Low Rank Tensor Learning for Spatiotemporal Traffic Data Modeling

Spatiotemporal Traffic Data Imputation (STDI)

Gong Wenwu

Southern University of Science and Technology

June 27, 2023

Outline

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

Traffic Time Series

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

- Multivariate: matrix form, i.e., spatial locations $\times$ time slots, such as Guangzhou speed traffic data [1]
- Multidimensional: tensor form, i.e., zones $\times$ zones $\times$ time slots, such as Abilene data [2]
- Sparsity: insufficient sampling[3]
- Time-varying behavior: time series model[4] and Bayesian factorization[5],[6]
- Spatial similarity

---

[1] https://github.com/xinychen/transdim

[2] https://doi.org/10.5281/zenodo

[3] Xinyu Chen et al. "Low-Rank Autoregressive Tensor Completion for Spatiotemporal Traffic Data Imputation". In: *IEEE Transactions on Intelligent Transportation Systems* (2021), pp. 1–10.

[4] Hsiang-Fu Yu, Nikhil Rao, and Inderjit S. Dhillon. "Temporal regularized matrix factorization for high-dimensional time series prediction". In: *NIPS'16 Proceedings of the 30th International Conference on Neural Information Processing Systems*. Vol. 29. 2016, pp. 847–855.

[5] Liang Xiong et al. "Temporal Collaborative Filtering with Bayesian Probabilistic Tensor Factorization". In: *Proceedings of the 2010 SIAM International Conference on Data Mining (SDM)*. Dec. 2010, pp. 211–222.

[6] Xinyu Chen and Lijun Sun. "Bayesian Temporal Factorization for Multidimensional Time Series Prediction". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021), pp. 1–1.

Traffic data form

For any given partially observed traffic matrix Y whose columns IJ correspond to time slots and rows M correspond to sensors

$$\boldsymbol{Y} = \begin{bmatrix} | & | & & | \\ \boldsymbol{y}_1 & \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_{IJ} \\ | & | & & | \end{bmatrix} \in \mathbb{R}^{M \times (IJ)}$$

- Matrix (road segment $\times$ time series)
- Third-order tensor (road segment $\times$ day $\times$ time point)
- Fourth-order tensor (spatiotemporal)

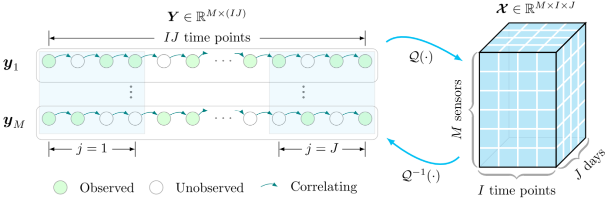
1. Spatiotemporal Traffic Data Modeling
   - Traffic Data Behaviors
   - Tensorization transformer
   - Low-rank tensor learning
2. Related works
3. Baselines
4. Applications

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

Tensorization operator

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

Converts the traffic time series matrix into a third-order tensor[7]

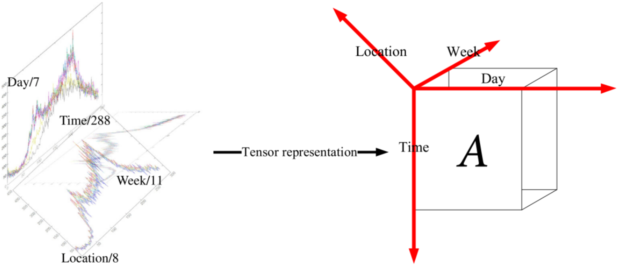


[7] Xinyu Chen et al. "Low-Rank Autoregressive Tensor Completion for Spatiotemporal Traffic Data Imputation". In: *IEEE Transactions on Intelligent Transportation Systems* (2021), pp. 1–10.

Tensorization operator

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

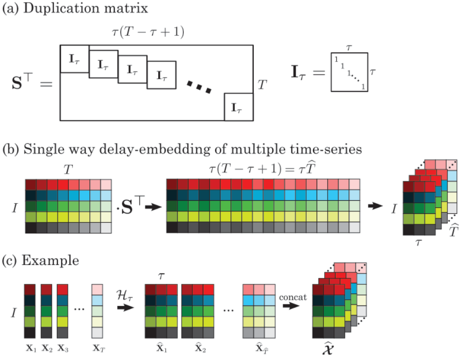
Converts the traffic time series matrix into a fourth-order tensor[8]



[8] 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.

Tensorization operator

Applies MDT[9] /Hankelization[10] operator along the spatial and temporal modes of traffic matrix



[9] Farnaz Sedighin et al. "Matrix and Tensor Completion in Multiway Delay Embedded Space Using Tensor Train, With Application to Signal Reconstruction". In: *IEEE Signal Processing Letters* 27 (2020), pp. 810–814.

[10] Qiquan Shi et al. *Block Hankel Tensor ARIMA for Multiple Short Time Series Forecasting*. 2020. arXiv: 2002.12135 [cs.LG].

Time line: tensor rank minimization

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

Low-rank tensor learning [11]

- **Candès & Recht'09:** Convex nuclear norm minimization for matrix completion.

$$\min_{\boldsymbol{X}} \; \|\boldsymbol{X}\|_* \quad \text{s.t. } \mathcal{P}_\Omega(\boldsymbol{X}) = \mathcal{P}_\Omega(\boldsymbol{Y})$$

- **Cai, Candès & Shen'10:** Singular value thresholding algorithm.

$$\begin{cases} \boldsymbol{X}^\ell = \mathcal{D}_\tau\left(\boldsymbol{Z}^{\ell-1}\right) \\ \boldsymbol{Z}^\ell = \boldsymbol{Z}^{\ell-1} + \delta_\ell \mathcal{P}_\Omega\left(\boldsymbol{Y} - \boldsymbol{X}^\ell\right) \end{cases}$$

- **Zhang et al.'12:** Nonconvex truncated nuclear norm minimization.
- **Liu et al.'13:** Convex nuclear norm minimization for tensor completion.

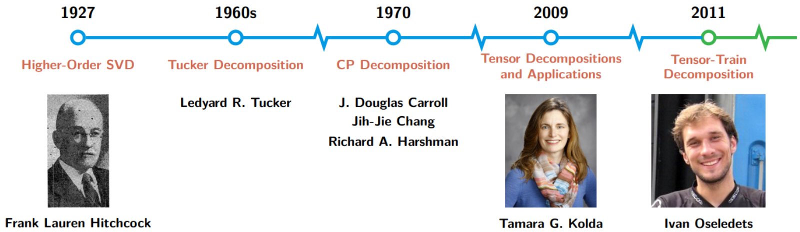
$$\min_{\boldsymbol{\mathcal{X}}} \; \|\boldsymbol{\mathcal{X}}\|_* \quad \text{s.t. } \mathcal{P}_\Omega(\boldsymbol{\mathcal{X}}) = \mathcal{P}_\Omega(\boldsymbol{\mathcal{Y}})$$

- **Lu, Peng & Wei'19:** Tensor nuclear norm induced by linear transform.

[11] https://xinychen.github.io/slides/phd_project_22summer.pdf

Time line: low-rank tensor approximation

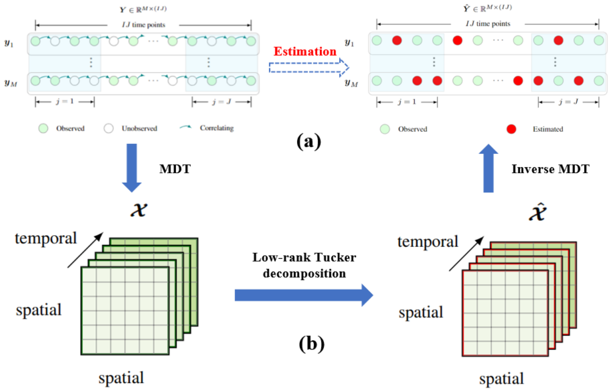
Low-rank tensor learning [12]



[12] https://xinychen.github.io/slides/phd_project_22summer.pdf

Framework

Tucker-based low-rank tensor approximation



1. Spatiotemporal Traffic Data Modeling
   - Traffic Data Behaviors
   - Tensorization transformer
   - Low-rank tensor learning
2. Related works
   - Low-rankness
   - Spatiotemporal priors
3. Baselines
   - Low-rankness
   - Temporal variation
   - Spatial regularization
4. Applications
   - Results Visualization

Nuclear norm minimization

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

Tensor representation : reshapes traffic time series into a third-order tensor[13],[14]; reshapes into fourth-order tensor[15],[16]

- Convex nuclear norm minimization
- Tubal rank minimization

13 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.

14 Kun Xie et al. "Accurate Recovery of Internet Traffic Data: A Sequential Tensor Completion Approach". In: *IEEE/ACM Transactions on Networking* 26.2 (2018), pp. 793–806.

15 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.

16 Xudong Wang et al. "Low-Rank Hankel Tensor Completion for Traffic Speed Estimation". In: *IEEE Transactions on Intelligent Transportation Systems* 24.5 (2023), pp. 4862–4871.

Low-rank tensor approximation

- Low-Tucker rank: approximate spatiotemporal traffic data using bilinear or multi-linear factorization models with a predefined rank
  - Regularization[17],[18]
  - Toeplitz matrix[19]
  - Spatio-temporal predictor[20]
- Tensor factorization[21] and Tensor-train decomposition[22]

---

[17] Xinyu Chen, Zhaocheng He, and Jiawei Wang. "Spatial-temporal traffic speed patterns discovery and incomplete data recovery via SVD-combined tensor decomposition". In: *Transportation Research Part C: Emerging Technologies* 86 (2018), pp. 59–77.

[18] Yuankai Wu et al. "A Fused CP Factorization Method for Incomplete Tensors". In: *IEEE Transactions on Neural Networks and Learning Systems* 30.3 (2019), pp. 751–764.

[19] Yang Wang et al. "Traffic Data Reconstruction via Adaptive Spatial-Temporal Correlations". In: *IEEE Transactions on Intelligent Transportation Systems* 20.4 (2019), pp. 1531–1543.

[20] Koh Takeuchi, Hisashi Kashima, and Naonori Ueda. "Autoregressive Tensor Factorization for Spatio-Temporal Predictions". In: *2017 IEEE International Conference on Data Mining (ICDM)*. 2017, pp. 1105–1110.

[21] Pan Zhou et al. "Tensor Factorization for Low-Rank Tensor Completion". In: *IEEE Transactions on Image Processing* 27.3 (2018), pp. 1152–1163.

[22] Zhiyuan Zhang et al. "A tensor train approach for internet traffic data completion". In: *Annals of Operations Research* 06 (2021), pp. 73–84.

1 Spatiotemporal Traffic Data Modeling

2 Related works

Low-rankness

Spatiotemporal priors

3 Baselines

4 Applications

Temporal variation

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

- Time series model:
  - Temporal matrix factorization[23]
  - AR/VAR model[24,25]
- Bayesian model: Matrix[26], Tensor[27,28] factorization

---

[23] Hsiang-Fu Yu, Nikhil Rao, and Inderjit S. Dhillon. "Temporal regularized matrix factorization for high-dimensional time series prediction". In: *NIPS'16 Proceedings of the 30th International Conference on Neural Information Processing Systems*. Vol. 29. 2016, pp. 847–855.

[24] Xinyu Chen et al. "Low-Rank Autoregressive Tensor Completion for Spatiotemporal Traffic Data Imputation". In: *IEEE Transactions on Intelligent Transportation Systems* (2021), pp. 1–10.

[25] Xinyu Chen et al. "Scalable low-rank tensor learning for spatiotemporal traffic data imputation". In: *Transportation Research Part C: Emerging Technologies* 129 (2021), p. 103226.

[26] Ruslan Salakhutdinov and Andriy Mnih. "Bayesian Probabilistic Matrix Factorization Using Markov Chain Monte Carlo". In: *Proceedings of the 25th International Conference on Machine Learning*. 2008, pp. 880–887.

[27] Liang Xiong et al. "Temporal Collaborative Filtering with Bayesian Probabilistic Tensor Factorization". In: *Proceedings of the 2010 SIAM International Conference on Data Mining (SDM)*. Dec. 2010, pp. 211–222.

[28] Xinyu Chen and Lijun Sun. "Bayesian Temporal Factorization for Multidimensional Time Series Prediction". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021), pp. 1–1.

Spatial similarity

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

- Graph convolutional neural networks[29]
- Graph-based regularization[30], graph Laplacian matrix[31]
- Graph Fourier transform[32] (graph-tensor product)

[29] Zahraa Al Sahili and Mariette Awad. *Spatio-Temporal Graph Neural Networks: A Survey*. 2023.

[30] Hsiang-Fu Yu, Nikhil Rao, and Inderjit S. Dhillon. "Temporal regularized matrix factorization for high-dimensional time series prediction". In: *NIPS'16 Proceedings of the 30th International Conference on Neural Information Processing Systems*. Vol. 29. 2016, pp. 847–855.

[31] Yang Wang et al. "Traffic Data Reconstruction via Adaptive Spatial-Temporal Correlations". In: *IEEE Transactions on Intelligent Transportation Systems* 20.4 (2019), pp. 1531–1543.

[32] Lei Deng et al. "Graph Spectral Regularized Tensor Completion for Traffic Data Imputation". In: *IEEE Transactions on Intelligent Transportation Systems* 23.8 (2022), pp. 10996–11010.

STH-LRTC[33]

- Task: Spatiotemporal Traffic Data (Matrix) imputation (STDI) using sparse observations (high-level missing)
- Novelty: fourth-order Hankel structured tensor, balanced spatiotemporal unfolding to approximate the tensor rank

$$\min_{\mathcal{X}} \|\mathcal{X}_{\square}\|_{r,*}$$
$$\text{s.t. } \begin{cases} \boldsymbol{X} = \mathcal{H}_{\tau_s,\tau_t}(\boldsymbol{Z}), \\ \boldsymbol{Z}_{\Omega} = \boldsymbol{Y}_{\Omega} \end{cases}$$

where

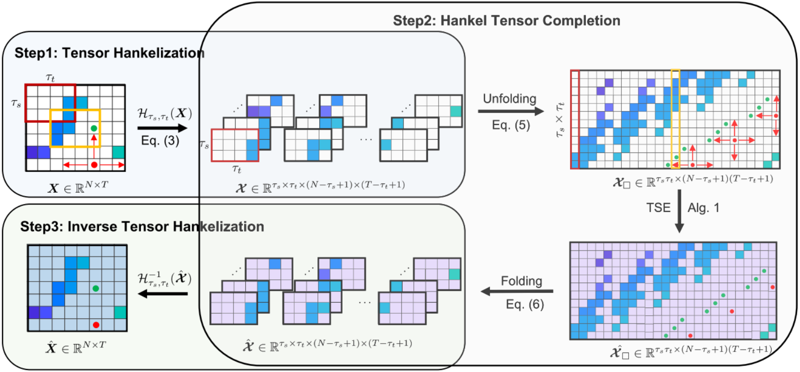
$$\boldsymbol{\mathcal{X}}_{\square} = \text{reshape}(\mathcal{X}, [p, q]),$$

$p = \tau_s \times \tau_t$ and $q = (N - \tau_s + 1) \times (T - \tau_t + 1)$.

[33] Xudong Wang et al. "Low-Rank Hankel Tensor Completion for Traffic Speed Estimation". In: *IEEE Transactions on Intelligent Transportation Systems* 24.5 (2023), pp. 4862–4871.

STH-LRTC

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE



1. Spatiotemporal Traffic Data Modeling
2. Related works
3. Baselines
   - Low-rankness
   - Temporal variation
   - Spatial regularization
4. Applications

LATC[35]

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

- Task: STDI, three Missing scenarios
- Novelty: temporal variation (AR)

$$\begin{aligned}&\min_{\mathcal{X},\boldsymbol{Z},\boldsymbol{A}} \|\mathcal{X}\|_{r,*} + \frac{\lambda}{2}\|\boldsymbol{Z}\|_{\boldsymbol{A},\mathcal{H}}\\ &s.t.\ \mathcal{X} = \mathcal{Q}(\boldsymbol{Z}),\ \mathcal{P}_\Omega(\boldsymbol{Z}) = \mathcal{P}_\Omega(\boldsymbol{Y}),\end{aligned}$$

- Extension: slice matrix SVD[34]

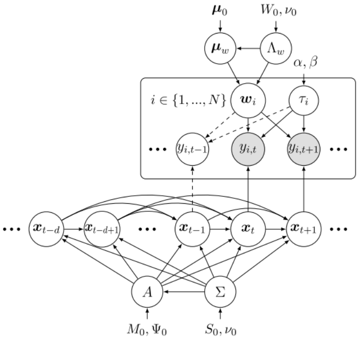
$$\begin{aligned}&\min_{\mathcal{X},\boldsymbol{Z}} \|\boldsymbol{\mathcal{X}}\|_* + \frac{\lambda}{2}\sum_t \|\boldsymbol{z}_t - \boldsymbol{z}_{t-1}\|_2^2\\ &\text{s.t.} \begin{cases}\mathcal{X} = \mathcal{Q}(\boldsymbol{Z}),\\ \mathcal{P}_\Omega(\boldsymbol{Z}) = \mathcal{P}_\Omega(\boldsymbol{Y})\end{cases}\end{aligned}$$

[34] Xinyu Chen et al. "Scalable low-rank tensor learning for spatiotemporal traffic data imputation". In: *Transportation Research Part C: Emerging Technologies* 129 (2021), p. 103226.

[35] Xinyu Chen et al. "Low-Rank Autoregressive Tensor Completion for Spatiotemporal Traffic Data Imputation". In: *IEEE Transactions on Intelligent Transportation Systems* (2021), pp. 1–10.

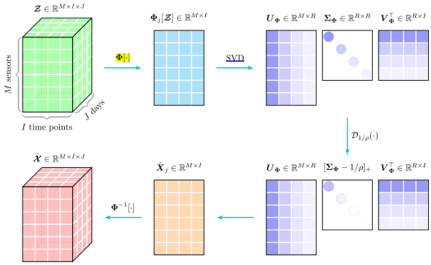
BGCP[36]

- General idea: Integrates low-rank matrix/tensor factorization and vector autoregressive (VAR) process into a single probabilistic graphical model
- Novelty: Bayesian: perform probabilistic predictions and produce uncertainty estimates



[36] Xinyu Chen and Lijun Sun. "Bayesian Temporal Factorization for Multidimensional Time Series Prediction". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021), pp. 1–1.

LSTC-Tubal[37]



Algorithm 1: imputer$(\boldsymbol{Y}, \rho, \lambda)$

Initialize $\boldsymbol{\mathcal{T}}^0$ as zeros. Set $\mathcal{P}_\Omega(\boldsymbol{Z}^0) = \mathcal{P}_\Omega(\boldsymbol{Y})$ and $\ell = 0$.
while *not converged* **do**
  Update ρ by $\rho = \min\{1.05 \times \rho, \rho_{\max}\}$;
  for $j = 1$ **to** J **do**
    Update $\boldsymbol{X}_j^{\ell+1}$ by Eq. (13);
  Update $\boldsymbol{Z}^{\ell+1}$ by Eq. (20);
  Update $\boldsymbol{\mathcal{T}}^{\ell+1}$ by Eq. (8);
  Transform the observation information by letting $\mathcal{P}_\Omega(\boldsymbol{Z}^{\ell+1}) = \mathcal{P}_\Omega(\boldsymbol{Y})$;
  $\ell := \ell + 1$;
return recovered matrix $\hat{\boldsymbol{X}}$.

[37] Xinyu Chen et al. "Scalable low-rank tensor learning for spatiotemporal traffic data imputation". In: *Transportation Research Part C: Emerging Technologies* 129 (2021), p. 103226.

TT-SVD[38]

明德求是 日新自强
VIRTUE | TRUTH | ADVANCE

- General idea: Toeplitz matrices for unfolding tensor
- Novelty: closed-form solutions

[38] Zhiyuan Zhang et al. "A tensor train approach for internet traffic data completion". In: *Annals of Operations Research* 06 (2021), pp. 73–84.

LCR[39]

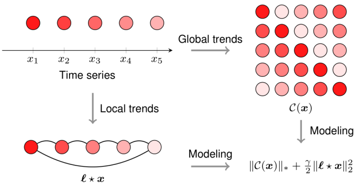
- General idea: Characterizing both global and local trends in STDI
- Novelty: Laplacian kernel to temporal regularization for characterizing local trends; a fast Fourier transform (FFT) solution in a relatively low time complexity

Laplacian Convolutional Representation (LCR)

For any partially observed time series $\boldsymbol{y} \in \mathbb{R}^T$ with observed index set Ω, LCR utilizes circulant matrix and Laplacian kernel to characterize **global and local trends** in time series, respectively, i.e.,

$$\begin{aligned} \min_{\boldsymbol{x}} \; & \|\mathcal{C}(\boldsymbol{x})\|_* + \gamma \cdot \mathcal{R}_\tau(\boldsymbol{x}) \\ \text{s.t. } & \|\mathcal{P}_\Omega(\boldsymbol{x} - \boldsymbol{y})\|_2 \leq \epsilon \end{aligned}$$

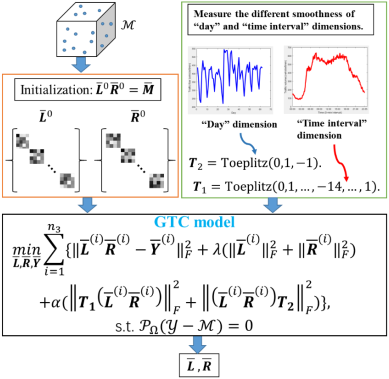
where $\mathcal{C} : \mathbb{R}^T \to \mathbb{R}^{T \times T}$ denotes the circulant operator. $\|\cdot\|_*$ denotes the nuclear norm of matrix, namely, the sum of singular values.



[39] Xinyu Chen et al. *Laplacian Convolutional Representation for Traffic Time Series Imputation*. 2022. arXiv: 2212.01529 [cs.LG].

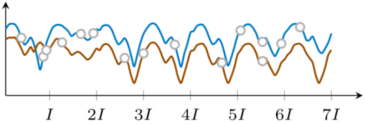
GT-SVD[40]

- General idea: exploits the underlying topological structure of the traffic and constructs temporal regularization
- Novelty: introduces graph Fourier transform and adopts the high-order Toeplitz matrices

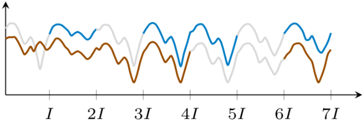


[40] Lei Deng et al. "Graph Spectral Regularized Tensor Completion for Traffic Data Imputation". In: *IEEE Transactions on Intelligent Transportation Systems* 23.8 (2022), pp. 10996–11010.

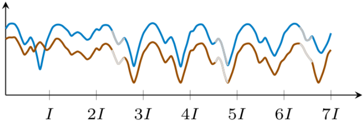
Missing data pattern[41]



(a) Random missing (RM).



(b) Non-random missing (NM).

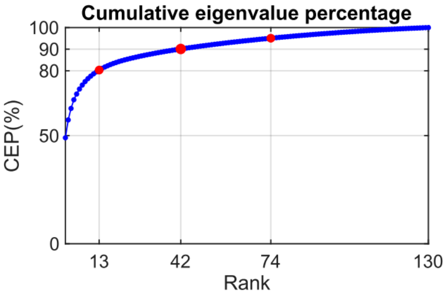


(c) Blackout missing (BM).

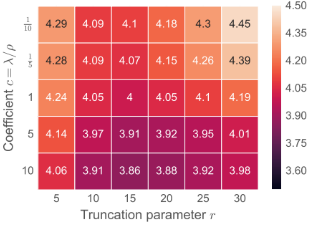
[41] Xinyu Chen et al. “Low-Rank Autoregressive Tensor Completion for Spatiotemporal Traffic Data Imputation”. In: *IEEE Transactions on Intelligent Transportation Systems* (2021), pp. 1–10.

Cumulative eigenvalue percentage

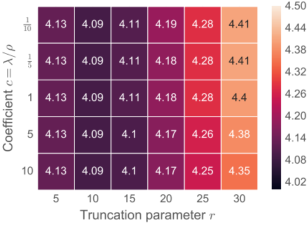
A few leading singular values have a significant contribution, e.g., the first 42 singular values cover 90% of all singular values, showing the low-rank feature.



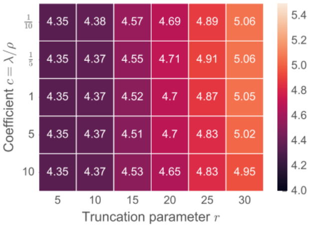
Heatmaps of imputation RMSE values



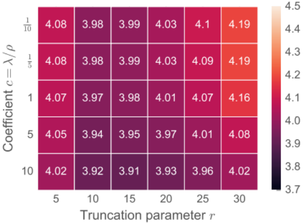
(c) 90%, RM.



(d) 30%, NM.

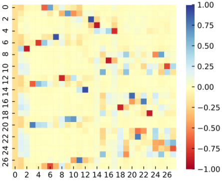


(e) 70%, NM.

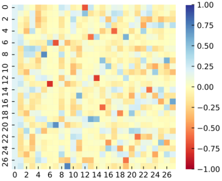


(f) 30%, BM.

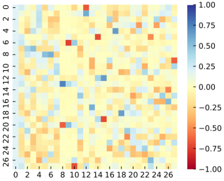
Heatmaps of unitary transform matrices (Slice matrix)



(a) The initialized one.



(b) At the 10th iteration.

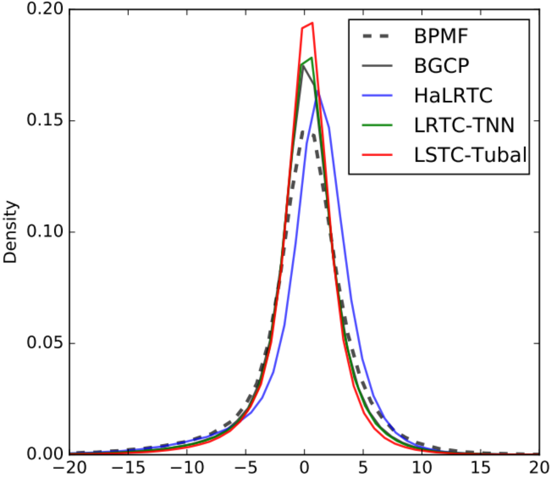


(c) At the 20th iteration.

Performance comparison

Data	Missing	LATC	LAMC	LRTC-TNN	BTMF	SPC
(G)	30%, RM	**5.71/2.54**	9.51/4.04	6.99/3.00	7.54/3.27	7.37/5.06
	70%, RM	**7.22/3.18**	10.40/4.37	8.38/3.59	8.75/3.73	8.91/4.44
	90%, RM	**9.11/3.86**	11.65/4.79	9.55/4.05	10.02/4.21	10.60/4.85
	30%, NM	9.63/4.09	10.11/4.23	9.61/**4.07**	10.32/4.33	**9.13**/5.29
	70%, NM	10.37/4.35	11.15/4.60	**10.36/4.34**	11.36/4.85	11.15/5.17
	30%, BM-6	**9.23/3.91**	12.15/5.17	9.45/3.97	12.43/7.04	11.14/5.13
(H)	30%, RM	19.12/24.97	22.65/42.94	**18.87/24.90**	22.37/28.66	19.82/26.21
	70%, RM	20.25/28.25	25.30/51.26	**20.07/28.13**	25.65/32.23	21.02/31.91
	90%, RM	24.32/**34.44**	32.30/66.13	**23.46**/35.84	31.51/46.24	24.97/49.68
	30%, NM	**19.93/47.38**	22.93/67.08	19.94/50.12	25.61/77.00	27.46/68.56
	70%, NM	24.30/47.30	29.23/63.95	**23.88/45.06**	34.50/70.11	46.86/98.81
	30%, BM-6	21.93/28.64	30.78/66.03	**21.40/27.83**	52.15/57.61	22.49/37.53
(S)	30%, RM	**4.90/3.16**	5.98/3.73	4.99/3.20	5.91/3.72	5.92/3.62
	70%, RM	**5.96/3.71**	8.02/4.70	6.10/3.77	6.47/3.98	7.38/4.30
	90%, RM	**7.47/4.51**	10.56/5.91	8.08/4.80	8.17/4.81	9.75/5.31
	30%, NM	7.11/4.33	6.99/4.25	**6.85/4.21**	9.26/5.36	8.87/4.99
	70%, NM	9.46/5.42	9.75/5.60	**9.23/5.35**	10.47/6.15	11.32/5.92
	30%, BM-12	**9.44/5.36**	27.05/13.66	9.52/5.41	14.33/13.60	11.30/5.84

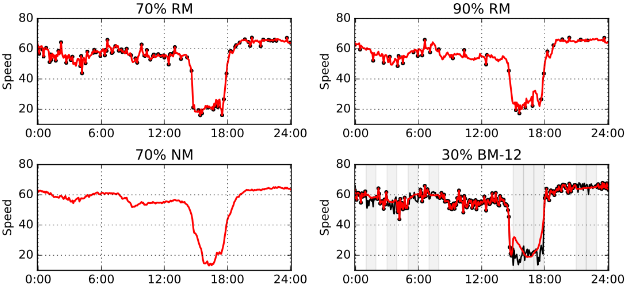
Residuals



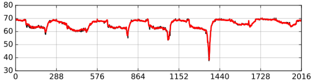
BPMF
BGCP
HaLRTC
LRTC-TNN
LSTC-Tubal
Density
0.20
0.15
0.10
0.05
0.00
−20
−15
−10
−5
0
5
10
15
20

Imputed values

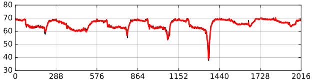
明德求是 日新自强
VIRTUE | TRUTH | ADVANCE



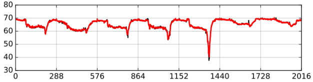
Imputed values



(a) The 1st time series.

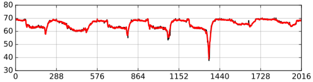


(b) The 2nd time series.

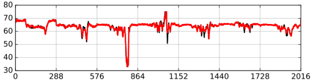


(c) The 3rd time series.

(d) The 4th time series.



(e) The 5th time series.



(f) The 6th time series.

Thank You!