Literature Review: Tensor Completion Based on Nuclear Norm

Tensor Completion for Estimating Missing Values in Visual Data

Gong Wenwu

Southern University of Science and Technology

December 29, 2021

Outline

Tensor Completion for Low Rank Data[1]

Author: Ji Liu, Department of Computer Sciences at University Wisconsin-Madison (http://pages.cs.wisc.edu/~ji-liu/), Directors of Kwai Seattle AI lab and FeDa Lab at Kwai Inc (Now)

Contributions: Extensions of low rank matrix completion

- Definition of the trace norm for tensors
- Algorithms for the low rank tensors completion
  - Simple low rank tensor completion (SiLRTC): employs a relaxation technique to separate the dependent relationships and uses the block coordinate descent method to achieve a globally optimal solution
  - High accuracy low rank tensor completion (HaLRTC): applies the alternating direction method of multipliers (ADMM) algorithm to solve low rank tensor estimation
  - Fast low rank tensor completion (FaLRTC): utilizes a smoothing scheme to transform the original nonsmooth problem into a smooth problem
- Comparison with different completion algorithms

[1] Ji Liu et al. "Tensor Completion for Estimating Missing Values in Visual Data". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.1 (2013), pp. 208–220.

Abstract

- Tensor completion is formulated as a convex optimization problem[2] (Extension of matrix completion)
- Three algorithms (based on the trace norm) are proposed to estimate missing values in tensors of visual data (Any other algorithms?)
- The algorithms work even with a small amount of samples (Missing ratio is high) and it can propagate structure to fill larger missing regions
- FaLTRC and HaLRTC are more efficient than SiLRTC and between FaLRTC and HaLRTC the former is more efficient to obtain a low accuracy solution and the latter is preferred if a high-accuracy solution is desired
- Experiments show potential applications with different types of data: images and videos (traffic data ?)

[2] Ryota Tomioka, Kohei Hayashi, and Hisashi Kashima. "Estimation of low-rank tensors via convex optimization". In: *arXiv:1010.0789* (2010).

Background

- Missing value estimation in computer version and graphics: image in-painting, video decoding and video in-painting.
- Core problem: Build up the relationship between the known elements and the unknown ones. (Capture the global information in the data)
- How to do? Low rank approximation[3]: minimize the trace norm of matrices to approximate the rank of matrices[4].(Data is of low rank and matrix based)
- Gap: Low Rank Tensor Completion[5] based on the trace norm of tensor. (a convex but nondifferentiable optimization problem: interdependent matrix trace norm terms)

---

[3] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization". In: *SIAM Review* 52.3 (2010), pp. 471–501. DOI: 10.1137/070697835.

[4] Jian-Feng Cai, Emmanuel J. Candès, and Zuowei Shen. "A Singular Value Thresholding Algorithm for Matrix Completion". In: *SIAM Journal on Optimization* 20.4 (2010), pp. 1956–1982. DOI: 10.1137/080738970.

[5] Ryota Tomioka, Kohei Hayashi, and Hisashi Kashima. "Estimation of low-rank tensors via convex optimization". In: *arXiv:1010.0789* (2010).

Notation

- Matrix X, Entries x_{ij}, Matrix SVD $X = U\Sigma V\top$
- $\Sigma(X)$ (vector) singular values of X in descending order, $\sigma_i(X)$ i-th largest singular value
- Frobenius norm $\|X\|_F := \left(\sum_{i,j} |x_{ij}|^2\right)^{\frac{1}{2}}$,
  Spectral norm $\|X\| := \sigma_1(X)$,
  Trace norm $\|X\|_{tr} := \sum_i \sigma_i(X)$
- Ω an index set for observed value
- n -mode tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_n}$. Its elements are denoted as $x_{i_1,\ldots,i_n}$, where $1 \le i_k \le I_k, 1 \le k \le n$, I_k is the size of k-th dimension.

Notation

'Shrinkage' operator $\mathbf{D}_\tau(X)$ is defined as

$$\mathbf{D}_\tau(X) := U\mathbf{D}_\tau(\Sigma)V^\top = U\Sigma_\tau V^\top$$

where $\mathbf{D}_\tau(\Sigma) = \Sigma_\tau = \operatorname{diag}\left(\max\left(\sigma_i - \tau, 0\right)\right) = \operatorname{diag}\left(\{\sigma_i - \tau\}_+\right)$.
Soft-thresholding rule is applied to the singular values of matrix X.
'Truncate' operator $\mathbf{T}_\tau(X)$ is defined as

$$\mathbf{T}_\tau(X) = U\Sigma_{\bar{\tau}}V^\top,$$

where $\Sigma_{\bar{\tau}} = \operatorname{diag}\left(\min\left(\sigma_i, \tau\right)\right)$ and $X = \mathbf{T}_\tau(X) + \mathbf{D}_\tau(X)$.
Unfold a tensor into a matrix

$$\operatorname{unfold}_i(\mathcal{X}) := \mathcal{X}_{(i)} := X_{(i)} \in \mathbb{R}^{I_i \times (I_1 \ldots I_{i-1} I_{i+1} \ldots I_n)}$$

The inverse of the unfolding operation is written

$$fold_i\left(\operatorname{unfold}_i(\mathcal{X})\right) = \mathcal{X}.$$

Tensor Completion

- Define the trace norm for the general tensor case:

$$\|\mathcal{X}\|_{tr} := \sum_{i=1}^{n} \alpha_i \left\|\mathcal{X}_{(i)}\right\|_{tr} \quad \textit{trace norm}$$

  where α_is are constants satisfying $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$. In essence, the trace norm of a tensor is a convex combination of the trace norms of all matrices unfolded ($\mathcal{X}_{(i)}$) along each mode.

- The tensor completion optimization problem can be written as

$$\begin{aligned} &\min_{\mathcal{X}} : \textstyle\sum_{i=1}^{n} \alpha_i \left\|\mathcal{X}_{(i)}\right\|_{tr} \\ &s.t. : \mathcal{X}_{\Omega} = \mathcal{T}_{\Omega} \ (noiseless) \end{aligned} \qquad \textit{nonsmooth convex problem}$$

  where $\mathcal{X}, \mathcal{T}$ are n-mode tensors with identical size in each mode. Difficult: the matrices share the same entries and cannot be optimized independently (multiple dependent nonsmooth terms)

Matrix Completion

The low rank matrix completion (nonconvex optimization):

$$\min_X : \operatorname{rank}(X) \quad s.t. : X_\Omega = M_\Omega(noiseless), \quad Nonconvex$$

where $X, M \in \mathbb{R}^{p\times q}$. The missing elements of X are determined such that the rank of the matrix X is as small as possible.
This leads to the following trace norm convex optimization problem (tightest convex envelop) for matrix completion:

$$\min_X : \|X\|_{tr}\ (trace\ norm\ , subgradient) \quad s.t. : X_\Omega = M_\Omega. \quad Convex$$

The trace norm minimization problem can be solved by Block coordinate descent iterative method[6], singular value thresholding algorithm[7], alternating direction method of multipliers (ADMM)[8]

[6] Shiqian Ma, Donald Goldfarb, and Lifeng Chen. "Fixed point and Bregman iterative methods for matrix rank minimization". In: *Mathematical Programming* 128 (2011), pp. 321–353. DOI: 10.1007/s10107-009-0306-5.

[7] Jian-Feng Cai, Emmanuel J. Candès, and Zuowei Shen. "A Singular Value Thresholding Algorithm for Matrix Completion". In: *SIAM Journal on Optimization* 20.4 (2010), pp. 1956–1982. DOI: 10.1137/080738970.

[8] Zhouchen Lin, Minming Chen, and Yi Ma. "The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices". In: *arXiv:1009.5055v3* (2010).

Baseline Methods for Tensor Completion[9]

- Tucker model: tensor factorization

$$\min_{\mathcal{X},C,U_1,\ldots,U_n} : \frac{1}{2}\|\mathcal{X}-C\times_1 U_1\times_2 U_2\times_3\cdots\times_n U_n\|_F^2 \quad s.t.: \mathcal{X}_\Omega = \mathcal{T}_\Omega$$

- Parafac model: Parafac model-based decomposition

$$\min_{\mathcal{X},U_1,U_2,\ldots,U_n} : \frac{1}{2}\|\mathcal{X} - U_1\circ U_2\circ\cdots\circ U_n\|_F^2 \quad s.t.: \mathcal{X}_\Omega = \mathcal{T}_\Omega$$

- SVD: consider the tensor as multiple matrices and force the unfolding matrix along each mode of the tensor to be low rank as follows:

$$\min_{\mathcal{X},M_1,M_2,\ldots,M_n} : \frac{1}{2}\sum_{i=1}^{n}\left\|\mathcal{X}_{(i)} - M_i\right\|_F^2 \quad (\textit{additional matrices } M_i)$$

$$\text{s.t.} \quad : \mathcal{X}_\Omega = \mathcal{T}_\Omega, \quad \operatorname{rank}(M_i) \le r_i \quad i = 1,\ldots,n,$$

where $M_i \in \mathbb{R}^{I_i\times(\prod_{k\neq i} I_k)}$ and $\mathcal{T},\mathcal{X} \in \mathbb{R}^{I_1\times\cdots\times I_n}$.

[9] Tamara G. Kolda and Brett W. Bader. "Tensor Decompositions and Applications". In: *SIAM REVIEW* 51.3 (2009), pp. 455–500.

Singular value thresholding (SVT) algorithm[10]

Goal: SVT algorithm is used to approximately minimize the trace norm of a matrix under convex constraints

Problem:

$$\min \|\boldsymbol{X}\|_{tr} \quad s.t.: \boldsymbol{X}_\Omega = \boldsymbol{M}_\Omega$$

$$\Leftrightarrow \min_{X \in \mathbb{R}^{p \times q}} : \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{M}\|_F^2 + \tau \|\boldsymbol{X}\|_{tr}$$

Then the Singular Value Shrinkage Operator $\mathbf{D}_\tau(\boldsymbol{Y})$ obeys:

$$\mathcal{D}_\tau(\boldsymbol{Y}) = \arg\min_{\boldsymbol{X}} \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{Y}\|_F^2 + \tau\|\boldsymbol{X}\|_{tr}$$

where Y is the initial matrix and need to update.

Algorithm: Fix $\tau > 0$ and a sequence $\{\delta_k\}$ of positive step sizes. Starting with $\boldsymbol{Y}^0$, the algorithm inductively defines

$$\begin{aligned} \boldsymbol{X}^k &= \mathbf{D}_\tau(\boldsymbol{Y}) = \text{shrink}\left(\boldsymbol{Y}^{k-1}, \tau\right) \\ \boldsymbol{Y}^k &= \boldsymbol{Y}^{k-1} + \delta_k \left(\boldsymbol{M} - \boldsymbol{X}^k\right)_\Omega \end{aligned}$$

until stopping criterion $\left\|\left(\boldsymbol{M} - \boldsymbol{X}^k\right)_\Omega\right\|_F / \|\boldsymbol{M}_\Omega\|_F \leq \epsilon$ is reached.

[10] Jian-Feng Cai, Emmanuel J. Candès, and Zuowei Shen. "A Singular Value Thresholding Algorithm for Matrix Completion". In: *SIAM Journal on Optimization* 20.4 (2010), pp. 1956–1982. DOI: 10.1137/080738970.

Alternating direction method of multipliers (ADMM)[11]

Goal: ADMM algorithm has the superior convergence property to solve Robust PCA problem (matrix completion).

Problem:

$$\min f(X), \quad s.t.: h(X) = 0$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^m$. One may define the augmented Lagrangian function:

$$L(X, Y, \mu) = f(X) + \langle Y, h(X) \rangle + \frac{\mu}{2} \|h(X)\|_F^2$$

where μ is a positive scalar.

Algorithm:

$$\begin{aligned}
&solve\ X_{k+1} = \arg\min_X L\left(X, Y_k, \mu_k\right) \\
&Y_{k+1} = Y_k + \mu_k h\left(X_{k+1}\right) \\
&updata\ \mu_k\ to\ \mu_{k+1}
\end{aligned}$$

[11] Zhouchen Lin, Minming Chen, and Yi Ma. "The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices". In: *arXiv:1009.5055v3* (2010).

Nonsmooth convex optimization

Problem Definition:

$$Minimize_{x \in \mathbb{R}^n} \quad f(\boldsymbol{x}) \quad \textit{nonsmooth convex problem}$$

where f is convex but not always differentiable.

- Subgradient methods yield ε-accuracy in $O\left(\frac{1}{\varepsilon^2}\right)$ $iterations$
- Nesterov's smoothing: if f is smooth, then accelerated GD yields ε -accuracy in $O\left(\frac{1}{\sqrt{\varepsilon}}\right)$ $iterations$
  - approximate the nonsmooth objective by a smooth function
  - solve the smooth problem and use its solution to approximate the original problem

Smooth approximation: Convex function f is called (α, β) smoothable if, for any $\mu > 0, \exists$ convex function f_μ s.t.

- Approximation accuracy: $f_\mu(\boldsymbol{x}) \leq f(\boldsymbol{x}) \leq f_\mu(\boldsymbol{x}) + \beta\mu, \forall \boldsymbol{x}$
- Smoothness: $f_\mu \ is \ \frac{\alpha}{\mu}$

  where μ trade-off between approximation accuracy and smoothness

f_μ is a $\frac{1}{\mu}$ smooth approximation of f with parameters (α, β).

Nonsmooth optimization

Approximation methods: Moreau envelope and Conjugation

- The Moreau envelope (or Moreau-Yosida regularization) of a convex function f with parameter $\mu > 0$ is defined as

$$M_{\mu f}(\boldsymbol{x}) := \inf_{\boldsymbol{z}} \left\{ f(\boldsymbol{z}) + \frac{1}{2\mu} \|\boldsymbol{x} - \boldsymbol{z}\|_2^2 \right\}$$

  where $M_{\mu f}$ is a smoothed or regularized form of f.
  Minimizing f and minimizing $M_{\mu f}$ are equivalent.

- Suppose $f = g^*$, namely, $f(\boldsymbol{x}) = \sup_{\boldsymbol{z}} \{\langle \boldsymbol{z}, \boldsymbol{x} \rangle - g(\boldsymbol{z})\}$ One can build a smooth approximation of f by adding a strongly convex component to its dual, namely, $f_\mu(\boldsymbol{x}) = \sup_{\boldsymbol{z}} \{\langle \boldsymbol{z}, \boldsymbol{x} \rangle - g(\boldsymbol{z}) - \mu d(\boldsymbol{z})\} = (g + \mu d)^*(\boldsymbol{x})$ for some 1-strongly convex and continuous function $d \geq 0$ (called proximity function).

Examples: l_1-norm (Huber function approximate)
trace norm (Research problem)

Equivalent convex optimization problem

- Recall that: tensor completion optimization problem

$$\begin{aligned} &\min_{\mathcal{X}} : \sum_{i=1}^{n} \alpha_i \left\| \mathcal{X}_{(i)} \right\|_{tr} \\ &s.t. : \mathcal{X}_\Omega = \mathcal{T}_\Omega \ (noiseless) \end{aligned} \qquad nonsmooth\ convex\ problem$$

- Introduce additional matrices $M_1, \ldots, M_n$ to obtain equivalent formulation (Omitted) and relax the $M_i = \mathcal{X}_{(i)}$ by $\left\| M_i - \mathcal{X}_{(i)} \right\|_F^2 \leq d_i$ (interdependent terms have been split and can be solved independently):

$$\begin{aligned} &\min_{\mathcal{X}, M_i} : \sum_{i=1}^{n} \alpha_i \left\| M_i \right\|_{tr} \\ &s.t. : \left\| \mathcal{X}_{(i)} - M_i \right\|_F^2 \leq d_i \ for \ i = 1, \ldots, n \quad and \quad \mathcal{X}_\Omega = \mathcal{T}_\Omega \end{aligned}$$

  $d_i (> 0)$ is a threshold that could be defined by the user.

- This convex but nondifferentiable optimization problem can be converted to an equivalent formulation for certain positive values of β_is :

$$\min_{\mathcal{X}, M_i} : \sum_{i=1}^{n} \alpha_i \left\| M_i \right\|_* + \frac{\beta_i}{2} \left\| \mathcal{X}_{(i)} - M_i \right\|_F^2 \quad s.t. : \mathcal{X}_\Omega = \mathcal{T}_\Omega$$

Algorithm : block coordinate descent framework

- Computing $\mathcal{X}$. The optimal $\mathcal{X}$ with all other variables fixed is given by solving the following subproblem:

$$\begin{array}{c} \min_{\mathcal{X}} : \sum_{i=1}^{n} \frac{\beta_i}{2} \left\| M_i - \mathcal{X}_{(i)} \right\|_F^2 \\ \text{s.t.: } \mathcal{X}_{\Omega} = \mathcal{T}_{\Omega} \end{array}$$

Then the solution is given by

$$\mathcal{X}_{i_1,\ldots,i_n} = \begin{array}{ll} \left(\frac{\sum_i \beta_i \text{fold}_i(M_i)}{\sum_i \beta_i} \right) & (\mathrm{i}_1, \ldots, \mathrm{i}_\mathrm{n}) \notin \Omega; \\ \mathcal{T}_{i_1,\ldots,i_n} & (\mathrm{i}_1, \ldots, \mathrm{i}_\mathrm{n}) \in \Omega \end{array}$$

- Computing M_i. M_i is the optimal solution of the following problem:

$$\min_{M_i} : \frac{\beta_i}{2} \left\| M_i - \mathcal{X}_{(i)} \right\|_F^2 + \alpha_i \left\| M_i \right\|_{tr} \equiv \frac{1}{2} \left\| M_i - \mathcal{X}_{(i)} \right\|_F^2 + \frac{\alpha_i}{\beta_i} \left\| M_i \right\|_{tr}$$

Singular value thresholding (SVT) algorithm !!!

ADMM framework

- Tensor completion optimization problem can be equivalent to

$$\min_{\mathcal{X},\mathcal{M}_1,\ldots,\mathcal{M}_n} : \sum_{i=1}^{n} \alpha_i \left\|\mathcal{M}_{i(i)}\right\|_{tr}$$
$$s.t. : \mathcal{X}_\Omega = \mathcal{T}_\Omega \; and \; \mathcal{X} = \mathcal{M}_i, \;\; i = 1, \ldots, n.$$

by replacing the dummy matrices M_i s by their tensor versions

- Define the augmented Lagrangian function as follows:

$$\begin{aligned} L_\rho\left(\mathcal{X}, \mathcal{M}_1, \ldots, \mathcal{M}_n, \mathcal{Y}_1, \ldots, \mathcal{Y}_n\right) = & \sum_{i=1}^{n} \alpha_i \left\|\mathcal{M}_{i(i)}\right\|_{tr} \\ & + \langle \mathcal{X} - \mathcal{M}_i, \mathcal{Y}_i \rangle \\ & + \frac{\rho}{2} \left\|\mathcal{M}_i - \mathcal{X}\right\|_F^2 \end{aligned}$$

Then update $\mathcal{M}_i$s, $\mathcal{X}$, and $\mathcal{Y}_i$s as follows:

- $\left\{\mathcal{M}_1^{k+1}, \ldots, \mathcal{M}_n^{k+1}\right\} = \operatorname{argmin}_{\mathcal{M}_1,\ldots,\mathcal{M}_n} :$ $L_\rho\left(\mathcal{X}^k, \mathcal{M}_1, \ldots, \mathcal{M}_n, \mathcal{Y}_1^{k+1}, \ldots, \mathcal{Y}_n^{k+1}\right)$
- $\mathcal{X}^{k+1} = \arg\min_{\mathcal{X}\in Q} :$ $L_\rho\left(\mathcal{X}, \mathcal{M}_1^{k+1}, \ldots, \mathcal{M}_n^{k+1}, \mathcal{Y}_1^{k}, \ldots, \mathcal{Y}_n^{k+1}\right)$
- $\mathcal{Y}_i^{k+1} = \mathcal{Y}_i^k - \rho\left(\mathcal{M}_i^{k+1} - \mathcal{X}^{k+1}\right)$

Algorithm illustration

Algorithm 4. HaLRTC: High Accuracy Low Rank Tensor Completion

Input: $\mathcal{X}$ with $\mathcal{X}_\Omega = \mathcal{T}_\Omega$, ρ, and K

Output: $\mathcal{X}$

1: Set $\mathcal{X}_\Omega = \mathcal{T}_\Omega$ and $\mathcal{X}_{\bar{\Omega}} = 0$.

2: **for** $k = 0$ to K **do**

3: **for** $i = 1$ to n **do**

4: $$\mathcal{M}_i = \text{fold}_i\left[\mathbf{D}_{\frac{\alpha_i}{\rho}}\left(\mathcal{X}_{(i)} + \frac{1}{\rho}\mathcal{Y}_{i(i)}\right)\right]$$

5: **end for**

6: $$\mathcal{X}_\Omega = \frac{1}{n}\left(\sum_{i=1}^{n}\mathcal{M}_i - \frac{1}{\rho}\mathcal{Y}_i\right)_{\bar{\Omega}}$$

7: $$\mathcal{Y}_i = \mathcal{Y}_i - \rho(\mathcal{M}_i - \mathcal{X})$$

8: **end for**

Nesterov's smoothing framework

Tensor completion optimization problem can be equivalent to

$$\min_{\mathcal{X}} f_\mu(\mathcal{X}) := \min_{\mathcal{X}} \sum_{i=1}^{n} \max_{\|\mathcal{Y}_{i(i)}\|\leq 1} : \alpha_i \langle \mathcal{X}, \mathcal{Y}_i \rangle - \frac{\mu_i}{2} \|\mathcal{Y}\|_F^2$$
$$s.t. : \mathcal{X}_\Omega = \mathcal{T}_\Omega$$

by introducing n dual variables $\mathcal{Y}_1, \ldots, \mathcal{Y}_n \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_n}$ and n positive constants $\mu_1, \ldots, \mu_n$.

$$\mathcal{Z}^{k+1} = \mathcal{Z}^k - \frac{\theta^{k+1}}{L^k} \nabla f_\mu \left(\mathcal{W}^{k+1}\right),$$
$$\left(\nabla f_\mu \left(\mathcal{W}^{k+1}\right)\right)_{i_1,\ldots,i_n} = \left(\sum_i \frac{(\alpha_i)^2}{\mu_i} \mathbf{T}_{\frac{\mu_i}{\alpha_i}} \left(\mathcal{W}_{(i)}^{k+1}\right)\right)_\Omega$$

Algorithm illustration

Algorithm 3. FaLRTC: Fast Low Rank Tensor Completion

Input: $c \in (0,1)$, $\mathcal{X}$ with $\mathcal{X}_\Omega = \mathcal{T}_\Omega$, K, μ_is, and L.

Output: $\mathcal{X}$

```
Initialize Z = W = X, L' = L, and B = 0
for k = 0 to K do
    while true do
        θ = L/(2L') (1 + sqrt(1 + 4L'B));
        W = (θ/L)/(B+θ/L) Z + B/(B+θ/L) X;
        if f_μ(X) ≤ f_μ(W) − ||∇f_μ(W)||_F^2 / 2L' then
            break;
        end if
        X' = W − ∇f_μ(W)/L';
        if f_μ(X') ≤ f_μ(W) − ||∇f_μ(W)||_F^2 / 2L' then
            X = X';
            break;
        end if
        L' = L'/c;
    end while
    L = L';
    Z = Z − (θ/L) ∇f_μ(W);
    B = B + θ/L;
end for
```

$$\theta = \frac{L}{2L'}\left(1 + \sqrt{1 + 4L'B}\right); \qquad \mathcal{W} = \frac{\theta/L}{B+\theta/L}\mathcal{Z} + \frac{B}{B+\theta/L}\mathcal{X};$$

$$f_\mu(\mathcal{X}) \le f_\mu(\mathcal{W}) - \|\nabla f_\mu(\mathcal{W})\|_F^2/2L'$$

$$\mathcal{X}' = \mathcal{W} - \nabla f_\mu(\mathcal{W})/L'; \qquad L' = L'/c;$$

$$L = L'; \quad \mathcal{Z} = \mathcal{Z} - \frac{\theta}{L}\nabla f_\mu(\mathcal{W}); \quad B = B + \frac{\theta}{L};$$

Model comparison

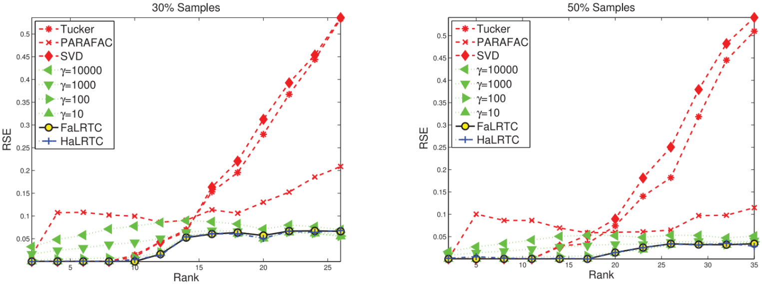


Figure 1: The RSE comparison on the synthetic data

Tensor completion and matrix completion

RSE Comparison (10^{-4}), Size $20 \times 20 \times 20$

Samples	MC1	MC2	MC3	TC
25%	1663	1782	1685	34
40%	247	258	241	2
60%	0	0	0	0

RSE Comparison (10^{-4}), Size $20 \times 20 \times 20 \times 20$

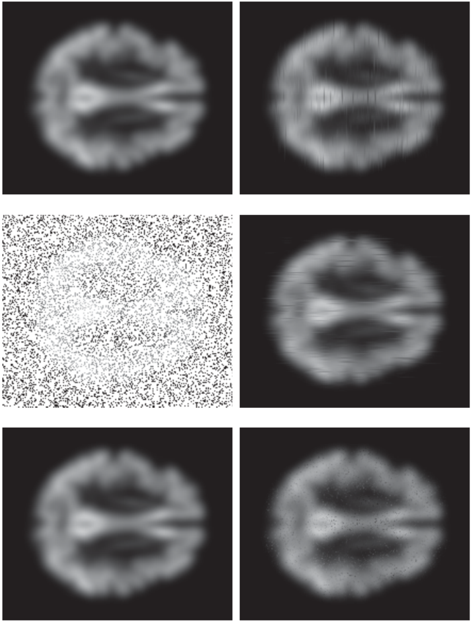
Samples	MC1	MC2	MC3	MC4	TC
20%	1875	1763	2011	1804	3
40%	92	102	97	88	0
60%	0	1	0	0	0

RSE Comparison (10^{-4}), Size $20 \times 20 \times 20 \times 20 \times 20$

Samples	MC1	MC2	MC3	MC4	MC5	TC
15%	1874	1830	1663	1502	1688	50
40%	125	119	131	114	136	0
60%	0	0	0	0	0	0

Figure 2: The RSE comparison on the synthetic data (MC_i represents the unfold tensor along the i th mode into a matrix structure $\mathcal{T}_{(i)}$)

Tensor completion and matrix completion (MRI Data)



Efficiency Comparison

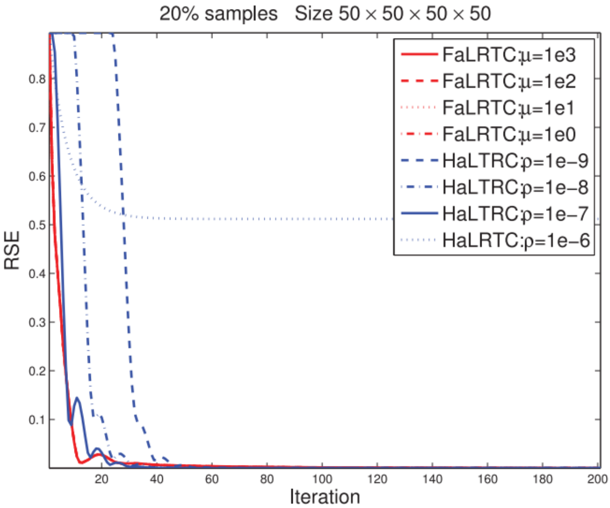


Figure 3: The RSE curves of the FaLRTC algorithm and the HaLRTC algorithm in terms of the number of iterations and the computation time

Missing data in images

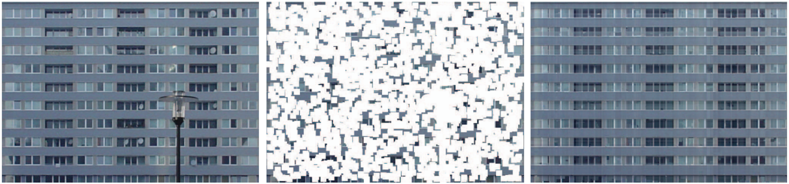


Figure 4: Facade in-painting. The left image is the original image; we select the lamp and satellite dishes, together with a large set of randomly positioned squares, as the missing parts, shown in white in the middle image; the right image is the result of the proposed LTRC algorithm.

Missing data in video

Figure 5: Video completion. The left image (one frame of the video) is the original; we randomly select pixels for removal, shown in white in the middle image; the right image is the result of the proposed LTRC algorithm.

Missing data in traffic data[12] (spatiotemporal)

- Application of HaLRTC algorithm
- Traffic data analysis: reshape tensor How to do ?

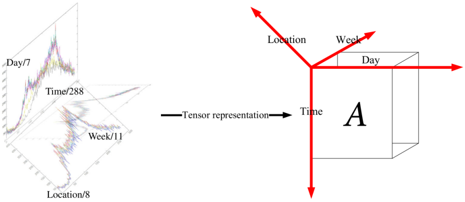


Figure 6: Tensor representation can capture the hidden weekly, daily, spatial correlation of traffic flow data.

- Correlation analysis of traffic data (similarity coefficient) to determine the parameters of tensor completion
- Two experiments design: all sensors are missing, impute missing data for local location independently

[12] Bin Ran et al. "Tensor based missing traffic data completion with spatial–temporal correlation". In: *Physica A: Statistical Mechanics and its Applications* 446 (2016), pp. 54–63.

Visual data recovery[13]

Contributions: Tensor train rank along with spatial and temporal regularization (Novel Model), Three solver subproblem Proximal Alternating Minimization (PAM) algorithm, Convergence analysis

- Tensor property: Sparsity in spatial and temporal mode (High order tensor transformer: KA)
- PAM algorithm with three subproblems are used to solve TT-Framelet model (Tensor decomposition and regularization)
- A detailed induction of PAM solver
- A novel tensor completion model (TT-framelet) by simultaneously exploiting the global low-rankness and local smoothness of visual data
- Use low-rank matrix factorization to characterize the global low-rankness; framelet and total variation regularization to enhance the local smoothness PAM solver
- Algorithm convergence analysis

[13] Jing-Hua Yang et al. "Tensor train rank minimization with hybrid smoothness regularization for visual data recovery". In: *Applied Mathematical Modelling* 81 (2020), pp. 711–726.

Missing data in traffic data[15] (spatiotemporal)

- Truncated Nuclear Norm New formulation for tensor rank approximation (Same as Tensor Train Rank[14])

$$\min_{\mathrm{M},\mathcal{X}_1,\mathcal{X}_2,\mathcal{X}_3} \sum_{k=1}^{3} \alpha_k \left\|\mathcal{X}_{k(k)}\right\|_{r_k,*}$$
$$s.t. \begin{cases} \mathcal{X}_k = \mathcal{M}, k = 1,2,3, \\ \mathcal{P}_\Omega(\mathcal{M}) = \mathcal{P}_\Omega(\mathcal{Y}), \end{cases}$$

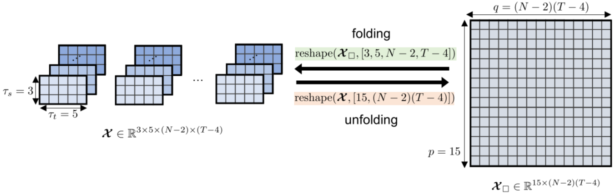
- LRTC-TNN (Extension of matrix form)
  `https://nbviewer.org/github/xinychen/transdim/blob/master/experiments/Imputation-LRTC-TNN.ipynb`
- Gaps: How to reshape traffic data (Traffic data analysis)? Roughness prior, How to tune the hyper-parameters (θ)

[14] Jing-Hua Yang et al. "Tensor train rank minimization with hybrid smoothness regularization for visual data recovery". In: *Applied Mathematical Modelling* 81 (2020), pp. 711–726.

[15] Xinyu Chen, Jinming Yang, and Lijun Sun. "A nonconvex low-rank tensor completion model for spatiotemporal traffic data imputation". In: *Transportation Research Part C: Emerging Technologies* 117 (2020), p. 102673.

Missing data in Traffic Speed Estimation[16] (spatiotemporal)

- Truncated Nuclear Norm New formulation for tensor rank approximation in traffic speed estimation
- Spatiotemporal Hankel tensor transformation



[16] Xudong Wang et al. *Low-Rank Hankel Tensor Completion for Traffic Speed Estimation*. 2021. arXiv: 2105.11335.

Thank You!