Identifiability in matrix sparse factorization

Léon Zheng

leon.zheng@ens-lyon.fr

M2 Internship, Inria DANTE, LIP (ENS de Lyon)

Supervisor: Rémi Gribonval (Inria DANTE / LIP)

October 9, 2020

INVENTEURS DU MONDE NUMÉRIQUE

Overview

Overview

Motivation: algorithm for matrix sparse factorization

Given a matrix $\boldsymbol{Z}$, we want to find some sparse factors $(\boldsymbol{X}_\ell)_{\ell=1}^{L}$ such that:

$$\boldsymbol{Z} \approx \boldsymbol{X}_1 \boldsymbol{X}_2 ... \boldsymbol{X}_L.$$

Motivation: algorithm for matrix sparse factorization

Given a matrix $\boldsymbol{Z}$, we want to find some sparse factors $(\boldsymbol{X}_\ell)_{\ell=1}^L$ such that:

$$\boldsymbol{Z} \approx \boldsymbol{X}_1 \boldsymbol{X}_2 \ldots \boldsymbol{X}_L.$$

Optimization problem

Let $\boldsymbol{Z}$ be an observed matrix, and $(\mathcal{E}_\ell)_{\ell=1}^L$ some sparsity constraint sets. We want to solve [Le Magoarou et al., 2016]:

$$\underset{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_L}{\text{Minimize}} \quad \underbrace{\|\boldsymbol{Z} - \prod_{\ell=1}^{L} \boldsymbol{X}_\ell\|^2}_{\text{Data-fidelity}} \quad + \quad \underbrace{\sum_{\ell=1}^{L} g_{\mathcal{E}_\ell}(\boldsymbol{X}_\ell)}_{\text{Sparsity-inducing penalty}} \quad . \tag{1}$$

Motivation: algorithm for matrix sparse factorization

Given a matrix $\boldsymbol{Z}$, we want to find some sparse factors $(\boldsymbol{X}_\ell)_{\ell=1}^{L}$ such that:

$$\boldsymbol{Z} \approx \boldsymbol{X}_1 \boldsymbol{X}_2 \ldots \boldsymbol{X}_L.$$

Optimization problem

Let $\boldsymbol{Z}$ be an observed matrix, and $(\mathcal{E}_\ell)_{\ell=1}^{L}$ some sparsity constraint sets. We want to solve [Le Magoarou et al., 2016]:

$$\underset{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_L}{\text{Minimize}} \quad \underbrace{\|\boldsymbol{Z} - \prod_{\ell=1}^{L} \boldsymbol{X}_\ell\|^2}_{\text{Data-fidelity}} + \underbrace{\sum_{\ell=1}^{L} g_{\mathcal{E}_\ell}(\boldsymbol{X}_\ell)}_{\text{Sparsity-inducing penalty}} \quad . \tag{1}$$

Applications:

- Fast transforms
- Sparse neural networks

Motivation: algorithm for matrix sparse factorization

Given a matrix $\boldsymbol{Z}$, we want to find some sparse factors $(\boldsymbol{X}_\ell)_{\ell=1}^L$ such that:

$$\boldsymbol{Z} \approx \boldsymbol{X}_1 \boldsymbol{X}_2 \dots \boldsymbol{X}_L.$$

Optimization problem

Let $\boldsymbol{Z}$ be an observed matrix, and $(\mathcal{E}_\ell)_{\ell=1}^L$ some sparsity constraint sets. We want to solve [Le Magoarou et al., 2016]:

$$\underset{\boldsymbol{X}_1,\dots,\boldsymbol{X}_L}{\text{Minimize}} \quad \underbrace{\|\boldsymbol{Z} - \prod_{\ell=1}^{L} \boldsymbol{X}_\ell\|^2}_{\text{Data-fidelity}} + \underbrace{\sum_{\ell=1}^{L} g_{\mathcal{E}_\ell}(\boldsymbol{X}_\ell)}_{\text{Sparsity-inducing penalty}} \quad . \tag{1}$$

Applications:

- Fast transforms
- Sparse neural networks

Difficulties:

- Nonconvex optimization
- Combinatorial issues

Motivation: algorithm for matrix sparse factorization

In matrix sparse factorization, what are the conditions which guarantee successful recovery of the sparse factors?

Motivation: algorithm for matrix sparse factorization

In matrix sparse factorization, what are the conditions which guarantee successful recovery of the sparse factors?

- This is still an open question.

Motivation: algorithm for matrix sparse factorization

In matrix sparse factorization, what are the conditions which guarantee successful recovery of the sparse factors?

- This is still an open question.
- It leads to the question of identifiability, which is about the uniqueness of the sparse factors in the recovery.

Analogy with linear sparse recovery [Foucart et al., 2017]

Linear sparse recovery problem

Recover a signal $\boldsymbol{x} \in \mathbb{C}^N$ from an observed data $\boldsymbol{y} \in \mathbb{C}^m$, given the linear model:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

Sparsity assumption on the signal $\boldsymbol{x}$: allows reconstruction when $m < N$.

Analogy with linear sparse recovery [Foucart et al., 2017]

Linear sparse recovery problem

Recover a signal $\boldsymbol{x} \in \mathbb{C}^N$ from an observed data $\boldsymbol{y} \in \mathbb{C}^m$, given the linear model:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

Sparsity assumption on the signal $\boldsymbol{x}$: allows reconstruction when $m < N$.

Algorithms for sparse recovery:
→ optimization methods, greedy methods, thresholding-based methods

Analogy with linear sparse recovery [Foucart et al., 2017]

Linear sparse recovery problem

Recover a signal $\boldsymbol{x} \in \mathbb{C}^N$ from an observed data $\boldsymbol{y} \in \mathbb{C}^m$, given the linear model:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

Sparsity assumption on the signal $\boldsymbol{x}$: allows reconstruction when $m < N$.

Algorithms for sparse recovery:
$\rightarrow$ optimization methods, greedy methods, thresholding-based methods

Conditions for the success of these algorithms?

Conditions for which the signal $\boldsymbol{x}$ is identifiable, *i.e.*, it is the unique solution of the sparse recovery problem, when we observe $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$?

Analogy with linear sparse recovery [Foucart et al., 2017]

Linear sparse recovery problem

Recover a signal $\boldsymbol{x} \in \mathbb{C}^N$ from an observed data $\boldsymbol{y} \in \mathbb{C}^m$, given the linear model:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

Sparsity assumption on the signal $\boldsymbol{x}$: allows reconstruction when $m < N$.

Algorithms for sparse recovery:
→ optimization methods, greedy methods, thresholding-based methods

Conditions for the success of these algorithms?

Conditions for which the signal $\boldsymbol{x}$ is identifiable, *i.e.*, it is the unique solution of the sparse recovery problem, when we observe $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$?

→ Identifiability is well studied for linear inverse problems [Foucart et al., 2017], but not for multilinear inverse problems, like matrix sparse factorization.

Problem formulation

We focus on matrix sparse factorization with two factors.

Problem formulation

We focus on matrix sparse factorization with two factors.

Objective: find conditions of identifiability

Let $\boldsymbol{Z} \in \mathbb{C}^{n \times m}$ be a matrix. Consider the bilinear inverse problem:

$$\begin{aligned} \text{find} \quad & (\boldsymbol{X}, \boldsymbol{Y}) \\ \text{such that} \quad & \boldsymbol{X}\boldsymbol{Y} = \boldsymbol{Z}, \\ & \boldsymbol{X}, \boldsymbol{Y} \text{ are } \underline{\text{sparse}}. \end{aligned} \tag{2}$$

Under which conditions the solution is unique, up to equivalence relations?

Problem formulation

We focus on matrix sparse factorization with two factors.

Objective: find conditions of identifiability

Let $\boldsymbol{Z} \in \mathbb{C}^{n \times m}$ be a matrix. Consider the bilinear inverse problem:

$$\begin{aligned} \text{find} \quad & (\boldsymbol{X}, \boldsymbol{Y}) \\ \text{such that} \quad & \boldsymbol{X}\boldsymbol{Y} = \boldsymbol{Z}, \\ & \boldsymbol{X}, \boldsymbol{Y} \text{ are } \underline{\text{sparse}}. \end{aligned} \tag{2}$$

Under which conditions the solution is unique, up to equivalence relations?

- Sparsity:

- Equivalence relations:

Problem formulation

We focus on matrix sparse factorization with two factors.

Objective: find conditions of identifiability

Let $\boldsymbol{Z} \in \mathbb{C}^{n \times m}$ be a matrix. Consider the bilinear inverse problem:

$$\begin{aligned} \text{find} \quad & (\boldsymbol{X}, \boldsymbol{Y}) \\ \text{such that} \quad & \boldsymbol{X}\boldsymbol{Y} = \boldsymbol{Z}, \\ & \boldsymbol{X}, \boldsymbol{Y} \text{ are } \underline{\text{sparse}}. \end{aligned} \tag{2}$$

Under which conditions the solution is unique, up to equivalence relations?

- Sparsity: a matrix is sparse if its support is *allowed*. We choose what are the allowed supports.
- Equivalence relations:

Problem formulation

We focus on matrix sparse factorization with two factors.

Objective: find conditions of identifiability

Let $\boldsymbol{Z} \in \mathbb{C}^{n \times m}$ be a matrix. Consider the bilinear inverse problem:

$$\begin{aligned} \text{find} \quad & (\boldsymbol{X}, \boldsymbol{Y}) \\ \text{such that} \quad & \boldsymbol{X}\boldsymbol{Y} = \boldsymbol{Z}, \\ & \boldsymbol{X}, \boldsymbol{Y} \text{ are } \underline{\text{sparse}}. \end{aligned} \tag{2}$$

Under which conditions the solution is unique, up to equivalence relations?

- Sparsity: a matrix is sparse if its support is *allowed*. We choose what are the allowed supports.
- Equivalence relations: scaling + permutations, because

$$\boldsymbol{X}\boldsymbol{Y} = (\boldsymbol{X}\boldsymbol{D})(\boldsymbol{D}^{-1}\boldsymbol{Y}) = (\boldsymbol{X}\boldsymbol{P})(\boldsymbol{P}^T\boldsymbol{Y})$$

  where $\boldsymbol{D}$ is a diagonal matrix, and $\boldsymbol{P}$ is a permutation matrix.

Contributions

1. Characterization of fixed-support identifiability
2. Characterization of right identifiability

We observe $\boldsymbol{Z} := \boldsymbol{X}\boldsymbol{Y}$.

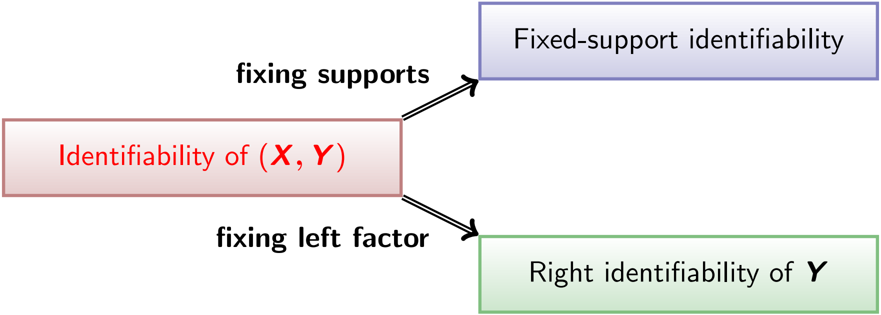


Figure: Deriving necessary conditions of identifiability by considering two problem variations

Contributions

1. Characterization of fixed-support identifiability
2. Characterization of right identifiability

We observe $\boldsymbol{Z} := \boldsymbol{X}\boldsymbol{Y}$.

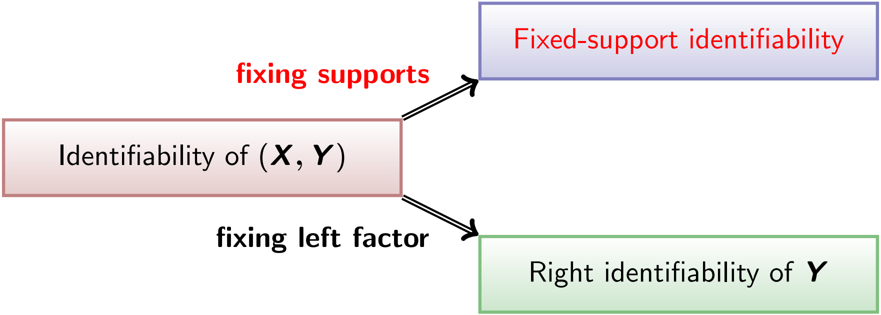


Figure: Deriving necessary conditions of identifiability by considering two problem variations

Contributions

1. Characterization of fixed-support identifiability
2. Characterization of right identifiability

We observe $\boldsymbol{Z} := \boldsymbol{X}\boldsymbol{Y}$.

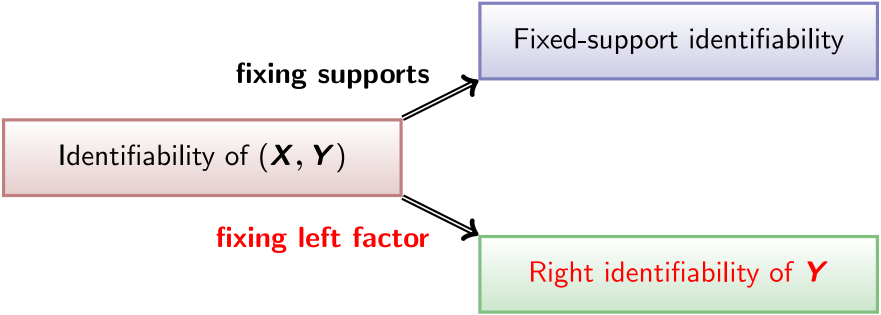


Figure: Deriving necessary conditions of identifiability by considering two problem variations

Overview

Fixed-support identifiability definition

Consider $(\boldsymbol{S_X}, \boldsymbol{S_Y})$ a fixed pair of supports.

Fixed-support identifiability definition

Consider $(\boldsymbol{S_X}, \boldsymbol{S_Y})$ a fixed pair of supports.

Example:

$$(\boldsymbol{S_X}, \boldsymbol{S_Y}) := \left(\begin{pmatrix} \boxed{\star \;\; \star} \\ 0 \;\; 0 \end{pmatrix}, \begin{pmatrix} 0 & \boxed{\star} \\ 0 & \boxed{\star} \end{pmatrix} \right)$$

$$(\boldsymbol{X_1}, \boldsymbol{Y_1}) := \left(\begin{pmatrix} \boxed{\mathbf{1} \;\; \mathbf{2}} \\ 0 \;\; 0 \end{pmatrix}, \begin{pmatrix} 0 & \boxed{\mathbf{3}} \\ 0 & \boxed{\mathbf{4}} \end{pmatrix} \right)$$

(a) Allowed supports

$$(\boldsymbol{X_2}, \boldsymbol{Y_2}) := \left(\begin{pmatrix} \boxed{\mathbf{2} \;\; 0} \\ 0 \;\; 0 \end{pmatrix}, \begin{pmatrix} 0 & \boxed{\mathbf{2}} \\ \mathbf{1} & \boxed{\mathbf{1}} \end{pmatrix} \right)$$

(b) Not allowed supports

Fixed-support identifiability definition

Consider $(\boldsymbol{S_X}, \boldsymbol{S_Y})$ a fixed pair of supports.

<u>Example</u>:

$$(\boldsymbol{S_X}, \boldsymbol{S_Y}) := \left(\begin{array}{cc} \boxed{\star \quad \star} \\ 0 \quad 0 \end{array} \right), \left(\begin{array}{c|c} 0 & \star \\ 0 & \star \end{array} \right)$$

$$(\boldsymbol{X_1}, \boldsymbol{Y_1}) := \left(\begin{array}{cc} \boxed{\mathbf{1} \quad \mathbf{2}} \\ 0 \quad 0 \end{array} \right), \left(\begin{array}{c|c} 0 & \mathbf{3} \\ 0 & \mathbf{4} \end{array} \right)$$

(a) Allowed supports

$$(\boldsymbol{X_2}, \boldsymbol{Y_2}) := \left(\begin{array}{cc} \boxed{\mathbf{2} \quad 0} \\ 0 \quad 0 \end{array} \right), \left(\begin{array}{c|c} 0 & \mathbf{2} \\ \mathbf{1} & \mathbf{1} \end{array} \right)$$

(b) Not allowed supports

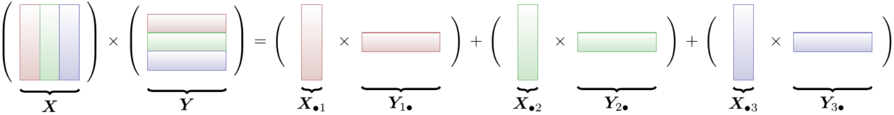
Definition: identifiability of $(\boldsymbol{S_X}, \boldsymbol{S_Y})$

Every pair $(\boldsymbol{X}, \boldsymbol{Y})$ with a support equal to $(\boldsymbol{S_X}, \boldsymbol{S_Y})$ is the unique solution (up to equivalence) for the factorization of $\boldsymbol{Z} := \boldsymbol{XY}$ into two factors supported by $(\boldsymbol{S_X}, \boldsymbol{S_Y})$.

$\rightarrow$ We will give here a characterization of this property.

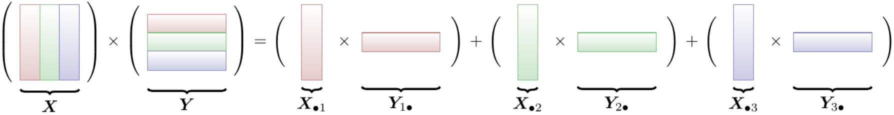
Rank 1 contributions representation

Let $(\boldsymbol{X}, \boldsymbol{Y})$ be a pair of factor.



Rank 1 contributions representation

Let $(\boldsymbol{X}, \boldsymbol{Y})$ be a pair of factor.

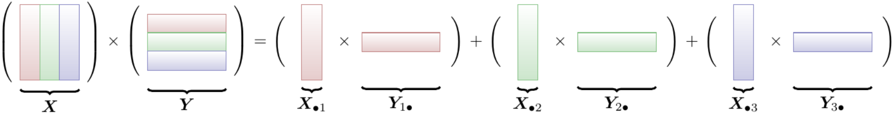


Definition

$(\boldsymbol{X}, \boldsymbol{Y})$ is represented by $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$, where r is the number of columns in $\boldsymbol{X}$ (or rows in $\boldsymbol{Y}$).

Rank 1 contributions representation

Let $(\boldsymbol{X}, \boldsymbol{Y})$ be a pair of factor.



Definition

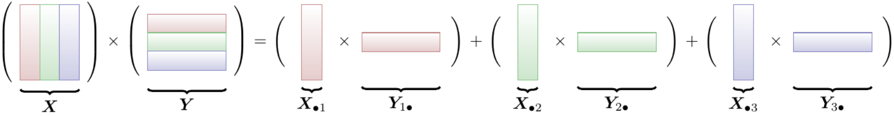
$(\boldsymbol{X}, \boldsymbol{Y})$ is represented by $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$, where r is the number of columns in $\boldsymbol{X}$ (or rows in $\boldsymbol{Y}$).

Lemma

Identifiability of $(\boldsymbol{X}, \boldsymbol{Y}) \iff$ *Identifiability of* $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$

Rank 1 contributions representation

Let $(\boldsymbol{X}, \boldsymbol{Y})$ be a pair of factor.



Definition

$(\boldsymbol{X}, \boldsymbol{Y})$ is represented by $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$, where r is the number of columns in $\boldsymbol{X}$ (or rows in $\boldsymbol{Y}$).

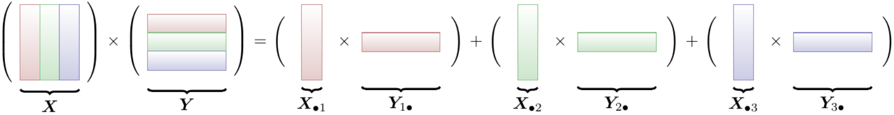
Lemma

Identifiability of $(\boldsymbol{X}, \boldsymbol{Y}) \iff$ *Identifiability of* $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$

$\rightarrow$ [Le Magoarou, 2016] used this representation to show that the butterfly factorization of the Discrete Fourier Transform matrix is identifiable.

Rank 1 contributions representation

Let $(\boldsymbol{X}, \boldsymbol{Y})$ be a pair of factor.



Definition

$(\boldsymbol{X}, \boldsymbol{Y})$ is represented by $(\boldsymbol{X}_{\bullet i} \boldsymbol{Y}_{i \bullet})_{i=1}^{r}$, where r is the number of columns in $\boldsymbol{X}$ (or rows in $\boldsymbol{Y}$).

Lemma

Identifiability of $(\boldsymbol{X}, \boldsymbol{Y}) \iff$ *Identifiability of* $(\boldsymbol{X}_{\bullet i} \boldsymbol{Y}_{i \bullet})_{i=1}^{r}$

$\rightarrow$ We are implicitly using lifting ideas, inspired by [Choudhary et al., 2014], [Malgouyres et al., 2016]. The lifting operator is $\mathscr{S} : (\boldsymbol{C_i})_{i=1}^{r} \mapsto \sum_{i=1}^{r} \boldsymbol{C_i}$.

Identifiability of the rank 1 contributions?

We now observe $\boldsymbol{Z} := \boldsymbol{X}\boldsymbol{Y}$.

Identifiability of $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$ from the observation $\boldsymbol{Z}$?

Identifiability of the rank 1 contributions?

We now observe $\boldsymbol{Z} := \boldsymbol{X}\boldsymbol{Y}$.

Identifiability of $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$ from the observation $\boldsymbol{Z}$?
$\rightarrow$ We have $\sum_{i=1}^{r}\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet} = \boldsymbol{Z}$.

$$\begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & \mathbf{2} & 0 & \mathbf{4} \\ 0 & \mathbf{3} & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{2} & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ \mathbf{1} & \mathbf{4} & \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{5} & \mathbf{5} & \mathbf{6} \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix}$$

Identifiability of the rank 1 contributions?

We now observe $\boldsymbol{Z} := \boldsymbol{X}\boldsymbol{Y}$.

Identifiability of $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$ from the observation $\boldsymbol{Z}$?
$\rightarrow$ We have $\sum_{i=1}^{r} \boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet} = \boldsymbol{Z}$.

$$\begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & \mathbf{2} & 0 & \mathbf{4} \\ 0 & \mathbf{3} & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{2} & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ \mathbf{1} & \mathbf{4} & \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{5} & \mathbf{5} & \mathbf{6} \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix}$$

Idea

Complete each rank 1 contribution from the entries not covered by the other rank 1 contributions.

Example

We know: the observed matrix $\boldsymbol{Z}$, and the supports of the rank 1 contributions $((\boldsymbol{S_X})_{\bullet i}(\boldsymbol{S_Y})_{i\bullet})_{i=1}^{r}$.
We want: to reconstruct the rank 1 contributions $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$.

$$\boldsymbol{Z} = \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ \mathbf{1} & \mathbf{4} & \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{5} & \mathbf{5} & \mathbf{6} \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & ? & 0 & ? \\ 0 & ? & 0 & ? \\ 0 & ? & 0 & ? \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & ? & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ ? & ? & ? & 0 \\ ? & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure: How to reconstruct the rank 1 contributions?

Example

We know: the observed matrix $\boldsymbol{Z}$, and the supports of the rank 1 contributions $((\boldsymbol{S_X})_{\bullet i}(\boldsymbol{S_Y})_{i\bullet})_{i=1}^r$.
We want: to reconstruct the rank 1 contributions $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^r$.

$$\begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ \mathbf{1} & \mathbf{4} & \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{5} & \mathbf{5} & \mathbf{6} \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & ? & 0 & \mathbf{4} \\ 0 & ? & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & ? & \mathbf{3} & 0 \\ \mathbf{1} & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure: We show in color the "observable" entries. The red contribution is completable from its observable entries.

Example

We know: the observed matrix $\boldsymbol{Z}$, and the supports of the rank 1 contributions $((\boldsymbol{S_X})_{\bullet i}(\boldsymbol{S_Y})_{i\bullet})_{i=1}^{r}$.
We want: to reconstruct the rank 1 contributions $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$.

$$\begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ \mathbf{1} & \mathbf{4} & \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{5} & \mathbf{5} & \mathbf{6} \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & ? & 0 & \mathbf{4} \\ 0 & ? & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & ? & \mathbf{3} & 0 \\ \mathbf{1} & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & \mathbf{2} & 0 & \mathbf{4} \\ 0 & \mathbf{3} & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & ? & \mathbf{3} & 0 \\ \mathbf{1} & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure: This “uncovers” entries in the green contribution.

Example

<u>We know</u>: the observed matrix $\boldsymbol{Z}$, and the supports of the rank 1 contributions $((\boldsymbol{S_X})_{\bullet i}(\boldsymbol{S_Y})_{i\bullet})_{i=1}^{r}$.
<u>We want</u>: to reconstruct the rank 1 contributions $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$.

$$\begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ \mathbf{1} & \mathbf{4} & \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{5} & \mathbf{5} & \mathbf{6} \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & ? & 0 & \mathbf{4} \\ 0 & ? & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & ? & \mathbf{3} & 0 \\ \mathbf{1} & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & \mathbf{2} & 0 & \mathbf{4} \\ 0 & \mathbf{3} & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & 0 \\ \mathbf{1} & \mathbf{2} & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure: Now it is possible to complete the green contribution.

Example

We know: the observed matrix $\boldsymbol{Z}$, and the supports of the rank 1 contributions $((\boldsymbol{S_X})_{\bullet i}(\boldsymbol{S_Y})_{i\bullet})_{i=1}^r$.
We want: to reconstruct the rank 1 contributions $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^r$.

$$\begin{pmatrix} 0 & 1 & 0 & 2 \\ 1 & 4 & 3 & 4 \\ 1 & 5 & 5 & 6 \\ 0 & 0 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 2 \\ 0 & ? & 0 & 4 \\ 0 & ? & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & ? & 3 & 0 \\ 1 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 & 2 \\ 0 & 2 & 0 & 4 \\ 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 0 \\ 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure: This “uncovers” entries in the blue contribution.

Example

We know: the observed matrix $\boldsymbol{Z}$, and the supports of the rank 1 contributions $((\boldsymbol{S_X})_{\bullet i}(\boldsymbol{S_Y})_{i\bullet})_{i=1}^{r}$.
We want: to reconstruct the rank 1 contributions $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$.

$$\begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ \mathbf{1} & \mathbf{4} & \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{5} & \mathbf{5} & \mathbf{6} \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & ? & 0 & \mathbf{4} \\ 0 & ? & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & ? & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & ? & \mathbf{3} & 0 \\ \mathbf{1} & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{2} \\ 0 & \mathbf{2} & 0 & \mathbf{4} \\ 0 & \mathbf{3} & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{2} & 0 \\ 0 & 0 & \mathbf{3} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure: Therefore, $(\boldsymbol{X}_{\bullet i}\boldsymbol{Y}_{i\bullet})_{i=1}^{r}$ are identifiable from the observation $\boldsymbol{Z}$.

Iterative completability from observable supports

Let $\boldsymbol{S}$ be a rank 1 support (= support of a rank 1 matrix).

Definition: $\boldsymbol{S}$ is completable from $\boldsymbol{S}' \subseteq \boldsymbol{S}$

We can complete any rank 1 matrix $\boldsymbol{M}$ with a support equal to $\boldsymbol{S}$, by observing only its entries on $\boldsymbol{S}'$.

Iterative completability from observable supports

Let $\boldsymbol{S}$ be a rank 1 support (= support of a rank 1 matrix).

Definition: $\boldsymbol{S}$ is completable from $\boldsymbol{S'} \subseteq \boldsymbol{S}$

We can complete any rank 1 matrix $\boldsymbol{M}$ with a support equal to $\boldsymbol{S}$, by observing only its entries on $\boldsymbol{S'}$.

Let $\boldsymbol{S_1}, \ldots, \boldsymbol{S_r}$ be r rank 1 supports.

Definition: iterative completability of $(\boldsymbol{S_i})_{i=1}^r$

- The rank 1 supports $\boldsymbol{S_i}$ for $i \in [\![1; r]\!]$ can be completed one by one from its observable support:

$$\boldsymbol{S_i} \setminus \bigcup_{i' \in [\![r]\!] \setminus \{i\}} \boldsymbol{S_{i'}}.$$

- When the i-th rank 1 support is completable from its observable support, we repeat with $(\boldsymbol{S_{i'}})_{i \neq i'}$.

Iterative completability from observable supports

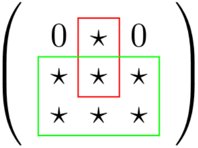


Figure: This example is iteratively completable.

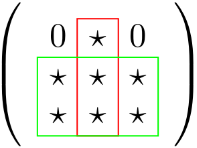


Figure: This example is not iteratively completable.

Fixed-support identifiability characterization

Theorem

For $r = 2$, $(\boldsymbol{S_X}, \boldsymbol{S_Y})$ is identifiable if, and only if, the supports of its rank 1 contributions are iteratively completable.

Remark: Sufficiency is true for all r.

Fixed-support identifiability characterization

Theorem

For $r = 2$, $(\boldsymbol{S_X}, \boldsymbol{S_Y})$ is identifiable if, and only if, the supports of its rank 1 contributions are iteratively completable.

Remark: Sufficiency is true for all r. Necessity is false for $r \geq 3$.

$$\begin{pmatrix} 0 & \star & \star & 0 \\ \star & \star & \star & 0 \\ \star & \star & \star & \star \\ \star & \star & \star & \star \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \star & ? & 0 & 0 \\ \star & ? & 0 & 0 \\ \star & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \star & \star & 0 \\ 0 & ? & \star & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & ? & \star \\ 0 & ? & \star & \star \end{pmatrix}$$

Figure: Counterexample showing that iterative completability is not a necessary condition for fixed-support identifiability.

$\rightarrow$ This leads to the notion of iterative partial completability (future work).

Overview

Some right identifiability results

Consider $\boldsymbol{X}$ a fixed left factor, and Ω_R a family of allowed right supports.

Theorem

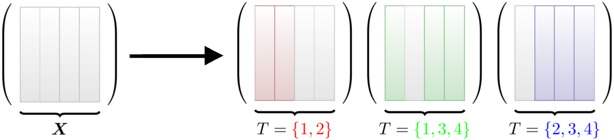
Suppose that $\boldsymbol{X}$ non-degenerate, and Ω_R is stable by inclusion. Then the following assertions are equivalent:

1. Ω_R *is right identifiable for* $\boldsymbol{X}$;
2. *the columns of* $\boldsymbol{X}$ *indexed by* T *are linearly independent, for all*

$$T \in \mathcal{T}(\Omega_R)$$

where $\mathcal{T}(\Omega_R)$ *is a collection of indices subsets determined by* Ω_R.

Example: for a specific Ω_R, we can have $\mathcal{T} = \{\{1,2\},\{1,3,4\},\{2,3,4\}\}$.



Some right identifiability results

Consider $\boldsymbol{X}$ a fixed left factor, and Ω_R a family of allowed right supports.

Theorem

Suppose that $\boldsymbol{X}$ non-degenerate, and Ω_R is stable by inclusion. Then the following assertions are equivalent:

1. *Ω_R is right identifiable for $\boldsymbol{X}$;*
2. *the columns of $\boldsymbol{X}$ indexed by T are linearly independent, for all*

$$T \in \mathcal{T}(\Omega_R)$$

where $\mathcal{T}(\Omega_R)$ is a collection of indices subsets determined by Ω_R.

Example (Family of right supports l-sparse by row)

Condition: all the columns of $\boldsymbol{X}$ are linearly independent.

Some right identifiability results

Consider $\boldsymbol{X}$ a fixed left factor, and Ω_R a family of allowed right supports.

Theorem

Suppose that $\boldsymbol{X}$ non-degenerate, and Ω_R is stable by inclusion. Then the following assertions are equivalent:

1. *Ω_R is right identifiable for $\boldsymbol{X}$;*
2. *the columns of $\boldsymbol{X}$ indexed by T are linearly independent, for all*

$$T \in \mathcal{T}(\Omega_R)$$

where $\mathcal{T}(\Omega_R)$ is a collection of indices subsets determined by Ω_R.

Example (Family of right supports k-sparse by column)

Condition: every subset of $2k$ columns of $\boldsymbol{X}$ is linearly independent.

$\rightarrow$ Similar result in compressive sensing literature [Foucart et al., 2017].

Overview

Conclusion

Summary

1. Fixed-support identifiability: with **rank 1 matrix completion** conditions.
2. Right identifiability: with **linear independence** of specific subsets of columns in the left factor.

Conclusion

Summary

1. Fixed-support identifiability: with **rank 1 matrix completion** conditions.
2. Right identifiability: with **linear independence** of specific subsets of columns in the left factor.

Open questions

- Fixed-support identifiability: characterization with **iterative partial completability**?
- Finding sufficient conditions of **generic identifiability**? Necessary and sufficient conditions?

References

Sunav Choudhary and Urbashi Mitra (2014)
Identifiability scaling laws in bilinear inverse problems
arXiv preprint arXiv:1402.2637

Luc Le Magoarou (2016)
Efficient matrices for signal processing and machine learning
Theses, INSA de Rennes.

Luc Le Magoarou, Rémi Gribonval (2016)
Flexible multilayer sparse approximations of matrices and applications.
IEEE Journal of Selected Topics in Signal Processing 10.4 (2016): 688-700.

François Malgouyres and Joseph Landsberg (2016)
On the identifiability and stable recovery of deep/multi-layer structured matrix factorization
2016 IEEE Information Theory Workshop (ITW). IEEE, 2016.

Simon Foucart, and Holger Rauhut (2017)
A mathematical introduction to compressive sensing.
Bull. Am. Math 54 (2017): 151-165.

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]

Given a bilinear mapping $\boldsymbol{S} : (\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y})$, derive $\mathscr{S} : \boldsymbol{W} \mapsto \mathscr{S}(\boldsymbol{W})$, with the identity: $\mathscr{S}(\boldsymbol{x}\boldsymbol{y}^T) = \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y})$. Then:

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]

Given a bilinear mapping $\boldsymbol{S} : (\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y})$, derive $\mathscr{S} : \boldsymbol{W} \mapsto \mathscr{S}(\boldsymbol{W})$, with the identity: $\mathscr{S}(\boldsymbol{x}\boldsymbol{y}^T) = \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y})$. Then:

$$\begin{aligned} &\text{find} && (\boldsymbol{x}, \boldsymbol{y}) \\ &\text{such that} && \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{z}, \\ & && (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{K}. \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} &\text{minimize} && \text{rank}(\boldsymbol{W}) \\ &\text{such that} && \mathscr{S}(\boldsymbol{W}) = \boldsymbol{z}, \\ & && \boldsymbol{W} \in \mathcal{K}'. \end{aligned}$$

where $\mathcal{K}' \cap \{\text{matrix } \boldsymbol{W} \text{ with rank at most } 1\} = \{\boldsymbol{x}\boldsymbol{y}^T \mid (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{K}\}$.

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]

Given a bilinear mapping $\boldsymbol{S} : (\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y})$, derive $\mathscr{S} : \boldsymbol{W} \mapsto \mathscr{S}(\boldsymbol{W})$, with the identity: $\mathscr{S}(\boldsymbol{x}\boldsymbol{y}^T) = \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y})$. Then:

$$\begin{aligned} &\text{find} && (\boldsymbol{x}, \boldsymbol{y}) \\ &\text{such that} && \boldsymbol{S}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{z}, \\ &&& (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{K}. \end{aligned} \iff \begin{aligned} &\text{minimize} && \text{rank}(\boldsymbol{W}) \\ &\text{such that} && \mathscr{S}(\boldsymbol{W}) = \boldsymbol{z}, \\ &&& \boldsymbol{W} \in \mathcal{K}'. \end{aligned}$$

where $\mathcal{K}' \cap \{\text{matrix } \boldsymbol{W} \text{ with rank at most } 1\} = \{\boldsymbol{x}\boldsymbol{y}^T \mid (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{K}\}$.

Proposition (Identifiability characterization [Choudhary et al., 2014])

$\text{Ker}\, \mathscr{S} \cap \{\textit{matrix } \boldsymbol{W} \textit{ with rank at most } 2\} \cap (\mathcal{K}' - \mathcal{K}') = \{0\}$.

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]
2. **Tensorial lifting** for multilayer matrix sparse factorization [Malgouyres et al., 2016]

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]
2. **Tensorial lifting** for multilayer matrix sparse factorization [Malgouyres et al., 2016]
3. Identifiability of butterfly factorization in Discrete Fourier Transform matrix, with **matrix completability conditions** [Le Magoarou, 2016]

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]
2. **Tensorial lifting** for multilayer matrix sparse factorization [Malgouyres et al., 2016]
3. Identifiability of butterfly factorization in Discrete Fourier Transform matrix, with **matrix completability conditions** [Le Magoarou, 2016]

Notation: $\omega = exp(i\frac{2\pi}{N})$. Here, for instance, $N = 4$.

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega^1 & \omega^2 & \omega^3 \\ 1 & \omega^2 & 1 & \omega^2 \\ 1 & \omega^3 & \omega^2 & \omega^1 \end{pmatrix}}_{\text{DFT matrix } N \times N} = \underbrace{\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \omega^1 \\ 1 & 0 & \omega^2 & 0 \\ 0 & 1 & 0 & \omega^3 \end{pmatrix}}_{\frac{N}{2}\text{-sparse by column}} \underbrace{\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & \omega^2 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & \omega^2 \end{pmatrix}}_{\text{2-sparse by row}}$$

Left support: $\frac{N}{2}$-sparse by column. Right support: 2-sparse by row.

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]
2. **Tensorial lifting** for multilayer matrix sparse factorization [Malgouyres et al., 2016]
3. Identifiability of butterfly factorization in Discrete Fourier Transform matrix, with **matrix completability conditions** [Le Magoarou, 2016]

Rank 1 matrix completability:

$$\boldsymbol{M} = \begin{pmatrix} 0 & \star & ? & ? \\ 0 & \star & ? & ? \\ 0 & \star & \star & \star \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure: Can we complete missing entries (?) from observable entries ($\star$)? The rank of $\boldsymbol{M}$ is at most 1.

Extra: existing identifiability results

1. **Lifting** for identifiability in generic bilinear inverse problems [Choudhary et al., 2014]
2. **Tensorial lifting** for multilayer matrix sparse factorization [Malgouyres et al., 2016]
3. Identifiability of butterfly factorization in Discrete Fourier Transform matrix, with **matrix completability conditions** [Le Magoarou, 2016]

Main issue

No general conditions easy to verify for identifiability in matrix sparse factorization.

Extra: equivalence relation? sparsity?

Equivalent pairs of factors

$(\boldsymbol{X}, \boldsymbol{Y}) \sim (\boldsymbol{A}, \boldsymbol{B})$ if $\boldsymbol{X}\boldsymbol{P}\boldsymbol{D} = \boldsymbol{A}$ and $\boldsymbol{D}^{-1}\boldsymbol{P}^T\boldsymbol{Y} = \boldsymbol{B}$, with:

- $\boldsymbol{D}$ a scaling matrix (diagonal, nonzero diagonal entries);
- $\boldsymbol{P}$ a permutation matrix.

Extra: equivalence relation? sparsity?

Equivalent pairs of factors

$(\boldsymbol{X}, \boldsymbol{Y}) \sim (\boldsymbol{A}, \boldsymbol{B})$ if $\boldsymbol{XPD} = \boldsymbol{A}$ and $\boldsymbol{D}^{-1}\boldsymbol{P}^T\boldsymbol{Y} = \boldsymbol{B}$, with:

- $\boldsymbol{D}$ a scaling matrix (diagonal, nonzero diagonal entries);
- $\boldsymbol{P}$ a permutation matrix.

Family of allowed supports

Let Ω be a subset of supports. $\boldsymbol{M} \in \mathbb{C}^{p \times q}$ is sparse $\iff$ $\mathrm{supp}(\boldsymbol{M}) \in \Omega$.

Extra: equivalence relation? sparsity?

Equivalent pairs of factors

$(\boldsymbol{X}, \boldsymbol{Y}) \sim (\boldsymbol{A}, \boldsymbol{B})$ if $\boldsymbol{XPD} = \boldsymbol{A}$ and $\boldsymbol{D}^{-1}\boldsymbol{P}^T\boldsymbol{Y} = \boldsymbol{B}$, with:

- $\boldsymbol{D}$ a scaling matrix (diagonal, nonzero diagonal entries);
- $\boldsymbol{P}$ a permutation matrix.

Family of allowed supports

Let Ω be a subset of supports. $\boldsymbol{M} \in \mathbb{C}^{p \times q}$ is sparse $\iff$ $\text{supp}(\boldsymbol{M}) \in \Omega$.

Support of a matrix $\boldsymbol{M} \in \mathbb{C}^{p \times q}$ as a binary matrix

Denote $\text{supp}(\boldsymbol{M}) \in \{0, 1\}^{p \times q}$ where $\text{supp}(\boldsymbol{M})_{ij} = 1 \iff \boldsymbol{M}_{ij} \neq 0$.

Extra: equivalence relation? sparsity?

Equivalent pairs of factors

$(\boldsymbol{X}, \boldsymbol{Y}) \sim (\boldsymbol{A}, \boldsymbol{B})$ if $\boldsymbol{XPD} = \boldsymbol{A}$ and $\boldsymbol{D}^{-1}\boldsymbol{P}^T\boldsymbol{Y} = \boldsymbol{B}$, with:

- $\boldsymbol{D}$ a scaling matrix (diagonal, nonzero diagonal entries);
- $\boldsymbol{P}$ a permutation matrix.

Family of allowed supports

Let Ω be a subset of supports. $\boldsymbol{M} \in \mathbb{C}^{p \times q}$ is sparse $\iff$ $\text{supp}(\boldsymbol{M}) \in \Omega$.

Support of a matrix $\boldsymbol{M} \in \mathbb{C}^{p \times q}$ as a binary matrix

Denote $\text{supp}(\boldsymbol{M}) \in \{0, 1\}^{p \times q}$ where $\text{supp}(\boldsymbol{M})_{ij} = 1 \iff \boldsymbol{M}_{ij} \neq 0$.

Family of allowed pairs of supports

Let $\hat{\Omega}$ be a subset of pairs of supports. $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathbb{C}^{n \times r} \times \mathbb{C}^{r \times m}$ is sparse $\iff$ $(\text{supp}(\boldsymbol{X}), \text{supp}(\boldsymbol{Y})) \in \hat{\Omega}$.

Extra: definition of identifiability

Consider $\hat{\Omega}$ a family of allowed pairs of supports.

Definition: identifiability of $\hat{\Omega}$

For all $(\boldsymbol{X}, \boldsymbol{Y}), (\boldsymbol{A}, \boldsymbol{B})$ with allowed support in Ω, we have:

$$\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{B} \Rightarrow (\boldsymbol{X}, \boldsymbol{Y}) \sim (\boldsymbol{A}, \boldsymbol{B}).$$

Problem formulation: under which condition $\hat{\Omega}$ is identifiable?

Extra: right identifiability is a necessary condition

Given $\hat{\Omega}$ a family of allowed pairs of supports, and $\boldsymbol{X}$ a left factor, denote:

$$\Omega_R(\boldsymbol{X}) := \{\boldsymbol{S_Y} \mid (\text{supp}(\boldsymbol{X}), \boldsymbol{S_Y}) \in \hat{\Omega}\}.$$

Extra: right identifiability is a necessary condition

Given $\hat{\Omega}$ a family of allowed pairs of supports, and $\boldsymbol{X}$ a left factor, denote:

$$\Omega_R(\boldsymbol{X}) := \{\boldsymbol{S_Y} \mid (\text{supp}(\boldsymbol{X}), \boldsymbol{S_Y}) \in \hat{\Omega}\}.$$

Lemma

If $\hat{\Omega}$ is identifiable, then for all left factors $\boldsymbol{X}$, $\Omega_R(\boldsymbol{X})$ is right identifiable for $\boldsymbol{X}$.

Extra: right identifiability is a necessary condition

Given $\hat{\Omega}$ a family of allowed pairs of supports, and $\boldsymbol{X}$ a left factor, denote:

$$\Omega_R(\boldsymbol{X}) := \{\boldsymbol{S_Y} \mid (\text{supp}(\boldsymbol{X}), \boldsymbol{S_Y}) \in \hat{\Omega}\}.$$

Lemma

If $\hat{\Omega}$ is identifiable, then for all left factors $\boldsymbol{X}$, $\Omega_R(\boldsymbol{X})$ is right identifiable for $\boldsymbol{X}$.

Definition: right identifiability of $\Omega_R(\boldsymbol{X})$ for $\boldsymbol{X}$

For all $\boldsymbol{Y}, \boldsymbol{B}$ with allowed support in $\Omega_R(\boldsymbol{X})$, we have:

$$\boldsymbol{XY} = \boldsymbol{XB} \Rightarrow (\boldsymbol{X}, \boldsymbol{Y}) \sim (\boldsymbol{X}, \boldsymbol{B}).$$

Extra: lifting principle

Lifting operator:

$$\mathscr{S} : (\boldsymbol{X_i})_{i=1}^{r} \mapsto \sum_{i=1}^{r} \boldsymbol{X_i}$$

Extra: lifting principle

Lifting operator:

$$\mathscr{S} : (\boldsymbol{X_i})_{i=1}^{r} \mapsto \sum_{i=1}^{r} \boldsymbol{X_i}$$

Proposition

$(\boldsymbol{S_X}, \boldsymbol{S_Y})$ *is identifiable if, and only if,*

Extra: lifting principle

Lifting operator:

$$\mathscr{S} : (\boldsymbol{X_i})_{i=1}^r \mapsto \sum_{i=1}^r \boldsymbol{X_i}$$

Proposition

$(\boldsymbol{S_X}, \boldsymbol{S_Y})$ is identifiable if, and only if,

$$\mathrm{Ker}(\mathscr{S}) \cap \prod_{i=1}^r (\Sigma_{\boldsymbol{S_i},1} - \Sigma_{\boldsymbol{S_i},1}) = \{0\}, \tag{3}$$

where $\boldsymbol{S_i} := (\boldsymbol{S_X})_{\bullet i}(\boldsymbol{S_Y})_{i\bullet}$ is the i-th rank 1 support of $(\boldsymbol{S_X}, \boldsymbol{S_Y})$, and:

$$\Sigma_{\boldsymbol{S_i},1} := \{\textit{matrix with rank at most 1, with a support equal to } \boldsymbol{S_i}\}.$$

Extra: iterative partial completability

Counterexample

Iterative completability is not a necessary condition for fixed-support identifiability when $r \geq 3$.

$$\begin{pmatrix} 0 & \star & \star & 0 \\ \star & \star & \star & 0 \\ \star & \star & \star & \star \\ \star & \star & \star & \star \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \star & ? & 0 & 0 \\ \star & ? & 0 & 0 \\ \star & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \star & \star & 0 \\ 0 & ? & \star & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & ? & \star \\ 0 & ? & \star & \star \end{pmatrix}$$

Figure: This example is not iteratively completable from observable supports.

Counterexample

Iterative completability is not a necessary condition for fixed-support identifiability when $r \geq 3$.

$$\begin{pmatrix} 0 & \mathbf{1} & \mathbf{2} & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{2} & 0 \\ \mathbf{2} & \mathbf{6} & \mathbf{5} & \mathbf{6} \\ \mathbf{3} & \mathbf{5} & \mathbf{2} & \mathbf{4} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & ? & 0 & 0 \\ \mathbf{2} & ? & 0 & 0 \\ \mathbf{3} & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \mathbf{1} & \mathbf{2} & 0 \\ 0 & ? & \mathbf{2} & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & ? & \mathbf{6} \\ 0 & ? & \mathbf{2} & \mathbf{4} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & ? & 0 & 0 \\ \mathbf{2} & ? & 0 & 0 \\ \mathbf{3} & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \mathbf{1} & \mathbf{2} & 0 \\ 0 & \mathbf{1} & \mathbf{2} & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & \mathbf{3} & \mathbf{6} \\ 0 & ? & \mathbf{2} & \mathbf{4} \end{pmatrix}$$

Figure: However, we can complete partially green and blue contributions.

Extra: iterative partial completability

Counterexample

Iterative completability is not a necessary condition for fixed-support identifiability when $r \geq 3$.

$$\begin{pmatrix} 0 & \mathbf{1} & \mathbf{2} & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{2} & 0 \\ \mathbf{2} & \mathbf{6} & \mathbf{5} & \mathbf{6} \\ \mathbf{3} & \mathbf{5} & \mathbf{2} & \mathbf{4} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & ? & 0 & 0 \\ \mathbf{2} & ? & 0 & 0 \\ \mathbf{3} & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \mathbf{1} & \mathbf{2} & 0 \\ 0 & ? & \mathbf{2} & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & ? & \mathbf{6} \\ 0 & ? & \mathbf{2} & \mathbf{4} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{1} & 0 & 0 \\ \mathbf{2} & ? & 0 & 0 \\ \mathbf{3} & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \mathbf{1} & \mathbf{2} & 0 \\ 0 & \mathbf{1} & \mathbf{2} & 0 \\ 0 & ? & \mathbf{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & \mathbf{3} & \mathbf{6} \\ 0 & ? & \mathbf{2} & \mathbf{4} \end{pmatrix}$$

Figure: This "uncovers" entries in red and green contributions.

Extra: iterative partial completability

Counterexample

Iterative completability is not a necessary condition for fixed-support identifiability when $r \geq 3$.

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & ? & 0 & 0 \\ 2 & ? & 0 & 0 \\ 3 & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & ? & 2 & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & ? & 6 \\ 0 & ? & 2 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & 3 & 6 \\ 0 & ? & 2 & 4 \end{pmatrix}$$

Figure: Then, red and green contributions are completable.

Extra: iterative partial completability

Counterexample

Iterative completability is not a necessary condition for fixed-support identifiability when $r \geq 3$.

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 6 & 5 & 6 \\ 3 & 5 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & ? & 0 & 0 \\ 2 & ? & 0 & 0 \\ 3 & ? & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & ? & 2 & 0 \\ 0 & ? & ? & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & ? & ? & 6 \\ 0 & ? & 2 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 3 & 6 \\ 0 & 2 & 2 & 4 \end{pmatrix}$$

Figure: We finally complete blue contribution.