# Low-Rank Matrix Recovery:
# The Matrix-Analogue of Compressed Sensing

Sven Müelich

Institute of Communications Engineering, Ulm University, Germany

Workshop "Compressive Sensing basierte Kryptographie"
Ulm, July 7, 2016

# Outline

# Low-Rank Matrix Recovery (LRMR)

### *Compressed Sensing (Donoho, Candès, Tao 2006)*

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x},$$

- $\boldsymbol{b} \in \mathbb{C}^m$ (known)
- $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ (known)
- $\boldsymbol{x} \in \mathbb{C}^n$ sparse (unknown)

$(n \ll m)$

### *LRMR (Candès, Tao, Recht 2009)*

$$\boldsymbol{b} = \mathcal{A}(\boldsymbol{X}),$$

- $\boldsymbol{b} \in \mathbb{C}^p$ (known)
- $\mathcal{A} : \mathbb{C}^{m \times n} \to \mathbb{C}^p$ linear (known)
- $\boldsymbol{X} \in \mathbb{C}^{m \times n}$ low rank (unknown)

$(p \ll n \cdot m)$

### *Hamming Metric Coding Problem*

$$\boldsymbol{s} = \boldsymbol{H}(\boldsymbol{c} + \boldsymbol{e}) = \boldsymbol{H}\boldsymbol{e}$$

- $\boldsymbol{s} \in K^{n-k}$ syndrome (known)
- $\boldsymbol{H} \in K^{n-k \times n}$ pc matrix (known)
- $\boldsymbol{c} \in K^n$ codeword (unknown)
- $\boldsymbol{e} \in K^n$, $\mathrm{wt}_{\mathrm{H}}(\boldsymbol{e})$ small (unknown)

### *Rank Metric Coding Problem*

$$\boldsymbol{s} = \boldsymbol{H}(\boldsymbol{c} + \boldsymbol{e}) = \boldsymbol{H}\boldsymbol{e}$$

- $\boldsymbol{s} \in L^{n-k}$ syndrome (known)
- $\boldsymbol{H} \in L^{n-k \times n}$ pc matrix (known)
- $\boldsymbol{c} \in L^n$ codeword (unknown)
- $\boldsymbol{e} \in L^n$, $\mathrm{wt}_{\mathrm{R}}(\boldsymbol{e})$ small (unknown)

$(L^n \simeq K^{m \times n})$

**Why do we need matrix recovery?**

- Size of data grows
- Observing matrices often impossible
- Applications want to process complete matrices
- Goal: Recover matrices from indirect or incomplete information

**(Recovery of) Low-rank matrices occurs in:**

- Collaborative Filtering (e.g. recommendation systems)
- Adjacency matrices (e.g. social networks)
- Machine Learning (e.g. multi-task learning, natural language processing)
- Distance matrices (e.g. nuclear magnetic resonance spectroscopy)
- Ensembles of signals (e.g. sensor networks)
- System identification
- Quantum state tomography
- ...

| | Compressed Sensing |
|---|---|
| Problem | $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$ |
| Naive | $\min \; \|\boldsymbol{x}\|_0$ s.t. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ |
| Better | $\min \; \|\boldsymbol{x}\|_1$ s.t. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ |
| How to solve? | Linear Programming |
| Algorithms | Basis Pursuit |
| | Greedy Algorithms |
| | Thresholding-based Algs. |

*Low-Rank Matrix Recovery (LRMR)*

$$\boldsymbol{b} = \mathcal{A}(\boldsymbol{X}),$$

- $\boldsymbol{b} \in \mathbb{C}^p$ (known)
- $\mathcal{A} : \mathbb{C}^{m \times n} \to \mathbb{C}^p$ linear (known)
- $\boldsymbol{X} \in \mathbb{C}^{m \times n}$ low rank (unknown)

$(p \ll n \cdot m)$

**Task:**

Reconstruct $\boldsymbol{X}$ from $\boldsymbol{b}$

**Algorithm:**

$$\min \operatorname{rank}(\boldsymbol{X}) \text{ subject to } \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}$$

**SVD of $\boldsymbol{X} \in \mathbb{C}^{m\times n}$:**

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^* = \sum_{l=1}^{n} \sigma_l \boldsymbol{u}_l \boldsymbol{v}_l^\mathsf{T}$$

- $n = \min\{n, m\}$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ (singular values of $\boldsymbol{X}$)
- $\boldsymbol{u}_l \in \mathbb{C}^m$ (left singular vectors)
- $\boldsymbol{v}_l \in \mathbb{C}^n$ (right singular vectors)

**Nuclear Norm:**

$$\|\boldsymbol{X}\|_* = \|\sigma(\boldsymbol{X})\|_1 = \sum_{l=1}^{n} \sigma_l(\boldsymbol{X})$$

- $\boldsymbol{\sigma} = \sigma(\boldsymbol{X}) = (\sigma_1, \ldots, \sigma_n)$

**Remarks:**

- Rank of $\boldsymbol{X}$ $\equiv$ $l_0$-norm of $\boldsymbol{\sigma}$
- Nuclear norm $\equiv$ $l_1$-norm of $\boldsymbol{\sigma}$

**Naive approach:**

$$\min \operatorname{rank}(\boldsymbol{X}) \text{ subject to } \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}$$

**Nuclear Norm Minimization:**

$$\min \|\boldsymbol{X}\|_* \text{ subject to } \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}$$

| | Compressed Sensing | Low-Rank Matrix Recovery |
|---|---|---|
| Problem | $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$ | $\boldsymbol{b} = \mathcal{A}(\boldsymbol{X})$ |
| Naive | min $\|\boldsymbol{x}\|_0$ s.t. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ | min rank$(\boldsymbol{X})$ s.t. $\mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}$ |
| Better | min $\|\boldsymbol{x}\|_1$ s.t. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ | min $\|\boldsymbol{X}\|_*$ s.t. $\mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}$ |
| How to solve? | Linear Programming | Semidefinite Programming |
| Algorithms | Basis Pursuit | Proximal Algorithm |
| | Greedy Algorithms | (Greedy Algorithms?) |
| | Thresholding-based Algs. | Thresholding-based Algs. |

NACHRICHTENTECHNIK Universität Ulm

### *Restricted Isometry Property RIP (Recht, Fazel, Parrilo)*

$\boldsymbol{b} = \mathcal{A}(\boldsymbol{X})$

$r$-restricted isometry constant: smallest number $\delta_r(\mathcal{A})$ such that

$$(1-\delta_r)\|\boldsymbol{X}\|_F \leq \|\mathcal{A}(\boldsymbol{X})\| \leq (1+\delta_r)\|\boldsymbol{X}\|_F$$

holds for all $\boldsymbol{X} \in \mathbb{C}^{m\times n}$ of rank $\leq r$.
(defined $\forall 1 \leq r \leq m$)

### *Theorem (Recht, Fazel, Parrilo)*

If

$$\delta_{2r} < 1 \quad \text{for some integer } r \geq 1$$

then $\boldsymbol{X}$ is the only matrix of rank $\leq r$ satisfying $\mathcal{A}(\boldsymbol{X}) = b$

$L/K$ Galois extension (i.e. normal and separable), $[L : K] =: m$

Galois group

$$\mathrm{Gal}\,(L/K) = \{\theta : L \to L \text{ automorphism s.t. } \theta(x) = x \;\; \forall x \in K\}$$

Assumption: $\mathrm{Gal}\,(L/K)$ is cyclic, $\theta$ generator

$L/K$ Galois extension, $\theta \in \mathrm{Gal}\,(L/K)$.

$$L[x;\theta] = \left\{ a = \sum_{i=0}^{d} a_i x^i \;:\; a_i \in L, d \in \mathbb{N} \right\}$$

Addition $(+)$ $\quad a + b = \sum_i (a_i + b_i) x^i$

Multiplication $(\cdot)$ $\quad a \cdot b = \sum_i \big( \sum_{j=0}^{i} a_j \theta^j (b_{i-j}) \big) x^i$ (non-commutative)

*Generalization of Linearized Polynomials*

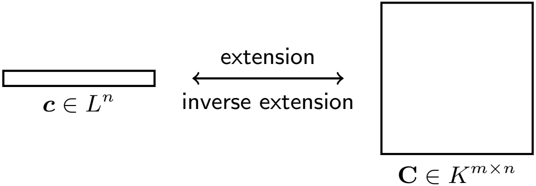
- Isomorphic to linearized polynomials in case $\mathbb{F}_{q^m}/\mathbb{F}_q$, $\theta = \cdot^q$

$L/K$ Galois extension, $\theta \in \mathrm{Gal}(L/K)$ generator.

***Definition***

$g_1, \ldots, g_n \in L$, linearly independent over $K$, $k \le n \le m = [L : K]$

$$\mathcal{C}_{\mathrm{G}}[n,k] = \left\{ \boldsymbol{c} = \left[ f(g_1), \ldots, f(g_n) \right] : f \in L[x;\theta] \wedge \deg f < k \right\} \subseteq L^n$$




**Rank Metric:** $\mathrm{wt}_{\mathrm{R}}(\boldsymbol{c}) = \mathrm{rank}(\mathbf{C})$, $\mathrm{d}_{\mathrm{R}}(\boldsymbol{c}_1, \boldsymbol{c}_2) = \mathrm{rank}(\mathbf{C}_1 - \mathbf{C}_2)$

***Theorem*** *(Augot, Loidreau, Robert)*

Minimum rank distance $d = \min_{\boldsymbol{c}_1 \neq \boldsymbol{c}_2} \mathrm{d}_{\mathrm{R}}(\boldsymbol{c}_1, \boldsymbol{c}_2) = n - k + 1$ (MRD)

**Decoding:** Augot, Loidreau, Robert (2013): $O(n^3)$
Müelich, Puchinger, Mödinger, Bossert (2016): $O(n^2)$

$\boldsymbol{E} \in K^{m \times n}$

(linear) $\downarrow$ Inverse extension in basis of $L$

$\boldsymbol{e} \in L^n$

(linear) $\downarrow$ Syndrome computation

$\boldsymbol{s} = \boldsymbol{e}\boldsymbol{H}^\mathsf{T} \in L^{n-k}$

(linear) $\downarrow$ Extension in basis of $L$

$\boldsymbol{b} \in K^{(n-k)m}$

$\boldsymbol{E} \in K^{m \times n}$

$\uparrow$ Decoding & ext. in basis of $L$

$\boldsymbol{r} = \boldsymbol{c} + \boldsymbol{e} \in L^n$

$\uparrow$ Find a solution of $\boldsymbol{r}\boldsymbol{H}^\mathsf{T} = \boldsymbol{s}$

$\boldsymbol{s} = \boldsymbol{e}\boldsymbol{H}^\mathsf{T} \in L^{n-k}$

$\uparrow$ Inverse extension in basis of $L$

$\boldsymbol{b} \in K^{(n-k)m}$

**Theorem** *Müelich, Puchinger, Bossert*

If $\mathrm{rank}(\boldsymbol{E}) \leq \frac{d-1}{2} = \frac{n-k}{2}$, $\boldsymbol{E}$ can be reconstructed from $\boldsymbol{b} = \mathcal{A}(\boldsymbol{E})$.

# LRMR: New Approach (Choice of $K$ and $L$)

$\boldsymbol{X} \in \mathbb{C}^{m \times n}$

↓ Rank-preserving mapping

$\boldsymbol{E} \in K^{m \times n}$

↓ Inverse extension in basis of $L$

$\boldsymbol{e} \in L^{n}$

↓ Syndrome computation

$\boldsymbol{s} = \boldsymbol{e}\boldsymbol{H}^{\top} \in L^{n-k}$

↓ Extension in basis of $L$

$\boldsymbol{b} \in K^{(n-k)m}$

**Needed:** $K \in \{\mathbb{R}, \mathbb{C}\}$. Possible $L$:

- $K = \mathbb{R}$: $L \in \{\mathbb{R}, \mathbb{C}\}$ $(m \leq 2)$
- $K = \mathbb{C}$: $L = \mathbb{C}$ $(m = 1)$

**Idea:** Choose $K$ to be a **dense** subfield of $\mathbb{R}$ or $\mathbb{C}$, e.g.,

| $K \subseteq$ | $\mathbb{R}$ | $\mathbb{C}$ |
|---|---|---|
| $K$ | $\mathbb{Q}$ | $\mathbb{Q}(\zeta_r)$ |
| $L$ | $\mathbb{Q}(\zeta_r)$ | Kummer extension |
| $m$ | $\varphi(r)$ | $r$ |

**Fundamental Work:**

Candès, Recht — Exact Matrix Completion via Convex Optimization (2009)
Foundations of Computational Mathematics, Vol. 9

Candès, Tao — The Power of Convex Relaxation:
Near-optimal Matrix Completion (2010 )
IEEE Transactions on Information Theory, Vol. 56

Recht, Fazel, Parrilo — Guaranteed Minimum-Rank Solutions of Linear
Matrix Equations via Nuclear Norm Minimization (2010)
SIAM Review, Vol. 52

**Overview Articles:**

Davenport, Romberg — An Overview of Low-Rank Matrix Recovery
from Incomplete Observations (2016)
arXiv preprint http://arxiv.org/abs/1601.06422

**CS and LRMR**

Fazel, Candès, Recht, Parillo — Compressed Sensing and Robust Recovery
of Low-Rank Matrices (2008)
42nd Asilomar Conference on
Signals, Systems and Computers

**Our Work:**

Müelich, Puchinger, Bossert — Low-Rank Matrix Recovery using
Gabidulin Codes in Characteristic Zero (2016)
Int. Workshop on Algebraic and Combinatorial
Coding Theory
nt.uni-ulm.de/mueelich → Publications