb{R}^{n \times m}$ be the vertex matrix of $\mathcal{P} \in \mathbb{R}^n$ containing m vertices of $\mathcal{P}$ in its columns.
- Let $\mathbf{\Pi} \in \mathbb{R}^{m \times m}$ denote a permutation matrix.

Brute-force Approach

- Let $\mathbf{V}_{\mathcal{P}} \in \mathbb{R}^{n \times m}$ be the vertex matrix of $\mathcal{P} \in \mathbb{R}^{n}$ containing m vertices of $\mathcal{P}$ in its columns.
- Let $\mathbf{\Pi} \in \mathbb{R}^{m \times m}$ denote a permutation matrix.
- $\mathcal{P}$ is identifiable $\iff$ $\{\mathbf{G} : \mathbf{G}\mathbf{V}_{\mathcal{P}} = \mathbf{V}_{\mathcal{P}}\mathbf{\Pi}, \mathbf{\Pi} \in \mathbb{R}^{m \times m}\}$ only contains signed permutation matrices.

Brute-force Approach

- Brute force approach requires search on all possible permutation matrices.

Brute-force Approach

- Brute force approach requires search on all possible permutation matrices.

Factorial Complexity !

Group Structure

Lemma

For a given polytope $\mathcal{P} \in \mathbb{R}^n$ with $\mathbf{V}_{\mathcal{P}} \in \mathbb{R}^{n \times m}$, the set $\{\mathbf{G} : \mathbf{G}\mathbf{V}_{\mathcal{P}} = \mathbf{V}_{\mathcal{P}}\mathbf{\Pi}, \mathbf{\Pi} \in \mathbb{R}^{m \times m}\}$ together with the matrix multiplication forms a group. We denote it with $\mathscr{G}(\mathcal{P})$

Group Structure

Theorem

Let $\mathit{Gen}(\mathscr{G}(\mathcal{P})) = \{\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_r\}$ be the generating set of $\mathscr{G}(\mathcal{P})$ for a given polytope $\mathcal{P} \in \mathbb{R}^n$. If all the elements of $\mathit{Gen}(\mathscr{G}(\mathcal{P}))$ are signed permutation matrices, then $\mathcal{P}$ is identifiable.

Group Structure

Theorem

Let $\mathit{Gen}(\mathscr{G}(\mathcal{P})) = \{\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_r\}$ be the generating set of $\mathscr{G}(\mathcal{P})$ for a given polytope $\mathcal{P} \in \mathbb{R}^n$. If all the elements of $\mathit{Gen}(\mathscr{G}(\mathcal{P}))$ are signed permutation matrices, then $\mathcal{P}$ is identifiable.

How to determine $\mathit{Gen}(\mathscr{G}(\mathcal{P}))$?

Group Structure

Theorem

Let $\mathit{Gen}(\mathscr{G}(\mathcal{P})) = \{\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_r\}$ be the generating set of $\mathscr{G}(\mathcal{P})$ for a given polytope $\mathcal{P} \in \mathbb{R}^n$. If all the elements of $\mathit{Gen}(\mathscr{G}(\mathcal{P}))$ are signed permutation matrices, then $\mathcal{P}$ is identifiable.

How to determine $\mathit{Gen}(\mathscr{G}(\mathcal{P}))$?

What is $|\mathit{Gen}(\mathscr{G}(\mathcal{P}))|$?

Graph Automorphism Group

Exploit Graph Automorphism Algorithms!

Graph Automorphism Group

Exploit Graph Automorphism Algorithms!

- For $\mathcal{P} \in \mathbb{R}^n$ with $\mathbf{V}_{\mathcal{P}} \in \mathbb{R}^{n \times m}$,

Graph Automorphism Group

Exploit Graph Automorphism Algorithms!

- For $\mathcal{P} \in \mathbb{R}^n$ with $\mathbf{V}_{\mathcal{P}} \in \mathbb{R}^{n \times m}$,
- Let $\mathbf{Q} = \mathbf{V}_{\mathcal{P}} \mathbf{V}_{\mathcal{P}}^T \in \mathbb{R}^{n \times n}$,

Graph Automorphism Group

Exploit Graph Automorphism Algorithms!

- For $\mathcal{P} \in \mathbb{R}^n$ with $\mathbf{V}_{\mathcal{P}} \in \mathbb{R}^{n\times m}$,
- Let $\mathbf{Q} = \mathbf{V}_{\mathcal{P}}\mathbf{V}_{\mathcal{P}}^T \in \mathbb{R}^{n\times n}$,
- Let $\mathbf{C} = \mathbf{V}_{\mathcal{P}}^T\mathbf{Q}^{-1}\mathbf{V}_{\mathcal{P}} \in \mathbb{R}^{m\times m}$

Graph Automorphism Group

Exploit Graph Automorphism Algorithms!

- For $\mathcal{P} \in \mathbb{R}^n$ with $\mathbf{V}_\mathcal{P} \in \mathbb{R}^{n \times m}$,
- Let $\mathbf{Q} = \mathbf{V}_\mathcal{P}\mathbf{V}_\mathcal{P}^T \in \mathbb{R}^{n \times n}$,
- Let $\mathbf{C} = \mathbf{V}_\mathcal{P}^T\mathbf{Q}^{-1}\mathbf{V}_\mathcal{P} \in \mathbb{R}^{m \times m}$
- Let $G_\mathcal{P}$ be the edge-colored complete graph where edge color from i-th node to j-th node is $\mathbf{V}_{\mathcal{P}_{:,i}}^T\mathbf{Q}^{-1}\mathbf{V}_{\mathcal{P}_{:,j}}$

Graph Automorphism Group

- Graph Automorphism Group:
  $\mathscr{G}(G_{\mathcal{P}}) = \{\boldsymbol{\Pi} \in \mathbb{R}^{m \times m} : \mathbf{C} = \boldsymbol{\Pi}^T \mathbf{C} \boldsymbol{\Pi}\}$

Graph Automorphism Group

- Graph Automorphism Group: $\mathscr{G}(G_{\mathcal{P}}) = \{\mathbf{\Pi} \in \mathbb{R}^{m \times m} : \mathbf{C} = \mathbf{\Pi}^T \mathbf{C} \mathbf{\Pi}\}$

$$\boxed{\mathscr{G}(G_{\mathcal{P}}) \cong \mathscr{G}(\mathcal{P}) \ (\cong \text{ denotes isomorphism})}$$

Graph Automorphism Group

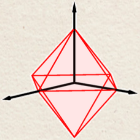
- Graph Automorphism Group:
$\mathscr{G}(G_{\mathcal{P}}) = \{\mathbf{\Pi} \in \mathbb{R}^{m\times m} : \mathbf{C} = \mathbf{\Pi}^T \mathbf{C} \mathbf{\Pi}\}$

$$\boxed{\mathscr{G}(G_{\mathcal{P}}) \cong \mathscr{G}(\mathcal{P}) \ (\cong \text{ denotes isomorphism})}$$

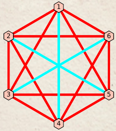
$$\boxed{|Gen(\mathscr{G}(G_{\mathcal{P}}))| \leq m - 1}$$ [2]

[2] Tommi Junttila and Petteri Kaski. "Engineering an Efficient Canonical Labeling Tool for Large and Sparse Graphs". In: *Proceedings of the Meeting on Algorithm Engineering & Expermiments*. New Orleans, Louisiana: Society for Industrial and Applied Mathematics, 2007, 135–149.

Graph Example



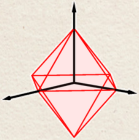
$\mathcal{B}_1$



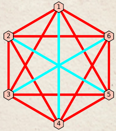
$G(\mathcal{B}_1)$

- For $\mathcal{B}_1 \in \mathbb{R}^3$, $\mathbf{V}_{\mathcal{B}_1} = \begin{bmatrix} \mathbf{I} & -\mathbf{I} \end{bmatrix}$

Graph Example



$\mathcal{B}_1$



$G(\mathcal{B}_1)$

- For $\mathcal{B}_1 \in \mathbb{R}^3$, $\mathbf{V}_{\mathcal{B}_1} = \begin{bmatrix} \mathbf{I} & -\mathbf{I} \end{bmatrix}$
- Coloring matrix $\mathbf{C}$ as

$$\mathbf{C}_{i,j} = \begin{cases} 0.5 & i = j, \\ -0.5 & |i-j| = 3, \\ 0 & \text{otherwise.} \end{cases}$$

Proposed Approach

Identifiability Characterization

- Construct $G_{\mathcal{P}}$
  from a given polytope $\mathcal{P} \in \mathbb{R}^n$

Proposed Approach

Identifiability Characterization

- Construct $G_{\mathcal{P}}$ from a given polytope $\mathcal{P} \in \mathbb{R}^n$
- Compute $Gen(\mathscr{G}(G_{\mathcal{P}}))$ using graph automorphism algorithm.

Proposed Approach

Identifiability Characterization

- Construct $G_{\mathcal{P}}$ from a given polytope $\mathcal{P} \in \mathbb{R}^n$
- Compute $Gen(\mathscr{G}(G_{\mathcal{P}}))$ using graph automorphism algorithm.
- Find $Gen(\mathscr{G}(\mathcal{P}))$ from $Gen(\mathscr{G}(G_{\mathcal{P}}))$.

Proposed Approach

Identifiability Characterization

- Construct $G_{\mathcal{P}}$ from a given polytope $\mathcal{P} \in \mathbb{R}^n$
- Compute $Gen(\mathscr{G}(G_{\mathcal{P}}))$ using graph automorphism algorithm.
- Find $Gen(\mathscr{G}(\mathcal{P}))$ from $Gen(\mathscr{G}(G_{\mathcal{P}}))$.
- Check if each element in $Gen(\mathscr{G}(\mathcal{P}))$ is signed permutation.

Numerical Example

- Random polytopes $\mathcal{P} = \{\mathbf{s} \in \mathbb{R}^n | \mathbf{A}\mathbf{s} \preceq \mathbf{b}\}$ up to dimension 10.

Numerical Example

- Random polytopes $\mathcal{P} = \{\mathbf{s} \in \mathbb{R}^n | \mathbf{A}\mathbf{s} \preceq \mathbf{b}\}$ up to dimension 10.
- Nonnegativity $\mathbf{s}_j \in [0, 1]$, or signed, i.e., $\mathbf{s}_j \in [-1, 1]$ constraints on the components are randomly decided.

Numerical Example

- Random polytopes $\mathcal{P} = \{\mathbf{s} \in \mathbb{R}^n | \mathbf{A}\mathbf{s} \preceq \mathbf{b}\}$ up to dimension 10.
- Nonnegativity $\mathbf{s}_j \in [0, 1]$, or signed, i.e., $\mathbf{s}_j \in [-1, 1]$ constraints on the components are randomly decided.
- Random number of sparsity constraint on sub-vectors with random length is decided $\left\| \begin{bmatrix} \mathbf{s}_{j_1^{(i)}} & \mathbf{s}_{j_2^{(i)}} & \cdots & \mathbf{s}_{j_{l_i}^{(i)}} \end{bmatrix} \right\|_1 \leq 1.$

Numerical Example

- Algorithm 1 utilizes graph automorphism algorithm.