# Deterministic Conditions for Subspace Identifiability from Incomplete Sampling

**Daniel L. Pimentel-Alarcón**,
Nigel Boston and Robert Nowak

University of Wisconsin-Madison

ISIT, 2015

# Outline

- Introduction
- What changes with missing data?
- Subspace Identifiability Problem
- Setup
- The Answer
- Application
- Conclusions

# Outline

- Introduction
- What changes with missing data?
- Subspace Identifiability Problem
- Setup
- The Answer
- Application
- Conclusions

# Introduction

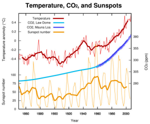
We have lots of data

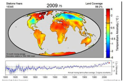
# Introduction

We have lots of data

# Introduction

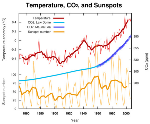
We have lots of data

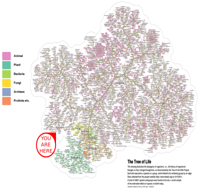


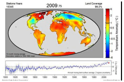


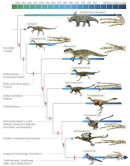
# Introduction

We have lots of data



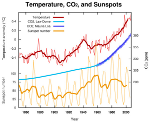


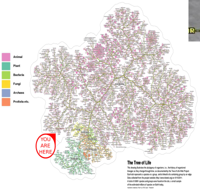


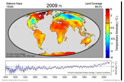


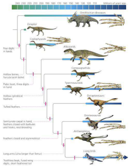
# Introduction

We have lots of data



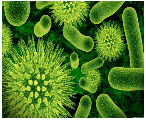
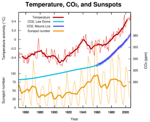


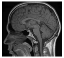


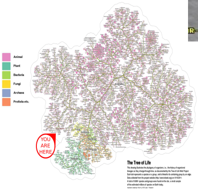


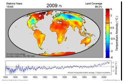
# Introduction

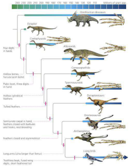
We have lots of data



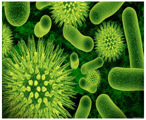
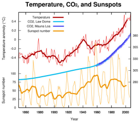



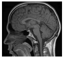





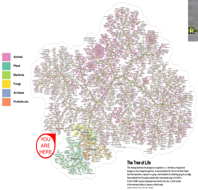
# Introduction

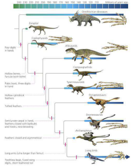
We have lots of data

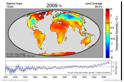













# Introduction

We have lots of data

# Introduction

We have lots of data

And we want to analyze it.

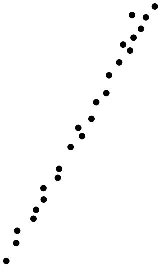
# Introduction

Linear Algebra is one of our favorite tools.

# Introduction

Linear Algebra is one of our favorite tools.

- Because data is often well-modeled by linear structures.

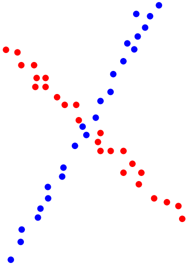


$$\begin{bmatrix} 1 & 2 & 1 & 3 & 2 & 1 & 3 & 1 & 2 & 2 \\ 2 & 4 & 2 & 6 & 4 & 2 & 6 & 2 & 4 & 4 \\ 3 & 6 & 3 & 9 & 6 & 3 & 9 & 3 & 6 & 6 \\ 1 & 2 & 1 & 3 & 2 & 1 & 3 & 1 & 2 & 2 \\ 2 & 4 & 2 & 6 & 4 & 2 & 6 & 2 & 4 & 4 \\ 3 & 6 & 3 & 9 & 6 & 3 & 9 & 3 & 6 & 6 \end{bmatrix}$$

# Introduction

Linear Algebra is one of our favorite tools.

- Because data is often well-modeled by linear structures.
- Or unions of linear structures.



$$\begin{bmatrix} 1 & 4 & 1 & 3 & 3 & 1 & 2 & 1 & 2 & 1 \\ 2 & 4 & 2 & 6 & 3 & 2 & 2 & 2 & 4 & 1 \\ 3 & 4 & 3 & 9 & 3 & 3 & 2 & 3 & 6 & 1 \\ 1 & 8 & 1 & 3 & 6 & 1 & 4 & 1 & 2 & 2 \\ 2 & 8 & 2 & 6 & 6 & 2 & 4 & 2 & 4 & 2 \\ 3 & 8 & 3 & 9 & 6 & 3 & 4 & 3 & 6 & 2 \end{bmatrix}$$

# Introduction

That's all very nice, but... often data is missing!

# Introduction

That's all very nice, but... often data is missing!

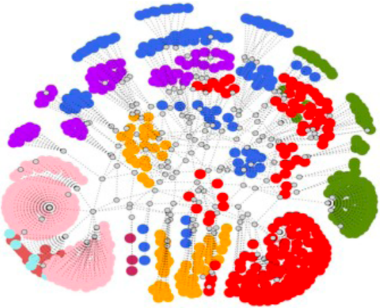
- Example: Vision.

Image: Hopkins 155 Dataset

# Introduction

Often data is missing!

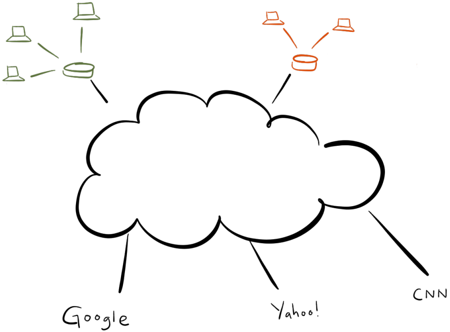
- Other example: Network topology estimation



# Introduction

Often data is missing!

- Other example: Network topology estimation



# Introduction

Often data is missing!

- Other example: Network topology estimation

$$\text{monitors}\left\{\underbrace{\begin{bmatrix} 1 & \cdot & \cdot & 3 & \cdot & 3 & \cdot & 1 & 2 & \cdot \\ 2 & \cdot & 2 & \cdot & \cdot & 6 & \cdot & \cdot & 4 & \cdot \\ \cdot & \cdot & 3 & \cdot & \cdot & 9 & \cdot & 3 & 6 & \cdot \\ 1 & \cdot & 1 & 3 & 6 & \cdot & 4 & 1 & 2 & 2 \\ \cdot & 8 & \cdot & \cdot & 6 & \cdot & 4 & \cdot & \cdot & \cdot \\ \cdot & 8 & \cdot & \cdot & \cdot & \cdot & 4 & \cdot & \cdot & 2 \end{bmatrix}}_{\text{IP's}}\right.$$

# Introduction

Often data is missing!

- Other example: Network topology estimation

$$\text{monitors}\left\{\underbrace{\begin{bmatrix} 1 & \cdot & \cdot & 3 & \cdot & 3 & \cdot & 1 & 2 & \cdot \\ 2 & \cdot & 2 & \cdot & \cdot & 6 & \cdot & \cdot & 4 & \cdot \\ \cdot & \cdot & 3 & \cdot & \cdot & 9 & \cdot & 3 & 6 & \cdot \\ 1 & \cdot & 1 & 3 & 6 & \cdot & 4 & 1 & 2 & 2 \\ \cdot & 8 & \cdot & \cdot & 6 & \cdot & 4 & \cdot & \cdot & \cdot \\ \cdot & 8 & \cdot & \cdot & \cdot & \cdot & 4 & \cdot & \cdot & 2 \end{bmatrix}}_{\text{IP's}}\right.$$

- We still want to analyze these datasets.

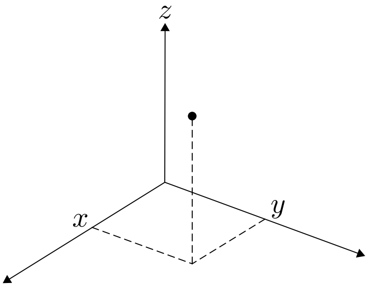
# Introduction

- We want to understand how things change when data is missing.

# Outline

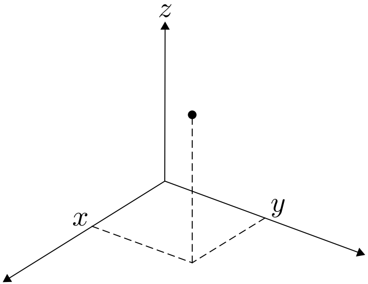
# What changes with missing data?

Say I give you one datapoint.



# What changes with missing data?

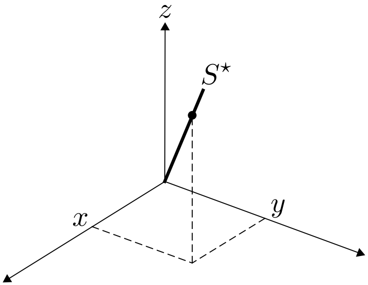
Say I give you one datapoint.



And I tell you it lies in a $1$-dimensional subspace $S^{\star}$.

# What changes with missing data?

Then you can uniquely identify $S^\star$.



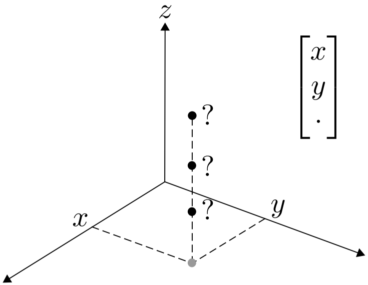
# What changes with missing data?

But what if data is missing?

# What changes with missing data?

But what if data is missing?

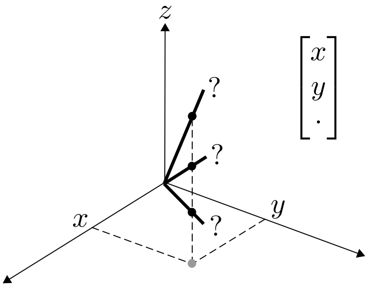
- Say I give you a point *without* the $z$ coordinate.



$$\begin{bmatrix} x \\ y \\ . \end{bmatrix}$$

# What changes with missing data?

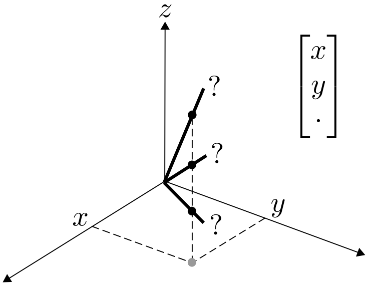
Then we cannot uniquely identify $S^\star$.



$$\begin{bmatrix} x \\ y \\ \cdot \end{bmatrix}$$

# What changes with missing data?

Then we cannot uniquely identify $S^\star$.



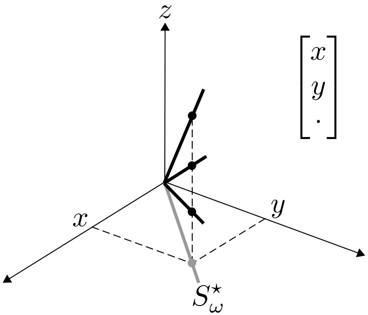
There are infinitely many *possible* subspaces.

# What changes with missing data?

Nevertheless, all those *possible* subspaces must satisfy one very important condition!

# What changes with missing data?

Nevertheless, all those *possible* subspaces must satisfy one very important condition!



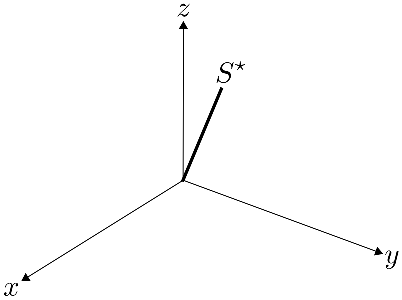
They must have the same canonical projection as $S^\star$.

# Outline

- Introduction ✓
- What changes with missing data? ✓
- Subspace Identifiability Problem
- Setup
- The Answer
- Application
- Conclusions

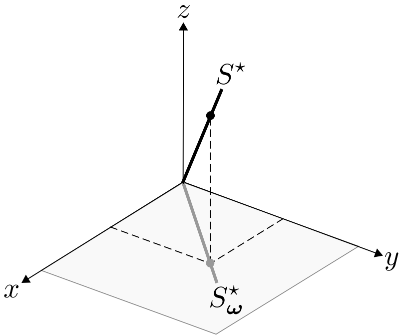
# Subspace Identifiability Problem

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



# Subspace Identifiability Problem

$S^\star_{\boldsymbol{\omega}} :=$ Canonical projection of $S^\star$.

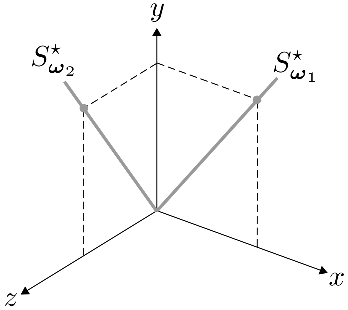


# Subspace Identifiability Problem

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

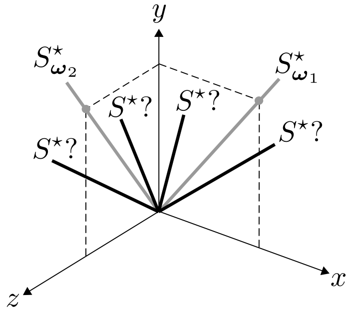
# Subspace Identifiability Problem

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



# Subspace Identifiability Problem

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

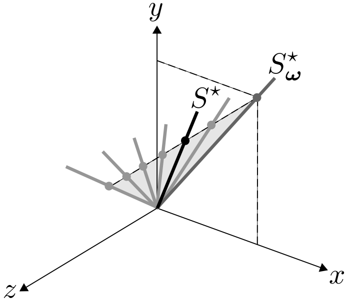


Can you uniquely determine $S^\star$ from this set of projections?

# Subspace Identifiability Problem

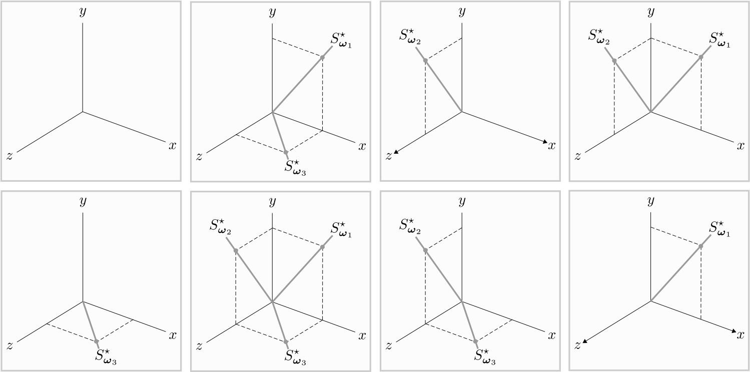
Is this even possible?

- There might be many subspaces that agree with the projections.



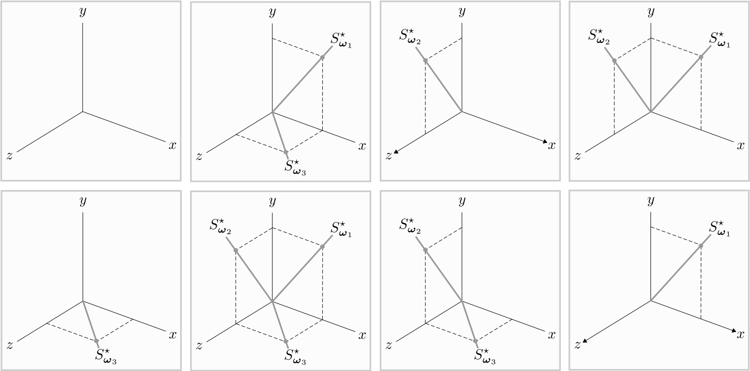
# Subspace Identifiability Problem

Well... sometimes you can, sometimes you can't.



# Subspace Identifiability Problem

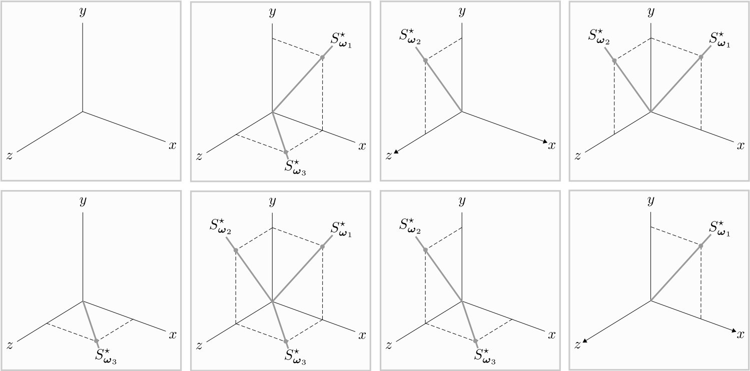
Well... sometimes you can, sometimes you can't.



Can you tell when?

# Subspace Identifiability Problem

Well... sometimes you can, sometimes you can't.



Can you tell when?

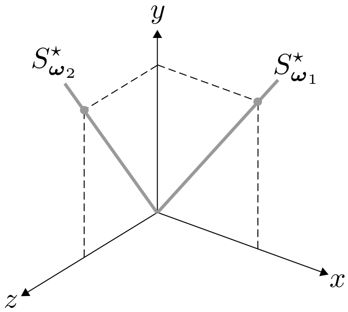
This is what we focused on: characterizing when can you identify $S^{\star}$ from its canonical projections.

# Outline

- Introduction ✓
- What changes with missing data? ✓
- Subspace Identifiability Problem ✓
- Setup
- The Answer
- Application
- Conclusions

# Setup

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



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

# Setup

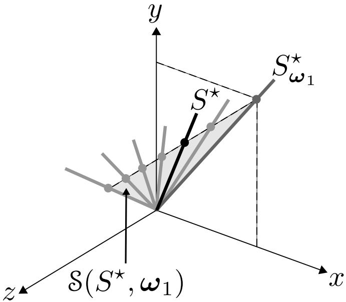
- $\mathrm{Gr}(r, \mathbb{R}^d) :=$ Grassmannian manifold of $r$-dimensional subspaces in $\mathbb{R}^d$.

# Setup

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

# 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

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

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

# Setup

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

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

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

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

# Outline

- Introduction ✓
- What changes with missing data? ✓
- Subspace Identifiability Problem ✓
- Setup ✓
- The Answer
- Application
- Conclusions

# The Answer

**Theorem (P.-A., Nowak, Boston, ISIT '15)**

*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.

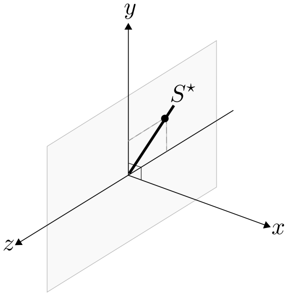
# The Answer

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 $\mathbf{\Omega}$,

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

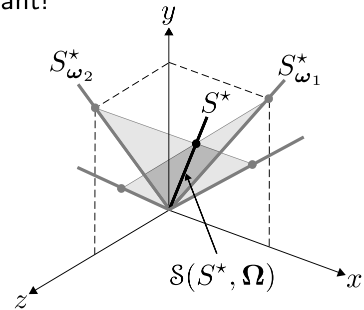
# The Answer

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 $\mathcal{S}(S^\star, \mathbf{\Omega})$ 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.

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

# Outline

- Introduction ✓
- What changes with missing data? ✓
- Subspace Identifiability Problem ✓
- Setup ✓
- The Answer ✓
- 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.

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

# Application

## Low-Rank Matrix Completion (LRMC)

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

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

- $\sim$ 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.

$$\mathbf{X}_{\boldsymbol{\Omega}} = \begin{bmatrix} 1 & \cdot & 3 & \cdot \\ 1 & 2 & \cdot & \cdot \\ \cdot & 2 & 3 & \cdot \\ \cdot & \cdot & \cdot & 4 \\ \cdot & \cdot & \cdot & 4 \end{bmatrix} \qquad \Rightarrow \qquad \mathbf{X} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

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

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

# 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:

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

# Application

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

- Maybe the real completion is:

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

And the real subspace is

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

# 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

Then we use our theorem!

- It can be used to validate the output of *any* LRMC algorithm!

# Application

Say we have:

$$\mathbf{X}_{\mathbf{\Omega}} = \begin{bmatrix} 1 & 2 & \cdot & \cdot & 2 & 1 & \cdot & \cdot & 2 & 2 \\ \cdot & \cdot & \cdot & 6 & \cdot & \cdot & 6 & \cdot & 4 & \cdot \\ 3 & 6 & 3 & \cdot & 6 & \cdot & 9 & 3 & \cdot & 6 \\ \cdot & \cdot & 1 & 3 & \cdot & 1 & \cdot & 1 & 2 & \cdot \\ 2 & 4 & \cdot & 6 & 4 & 2 & 6 & \cdot & 4 & 4 \\ \cdot & \cdot & 3 & \cdot & 6 & \cdot & \cdot & \cdot & \cdot & 6 \end{bmatrix}$$

# Application

Split $\mathbf{X}_{\mathbf{\Omega}}$ into:

$$\mathbf{X}_{\mathbf{\Omega}} = \left[ \underbrace{\begin{matrix} 1 & 2 & \cdot & \cdot & 2 & 1 & \cdot \\ \cdot & \cdot & \cdot & 6 & \cdot & \cdot & 6 \\ 3 & 6 & 3 & \cdot & 6 & \cdot & 9 \\ \cdot & \cdot & 1 & 3 & \cdot & 1 & \cdot \\ 2 & 4 & \cdot & 6 & 4 & 2 & 6 \\ \cdot & \cdot & 3 & \cdot & 6 & \cdot & \cdot \end{matrix}}_{\mathbf{X}_{\mathbf{\Omega}_1}} \; \underbrace{\begin{matrix} \cdot & 2 & 2 \\ \cdot & 4 & \cdot \\ 3 & \cdot & 6 \\ 1 & 2 & \cdot \\ \cdot & 4 & 4 \\ \cdot & \cdot & 6 \end{matrix}}_{\mathbf{X}_{\mathbf{\Omega}_2}} \right]$$

# Application

Split $\mathbf{X}_{\boldsymbol{\Omega}}$ into:

$$\mathbf{X}_{\boldsymbol{\Omega}} = \left[ \underbrace{\begin{matrix} 1 & 2 & \cdot & \cdot & 2 & 1 & \cdot \\ \cdot & \cdot & \cdot & 6 & \cdot & \cdot & 6 \\ 3 & 6 & 3 & \cdot & 6 & \cdot & 9 \\ \cdot & \cdot & 1 & 3 & \cdot & 1 & \cdot \\ 2 & 4 & \cdot & 6 & 4 & 2 & 6 \\ \cdot & \cdot & 3 & \cdot & 6 & \cdot & \cdot \end{matrix}}_{\mathbf{X}_{\boldsymbol{\Omega}_1}} \; \underbrace{\begin{matrix} \cdot & 2 & 2 \\ \cdot & 4 & \cdot \\ 3 & \cdot & 6 \\ 1 & 2 & \cdot \\ \cdot & 4 & 4 \\ \cdot & \cdot & 6 \end{matrix}}_{\mathbf{X}_{\boldsymbol{\Omega}_2}} \right]$$

Such that $\boldsymbol{\Omega}_2$ satisfies the subspace identifiability condition.

# Application

Use $\mathbf{X}_{\mathbf{\Omega}_1}$ to identify a candidate subspace $\widehat{S}$.

$$\left[\underbrace{\begin{matrix} 1 & 2 & \cdot & \cdot & 2 & 1 & \cdot \\ \cdot & \cdot & \cdot & 6 & \cdot & \cdot & 6 \\ 3 & 6 & 3 & \cdot & 6 & \cdot & 9 \\ \cdot & \cdot & 1 & 3 & \cdot & 1 & \cdot \\ 2 & 4 & \cdot & 6 & 4 & 2 & 6 \\ \cdot & \cdot & 3 & \cdot & 6 & \cdot & \cdot \end{matrix}}_{\mathbf{X}_{\mathbf{\Omega}_1}} \; \underbrace{\begin{matrix} \cdot & 2 & 2 \\ \cdot & 4 & \cdot \\ 3 & \cdot & 6 \\ 1 & 2 & \cdot \\ \cdot & 4 & 4 \\ \cdot & \cdot & 6 \end{matrix}}_{\mathbf{x}_{\mathbf{\Omega}_2}}\right]$$

$$\Downarrow$$

$$\widehat{S}$$

# Application

If $\widehat{S}$ is compatible with $\mathbf{X}_{\mathbf{\Omega}_2}$, then $\widehat{S} = S^\star$.

$$\left[\begin{array}{cc} \underbrace{\begin{matrix} 1 & 2 & \cdot & \cdot & 2 & 1 & \cdot \\ \cdot & \cdot & \cdot & 6 & \cdot & \cdot & 6 \\ 3 & 6 & 3 & \cdot & 6 & \cdot & 9 \\ \cdot & \cdot & 1 & 3 & \cdot & 1 & \cdot \\ 2 & 4 & \cdot & 6 & 4 & 2 & 6 \\ \cdot & \cdot & 3 & \cdot & 6 & \cdot & \cdot \end{matrix}}_{\mathbf{X}_{\mathbf{\Omega}_1}} & \underbrace{\begin{matrix} \cdot & 2 & 2 \\ \cdot & 4 & \cdot \\ 3 & \cdot & 6 \\ 1 & 2 & \cdot \\ \cdot & 4 & 4 \\ \cdot & \cdot & 6 \end{matrix}}_{\mathbf{X}_{\mathbf{\Omega}_2}} \end{array}\right]$$

# 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 ✓
- What changes with missing data? ✓
- Subspace Identifiability Problem ✓
- Setup ✓
- The Answer ✓
- 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.

Thanks.