Matrix Completion

Franz J. Király\textsuperscript{12}, Louis Theran\textsuperscript{2}, Ryota Tomioka\textsuperscript{3}

\textsuperscript{1} Machine Learning Group
Technische Universität Berlin

\textsuperscript{2} Discrete Geometry Group
Freie Universität Berlin

\textsuperscript{3} Department of Mathematical Informatics
The University of Tokyo
In this talk:

Combinatorial Algebraic Methods

for understanding

Matrix Completion

yielding

Identifiability Criteria and Competitive Algorithms
What is Matrix Completion?

**Generative model:**
- low rank matrix $A$ (unknown)
- entries are random values

Matrix completion
- **Input:** some entries of $A$ plus noise
  - rank of $A$
- **Output:** $A$

$$
\begin{pmatrix}
1 & 2 & 3 \\
3 & 6 & 9 \\
2 & 4 & 6
\end{pmatrix}
$$
Matrix Completion

**Definition:** masking

\[
\Omega : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{\alpha}
\]

\[
\begin{pmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{pmatrix}
\rightarrow
\begin{pmatrix}
  \cdot & a_{i_1j_1} & \cdot & \cdot \\
  \cdot & \cdot & a_{i_2j_2} & \cdot \\
  \cdot & a_{i_\alpha j_\alpha} & \cdot & \cdot \\
  \cdot & \cdot & \cdot & \cdot
\end{pmatrix}
\]

\[
E(\Omega) = \{(i_k, j_k) ; 1 \leq k \leq \alpha\} \quad \#E(\Omega) = \alpha
\]

**Question:** Given a masking \( \Omega \) and \( A \in \mathbb{C}^{m \times n} \) of rank \( r \), when can \( A \) be uniquely reconstructed from \( \Omega(A) \)?
Matrix Completion

**Question:** Given a masking $\Omega$ and $A \in \mathbb{C}^{m \times n}$ of rank $r$, when can $A$ be uniquely reconstructed from $\Omega(A)$?

**Definition:** $\mathcal{M} = \{ A \in \mathbb{C}^{m \times n} ; \text{rk} A \leq r \}$

**Note:** $\mathcal{M}$ is an algebraic variety, so-called determinantal variety because $\mathcal{M}$ is the solution set of all $(r+1) \times (r+1)$ minor equations $\det \ldots = 0$

**Sub-question:** When is $\Omega : \mathcal{M} \to \mathbb{C}^\alpha$ injective?
Matrix Completion

**Sub-question:** When is $\Omega : \mathcal{M} \to \mathbb{C}^\alpha$ injective?

$\mathcal{M} =$ set of matrices of rank $\leq r$

**Example:** $m = n = 2, r = 1$

$\Omega : \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

$a_{22} = \frac{a_{12}a_{21}}{a_{11}}$ undefined if $a_{11} = 0$

**General answer:** Only if $\#E(\Omega) = mn$.

**Proof:** Assume $\#E(\Omega) < mn$.
Take some non-zero matrix $A$ such that $\Omega(A) = 0$.
Then, $\lambda A \in \Omega^{-1}(\Omega(A))$ for all $\lambda \in \mathbb{C}$. 
The right question to ask

**Bad question:** When is $\Omega : \mathcal{M} \to \mathbb{C}^\alpha$ injective?

**Answer:** Only if $\# E(\Omega) = mn$.

But what if we allow a zero set of exceptions?

**Definition:** $\Omega$ is called **generically injective** or **identifiable** if $\Omega$ is injective when restricted to almost all matrices

i.e. the fiber $\Omega^{-1}(\Omega(A)) = \{A\}$ for $A \in U$, with $U$ dense in $\mathcal{M}$

($= \mathcal{M} \setminus U$ has zero measure)

**Definition:** $\Omega$ is called **generically finite** if $\Omega$ is $n$-to-one when restricted to almost all matrices
Parameterizing the masking

Better question: When is $\Omega : \mathcal{M} \to \mathbb{C}^\alpha$ generically injective?

Definition: mask of

$$
\Omega : \begin{pmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} \\
    a_{21} & a_{22} & \cdots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mn}
\end{pmatrix} \rightarrow \begin{pmatrix}
    \cdot & a_{i_1 j_1} & \cdot & \cdot \\
    \cdot & \cdot & a_{i_2 j_2} & \cdot \\
    \cdot & a_{i_{\alpha_1} j_{\alpha}} & \cdot & \cdot
\end{pmatrix}
$$

is matrix $M(\Omega) \in \{0, 1\}^{m \times n}$

such that $\Omega : A \mapsto M(\Omega) \circ A$

e.g. $M(\Omega) = \begin{pmatrix}
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & \vdots & \vdots & 0 \\
    0 & 1 & 0 & 0
\end{pmatrix}$
When can almost all matrices be reconstructed?

**Theorem:**
For a generic matrix $A$, the dimension, and cardinality of the fiber $\Omega^{-1}(\Omega(A))$ (=the set of possible reconstructions) depends only on $m, n, r$ and the mask $M(\Omega)$. Moreover, that dimension is minimal for generic $A$.

**Corollary:**
Whether $\Omega$ is generically injective depends only on $m, n, r$ and the mask $M(\Omega)$.

(generic low-rank)

reconstructability of a matrix depends only on mask! (= set of known entries)

with probability one, i.e., pathologies occur almost never
Intermezzo: Dimension

Let $X$ be an algebraic variety. (= e.g., subset of $\mathbb{C}^n$ defined by polynomial equations)

Let $P$ be a point on $X$.

$\dim_P X =$ degrees of freedom to move $P$ on $X$

$= \text{dimension of tangent space of } X \text{ at } P$

$\dim X = \dim_P X$ for generic/almost all $P$

(if ambiguous, take max)
The upper semicontinuity theorem

Upper semicontinuity theorem (Chevalley, Grothendieck): Let $\mathcal{M}$ be a variety, let $\mathcal{M} \to X$ be an algebraic map which is locally of finite type. Then, the map $A \mapsto \dim_A \Omega^{-1}(\Omega(A))$ is upper semicontinuous w.r.t. the Zariski topology.

Intuitively: Any randomly chosen $A$ has the same fiber dimension. Restricting the set from which you choose $A$ randomly by algebraic constraints can make the fiber dimension only bigger or equal.
When can almost all matrices be reconstructed?

**Theorem:**
For a generic matrix $A$, the dimension, and cardinality of the fiber $\Omega^{-1}(\Omega(A))$ (=the set of possible reconstructions) depends only on $m, n, r$ and the mask $M(\Omega)$. Moreover, that dimension is minimal for generic $A$.

**Corollary:**
Whether $\Omega$ is generically injective depends only on $m, n, r$ and the mask $M(\Omega)$.

**Proof:** $\Omega : \mathcal{M} \rightarrow \mathbb{C}^\alpha$ is algebraic.
$\mathcal{M}$ is irreducible.
Apply upper semicontinuity.
Bad and good masks

Two possibilities:

mask is identifiable

\[
\begin{pmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{pmatrix}
\]

ea consistent estimator for true matrix exists:

Approximate Algebra

mask is not identifiable

\[
\begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}
\]

no algorithm can identify true matrix

(except, by chance)

(i.e., “prophetic” choice of prior)
Which masks are identifiable?

**Theorem:**
\( \Omega \) is generically finite only if \( \#E(\Omega) \geq r(m + n - r) \)

**Proof:** \( \Omega : \mathcal{M} \rightarrow \mathbb{C}^\alpha \) is algebraic

\[
\dim \Omega^{-1}(\Omega(A)) = \dim \mathcal{M} - \dim \Omega(\mathcal{M})
\]

for generic \( A \) (generic fiber theorem)

\[
\dim \Omega(\mathcal{M}) \leq \#E(\Omega) \quad \dim \mathcal{M} = r(m + n - r)
\]

Thus \( \dim \Omega^{-1}(\Omega(A)) \geq r(m + n - r) - \#E(\Omega) \)

\( \Omega \) is generically finite if and only if \( \dim \Omega^{-1}(\Omega(A)) = 0 \).
Which masks are identifiable?

**Theorem:**
\[ \Omega \text{ is generically finite only if } \#E(\Omega) \geq r(m + n - r) \]

This bound is not strict!

**Example:** \( m = n = 3, r = 1 \)

- Identifiable:
  \[
  \begin{pmatrix}
  1 & 0 & 1 \\
  1 & 1 & 0 \\
  1 & 0 & 0 \\
  \end{pmatrix}
  \quad \begin{pmatrix}
  1 & 1 & 0 \\
  1 & 0 & 1 \\
  1 & 1 & 0 \\
  \end{pmatrix}
  \]

- Not identifiable:
  \[
  \begin{pmatrix}
  1 & 0 & 1 \\
  0 & 1 & 0 \\
  1 & 0 & 1 \\
  \end{pmatrix}
  \quad \begin{pmatrix}
  1 & 1 & 0 \\
  1 & 1 & 0 \\
  1 & 1 & 0 \\
  \end{pmatrix}
  \]
Necessary conditions

\[
\begin{pmatrix}
1 & 0 & 1 \\
1 & 1 & 0 \\
1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
1 & 1 & 0 \\
1 & 0 & 1 \\
1 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 0 \\
1 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 1 & 0 \\
1 & 1 & 0 \\
1 & 1 & 0
\end{pmatrix}
\]

\[\Rightarrow\text{Additional information on } \Omega \text{ is needed}\]

**Definition:**
mask defines bipartite graph \( G(\Omega) \)

\[
\begin{pmatrix}
1 & 0 & 1 \\
1 & 1 & 0 \\
1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 0 \\
1 & 0 & 1
\end{pmatrix}
\]

columns
rows

By **Theorem**, all information is encoded in \( G(\Omega) \)
Necessary conditions

Definition:
A graph is \(r\)-(edge)-connected if, after removing arbitrary \(r - 1\) edges, the graph is still connected.

Theorem:
\(\Omega\) is generically finite only if \(G(\Omega)\) is \(r\)-connected.

Sketch of proof:
If \(G(\Omega)\) is not \(r\)-connected, then
\[
M(\Omega) = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} + M' > r - (r - 1) \text{ degrees of freedom}
\]

\[
\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} < r \text{ non-zero entries}
\]

\[
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}
\]
Sufficient conditions

**Theorem:**
\( \Omega \) is generically injective if \( G(\Omega) \) is \( r \)-closable.

**Intuitive Algebraic Definition:**
\[
\begin{pmatrix}
\times & \times & \times \\
\times & \times & \times \\
\times & \times & \times \\
\end{pmatrix}
\]
solve \( \det A = 0 \) for \( a_{33} \)

**Semi-formal Graph-Theoretic Definition:**
Let \( K_{r+1,r+1}^- \) be the complete bipartite graph minus one edge. 
\( r \)-closure of a graph is obtained by repeatedly replacing the induced structure of \( K_{r+1,r+1}^- \) with \( K_{r+1,r+1} \) inside the graph.

If the \( r \)-closure of \( G(\Omega) \) is \( K_{m,n} \) then \( G(\Omega) \) is called \( r \)-closable.
Sufficient conditions

Theorem:
For $r = 1$ and $r = m - 1$, $r$-connectedness = $r$-closability
but not in general!

Proof:
$r = 1$
1-closable = connected = 1-connected

$r = m - 1$
gen.injective = at most 1 entry per column missing
⇒ $(m - 1)$-closable

$r = 2$

\[
\begin{pmatrix}
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 \\
1 & 1 & 1 & 0
\end{pmatrix}
\]
is a counterexample
i.e., finitely completable but not 2-closable
Matroidal Property

**Theorem:**
For every $r$, there exists a minimal family of “forbidden subgraphs"

$$\mathcal{C}_r = \{ G_1^{(r)}, G_2^{(r)}, \ldots, \}$$

s.t. $\Omega$ with $E(\Omega) = r(m + n - r)$ is generically finite if and only if $G(\Omega)$ contains no graph in $\mathcal{C}_r$

Moreover, each $G_i^{(r)}$ is $(r + 1)$-connected

(implies degrees $\geq r + 1$)

**Example:**

$$\mathcal{C}_1 = \{ \text{cycles of even length} \}$$

$= \{ \begin{array}{c} \begin{array}{c} \text{cycles of even length} \\ \end{array} \\ \end{array} \}$

$$K_{r+1,r+1} \in \mathcal{C}_r$$
Asymptotics and Probabilistics

Central question:
Take a random mask $M(\Omega) \in \{0, 1\}^{m \times n}$ where we uniformly, randomly allocate $k$ ones and $mn - k$ zeros. (= random observations in a low-rank model)

How probable is it for $\Omega$ to be generically identifiable/finite?

Previous results: (symmetric case $m = n$)

w.h.p, $\Omega$ is generically finite if

$k = O(n \log n)$ for incoherent input matrices

(Keshavan, Oh, Montanari, 2009) (Candès, Tao, 2009)

$k = O(n \log n)$ for generic matrices, estimated by numerics

(Singer, Cucuringu, 2009)

w.h.p, $\Omega$ is generically identifiable if

$k = O(n \log^2 n)$ for incoherent input matrices

(Candès, Tao, 2009)
Central question:
Take a random mask \( M(\Omega) \in \{0, 1\}^{m \times n} \) where we uniformly, randomly allocate \( k \) ones and \( mn - k \) zeros. (= random observations in a low-rank model)

How probable is it for \( \Omega \) to be generically identifiable/finite?

Theorem: (square case \( m = n \))

Exact threshold for \( \Omega \) being generically finite is

\[
    k = \Theta(n \log n)
\]

without conditioning on input matrix!

Proof: upper bounds and necessity of connectedness incoherence assumption can be removed
Phase transition is due to a shift in combinatorial properties of graph!

w.h.p., $G(\Omega)$ is connected, 2-connected, 3-connected, identifiable, 3-closable.
Approximate Algebra for Matrix Completion

How to use Algebra in noisy estimation?

Approximate Algebra provides tools to cope with noise.

Sketch of algorithm: (rank 1 example)

1. Find equations of form $a_{k\ell} = \frac{a_{i\ell}a_{kj}}{a_{ij}} = f(\text{known entries})$

2. \[
\begin{pmatrix}
  a_{ij} & \cdots & a_{kj} \\
  \vdots & \ddots & \vdots \\
  a_{i\ell} & \cdots & a_{k\ell}
\end{pmatrix}
\quad \text{approximately} \quad
\begin{pmatrix}
  a_{ij} & \cdots & a_{kj} \\
  \vdots & \ddots & \vdots \\
  a_{i\ell} & \cdots & a_{k\ell}
\end{pmatrix}
\]
Approximate Combinatorial Algebra yields competitive Matrix Completion method
Summary and Outlook

Combinatorial Algebra is a Conceptually New Approach to Matrix Completion which makes it approachable by Tools from:

- Classic and Combinatorial Algebraic Geometry
- Combinatorial and Probabilistic Graph Theory

The New Method yields Novel Theoretical Insights:

- Identifiability of the True Matrix depends only on Mask
- Combinatorial Conditions for Generic Identifiability
- Probabilistic Graph Theory explains the Asymptotics

The New Method allows Novel and Competitive Algorithms:

- Consistent and Accurate Estimator for Completion
- Polynomial Time Algorithm to determine Completability
Open Questions

Finitely Completable Masks
How to characterize the finitely completable masks for rank>1?

Relation to Optimization
How do the results on Trace Norm and SDP relate to this?

Sub-Completability
How to find and complete large completable sub-masks?
So: why Algebra?

Matrix Completion

is in its heart an estimation task with

Algebraic Combinatorial structure

which any useful method has to rely upon