Identifying Subspaces from Canonical Projections

Daniel L. Pimentel-Alarcón

Applied Algebra Seminar
November 13^{th}, 2014

Robert Nowak and Nigel Boston

Outline

- Introduction
- Problem Description
- Setup
- The Answer
- Sketch of the proof
- Application
- Conclusions

Outline

- Introduction
- Problem Description
- Setup
- The Answer
- Sketch of the proof
- Application
- Conclusions

Introduction

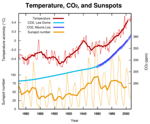
We have lots of data

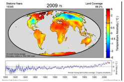
Introduction

We have lots of data

Introduction

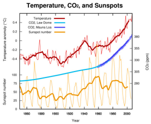
We have lots of data

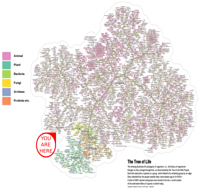


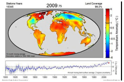


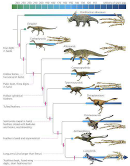
Introduction

We have lots of data



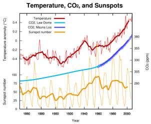


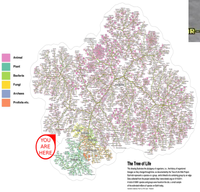


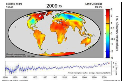


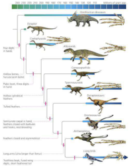
Introduction

We have lots of data



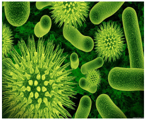


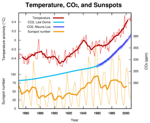


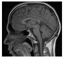


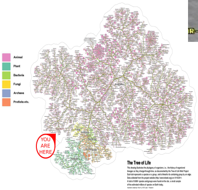
Introduction

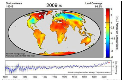
We have lots of data

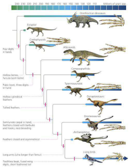






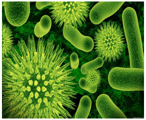


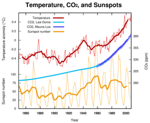


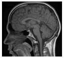


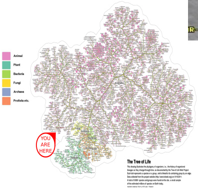
Introduction

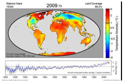
We have lots of data

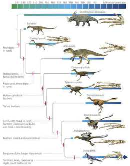






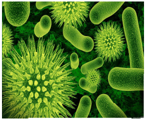


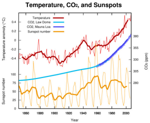


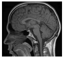


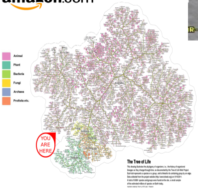
Introduction

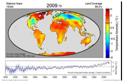
We have lots of data

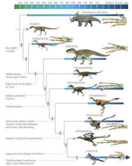












Introduction

But many times data is missing!

Introduction

But many times data is missing!
A typical example: Netflix.

Introduction

But many times data is missing!
A typical example: Netflix.

$$
movies \left\{ \underbrace{\begin{bmatrix}
4 & \cdot & 1 & 1 & 5 & 4 & 2 & 5 & 3 & \cdot \\
\cdot & 5 & \cdot & 1 & \cdot & \cdot & \cdot & 4 & \cdot & 5 \\
1 & \cdot & 2 & \cdot & 2 & 5 & 1 & \cdot & \cdot & 4 \\
2 & \cdot & 5 & 4 & \cdot & 1 & \cdot & 2 & \cdot & 4 \\
\cdot & 5 & 1 & \cdot & 5 & \cdot & \cdot & \cdot & 1 & \cdot \\
1 & 5 & \cdot & 2 & \cdot & 2 & 1 & \cdot & \cdot & 5
\end{bmatrix}}_{people} \right.
$$

Introduction

But many times data is missing!
A typical example: Netflix.

$$
movies \left\{ \underbrace{\begin{bmatrix}
4 & \cdot & 1 & 1 & 5 & 4 & 2 & 5 & 3 & \cdot \\
\cdot & 5 & \cdot & 1 & \cdot & \cdot & \cdot & 4 & \cdot & 5 \\
1 & \cdot & 2 & \cdot & 2 & 5 & 1 & \cdot & \cdot & 4 \\
2 & \cdot & 5 & 4 & \cdot & 1 & \cdot & 2 & \cdot & 4 \\
\cdot & 5 & 1 & \cdot & 5 & \cdot & \cdot & \cdot & 1 & \cdot \\
1 & 5 & \cdot & 2 & \cdot & 2 & 1 & \cdot & \cdot & 5
\end{bmatrix}}_{people} \right.
$$

Nobody has seen every movie.

Introduction

But if we knew *who* would like *what*...

Introduction

But if we knew *who* would like *what*...

$$
movies \left\{ \underbrace{\begin{bmatrix} 4 & 5 & 1 & 1 & 5 & 4 & 2 & 5 & 3 & 1 \\ 2 & 5 & 3 & 1 & 4 & 1 & 4 & 4 & 3 & 5 \\ 1 & 3 & 2 & 5 & 2 & 5 & 1 & 3 & 2 & 4 \\ 2 & 1 & 5 & 4 & 3 & 1 & 5 & 2 & 5 & 4 \\ 3 & 5 & 1 & 1 & 5 & 1 & 4 & 3 & 1 & 3 \\ 1 & 5 & 3 & 2 & 2 & 2 & 1 & 4 & 5 & 5 \end{bmatrix}}_{people} \right.
$$

Introduction

But if we knew *who* would like *what*...

$$movies \left\{ \underbrace{\begin{bmatrix} 4 & 5 & 1 & 1 & 5 & 4 & 2 & 5 & 3 & 1 \\ 2 & 5 & 3 & 1 & 4 & 1 & 4 & 4 & 3 & 5 \\ 1 & 3 & 2 & 5 & 2 & 5 & 1 & 3 & 2 & 4 \\ 2 & 1 & 5 & 4 & 3 & 1 & 5 & 2 & 5 & 4 \\ 3 & 5 & 1 & 1 & 5 & 1 & 4 & 3 & 1 & 3 \\ 1 & 5 & 3 & 2 & 2 & 2 & 1 & 4 & 5 & 5 \end{bmatrix}}_{people} \right.$$

We would be able to make good recommendations!

Introduction

Replace movies with whatever you want...

Introduction

Replace movies with whatever you want...

$$clothes \left\{ \underbrace{\begin{bmatrix} 4 & \cdot & 1 & 1 & 5 & 4 & 2 & 5 & 3 & \cdot \\ \cdot & 5 & \cdot & 1 & \cdot & \cdot & \cdot & 4 & \cdot & 5 \\ 1 & \cdot & 2 & \cdot & 2 & 5 & 1 & \cdot & \cdot & 4 \\ 2 & \cdot & 5 & 4 & \cdot & 1 & \cdot & 2 & \cdot & 4 \\ \cdot & 5 & 1 & \cdot & 5 & \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & 5 & \cdot & 2 & \cdot & 2 & 1 & \cdot & \cdot & 5 \end{bmatrix}}_{people} \right.$$

Introduction

Replace movies with whatever you want...

$$
clothes \left\{ \underbrace{\begin{bmatrix} 4 & \cdot & 1 & 1 & 5 & 4 & 2 & 5 & 3 & \cdot \\ \cdot & 5 & \cdot & 1 & \cdot & \cdot & \cdot & 4 & \cdot & 5 \\ 1 & \cdot & 2 & \cdot & 2 & 5 & 1 & \cdot & \cdot & 4 \\ 2 & \cdot & 5 & 4 & \cdot & 1 & \cdot & 2 & \cdot & 4 \\ \cdot & 5 & 1 & \cdot & 5 & \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & 5 & \cdot & 2 & \cdot & 2 & 1 & \cdot & \cdot & 5 \end{bmatrix}}_{people} \right.
$$

We would be able to make good adds!

Introduction

We want to analyze all that data.

Introduction

We want to analyze all that data.

- Linear algebra is one of our favorite tools...

Introduction

We want to analyze all that data.

- Linear algebra is one of our favorite tools...
  - We know lots about linear algebra.

Introduction

We want to analyze all that data.

- Linear algebra is one of our favorite tools...
  - We know lots about linear algebra.
- But what if data are missing?

Introduction

We want to analyze all that data.

- Linear algebra is one of our favorite tools...
  - We know lots about linear algebra.
- But what if data are missing?
  - ?????

Introduction

We want to analyze all that data.

- Linear algebra is one of our favorite tools...
  - We know lots about linear algebra.
- But what if data are missing?
  - ?????
- There is great interest on extending usage of linear algebra to incomplete datasets.

Introduction

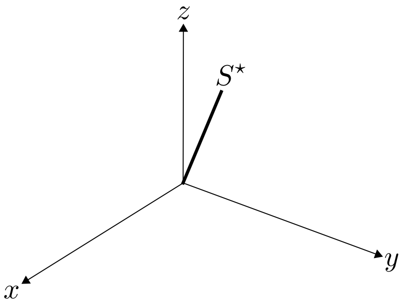
We want to analyze all that data.

- Linear algebra is one of our favorite tools...
  - We know lots about linear algebra.
- But what if data are missing?
  - ?????
- There is great interest on extending usage of linear algebra to incomplete datasets.
- That is what we are studying.

Outline

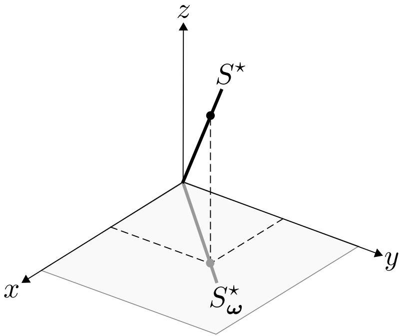
Problem description

$S^\star :=$ r-dimensional subspace of $\mathbb{R}^d$, $r < d$.



Problem description

$S^\star_{\boldsymbol{\omega}} :=$ Projection of $S^\star$ onto a canonical subspace.

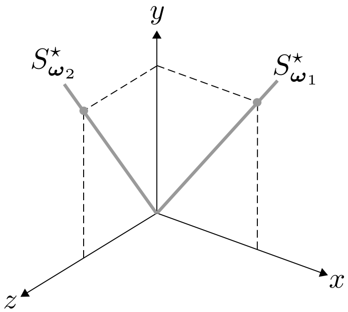


Problem description

Suppose I don't tell you $S^\star$...

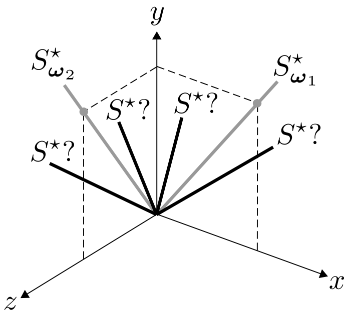
Problem description

Suppose I don't tell you $S^\star$... but I give you a set of projections of $S^\star$ onto some canonical subspaces.



Problem description

Suppose I don't tell you $S^\star$...but I give you a set of projections of $S^\star$ onto some canonical subspaces.

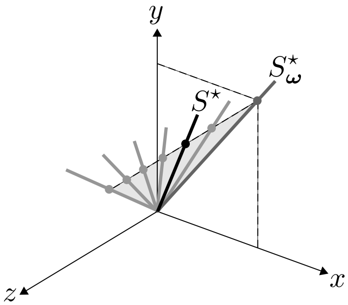


Can you uniquely determine $S^\star$ from this set of projections?

Problem description

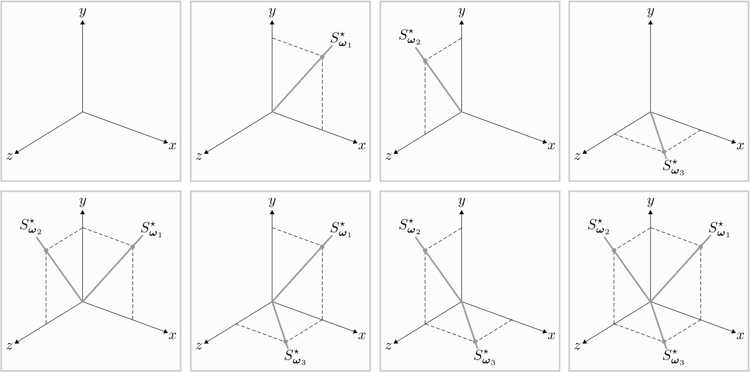
Is this even possible?

There might be many subspaces that agree with the projections.



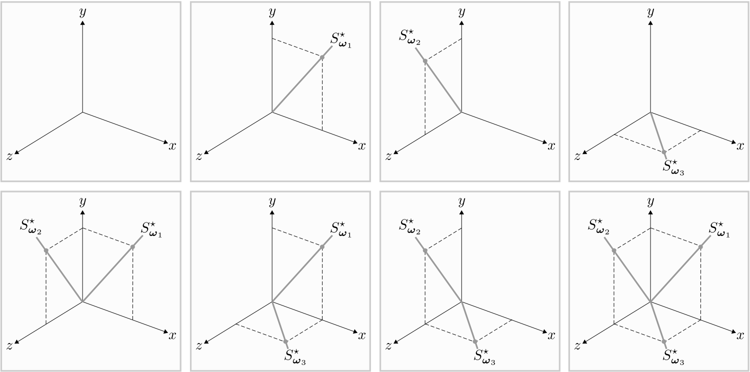
Problem description

Well... it depends on which set of projections I give you.



Problem description

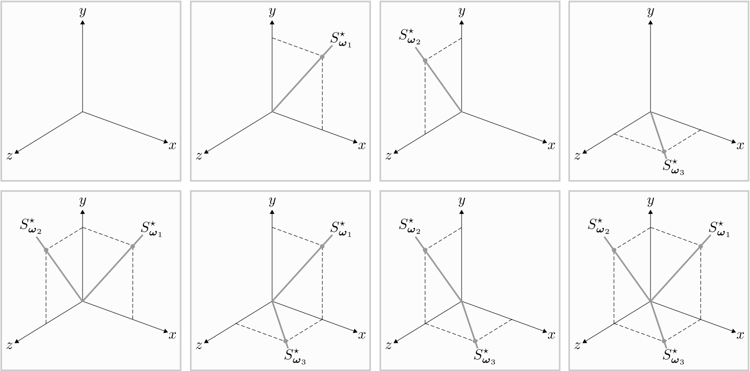
Well... it depends on which set of projections I give you.



Can you tell which are *the good sets*?

Problem description

Well... it depends on which set of projections I give you.



Can you tell which are *the good sets*?

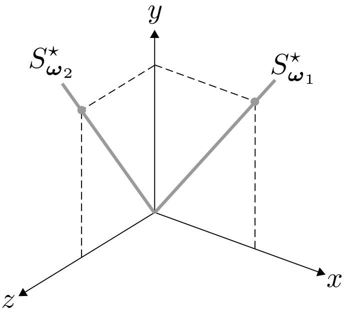
This is what we answer here: which are *the good sets*.

Outline

- Introduction
- Problem Description ✓
- Setup
- The Answer
- Sketch of the proof
- Application
- Conclusions

Setup

The columns of $\boldsymbol{\Omega}$ will index the given projections.



$$\boldsymbol{\Omega} = \begin{array}{c} \begin{array}{cc} \boldsymbol{\omega}_1 & \boldsymbol{\omega}_2 \end{array} \\ \left[\begin{array}{cc} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{array}\right] \end{array}$$

Setup

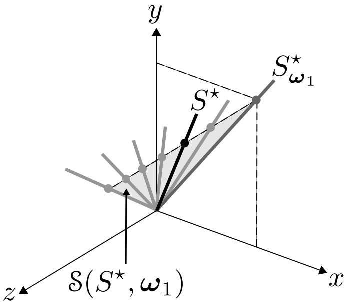
- $\mathrm{Gr}(r, \mathbb{R}^d) :=$ Grassmannian manifold of r-dimensional subspaces in $\mathbb{R}^d$.

Setup

- $\mathrm{Gr}(r, \mathbb{R}^d) :=$ Grassmannian manifold of r-dimensional subspaces in $\mathbb{R}^d$.
- $\mathcal{S}(S^\star, \boldsymbol{\Omega}) :=$ Set of r-dimensional subspaces that agree with $S^\star$ on $\boldsymbol{\Omega}$.

Setup

- $\mathrm{Gr}(r, \mathbb{R}^d) :=$ Grassmannian manifold of r-dimensional subspaces in $\mathbb{R}^d$.
- $\mathcal{S}(S^\star, \boldsymbol{\Omega}) :=$ Set of r-dimensional subspaces that agree with $S^\star$ on $\boldsymbol{\Omega}$.

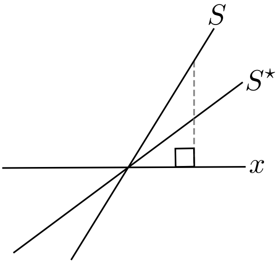


Setup

- $S^\star$ is r-dimensional.

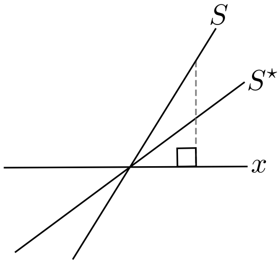
Setup

- $S^\star$ is r-dimensional.
- The projection of $S^\star$ onto $\leq r$ canonical coordinates gives no information about $S^\star$.



Setup

- $S^\star$ is r-dimensional.
- The projection of $S^\star$ onto $\leq r$ canonical coordinates gives no information about $S^\star$.



- $\Rightarrow$ Assume w.l.o.g. that all projections are onto $r + 1$ canonical coordinates.

Setup

- For any matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$:

$$\mathbf{\Omega}' = \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}}_{n(\mathbf{\Omega}') \;:=\; \#\text{columns}} \quad \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} m(\mathbf{\Omega}') := \#\text{nonzero rows}$$

Setup

- For any matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$:

$$\mathbf{\Omega}' = \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}}_{n(\mathbf{\Omega}') \;:=\; \#\text{columns}} \quad \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} m(\mathbf{\Omega}') := \#\text{nonzero rows}$$

- $d - r$ projections are *necessary*, so we will assume w.l.o.g.

$$n(\mathbf{\Omega}) = d - r.$$

Outline

- Introduction
- Problem Description ✓
- Setup ✓
- The Answer
- Sketch of the proof
- Application
- Conclusions

The Answer

Theorem (Pimentel-Alarcón, Nowak, Boston, '14)

For almost every $S^\star$*, with respect to the uniform measure over* $\mathrm{Gr}(r, \mathbb{R}^d)$*,* $S^\star$ *is the only subspace in* $\mathcal{S}(S^\star, \mathbf{\Omega})$ *if and only if for every matrix* $\mathbf{\Omega}'$ *formed with a subset of the columns in* $\mathbf{\Omega}$*,*

$$m(\mathbf{\Omega}') \;\geq\; n(\mathbf{\Omega}') + r.$$

The Answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \mathbf{\Omega})$ if and only if for every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

$$m(\mathbf{\Omega}') \ \geq \ n(\mathbf{\Omega}') + r.$$

The Answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \mathbf{\Omega})$ if and only if for every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

$$m(\mathbf{\Omega}') \;\geq\; n(\mathbf{\Omega}') + r.$$

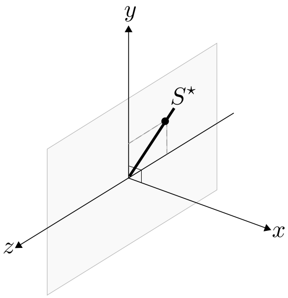
There is a set of measure zero of *bad* subspaces that we wouldn't identify.

The Answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \boldsymbol{\Omega})$ if and only if for every matrix $\boldsymbol{\Omega}'$ formed with a subset of the columns in $\boldsymbol{\Omega}$,

$$m(\boldsymbol{\Omega}') \;\geq\; n(\boldsymbol{\Omega}') + r.$$

There is a set of measure zero of *bad* subspaces that we wouldn't identify.



The Answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \mathbf{\Omega})$ if and only if for every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

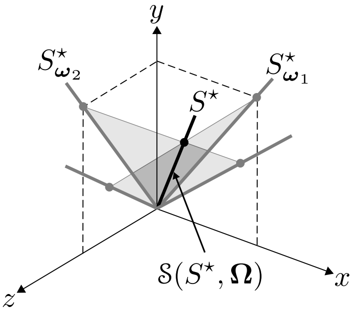
$$m(\mathbf{\Omega}') \ \geq \ n(\mathbf{\Omega}') + r.$$

The Answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \mathbf{\Omega})$ if and only if for every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

$$m(\mathbf{\Omega}') \ \geq \ n(\mathbf{\Omega}') + r.$$

This is what we want!



The answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \mathbf{\Omega})$ if and only if for every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

$$m(\mathbf{\Omega}') \ \geq \ n(\mathbf{\Omega}') + r.$$

The answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \mathbf{\Omega})$ if and only if for every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

$$m(\mathbf{\Omega}') \;\geq\; n(\mathbf{\Omega}') + r.$$

This is the answer!

The answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \mathbf{\Omega})$ if and only if for every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

$$m(\mathbf{\Omega}') \;\geq\; n(\mathbf{\Omega}') + r.$$

This is the answer!

Every subset of n columns of $\mathbf{\Omega}$ has at least $n + r$ nonzero rows.

The answer

For almost every $S^\star$, with respect to the uniform measure over $\mathrm{Gr}(r, \mathbb{R}^d)$, $S^\star$ is the only subspace in $\mathcal{S}(S^\star, \boldsymbol{\Omega})$ if and only if for every matrix $\boldsymbol{\Omega}'$ formed with a subset of the columns in $\boldsymbol{\Omega}$,

$$m(\boldsymbol{\Omega}') \;\geq\; n(\boldsymbol{\Omega}') + r.$$

This is the answer!

Every subset of n columns of $\boldsymbol{\Omega}$ has at least $n + r$ nonzero rows.

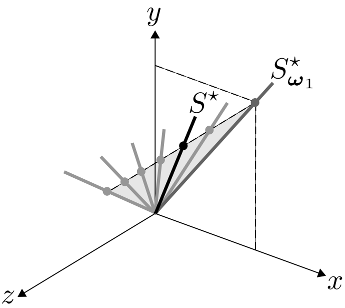
$$\boldsymbol{\Omega} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \;\Rightarrow\; \text{Check: } \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

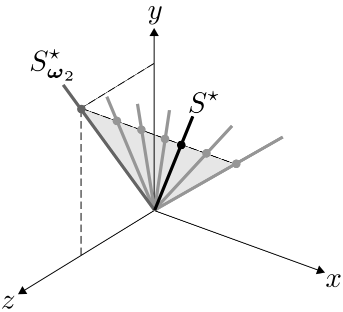
Outline

- Introduction
- Problem Description ✓
- Setup ✓
- The Answer ✓
- Sketch of the proof
- Application
- Conclusions

Sketch of the proof

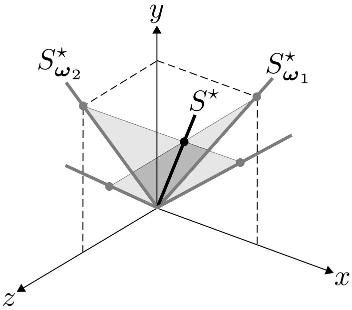
We will find the subspaces that agree with *each* projection.





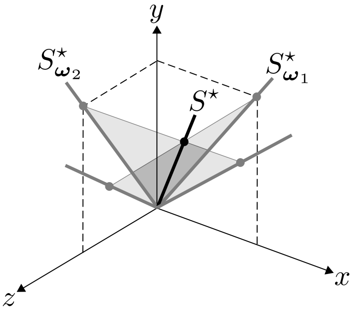
Sketch of the proof

Then find the intersection.



Sketch of the proof

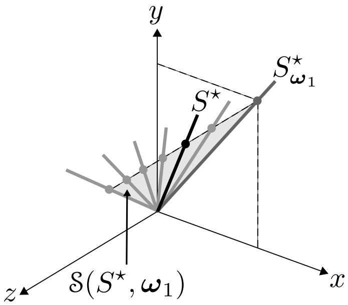
Then find the intersection.



If the intersection only contains one subspace, then ;)

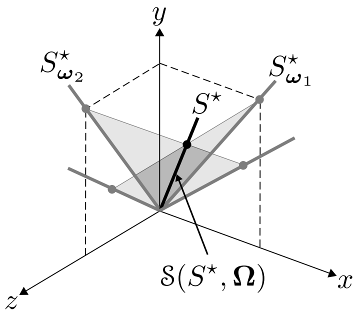
Sketch of the proof

$\mathcal{S}(S^\star, \boldsymbol{\omega}_i) :=$ Set of r-dimensional subspaces matching $S^\star$ on $\boldsymbol{\omega}_i$.



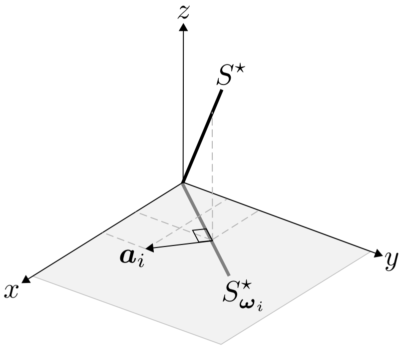
Sketch of the proof

$$\mathcal{S}(S^\star, \boldsymbol{\Omega}) = \bigcap_i \mathcal{S}(S^\star, \boldsymbol{\omega}_i).$$



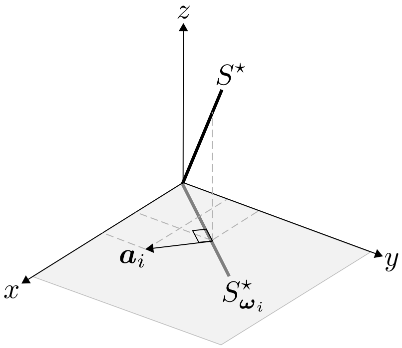
Sketch of the proof

$\boldsymbol{a}_i :=$ Vector orthogonal to the i^{th} projection.



Sketch of the proof

$\boldsymbol{a}_i :=$ Vector orthogonal to the i^{th} projection.



An entry in $\boldsymbol{a}_i$ is zero iff the corresponding entry in $\boldsymbol{\omega}_i$ is zero.

Sketch of the proof

One great thing:

- Every subspace in $\mathcal{S}(S^\star, \boldsymbol{\omega}_i)$ is orthogonal to $\boldsymbol{a}_i$.

Sketch of the proof

One great thing:

- Every subspace in $\mathcal{S}(S^\star, \boldsymbol{\omega}_i)$ is orthogonal to $\boldsymbol{a}_i$.

Cool! $\Rightarrow$

- Construct

$$\boldsymbol{A} = \left[\ \boldsymbol{a}_1 \ \middle|\ \cdots \ \middle|\ \boldsymbol{a}_N \ \right].$$

Sketch of the proof

One great thing:

- Every subspace in $\mathcal{S}(S^\star, \boldsymbol{\omega}_i)$ is orthogonal to $\boldsymbol{a}_i$.

Cool! $\Rightarrow$

- Construct

$$\boldsymbol{A} = \left[\ \boldsymbol{a}_1 \ \middle|\ \cdots \ \middle|\ \boldsymbol{a}_N\ \right].$$

- Every $S \in \mathcal{S}(S^\star, \boldsymbol{\Omega})$ must be contained in

$$\ker \boldsymbol{A}^\mathsf{T}.$$

Sketch of the proof

What we really want is to determine $\dim \ker \boldsymbol{A}^\mathsf{T}$.

Sketch of the proof

What we really want is to determine $\dim \ker \boldsymbol{A}^\mathsf{T}$.

- If $\dim \ker \boldsymbol{A}^\mathsf{T} > r$
  $\Rightarrow$ There are many subspaces that agree with the projections

Sketch of the proof

What we really want is to determine $\dim \ker \boldsymbol{A}^{\mathsf{T}}$.

- If $\dim \ker \boldsymbol{A}^{\mathsf{T}} > r$
  $\Rightarrow$ There are many subspaces that agree with the projections
- If $\dim \ker \boldsymbol{A}^{\mathsf{T}} = r$
  $\Rightarrow$ Only $S^{\star}$ will agree with the projections.

Sketch of the proof

What we really want is to determine $\dim \ker \boldsymbol{A}^\mathsf{T}$.

- If $\dim \ker \boldsymbol{A}^\mathsf{T} > r$
  $\Rightarrow$ There are many subspaces that agree with the projections
- If $\dim \ker \boldsymbol{A}^\mathsf{T} = r$
  $\Rightarrow$ Only $S^\star$ will agree with the projections. Moreover,

$$S^\star = \ker \boldsymbol{A}^\mathsf{T}$$

.

Sketch of the proof

- For any matrix $\boldsymbol{A}'$ formed with a subset of the columns in $\boldsymbol{A}$:

$$\boldsymbol{A}' = \underbrace{\begin{bmatrix} a_{11} & 0 & \\ a_{21} & a_{22} & \\ 0 & a_{32} & \\ 0 & 0 & \end{bmatrix}}_{n(\boldsymbol{A}') \,:=\, \#\text{columns}} \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} m(\boldsymbol{A}') := \#\text{nonzero rows}$$

Sketch of the proof

- For any matrix $\boldsymbol{A}'$ formed with a subset of the columns in $\boldsymbol{A}$:

$$\boldsymbol{A}' = \underbrace{\begin{bmatrix} a_{11} & 0 & \\ a_{21} & a_{22} & \\ 0 & a_{32} & \\ 0 & 0 & \end{bmatrix}}_{n(\boldsymbol{A}') \;:=\; \#\text{columns}} \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} m(\boldsymbol{A}') := \#\text{nonzero rows}$$

- We want $\dim\ker \boldsymbol{A}^{\mathsf{T}} = r$, so $\boldsymbol{A}$ better have $d - r$ linearly independent columns.

Sketch of the proof

We know how to deal with $\boldsymbol{A}$ using linear algebra!

- Through some technical details:

Lemma (Pimentel-Alarcón, Nowak, Boston, '14)

For almost every $S^\star$, the columns of $\boldsymbol{A}$ are linearly dependent if and only if $m(\boldsymbol{A}') < n(\boldsymbol{A}') + r$ for some matrix $\boldsymbol{A}'$ formed with a subset of the columns in $\boldsymbol{A}$.

Sketch of the proof

The zero entries of $\boldsymbol{\Omega}$ and $\boldsymbol{A}$ are in the same positions.

$$\boldsymbol{\Omega} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \Longleftrightarrow \quad \boldsymbol{A}' = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \\ 0 & 0 & a_{43} \end{bmatrix}$$

Sketch of the proof

The zero entries of $\boldsymbol{\Omega}$ and $\boldsymbol{A}$ are in the same positions.

$$\boldsymbol{\Omega} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \iff \boldsymbol{A}' = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \\ 0 & 0 & a_{43} \end{bmatrix}$$

Then

$$m(\boldsymbol{\Omega}') \geq n(\boldsymbol{\Omega}') + r \iff m(\boldsymbol{A}') \geq n(\boldsymbol{A}') + r$$

□

Sketch of the proof

To wrap up:

- Every $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

$$m(\mathbf{\Omega}') \geq n(\mathbf{\Omega}') + r$$

Sketch of the proof

To wrap up:

- Every $\boldsymbol{\Omega}'$ formed with a subset of the columns in $\boldsymbol{\Omega}$,

$$m(\boldsymbol{\Omega}') \geq n(\boldsymbol{\Omega}') + r$$

- Iff every $\boldsymbol{A}'$ formed with a subset of the columns in $\boldsymbol{A}$,

$$m(\boldsymbol{A}') \geq n(\boldsymbol{A}') + r$$

Sketch of the proof

To wrap up:

- Every $\boldsymbol{\Omega}'$ formed with a subset of the columns in $\boldsymbol{\Omega}$,

$$m(\boldsymbol{\Omega}') \geq n(\boldsymbol{\Omega}') + r$$

- Iff every $\boldsymbol{A}'$ formed with a subset of the columns in $\boldsymbol{A}$,

$$m(\boldsymbol{A}') \geq n(\boldsymbol{A}') + r$$

- Iff the columns in $\boldsymbol{A}$ are linearly independent, i.e.,

$$\dim \ker \boldsymbol{A}^\mathsf{T} = r$$

Sketch of the proof

To wrap up:

- Every $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$,

$$m(\mathbf{\Omega}') \geq n(\mathbf{\Omega}') + r$$

- Iff every $\boldsymbol{A}'$ formed with a subset of the columns in $\boldsymbol{A}$,

$$m(\boldsymbol{A}') \geq n(\boldsymbol{A}') + r$$

- Iff the columns in $\boldsymbol{A}$ are linearly independent, i.e.,

$$\dim \ker \boldsymbol{A}^{\mathsf{T}} = r$$

- Iff $S^{\star}$ is the only subspace in $\mathcal{S}(S^{\star}, \mathbf{\Omega})$. □

Outline

- Introduction
- Problem Description ✓
- Setup ✓
- The Answer ✓
- Sketch of the proof ✓
- Application
- Conclusions

Application

Low-Rank Matrix Completion (LRMC)

Application

Low-Rank Matrix Completion (LRMC)

- Given a subset of entries in a rank r matrix, exactly recover *all* of the missing entries.

$$\mathfrak{X}_{\mathbf{\Omega}} = \begin{bmatrix} 1 & \cdot & 3 & \cdot \\ 1 & 2 & \cdot & \cdot \\ \cdot & 2 & 3 & \cdot \\ \cdot & \cdot & \cdot & 4 \\ \cdot & \cdot & \cdot & 4 \end{bmatrix} \quad \Rightarrow \quad \hat{\mathfrak{X}} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

Application

Low-Rank Matrix Completion (LRMC)

- Given a subset of entries in a rank r matrix, exactly recover *all* of the missing entries.

$$\mathfrak{X}_{\boldsymbol{\Omega}} = \begin{bmatrix} 1 & \cdot & 3 & \cdot \\ 1 & 2 & \cdot & \cdot \\ \cdot & 2 & 3 & \cdot \\ \cdot & \cdot & \cdot & 4 \\ \cdot & \cdot & \cdot & 4 \end{bmatrix} \quad \Rightarrow \quad \hat{\mathfrak{X}} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

- $\sim$ Identifying the subspace spanned by the columns, $S^\star$.

Application

Low-Rank Matrix Completion (LRMC)

- Given a subset of entries in a rank r matrix, exactly recover *all* of the missing entries.

$$\mathfrak{X}_{\mathbf{\Omega}} = \begin{bmatrix} 1 & \cdot & 3 & \cdot \\ 1 & 2 & \cdot & \cdot \\ \cdot & 2 & 3 & \cdot \\ \cdot & \cdot & \cdot & 4 \\ \cdot & \cdot & \cdot & 4 \end{bmatrix} \quad \Rightarrow \quad \hat{\mathfrak{X}} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

- $\sim$ Identifying the subspace spanned by the columns, $S^\star$. Here

$$\hat{S} = \mathrm{span} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Application

How do we know we got the right completion (subspace)?

Application

How do we know we got the right completion (subspace)?

- Maybe the real completion is:

$$\mathfrak{X}_{\boldsymbol{\Omega}} = \begin{bmatrix} 1 & \cdot & 3 & \cdot \\ 1 & 2 & \cdot & \cdot \\ \cdot & 2 & 3 & \cdot \\ \cdot & \cdot & \cdot & 4 \\ \cdot & \cdot & \cdot & 4 \end{bmatrix} \quad \Rightarrow \quad \mathfrak{X} = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & 2 & 3 & 2 \\ 1 & 2 & 3 & 2 \\ 2 & 4 & 6 & 4 \\ 2 & 4 & 6 & 4 \end{bmatrix}$$

Application

How do we know we got the right completion (subspace)?

- Maybe the real completion is:

$$\mathcal{X}_{\mathbf{\Omega}} = \begin{bmatrix} 1 & \cdot & 3 & \cdot \\ 1 & 2 & \cdot & \cdot \\ \cdot & 2 & 3 & \cdot \\ \cdot & \cdot & \cdot & 4 \\ \cdot & \cdot & \cdot & 4 \end{bmatrix} \quad \Rightarrow \quad \mathcal{X} = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & 2 & 3 & 2 \\ 1 & 2 & 3 & 2 \\ 2 & 4 & 6 & 4 \\ 2 & 4 & 6 & 4 \end{bmatrix}$$

And the real subspace is

$$S^\star = \operatorname{span} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

Application

How do we know we got the right completion (subspace)?

Application

How do we know we got the right completion (subspace)?
Known results e.g. (Candès and Recht, '09)

Application

How do we know we got the right completion (subspace)?
Known results e.g. (Candès and Recht, '09)

- Require random observed entries.

Application

How do we know we got the right completion (subspace)?
Known results e.g. (Candès and Recht, '09)

- Require random observed entries.
  - May not be justified.

Application

How do we know we got the right completion (subspace)?
Known results e.g. (Candès and Recht, '09)

- Require random observed entries.
    - May not be justified.
- Require incoherence

Application

How do we know we got the right completion (subspace)?
Known results e.g. (Candès and Recht, '09)

- Require random observed entries.
  - May not be justified.
- Require incoherence
  - Sufficient, but not necessary condition.

Application

How do we know we got the right completion (subspace)?
Known results e.g. (Candès and Recht, '09)

- Require random observed entries.
  - May not be justified.
- Require incoherence
  - Sufficient, but not necessary condition.
  - Generally unverifiable or unjustified in practice.

Application

How do we know we got the right completion (subspace)?
Known results e.g. (Candès and Recht, '09)

- Require random observed entries.
  - May not be justified.
- Require incoherence
  - Sufficient, but not necessary condition.
  - Generally unverifiable or unjustified in practice.
- Work with high probability (if assumptions are met).

Application

How do we know we got the right completion (subspace)?
Known results e.g. (Candès and Recht, '09)

- Require random observed entries.
  - May not be justified.
- Require incoherence
  - Sufficient, but not necessary condition.
  - Generally unverifiable or unjustified in practice.
- Work with high probability (if assumptions are met).

What if these assumptions are not met? How can we validate a completion?

Application

Corollary (Pimentel-Alarcón, Nowak, Boston, '14)

Let the columns of $\mathfrak{X}$ *be drawn independently according to* μ*, an absolutely continuous distribution with respect to the Lebesgue measure on* $S^\star$*. Suppose* $\mathfrak{X}_{\mathbf{\Omega}}$ *can be partitioned into two sets of columns,* $\mathfrak{X}_{\mathbf{\Omega}_1}$ *and* $\mathfrak{X}_{\mathbf{\Omega}_2}$*, such that* $\mathbf{\Omega}_2$ *satisfies the conditions of the subspace identifiability theorem.*
Let $\hat{S}$ *be the output of running an LRMC algorithm on* $\mathfrak{X}_{\mathbf{\Omega}_1}$*.*
Then for almost every $S^\star$*, and almost surely with respect to* μ*,* $\mathfrak{X}_{\mathbf{\Omega}_2}$ *fits in* $\hat{S}$ *if and only if* $\hat{S} = S^\star$*.*

Application

Corollary (Pimentel-Alarcón, Nowak, Boston, '14)

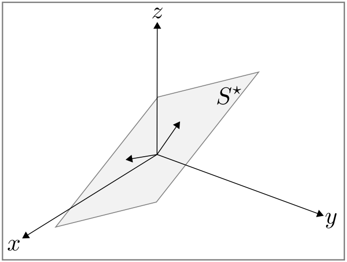
Let the columns of $\mathfrak{X}$ *be drawn independently according to* μ*, an absolutely continuous distribution with respect to the Lebesgue measure on* $S^\star$*. Suppose* $\mathcal{X}_\Omega$ *can be partitioned into two sets of columns,* $\mathcal{X}_{\Omega_1}$ *and* $\mathcal{X}_{\Omega_2}$*, such that* Ω_2 *satisfies the conditions of the subspace identifiability theorem.*
Let $\hat{S}$ *be the output of running an LRMC algorithm on* $\mathcal{X}_{\Omega_1}$*.*
Then for almost every $S^\star$*, and almost surely with respect to* μ*,* $\mathcal{X}_{\Omega_2}$ *fits in* $\hat{S}$ *if and only if* $\hat{S} = S^\star$*.*

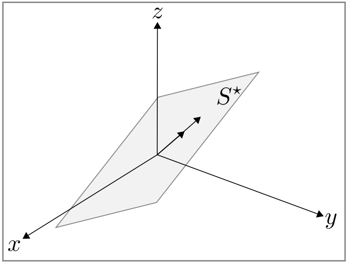
Application

Just to make sure we have enough *useful* data

Application

Just to make sure we have enough *useful* data





Application

Corollary (Pimentel-Alarcón, Nowak, Boston, '14)

Let the columns of $\mathfrak{X}$ be drawn independently according to μ, an absolutely continuous distribution with respect to the Lebesgue measure on $S^\star$. Suppose $\mathfrak{X}_{\mathbf{\Omega}}$ can be partitioned into two sets of columns, $\mathfrak{X}_{\mathbf{\Omega}_1}$ and $\mathfrak{X}_{\mathbf{\Omega}_2}$, such that $\mathbf{\Omega}_2$ satisfies the conditions of the subspace identifiability theorem.

Let $\hat{S}$ be the output of running an LRMC algorithm on $\mathfrak{X}_{\mathbf{\Omega}_1}$. Then for almost every $S^\star$, and almost surely with respect to μ, $\mathfrak{X}_{\mathbf{\Omega}_2}$ fits in $\hat{S}$ if and only if $\hat{S} = S^\star$.

Application

Corollary (Pimentel-Alarcón, Nowak, Boston, '14)

Let the columns of $\mathcal{X}$ be drawn independently according to μ, an absolutely continuous distribution with respect to the Lebesgue measure on $S^\star$. Suppose $\mathcal{X}_\Omega$ can be partitioned into two sets of columns, $\mathcal{X}_{\Omega_1}$ and $\mathcal{X}_{\Omega_2}$, such that Ω_2 satisfies the conditions of the subspace identifiability theorem.
Let $\hat{S}$ be the output of running an LRMC algorithm on $\mathcal{X}_{\Omega_1}$.
Then for almost every $S^\star$, and almost surely with respect to μ, $\mathcal{X}_{\Omega_2}$ fits in $\hat{S}$ if and only if $\hat{S} = S^\star$.

Application

Corollary (Pimentel-Alarcón, Nowak, Boston, '14)

Let the columns of $\mathcal{X}$ be drawn independently according to μ, an absolutely continuous distribution with respect to the Lebesgue measure on $S^\star$. Suppose $\mathcal{X}_\Omega$ can be partitioned into two sets of columns, $\mathcal{X}_{\Omega_1}$ and $\mathcal{X}_{\Omega_2}$, such that Ω_2 satisfies the conditions of the subspace identifiability theorem.
Let $\hat{S}$ be the output of running an LRMC algorithm on $\mathcal{X}_{\Omega_1}$.
Then for almost every $S^\star$, and almost surely with respect to μ, $\mathcal{X}_{\Omega_2}$ fits in $\hat{S}$ if and only if $\hat{S} = S^\star$.

Application

Corollary (Pimentel-Alarcón, Nowak, Boston, '14)

Let the columns of $\mathcal{X}$ be drawn independently according to μ, an absolutely continuous distribution with respect to the Lebesgue measure on $S^\star$. Suppose $\mathcal{X}_\Omega$ can be partitioned into two sets of columns, $\mathcal{X}_{\Omega_1}$ and $\mathcal{X}_{\Omega_2}$, such that Ω_2 satisfies the conditions of the subspace identifiability theorem. Let $\hat{S}$ be the output of running an LRMC algorithm on $\mathcal{X}_{\Omega_1}$. Then for almost every $S^\star$, and almost surely with respect to μ, $\boldsymbol{\mathcal{X}}_{\boldsymbol{\Omega}_2}$ fits in $\hat{S}$ if and only if $\hat{S} = S^\star$.

Application

In contrast, our results:

Application

In contrast, our results:

- Work for arbitrary observation schemes.

Application

In contrast, our results:

- Work for arbitrary observation schemes.
- Work for almost every subspace.

Application

In contrast, our results:

- Work for arbitrary observation schemes.
- Work for almost every subspace.
  - No incoherence assumption required.

Application

In contrast, our results:

- Work for arbitrary observation schemes.
- Work for almost every subspace.
  - No incoherence assumption required.
- Hold with probability 1.

Outline

- Introduction
- Problem Description ✓
- Setup ✓
- The Answer ✓
- Sketch of the proof ✓
- Application ✓
- Conclusions

Conclusions

Now we know that:

- It is possible to uniquely identify an r-dimensional subspace $S^\star$ from its projections onto $\mathbf{\Omega}$.

Conclusions

Now we know that:

- It is possible to uniquely identify an r-dimensional subspace $S^\star$ from its projections onto $\mathbf{\Omega}$.
- If and only if every subset of n columns of $\mathbf{\Omega}$ has at least $n + r$ nonzero rows.

Conclusions

Now we know that:

- It is possible to uniquely identify an r-dimensional subspace $S^\star$ from its projections onto $\mathbf{\Omega}$.
- If and only if every subset of n columns of $\mathbf{\Omega}$ has at least $n + r$ nonzero rows.
- Whence $S^\star = \ker \boldsymbol{A}^\mathsf{T}$.

Thanks.