Low Rank Tensor Completion and Applications

Zheng Tan, Wensi Ai, Siyu Huang, Zhengtong Liu

https://github.com/ZT220501/Math-270-Tensor-Completion

Introduction!

Low Rank Tensor Completion

- Tensor: $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$
- Operations on matrix can be generalized to tensor
- Denote tensor after Fourier transformation along the third dimension to be $\hat{\mathcal{X}}$
- Tubal rank: maximum rank of frontal slice after Fourier transform along the third dimension
- Sampling method: fully random, random tubes $\mathcal{X}(i, j, :)$
- Four classical/state of the art methods are used

Tensor Nuclear Norm Alternating Direction Method of Multiplier (TNN-ADMM)

Background

Nuclear norm:

Given $X = U\Sigma V^*$

$\text{rank}(X) = \|\mathbf{s}\|_0$, where $\mathbf{s} = (\sigma_1, \cdots, \sigma_n)$

Define $\|X\|_* = \|\mathbf{s}\|_1 = \sum_{i=1}^{N} \sigma_i$

Background

Lagrangian:

$$\min f(\mathbf{x}) \quad \text{s.t.} \quad A\mathbf{x} = \mathbf{b}$$

$$L(\mathbf{x}; \mathbf{u}) = f(\mathbf{x}) + \langle \mathbf{u}, A\mathbf{x} - \mathbf{b} \rangle$$

Lagrange dual problem:

$$d^* := \max_{\mathbf{u} \in \mathbb{R}^M} g(\mathbf{u}).$$

$$g(\mathbf{u}) \quad = \quad \inf_{\mathbf{x} \in \mathbb{R}^N} L(\mathbf{x}; \mathbf{u})$$

Augmented Lagrangian:

$$L_\rho(\mathbf{x}; \mathbf{u}) = f(\mathbf{x}) + \langle A\mathbf{x} - \mathbf{b}, \mathbf{u} \rangle + \frac{\rho}{2} \|A\mathbf{x} - \mathbf{b}\|_2^2$$

...

$$g_\rho = \inf_{\mathbf{x}} L_\rho(\mathbf{x}; \mathbf{u})$$

ADMM

Unconstrained formulation:

$$\lambda\|X\|_* + \frac{1}{2}\|A \circ X - M\|_F^2$$

Splitting form:

$$\lambda\|X\|_* + \frac{1}{2}\|A \circ Y - M\|_F^2 \quad \text{s.t.} \quad X = Y$$

Augmented Lagrangian:

$$\lambda\|X\|_* + \frac{1}{2}\|A \circ Y - M\|_F^2 + \rho\langle W, X - Y\rangle + \frac{\rho}{2}\|X - Y\|_F^2$$

$$\lambda\|X\|_* + \frac{1}{2}\|A \circ Y - M\|_F^2 + \rho\langle W, X - Y\rangle + \frac{\rho}{2}\|X - Y\|_F^2$$

ADMM

Updates x, y in an alternating fashion, such that each subproblem has a closed-form solution and can be calculated efficiently

X: $$\arg\min_X \lambda\|X\|_* + \frac{\rho}{2}\|X - Y + W\|_F^2$$

Y: $$\arg\min_Y \frac{1}{2}\|A \circ Y - M\|_F^2 + \frac{\rho}{2}\|X - Y + W\|_F^2$$

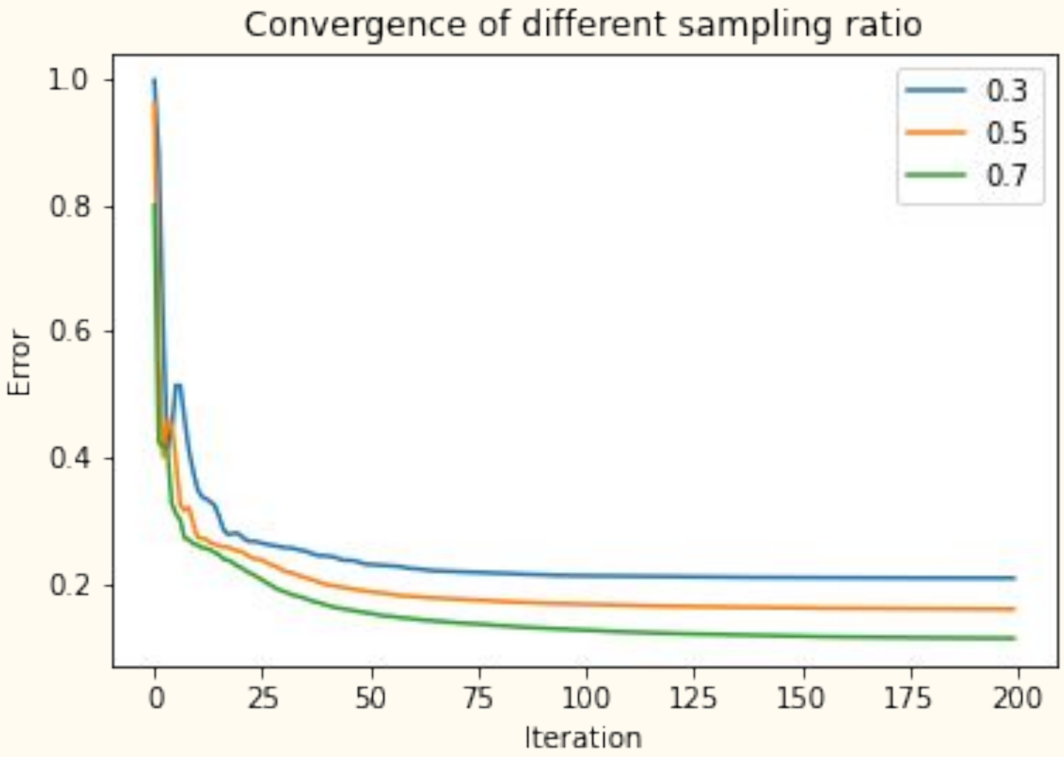
W: $$W^{k+1} = W^k + (X^{k+1} - Y^{k+1})$$

$$\arg\min_{X} \lambda \|X\|_* + \frac{\rho}{2} \|X - Y + W\|_F^2$$

Let $Y - W = U\Sigma V^*$, apply `shrink` on each diagonal element of Σ, i.e., $\bar{\mathbf{s}} = \text{shrink}(\mathbf{s}, \lambda/\rho)$, then the closed-form solution of X is given by $X = U\bar{\Sigma}V^*$.

$$\text{shrink}(\mathbf{v}, \mu) = \begin{cases} \mathbf{v} - \mu & \mathbf{v} > \mu \\ 0 & |\mathbf{v}| \leq \mu \\ \mathbf{v} + \mu & \mathbf{v} < -\mu \end{cases}$$

Experiment



Convergence of different sampling ratio
1.0
0.8
0.6
0.4
0.2
Error
0
25
50
75
100
125
150
175
200
Iteration
0.3
0.5
0.7

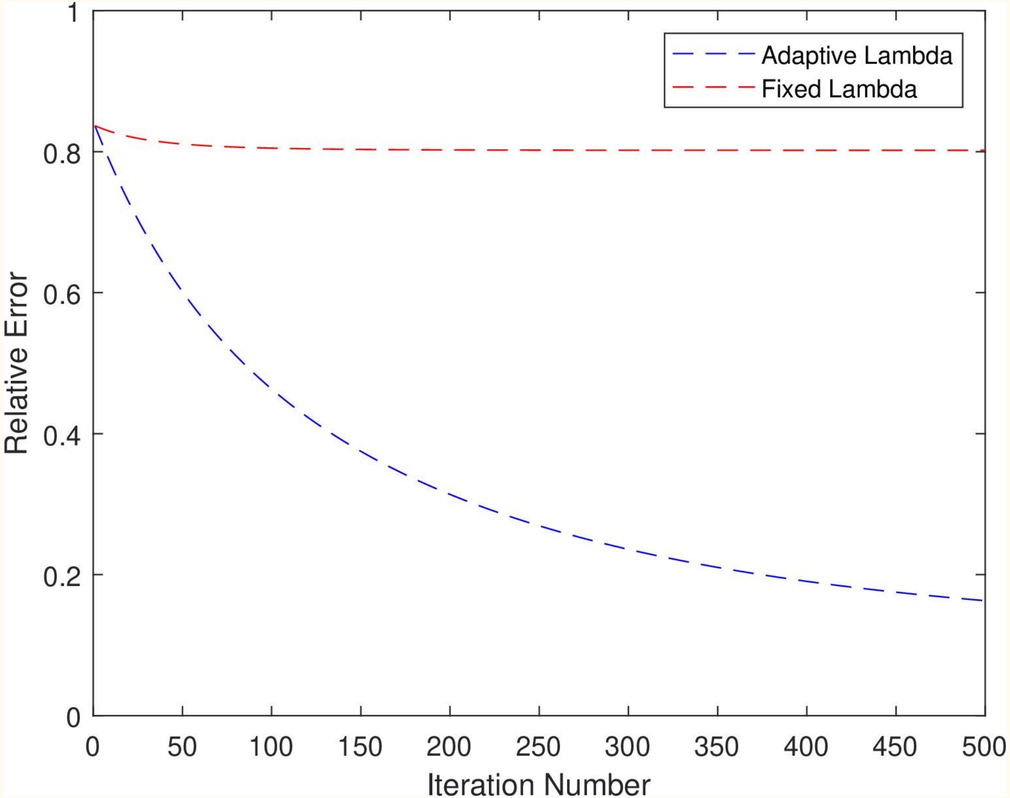
Average Rank Approximation

Ideas

- Given a complete tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, seek to minimize its averaged rank $\frac{1}{n_3}\sum_{i=1}^{n_3} rank(\hat{\mathcal{X}}(:,:,i))$ subject to $\mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{M})$
- Can be alternatively written as minimizing $\sum_{i,k} \delta(\hat{\mathcal{S}}(i,i,k))$ with the same constraint.
- Approximate the Kronecker delta by convex relaxation

$$\phi_\lambda(x) := \min\{1, \frac{1}{2\lambda}x^2\}$$

Synthetic Result



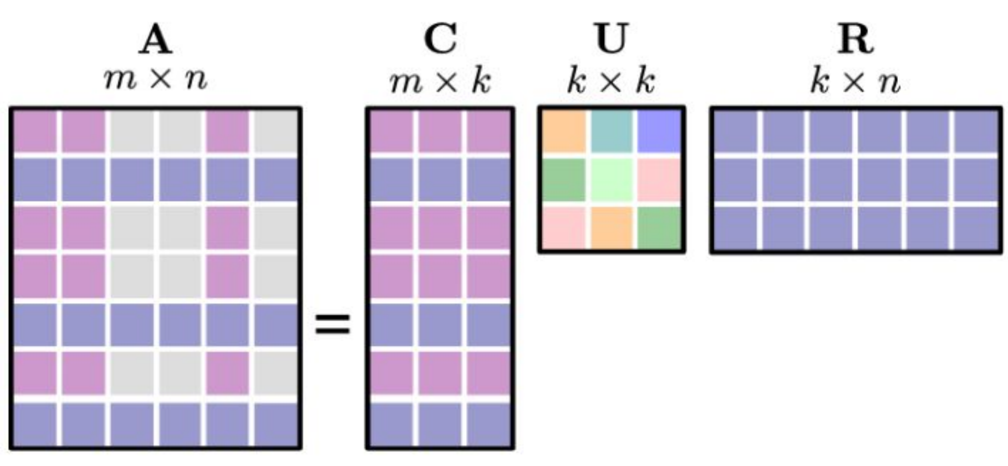
CUR Approximation

CUR-Based Low-Rank Approximation of Matrices

$$\text{Let } A \in \mathcal{R}^{m\times n}, \text{ and } C \in \mathcal{R}^{m\times k},\ R \in \mathcal{R}^{k\times n} \text{ are selected rows and columns, respectively.}$$

$$U \in \mathcal{R}^{k\times k} \text{ is the intersection matrix.}$$

$$A \cong CU^{+}R.$$

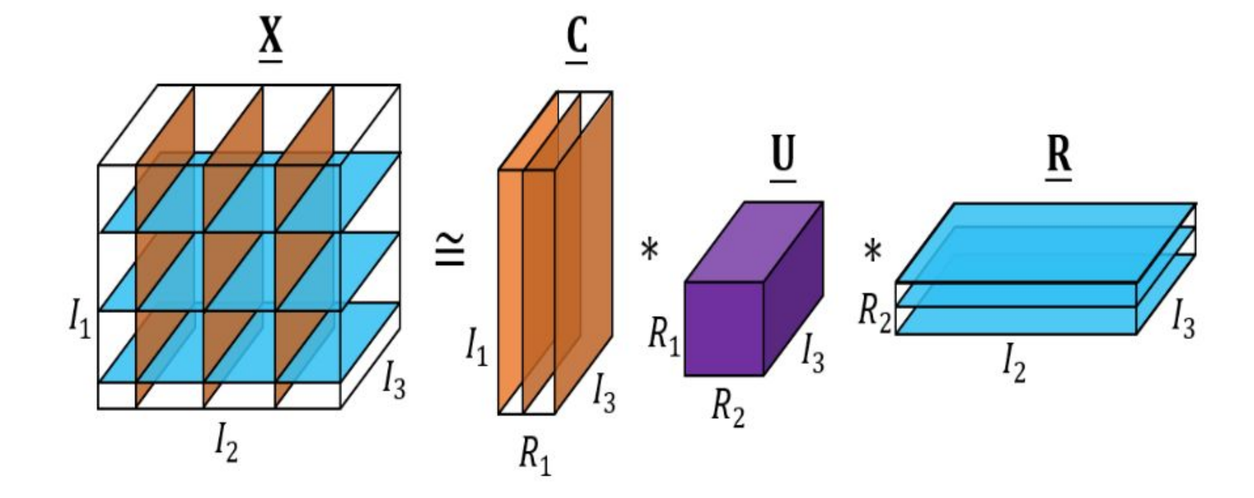


Extend CUR to Tensors…..

$Let\ \mathcal{X} \in \mathcal{R}^{I_1 \times I_2 \times I_3},\ and\ C \in \mathcal{R}^{I_1 \times L_1 \times I_3},\ R \in \mathcal{R}^{L_2 \times I_2 \times I_3}\ are\ lateral\ and\ horizontal\ slices,\ respectively.$

$$U \in \mathcal{R}^{L_1 \times L_2 \times I_3}\ is\ the\ intersection\ tensor.$$

$$\mathcal{X} \cong C\ * U^{+}\ * R,\ where\ *\ is\ tubal\ product.$$



Key Idea

Let X *be our data with missing entries, and* Ω *be the index set that denotes missing entries.*

$$Update\ \mathcal{X}\ by\ \mathcal{P}_{\Omega}(\mathcal{X}) = \mathcal{P}_{\Omega}(C * U^{+} * R),$$

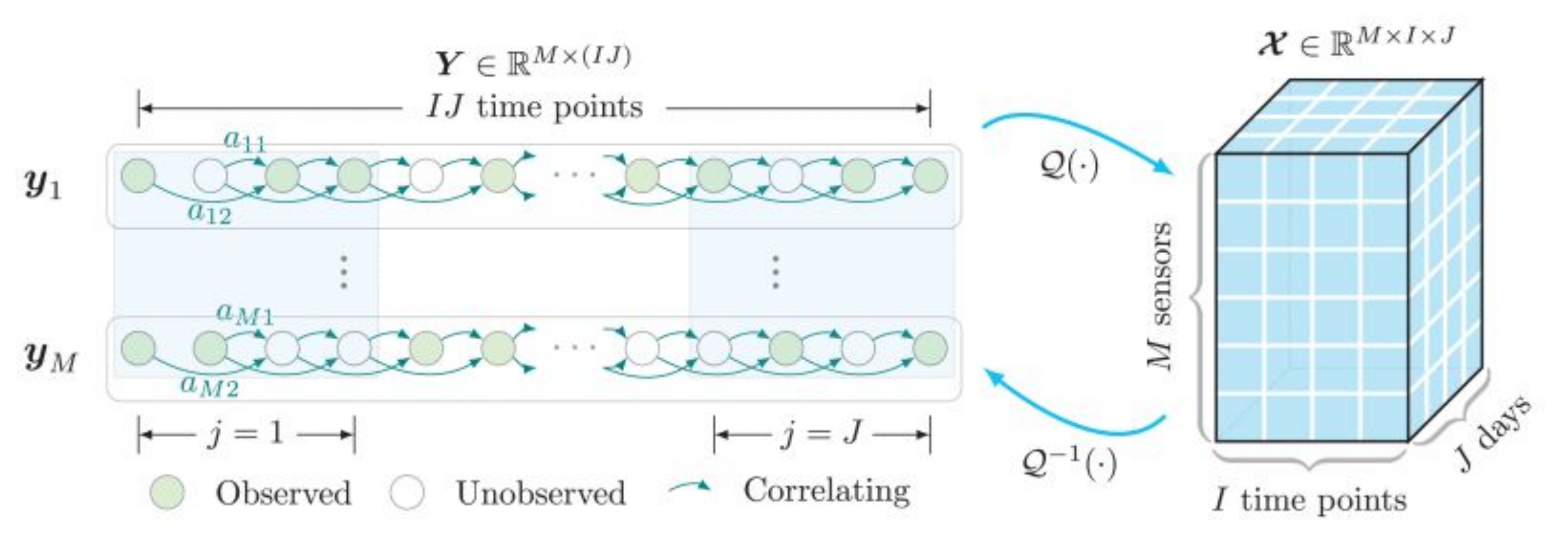
$$and\ keep\ \mathcal{P}_{\Omega^{\perp}}(\mathcal{X}))\ unchanged.$$

Low-rank Autoregressive Tensor Completion (LATC)

Key Ideas

1. Minimize the tensor rank for global consistency

2. Minimize temporal variation for local consistency (temporal consistency)



Math Formulation

The *temporal variation for local consistency* of a time series matrix $\mathbf{Z}$ with a coefficient matrix $\mathbf{A} \in \mathbb{R}^{M \times d}$ and a time lag set $\mathcal{H} = \{h_1, \cdots, h_d\}$ is defined as

$$\|\mathbf{Z}\|_{\mathbf{A},\mathcal{H}} = \sum_{m,t} (z_{m,t} - \sum_i a_{m,i} z_{m,t-h_i})^2$$

This regularization ensures that the individual time series z_m should be fitted in an autoregressive pattern, and z_m should be consistent with the previous d time series.

Math Formulation

The **Low-Rank Autogressive Tensor Completion (LATC)** refers to the optimization model

$$\min_{\mathcal{X},\mathbf{Z},\mathbf{A}} \|\mathcal{X}\|_{r,*} + \frac{\lambda}{2}\|\mathbf{Z}\|_{\mathbf{A},\mathcal{H}} \quad \text{s.t.} \mathcal{X} = \mathcal{Q}(\mathbf{Z}),\ \mathcal{P}_{\Omega}(\mathbf{Z}) = \mathcal{P}_{\Omega}(\mathbf{Y})$$

where $\|\mathcal{X}\|_{r,*}$ denotes the truncated nuclear norms with rank r; $\mathbf{Y} \in \mathbb{R}^{M\times(IJ)}$ denotes the partially observed time series matrix; $\mathcal{X} = \mathcal{Q}(\mathbf{Y}) \in \mathbb{R}^{M\times I\times J}$ is the tensorization of $\mathbf{Y}$ (M: number of sensors, I: the number of time points per day, J: number of days);

$\mathcal{P}_{\Omega}(\mathbf{Z}) = \mathcal{P}_{\Omega}(\mathbf{Y})$ ensures that the values of the observed entries stay the same, with

$$[\mathcal{P}_{\Omega}(\mathbf{Y})]_{m,n} = \begin{cases} y_{m,n}, & \text{if } (m,n) \in \Omega, \\ 0, & \text{otherwise,} \end{cases}$$

where $m = 1, \cdots, M$ and $n = 1 \cdots IJ$.

Math Formulation

Then the augmented Lagrangian function can be written as

$$\mathcal{L}(\mathcal{X}, \mathbf{Z}, \mathbf{A}, \mathcal{T}) = \|\mathcal{X}\|_{r,*} + \frac{\lambda}{2}\|\mathbf{Z}\|_{\mathbf{A},\mathcal{H}} + \frac{\rho}{2}\|\mathcal{X} - \mathcal{Q}(\mathbf{Z})\|_F^2 + \langle \mathcal{X} - \mathcal{Q}(\mathbf{Z}), \mathcal{T} \rangle$$

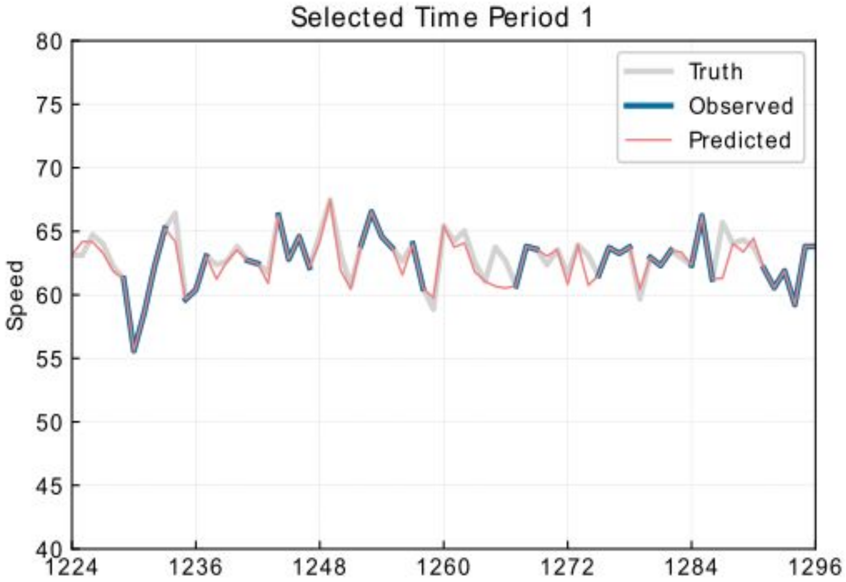
where ρ is the weight parameter of the added Frobenius norm penalty and $\mathcal{T} \in \mathbb{R}^{M \times I \times J}$ is the dual variable.

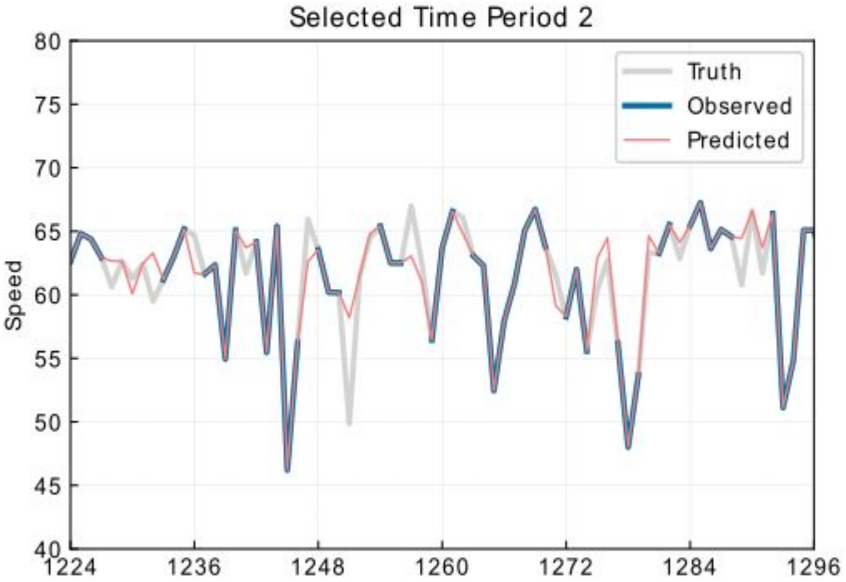
The constraint $\mathcal{P}_\Omega(\mathbf{Z}) = \mathcal{P}_\Omega(\mathbf{Y})$ is kept after each iteration to keep the observation consistency.

Then ADMM can help transform the augmented Lagrangian function into subproblems to update $\mathcal{X}$, $\mathbf{Z}$, and $\mathcal{T}$; the solution can be approximated iteratively.

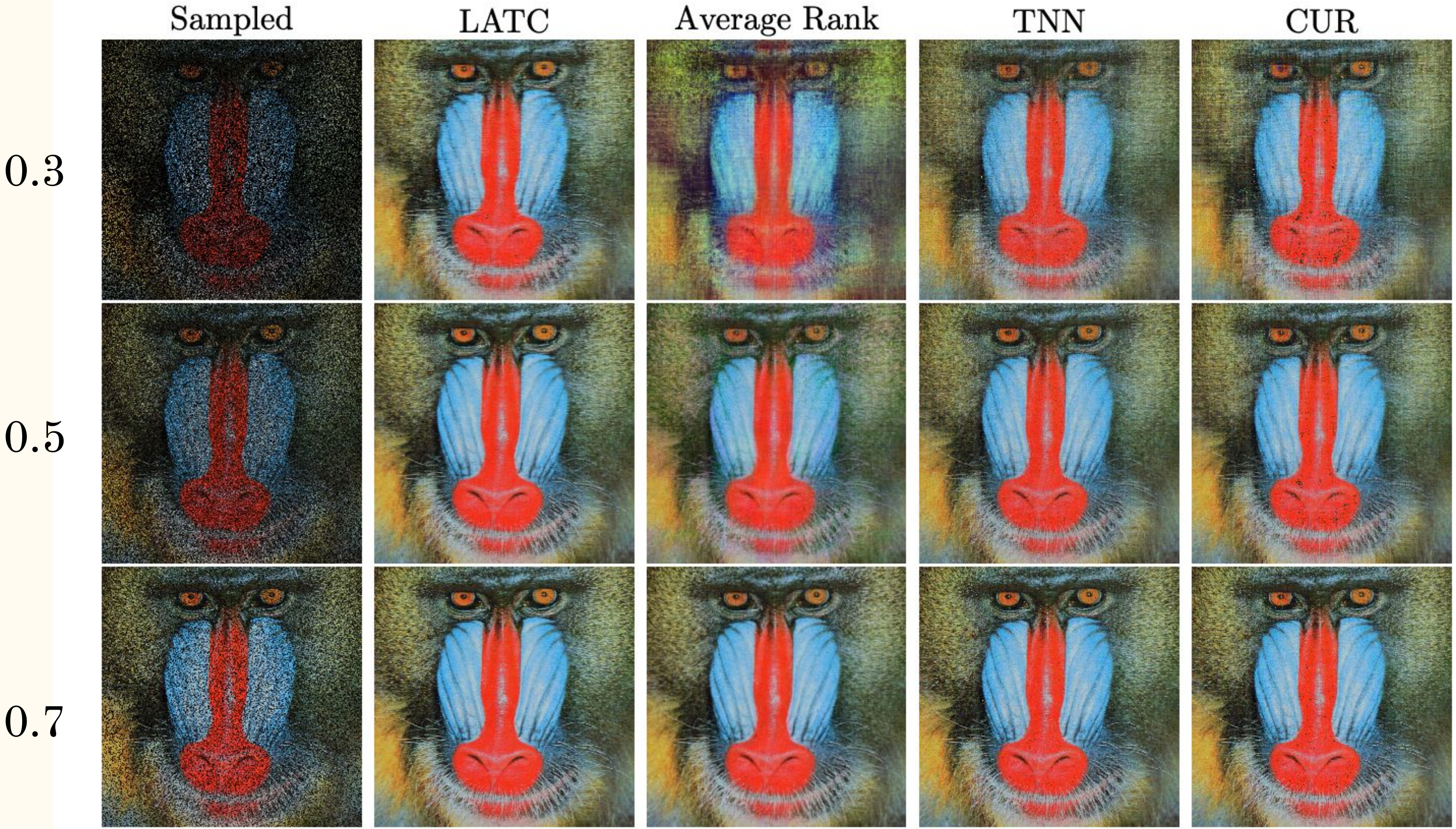
$$\begin{aligned}
\boldsymbol{\mathcal{X}}^{\ell+1,\nu+1} &:= \arg\min_{\boldsymbol{\mathcal{X}}} \ \mathcal{L}(\boldsymbol{\mathcal{X}}, \mathbf{Z}^{\ell+1,\nu}, \mathbf{A}^{\ell}, \boldsymbol{\mathcal{T}}^{\ell+1,\nu}), \\
\mathbf{Z}^{\ell+1,\nu+1} &:= \arg\min_{\mathbf{Z}} \ \mathcal{L}(\boldsymbol{\mathcal{X}}^{\ell+1,\nu+1}, \mathbf{Z}, \mathbf{A}^{\ell}, \boldsymbol{\mathcal{T}}^{\ell+1,\nu}), \\
\boldsymbol{\mathcal{T}}^{\ell+1,\nu+1} &:= \boldsymbol{\mathcal{T}}^{\ell+1,\nu} + \rho(\boldsymbol{\mathcal{X}}^{\ell+1,\nu+1} - \mathcal{Q}(\mathbf{Z}^{\ell+1,\nu+1})),
\end{aligned}$$

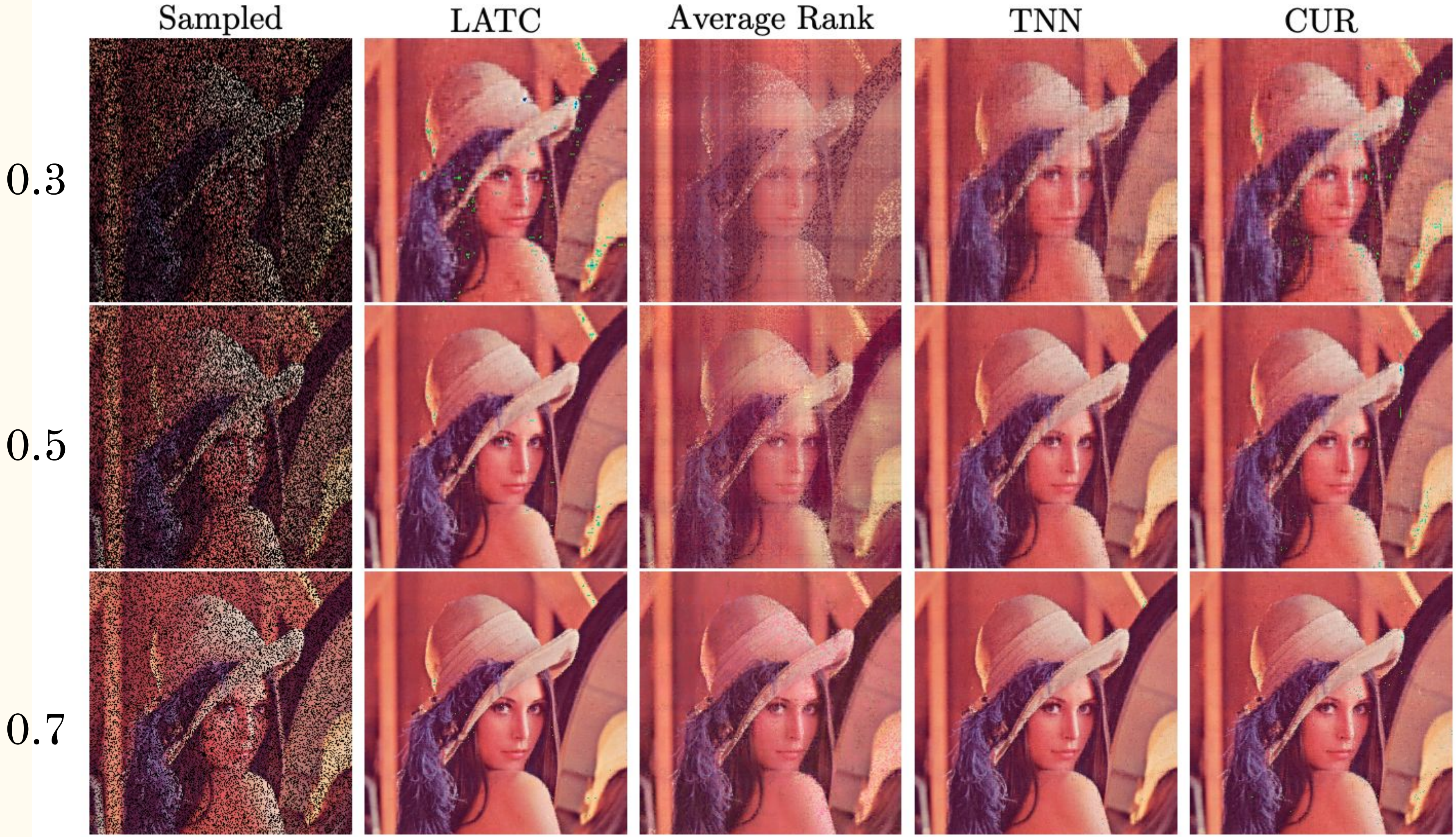
Application to Spatiotemporal Traffic Data





Benchmark Results!


Sampled
LATC
Average Rank
TNN
CUR
0.3
0.5
0.7

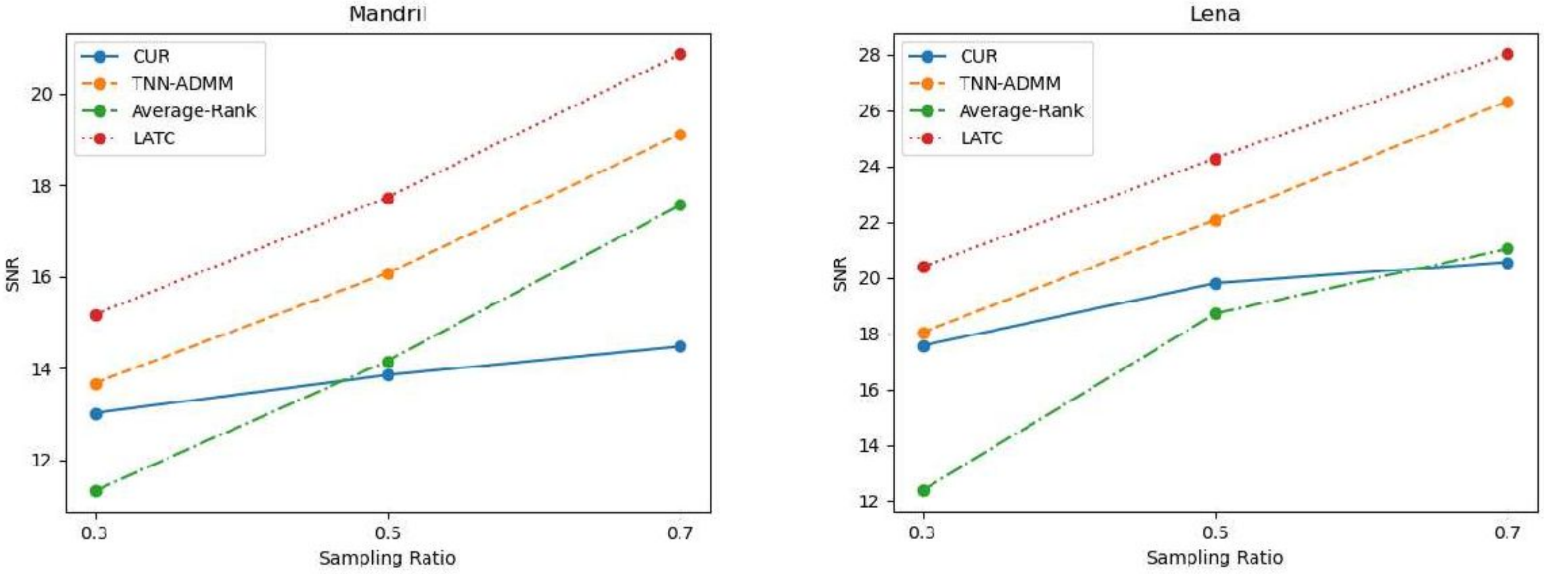


Sampled
LATC
Average Rank
TNN
CUR
0.3
0.5
0.7

Signal-to-Noise Ratio (SNR)

$$\text{SNR} := 20 \log_{10} \frac{||\mathcal{T}||_F}{||\tilde{\mathcal{T}} - \mathcal{T}||_F}$$

where $\tilde{\mathcal{T}}$ is the observed part, and $\mathcal{T}$ is the actual image tensor.



Thanks!