Subspace Projection Matrix Completion on Grassmann Manifold

Xinyue Shen and Yuantao Gu

Dept. EE, Tsinghua University, Beijing, China
http://gu.ee.tsinghua.edu.cn/

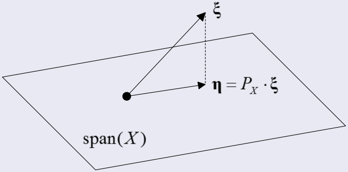
ICASSP 2015, Brisbane

Contents

Background

Subspace projection matrix

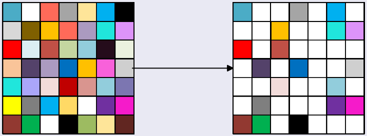
- $\{P_X := XX^\dagger | X \in \mathbb{R}^{N\times s}, \mathrm{rank}(X) = s\} \simeq \{\mathcal{S} \subset \mathbb{R}^N | \dim(\mathcal{S}) = s\}$
  One to one correspondence with subspaces
- $\mathrm{rank}(P_X) = s$
  Low rank matrix $\Longleftarrow$ Low dimensional subspace $s \ll N$
- Degree of freedom $= s(N - s)$
  Low degree of freedom $\Longleftarrow$ High dimensional subspace
- **More structures**: symmetric, non-zero eigenvalue 1, ...



Background

Matrix completion

- Uniformly random positions $\Omega_i, i = 1, \cdots, M$



- Low rank matrix
  - $M \geq O(N^{5/4} s \log N)$ [*Candès and Recht, 2008, Caltech*]
  - $M \geq O(N s \log^2 N \max\{\mu_1^2, \mu_0\})$ [*Recht, 2009, UW-Madison*]
  - ...

Background

Grassmann manifold of s dimensional subspaces in $\mathbb{R}^N$

$$\mathrm{Gr}_{N,s} := \frac{\mathcal{O}(N)}{\mathcal{O}(s)\mathcal{O}(N-s)} \tag{1}$$

- $s = 1$, $\mathbb{RP}^{N-1} = \{[\mathbf{x}] \in \mathbb{R}^N_* | \mathbf{x} \sim t\mathbf{x}, t \neq 0\}$
- $\dim(\mathrm{Gr}_{N,s}) = s(N-s)$
- Tangent space at $X \in \mathrm{Gr}_{N,s}$

$$T_X \mathrm{Gr}_{N,s} = \{\xi \in \mathbb{R}^{N\times s} : \xi^{\mathrm{T}} X = 0\} \tag{2}$$

- Inner product $\forall \eta, \xi \in T_X \mathrm{Gr}_{N,s}$

$$\langle \eta, \xi \rangle_X := \mathrm{trace}\left((X^{\mathrm{T}}X)^{-1}\eta^{\mathrm{T}}\xi\right) \tag{3}$$

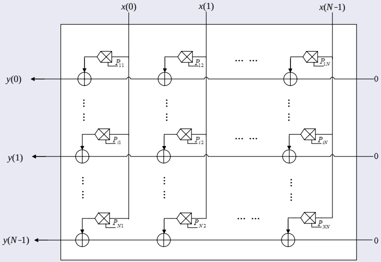
Motivation

Exploration: fewer samples?

- Structures: sparsity, low rank, low-dim manifolds...
- $M \geq O(s(N-s))$?

Subspace identification

- A demonstration: Multi-bandpass filter



Formulation

Setup

- $X_0 \in \mathbb{R}^{N\times s}$, dim(span($X_0$))=s
- $\mathcal{A} : \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}$

$$[\mathcal{A}(P_{X_0})]_{\Omega_i} := [P_{X_0}]_{\Omega_i},\ i = 1, 2, \cdots M,$$
$$[\mathcal{A}(P_{X_0})]_{i,j} := 0,\ (i, j) \notin \Omega$$

  $\Omega \subset \{1, 2, \cdots, N\} \times \{1, 2, \cdots, N\}$ uniformly randomly chosen, $|\Omega| = M < N^2$
- $Y := \mathcal{A}(P_{X_0})$

Formulation

Recovery

- $P_X = XX^\dagger = \bar{X}\bar{X}^{\mathrm{T}}, \bar{X} \in \mathrm{Gr}_{N,s}$

$$\hat{X} = \mathrm{argmin}_{X \in \mathrm{Gr}_{N,s}} \|\mathcal{A}(XX^{\mathrm{T}}) - Y\|_F^2 \quad (4)$$

- Cost function well-defined on $\mathrm{Gr}_{N,s}$
- Non-convex optimization
- Recovery error: projection distance

$$d_p(\mathrm{span}(\hat{X}), \mathrm{span}(X_0)) = \frac{1}{2}\|X_0X_0^{\mathrm{T}} - \hat{X}\hat{X}^{\mathrm{T}}\|_F^2 \quad (5)$$

Algorithm: GGDLS

Grassmann Manifold Gradient Descent Line Search Algorithm

Require: a smooth scalar field f on $\mathrm{Gr}_{N,s}$;
a smooth mapping $R : \mathcal{T}\mathrm{Gr}_{N,s} \to \mathrm{Gr}_{N,s}$;
scalars for the Armijo step size $\alpha > 0, \beta, \sigma \in (0, 1)$;
Input: Initial iterate $X_0 \in \mathrm{Gr}_{N,s}$;
Output: Sequence of iterates $\{X_k\}$.

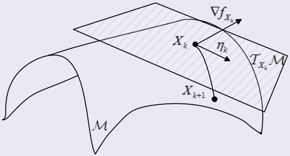
For $k = 0, 1, 2, \cdots$ **do:**

1.Compute the Euclidean gradient ∇f_{X_k} at X_k ;

2.Project ∇f_{X_k} onto the tangent space $\mathcal{T}_{X_k}\mathrm{Gr}_{N,s}$ to obtain η_k;

3.$X_{k+1} = R_{X_k}(t_k \eta_k)$, t_k is the Armijo step size;

Until: Stopping criterion satisfied;



Algorithm: GGDLS

- $[\nabla f_X]_{p,q} = \sum_{j:(p,j)\in\Omega}[P_X - Y]_{p,j}X_{j,q} + \sum_{i:(i,p)\in\Omega}[P_X - Y]_{i,p}X_{i,q}$

Definition

$\mathrm{grad} f(X)$ defined as the unique element in $T_X\mathcal{M}$ such that

$$\langle \mathrm{grad} f(X), \xi\rangle_X = \nabla f(X)[\xi], \quad \forall \xi \in T_X\mathcal{M} \tag{6}$$

- $-\mathrm{grad} f(X_k) = \eta_k = -(I_N - X_k X_k^{\mathrm{T}})\nabla f_{X_k}$

Definition

R defined by $R_X : \mathcal{T}_X\mathcal{M} \to \mathcal{M}$ such that

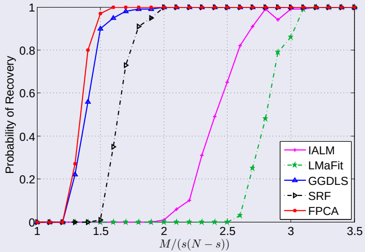
$$R_X(0) = X, \quad R_{X_*}(0) = \mathrm{id}_{T_X\mathcal{M}} \tag{7}$$

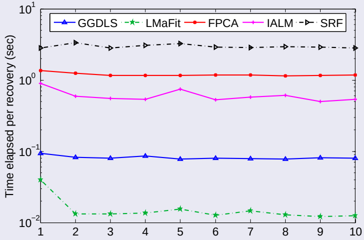
- R_X: Q factor of the QR decomposition of X

Numerical Experiment

Subspace projection matrix completion: noiseless

- $N = 100,\ s = 10,\ \alpha = 1, \beta = 0.8, \sigma = 0.01$, initial value from Y
- Left figure: $\|X_0 X_0^{\mathrm{T}} - \hat{X}\hat{X}^{\mathrm{T}}\|_F/\sqrt{s} < 10^{-2}$ for successful recovery, 100 trials per point
- Right figure: $\frac{M}{s(N-s)} = 3.5$

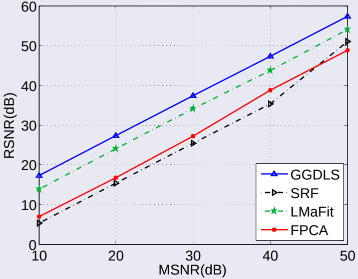




Numerical Experiment

Subspace projection matrix completion: noisy

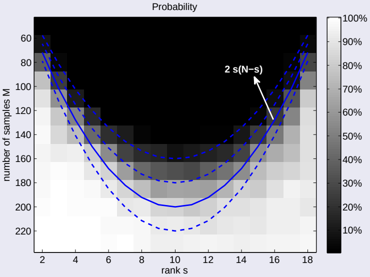
- $N = 100$, $s = 10$, $\alpha = 1, \beta = 0.8, \sigma = 0.01$, initial value from Y
- $\frac{M}{s(N-s)} = 3.5$
- 100 trials per point



Numerical Experiment

Subspace projection matrix completion: noiseless

- $N = 20,\ \alpha = 100, \beta = 0.8, \sigma = 0.01$, initial value from Y
- $\|X_0 X_0^{\mathrm{T}} - \hat{X}\hat{X}^{\mathrm{T}}\|_F / \sqrt{s} < 10^{-2}$ for successful recovery
- 200 trials per point



Theoretical Support: Subspace Projection Matrix Sensing

Theorem

For fixed $0 < \varepsilon < 1$ and $\beta > 0$, assume that $N \geq 3$. Let $\mathcal{A} : \mathbb{R}^{N \times N} \to \mathbb{R}^M$ be a random orthoprojector with

$$M \geq \left(\frac{2 + \beta}{\varepsilon^2 - \varepsilon^3/3} \right) O \left(s(N - s) \log \frac{N}{\varepsilon} \right). \tag{8}$$

If $M < N^2$, then with probability exceeding $1 - \mathrm{e}^{-c_1 \beta s(N-s)}$, in which c_1 is a universal constant, the following property holds for every pair of $P_X, P_Y \in \mathcal{P}_{N,s}$, $P_X \neq P_Y$

$$(1 - \varepsilon) \frac{\sqrt{M}}{N} \leq \frac{\|\mathcal{A}(P_X - P_Y)\|_2}{\|P_X - P_Y\|_F} \leq (1 + \varepsilon) \frac{\sqrt{M}}{N}. \tag{9}$$

- Restricted Isometry Property of Subspace Projection Matrix Under Random Compression, Xinyue Shen and Yuantao Gu, 2015, *IEEE Signal Processing Letters, 22(9), 1326 - 1330.*

Conclusion

Subspace projection matrices

- Structural data: on manifold
- Low rank: more structures, less number of samples
- High rank: low intrinsic dimension

GGLDS algorithm

- Optimize on a non-convex manifold
- Exploit the unique structure of subspace projection matrices