A Characterization of Deterministic Sampling Patterns for Low-Rank Matrix Completion

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

University of Wisconsin-Madison

Allerton, 2015

Outline

- Introduction
- When can we Low-Rank Matrix Complete?
- The Answer
- Implications
- Idea of the proof
- Conclusions
- Open questions (if time allows)

Outline

- Introduction
- When can we Low-Rank Matrix Complete?
- The Answer
- Implications
- Idea of the proof
- Conclusions

Introduction

We have lots of data

Introduction

We have lots of data

And we want to analyze it.

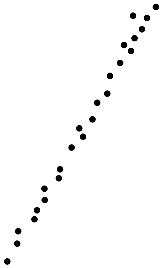
Introduction

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 subspaces.

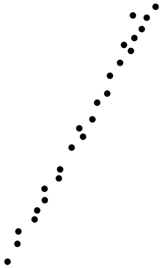


$$\begin{bmatrix} 1 & 2 & 1 & 3 & 2 & 1 & 3 & 1 & 2 & 2 \\ 2 & 4 & 2 & 6 & 4 & 2 & 6 & 2 & 4 & 4 \\ 3 & 6 & 3 & 9 & 6 & 3 & 9 & 3 & 6 & 6 \\ 1 & 2 & 1 & 3 & 2 & 1 & 3 & 1 & 2 & 2 \\ 2 & 4 & 2 & 6 & 4 & 2 & 6 & 2 & 4 & 4 \\ 3 & 6 & 3 & 9 & 6 & 3 & 9 & 3 & 6 & 6 \end{bmatrix}$$

Introduction

Linear Algebra is one of our favorite tools.

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



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

- We know how to find the subspace (e.g., using SVD).

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

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

- Example: Vision.

Image: Hopkins 155 Dataset

- We still want to find subspaces.

Introduction

Low-Rank Matrix Completion (LRMC) aims to find the subspace from incomplete datasets.

Introduction

Low-Rank Matrix Completion:

Introduction

Low-Rank Matrix Completion:

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

Introduction

Low-Rank Matrix Completion:

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

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

Introduction

Low-Rank Matrix Completion:

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

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

Introduction

Low-Rank Matrix Completion:

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

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

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

Introduction

Low-Rank Matrix Completion:

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

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

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

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

Outline

- Introduction ✓
- When can we Low-Rank Matrix Complete?
- The Answer
- Implications
- Idea of the proof
- Conclusions

When can we Low-Rank Matrix Complete?

- Existing theory (e.g. Candès and Recht, '09) essentially says:

When can we Low-Rank Matrix Complete?

- Existing theory (e.g. Candès and Recht, '09) essentially says:
  - If entries are observed uniformly at random.

When can we Low-Rank Matrix Complete?

- Existing theory (e.g. Candès and Recht, '09) essentially says:
  - If entries are observed uniformly at random.
    - May not be justified.

When can we Low-Rank Matrix Complete?

- Existing theory (e.g. Candès and Recht, '09) essentially says:
  - If entries are observed uniformly at random.
    - May not be justified.
  - If matrix is incoherent

When can we Low-Rank Matrix Complete?

- Existing theory (e.g. Candès and Recht, '09) essentially says:
  - If entries are observed uniformly at random.
    - May not be justified.
  - If matrix is incoherent
    - Sufficient, but not necessary condition.

When can we Low-Rank Matrix Complete?

- Existing theory (e.g. Candès and Recht, '09) essentially says:
  - If entries are observed uniformly at random.
    - May not be justified.
  - If matrix is incoherent
    - Sufficient, but not necessary condition.
    - Generally unverifiable or unjustified in practice.

When can we Low-Rank Matrix Complete?

- Existing theory (e.g. Candès and Recht, '09) essentially says:
  - If entries are observed uniformly at random.
    - May not be justified.
  - If matrix is incoherent
    - Sufficient, but not necessary condition.
    - Generally unverifiable or unjustified in practice.
  - Then with high probability you can complete the matrix.

When can we Low-Rank Matrix Complete?

- Existing theory (e.g. Candès and Recht, '09) essentially says:
  - If entries are observed uniformly at random.
    - May not be justified.
  - If matrix is incoherent
    - Sufficient, but not necessary condition.
    - Generally unverifiable or unjustified in practice.
  - Then with high probability you can complete the matrix.

  But what if these assumptions are not met?

When can we Low-Rank Matrix Complete?

- What makes a matrix *completable*?

When can we Low-Rank Matrix Complete?

- What makes a matrix *completable*?
- What conditions must a matrix satisfy?

Outline

- Introduction ✓
- When can we Low-Rank Matrix Complete? ✓
- The Answer
- Implications
- Idea of the proof
- Conclusions

The Answer

Setup

- $\mathbf{\Omega}$ will indicate the observed entries:

The Answer

Setup

- $\boldsymbol{\Omega}$ will indicate the observed entries:

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

The Answer

Setup

- $\mathbf{\Omega}$ will indicate the observed entries:

$$\mathbf{X}_{\mathbf{\Omega}} = \begin{bmatrix} 1 & \cdot & 3 & \cdot \\ 1 & 2 & \cdot & \cdot \\ \cdot & 2 & 3 & \cdot \\ \cdot & \cdot & \cdot & 4 \\ \cdot & \cdot & \cdot & 4 \end{bmatrix} \qquad \mathbf{\Omega} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The Answer

Setup

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

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

The Answer

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

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\mathbf{\Omega}}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\mathbf{\Omega}}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\mathbf{\Omega}}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

There is a set of measure zero of bad matrices for which this theorem does not apply.

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\mathbf{\Omega}}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

There is a set of measure zero of bad matrices for which this theorem does not apply.

$$\mathbf{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\mathbf{\Omega}}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

There is a set of measure zero of bad matrices for which this theorem does not apply.

$$\mathbf{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \mathbf{X}_{\mathbf{\Omega}} = \begin{bmatrix} 0 & \cdot & \cdot \\ 0 & 0 & \cdot \\ \cdot & 0 & 0 \\ \cdot & \cdot & 0 \end{bmatrix}$$

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\mathbf{\Omega}}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\mathbf{\Omega}}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

$$\mathbf{X}_{\mathbf{\Omega}} = \begin{bmatrix} 1 & 1 & 3 & \cdot \\ 1 & 2 & \cdot & 1 \\ 3 & \cdot & 5 & 4 \\ \cdot & 7 & 6 & 5 \end{bmatrix}$$

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X_\Omega}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

$$\mathbf{X_\Omega} = \begin{bmatrix} 1 & 1 & 3 & \cdot \\ 1 & 2 & \cdot & 1 \\ 3 & \cdot & 5 & 4 \\ \cdot & 7 & 6 & 5 \end{bmatrix} \qquad \Rightarrow \begin{cases} \mathbf{X} = \begin{bmatrix} 1 & 1 & 3 & 2 \\ 1 & 2 & 1 & 1 \\ 3 & 5 & 5 & 4 \\ 4 & 7 & 6 & 5 \end{bmatrix} \\ \\ \mathbf{X} = \begin{bmatrix} 1 & 1 & 3 & 10 \\ 1 & 2 & \frac{21}{13} & 1 \\ 3 & \frac{53}{9} & 5 & 4 \\ \frac{68}{19} & 7 & 6 & 5 \end{bmatrix} \end{cases}$$

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\mathbf{\Omega}}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\Omega}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

This is the answer!

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\boldsymbol{\Omega}}$ if and only if every matrix $\boldsymbol{\Omega}'$ formed with a subset of the columns in $\boldsymbol{\Omega}$ satisfies

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

This is the answer!

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

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_{\Omega}$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

This is the answer!

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

$$\mathbf{\Omega} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The Answer

For almost every $\mathbf{X}$, there exist at most finitely many rank-r completions of $\mathbf{X}_\Omega$ if and only if every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\mathbf{\Omega}$ satisfies

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

This is the answer!

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

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

The answer

Now we know when there are at most finitely many completions.

- Then what?

The Answer

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

If in addition $\mathbf{X}_{\mathbf{\Omega}}$ has an extra $(d-r)$ columns observed on $\hat{\mathbf{\Omega}}$, such that every matrix $\mathbf{\Omega}'$ formed with a subset of the columns in $\hat{\mathbf{\Omega}}$ satisfies

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

then $\mathbf{X}$ can be uniquely recovered from $\mathbf{X}_{\mathbf{\Omega}}$.

The Answer (in words)

If a matrix does not satisfy our sampling conditions, then you cannot complete it.

The Answer (in words)

If a matrix does not satisfy our sampling conditions, then you cannot complete it.

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

The Answer (in words)

If a matrix satisfies our sampling conditions, then you can finitely complete it.

The Answer (in words)

If a matrix satisfies our sampling conditions, then you can finitely complete it.

$$\mathbf{X}_{\mathbf{\Omega}} = \begin{bmatrix} 1 & 1 & 3 & \cdot \\ 1 & 2 & \cdot & 1 \\ 3 & \cdot & 5 & 4 \\ \cdot & 7 & 6 & 5 \end{bmatrix}$$

The Answer (in words)

If a matrix satisfies our sampling conditions, then you can finitely complete it.

$$\mathbf{X}_{\boldsymbol{\Omega}} = \begin{bmatrix} 1 & 1 & 3 & \cdot \\ 1 & 2 & \cdot & 1 \\ 3 & \cdot & 5 & 4 \\ \cdot & 7 & 6 & 5 \end{bmatrix}$$

Sometimes finitely completable = uniquely completable (e.g., rank$= 1$), but sometimes not.

The Answer (in words)

But just a few additional samples in a finitely completable matrix make it uniquely completable.

The Answer (in words)

But just a few additional samples in a finitely completable matrix make it uniquely completable.

$$\mathbf{X}_{\mathbf{\Omega}} = \begin{bmatrix} 1 & 1 & 3 & \cdot & -1 & 1 \\ 1 & 2 & \cdot & 1 & \cdot & -1 \\ 3 & \cdot & 5 & 4 & 3 & \cdot \\ \cdot & 7 & 6 & 5 & 5 & -2 \end{bmatrix}$$

The Answer (take home message)

In essence:

r **complete** columns (linearly independent) uniquely define an r-dimensional subspace.

The Answer (take home message)

In essence:

r **complete** columns (linearly independent) uniquely define an r-dimensional subspace.

$(r+1)(d-r)$ **incomplete** columns (observed in the right places) uniquely define an r-dimensional subspace.

Outline

- Introduction ✓
- When can we Low-Rank Matrix Complete? ✓
- The Answer ✓
- Implications
- Idea of the proof
- Conclusions

Implications (coherence)

- P.-A., Nowak, Boston (Allerton '15):
  - For almost every matrix, $\mathcal{O}(\max\{r, \log d\})$ uniform random entries per column are sufficient for completion.

Implications (coherence)

- P.-A., Nowak, Boston (Allerton '15):
  - For almost every matrix, $\mathcal{O}(\max\{r, \log d\})$ uniform random entries per column are sufficient for completion.
- Regardless of coherence!

Implications

- Our results tell us exactly which entries to observe.

Implications

- Our results tell us exactly which entries to observe.
  - We can now design Adaptive LRMC Algorithms.

Implications

- Our results tell us exactly which entries to observe.
  - We can now design Adaptive LRMC Algorithms.
- Help answer an important open question:

Implications

- Our results tell us exactly which entries to observe.
  - We can now design Adaptive LRMC Algorithms.
- Help answer an important open question:
  - The Sample Complexity of Subspace Clustering with Missing Data.

Implications

Validation criteria:

Implications

Validation criteria:

- Suppose you observe the right entries.

Implications

Validation criteria:

- Suppose you observe the right entries.
- Try to complete the matrix using any method.

Implications

Validation criteria:

- Suppose you observe the right entries.
- Try to complete the matrix using any method.
- If you find a rank-r completion, then it is the right completion.

Implications

Validation criteria:

- Suppose you observe the right entries.
- Try to complete the matrix using any method.
- If you find a rank-r completion, then it is the right completion.
- In lieu of coherence assumptions.

Implications

Validation criteria:

- Suppose you observe the right entries.
- Try to complete the matrix using any method.
- If you find a rank-r completion, then it is the right completion.
- In lieu of coherence assumptions.
- In lieu of uniform sampling assumptions.

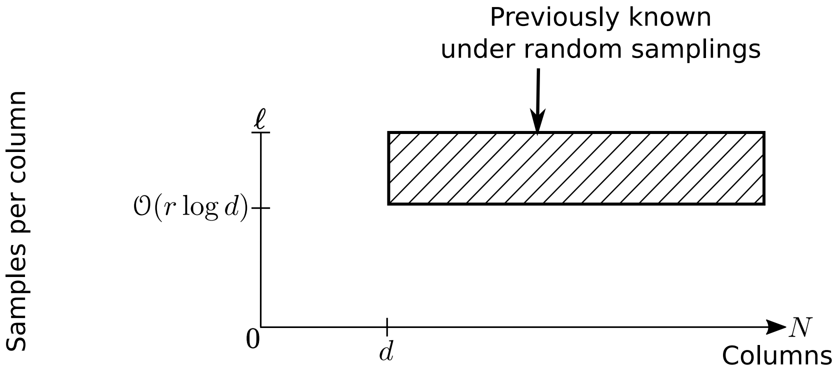
Implications

Validation criteria:

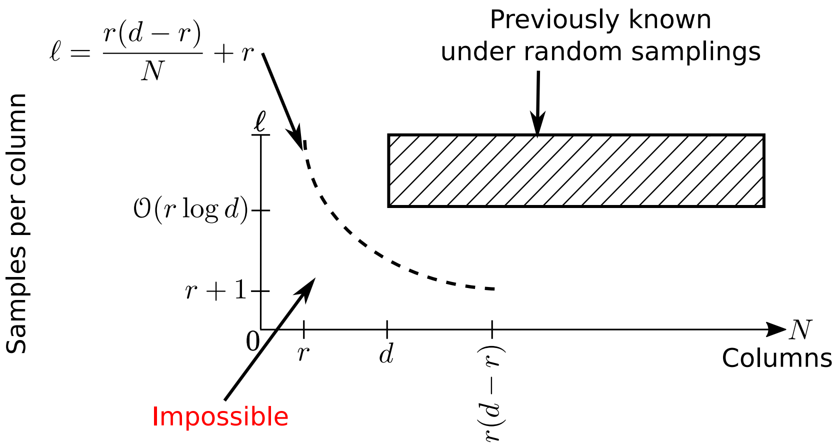
- Suppose you observe the right entries.
- Try to complete the matrix using any method.
- If you find a rank-r completion, then it is the right completion.
- In lieu of coherence assumptions.
- In lieu of uniform sampling assumptions.
- With probability 1 (as opposed to *with high probability*).

Implications: better understanding of sampling regimes

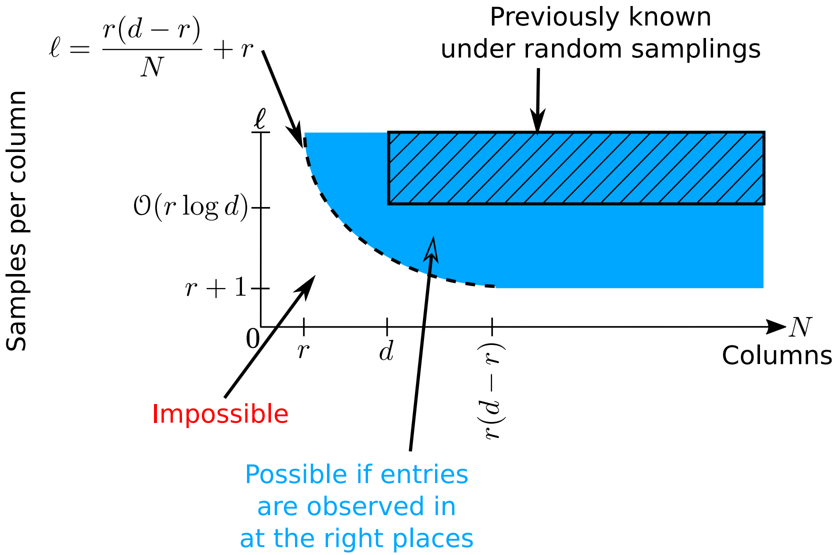
Implications: better understanding of sampling regimes



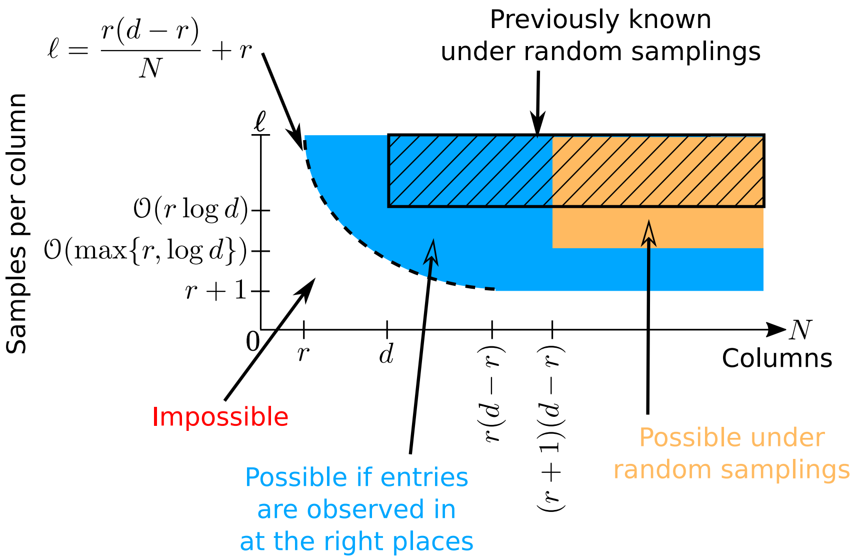
Implications: better understanding of sampling regimes



Implications: better understanding of sampling regimes



Implications: better understanding of sampling regimes



Outline

- Introduction ✓
- When can we Low-Rank Matrix Complete? ✓
- The Answer ✓
- Implications ✓
- Idea of the proof
- Conclusions

Idea of the proof

A column with $r + 1$ samples imposes one restriction on what the subspace may be.

Idea of the proof

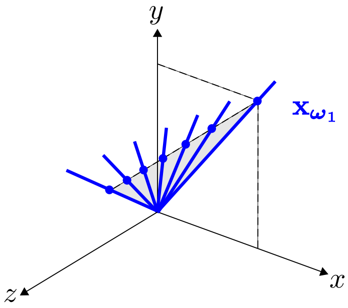
A column with $r+1$ samples imposes one restriction on what the subspace may be.

$$\begin{array}{cc} & \mathbf{x}_{\boldsymbol{\omega}_1} \\ \mathbf{X}_{\boldsymbol{\Omega}} \quad = & \left[\begin{array}{c} 1 \\ 1 \\ \cdot \end{array}\right] \end{array}$$

Idea of the proof

A column with $r+1$ samples imposes one restriction on what the subspace may be.

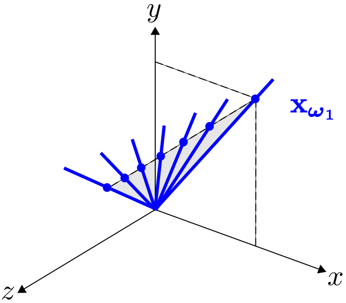
$$\mathbf{X}_{\mathbf{\Omega}} = \begin{array}{c} \mathbf{x}_{\boldsymbol{\omega}_1} \\ \begin{bmatrix} 1 \\ 1 \\ \cdot \end{bmatrix} \end{array}$$



Idea of the proof

A column with $r+1$ samples imposes one restriction on what the subspace may be.

$$\mathbf{X}_{\boldsymbol{\Omega}} = \begin{array}{c} \mathbf{x}_{\boldsymbol{\omega}_1} \\ \begin{bmatrix} 1 \\ 1 \\ \cdot \end{bmatrix} \end{array}$$

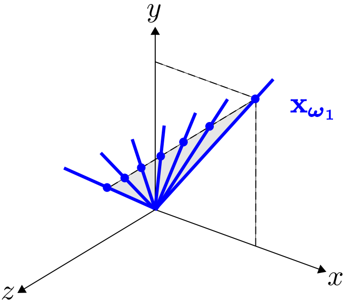


- A subspace S *fits* $\mathbf{x}_{\boldsymbol{\omega}_1} \iff f_1(S) = 0$.

Idea of the proof

A column with $r+1$ samples imposes one restriction on what the subspace may be.

$$\mathbf{X}_{\mathbf{\Omega}} \quad = \quad \begin{array}{c} \mathbf{x}_{\boldsymbol{\omega}_1} \\ \left[\begin{array}{c} 1 \\ 1 \\ \cdot \end{array}\right] \end{array}$$



- A subspace S *fits* $\mathbf{x}_{\boldsymbol{\omega}_1} \iff f_1(S) = 0$.
- This reduces one degree of freedom in the Grassmannian.

Idea of the proof

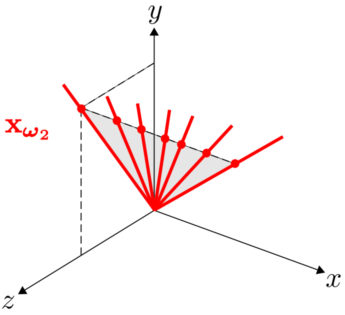
An other column with $r+1$ samples imposes an other restriction.

$$\mathbf{X}_{\boldsymbol{\Omega}} \quad = \quad \begin{array}{c} \mathbf{x}_{\boldsymbol{\omega}_2} \\ \left[\begin{array}{c} \cdot \\ 2 \\ 2 \end{array}\right] \end{array}$$

Idea of the proof

An other column with $r+1$ samples imposes an other restriction.

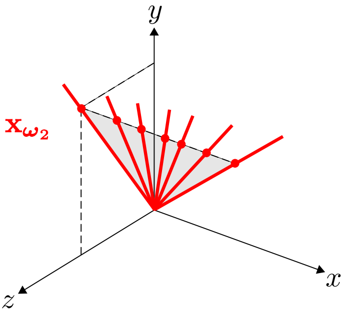
$$\mathbf{X}_{\mathbf{\Omega}} = \begin{array}{c} \mathbf{x}_{\boldsymbol{\omega}_2} \\ \begin{bmatrix} \cdot \\ 2 \\ 2 \end{bmatrix} \end{array}$$



Idea of the proof

An other column with $r+1$ samples imposes an other restriction.

$$\mathbf{X}_{\boldsymbol{\Omega}} = \begin{array}{c} \mathbf{x}_{\boldsymbol{\omega}_2} \\ \left[\begin{array}{c} \cdot \\ 2 \\ 2 \end{array}\right] \end{array}$$



- A subspace S *fits* $\mathbf{x}_{\boldsymbol{\omega}_2}$ $\iff$ $f_2(S) = 0$.

Idea of the proof

Each column with $r+1$ samples imposes one restriction.

Idea of the proof

Each column with $r+1$ samples imposes one restriction.

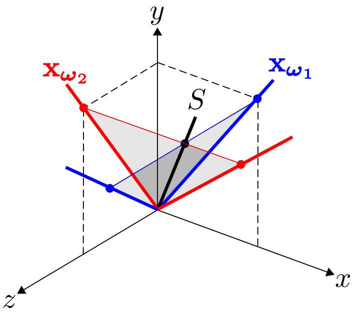
$$\mathbf{X}_{\mathbf{\Omega}} \quad = \quad \begin{array}{c} \begin{array}{cc} \mathbf{x}_{\boldsymbol{\omega}_1} & \mathbf{x}_{\boldsymbol{\omega}_2} \end{array} \\ \begin{bmatrix} 1 & \cdot \\ 1 & 2 \\ \cdot & 2 \end{bmatrix} \end{array}$$

Idea of the proof

Each column with $r+1$ samples imposes one restriction.

$$\mathbf{X}_{\boldsymbol{\Omega}} \quad = \quad \begin{array}{cc} \mathbf{x}_{\boldsymbol{\omega}_1} & \mathbf{x}_{\boldsymbol{\omega}_2} \\ \end{array}$$

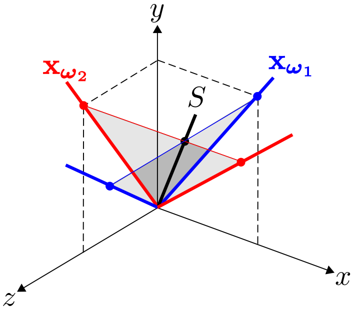
$$\left[\begin{array}{cc} 1 & \cdot \\ 1 & 2 \\ \cdot & 2 \end{array} \right]$$



Idea of the proof

Each column with $r+1$ samples imposes one restriction.

$$
\begin{array}{cc}
 & \begin{array}{cc} \mathbf{x}_{\omega_1} & \mathbf{x}_{\omega_2} \end{array} \\
\mathbf{X}_{\Omega} \;=\; & \left[\begin{array}{cc} 1 & \cdot \\ 1 & 2 \\ \cdot & 2 \end{array}\right]
\end{array}
$$



- A subspace S *fits* $\mathbf{X}_{\Omega}$ $\iff \left\{ \begin{array}{l} f_1(S) = 0 \\ f_2(S) = 0 \end{array} \right.$.

Idea of the proof

- Each column with $r + 1$ samples imposes one restriction:

$$f_1, f_2, \ldots, f_N.$$

Idea of the proof

- Each column with $r+1$ samples imposes one restriction:

$$f_1, f_2, \ldots, f_N.$$

- The Grassmannian has $r(d-r)$ degrees of freedom.

Idea of the proof

- Each column with $r+1$ samples imposes one restriction:

$$f_1, f_2, \ldots, f_N.$$

- The Grassmannian has $r(d-r)$ degrees of freedom.
- If we have $r(d-r)$ not redundant restrictions:

Idea of the proof

- Each column with $r+1$ samples imposes one restriction:

$$f_1, f_2, \ldots, f_N.$$

- The Grassmannian has $r(d-r)$ degrees of freedom.
- If we have $r(d-r)$ not redundant restrictions:
  - We can identify $S^\star$ up to finite choice.

Idea of the proof

- $f_i(S)$ only involves the variables (of S) corresponding to the nonzero rows of $\boldsymbol{\omega}_i$.

Idea of the proof

- $f_i(S)$ only involves the variables (of S) corresponding to the nonzero rows of $\boldsymbol{\omega}_i$.
- We want all sets of n polynomials to involve at least $n/r + r$ variables (otherwise they will be dependent)

Outline

- Introduction ✓
- When can we Low-Rank Matrix Complete? ✓
- The Answer ✓
- Implications ✓
- Idea of the proof ✓
- Conclusions

Conclusions

- In essence:

Conclusions

- In essence:

Our results tell us when will a set of incomplete vectors uniquely define an r-dimensional subspace (just as a set of r linearly independent complete vectors would).

Conclusions

- In essence:

Our results tell us when will a set of incomplete vectors uniquely define an r-dimensional subspace (just as a set of r linearly independent complete vectors would).

- This sheds new light on LRMC.

Thanks.

Outline

- Introduction ✓
- When can we Low-Rank Matrix Complete? ✓
- The Answer ✓
- Implications ✓
- Idea of the proof ✓
- Conclusions ✓
- Open questions

Open questions

It is one thing to be theoretically able to complete a matrix; an other one to complete it efficiently.

Open questions

It is one thing to be theoretically able to complete a matrix; an other one to complete it efficiently.

- P.-A., Nowak, Boston (Allerton '15):

Open questions

> It is one thing to be theoretically able to complete a matrix; an other one to complete it efficiently.

- P.-A., Nowak, Boston (Allerton '15):
  - New sampling regimes where you can theoretically complete a matrix.

Open questions

It is one thing to be theoretically able to complete a matrix; an other one to complete it efficiently.

- P.-A., Nowak, Boston (Allerton '15):
  - New sampling regimes where you can theoretically complete a matrix.
  - This may involve solving a complex system of polynomial equations!

Open questions

It is one thing to be theoretically able to complete a matrix; an other one to complete it efficiently.

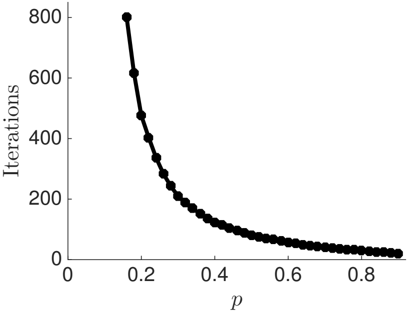
- P.-A., Nowak, Boston (Allerton '15):
  - New sampling regimes where you can theoretically complete a matrix.
  - This may involve solving a complex system of polynomial equations!
  - This is computationally prohibitive.

Open questions

Does missingness come at a price?

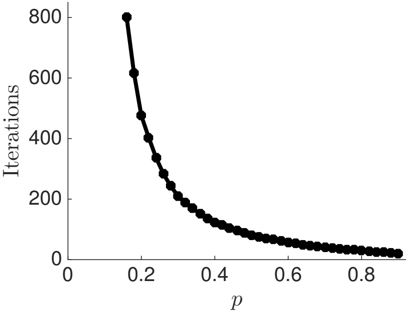
Open questions

Does missingness come at a price?



Open questions

Does missingness come at a price?



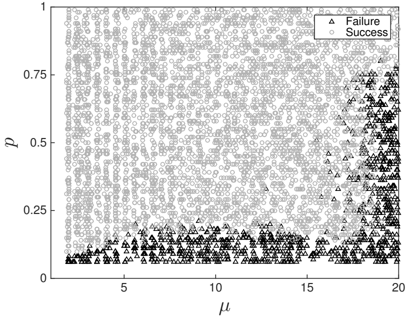
How much missing data can we handle and remain computationally efficient?

Open questions

Can practical algorithms complete coherent matrices?

Open questions

Can practical algorithms complete coherent matrices?

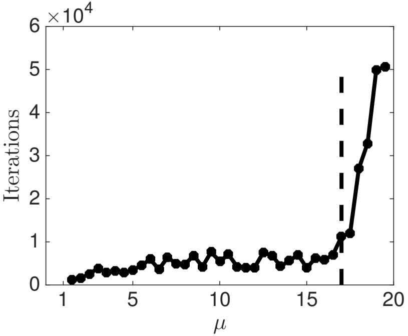


Open questions

Does coherence come at a price?

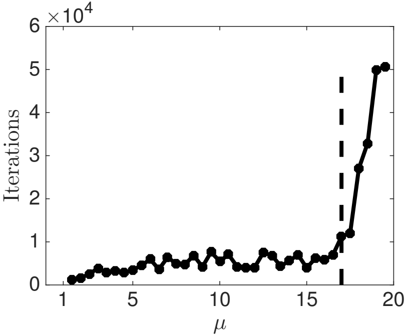
Open questions

Does coherence come at a price?



Open questions

Does coherence come at a price?



How much coherence can we handle and remain computationally efficient?

Thanks again!

Thanks again!

(this time I'm really done)