



**POLYTECHNIQUE  
MONTRÉAL**

UNIVERSITÉ  
D'INGÉNIERIE



# **Spatiotemporal Data Modeling Dynamic Patterns and Long-Range Correlations**

**Xinyu Chen**

University of Montreal, Canada

December 4, 2023

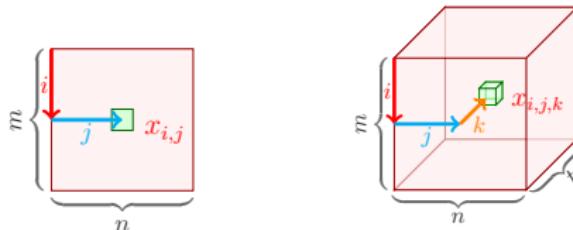
**Past work:** (Pattern discovery, traffic data imputation/forecasting)

- ❶ X. Chen, C. Zhang, X. Chen, N. Saunier, L. Sun (2023). Discovering dynamic patterns from spatiotemporal data with time-varying low-rank autoregression. *IEEE Transactions on Knowledge and Data Engineering*. Early access.
- ❷ X. Chen, Z. Cheng, N. Saunier, L. Sun (2022). Laplacian convolutional representation for traffic time series imputation. *arXiv preprint arXiv:2212.01529*.
- ❸ X. Chen, L. Sun (2022). Bayesian temporal factorization for multidimensional time series prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 44 (9): 4659-4673.

# Tensors

---

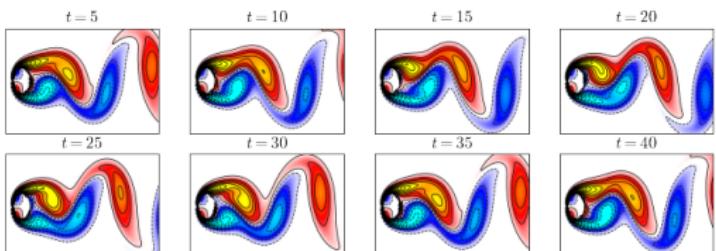
- What is tensor?  $\mathbf{X} \in \mathbb{R}^{m \times n}$  vs.  $\mathcal{X} \in \mathbb{R}^{m \times n \times t}$



- Tensors are everywhere!



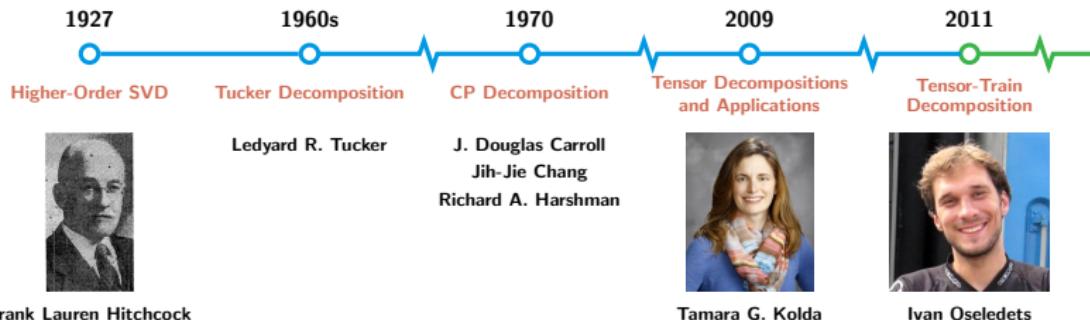
Color image with  
RGB channels



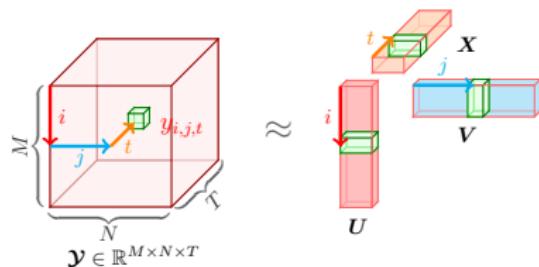
Dynamical system (fluid flow)

# Tensor Factorization

- Revisit tensor factorization (TF)



- **CP tensor factorization:** Factorize  $\mathcal{Y}$  into the combination of three rank- $R$  factor matrices (i.e., low-dimensional latent factors).



$$\begin{cases} y_{i,j,t} \approx \sum_{r=1}^R u_{i,r} v_{j,r} x_{t,r} \\ \mathcal{Y} \approx \sum_{r=1}^R \mathbf{u}_r \otimes \mathbf{v}_r \otimes \mathbf{x}_r \end{cases}$$

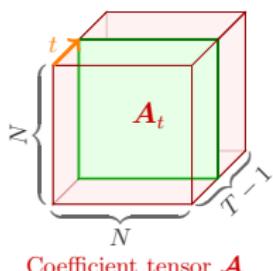
# Time-Varying Autoregression

- **Motivation:** Identify the dynamic patterns of time-varying systems.  
(e.g., fluid flow, climate variables, and human mobility)
- Given a sequence of time series snapshots  $\mathbf{y}_t \in \mathbb{R}^N$ ,  $t = 1, 2, \dots, T$ ,

$$\min_{\{\mathbf{A}_t\}} \frac{1}{2} \sum_t \|\mathbf{y}_t - \mathbf{A}_t \mathbf{y}_{t-1}\|_2^2$$

Time-varying autoregression

[Over-parameterization]  $\mathcal{O}(N^2(T-1))$  parameters vs.  $\mathcal{O}(NT)$  data.



Parameterize coefficients via TF:

$$\begin{aligned} \mathcal{A} &= \mathcal{G} \times_1 \mathbf{W} \times_2 \mathbf{V} \times_3 \mathbf{X} \\ \Rightarrow \quad \mathbf{A}_t &= \mathcal{G} \times_1 \mathbf{W} \times_2 \mathbf{V} \times_3 \mathbf{x}_t^\top \end{aligned}$$

**Significance:** Spatial models/patterns  $\mathbf{W}$  & temporal modes/patterns  $\mathbf{X}$

# Time-Varying Autoregression

- Optimization problem:

$$\min_{\mathbf{W}, \mathcal{G}, \mathbf{V}, \mathbf{x}} \frac{1}{2} \sum_{t=2}^T \left\| \mathbf{y}_t - (\mathcal{G} \times_1 \mathbf{W} \times_2 \mathbf{V} \times_3 \mathbf{x}_t^\top) \mathbf{y}_{t-1} \right\|_2^2$$

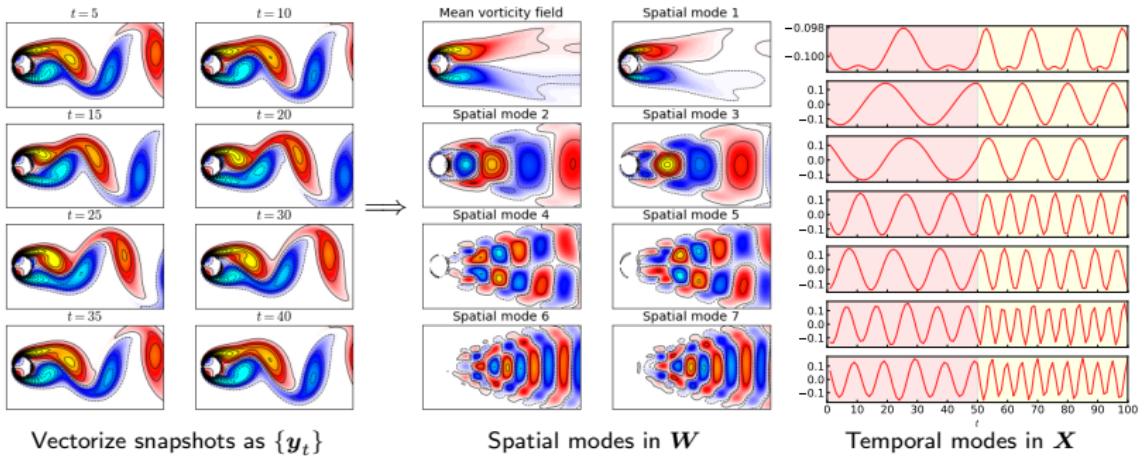
- Alternating minimization:

## Algorithm (Chen et al.'23)

Repeat

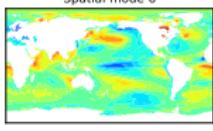
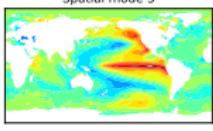
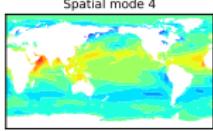
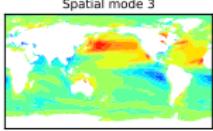
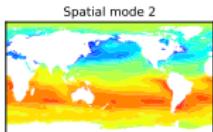
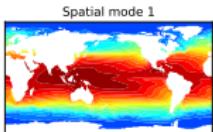
- Estimate  $\mathcal{G}$  by least squares (LS)
- Estimate  $\mathbf{W}$  by LS
- Estimate  $\mathbf{V}$  by conjugate gradient (CG, Golub & Van Loan'13)
- **For**  $t \in \{2, 3, \dots, T\}$  **do**
  - Estimate  $\mathbf{x}_t$  by LS

- **Multi-resolution fluid flow dataset** (the first 50 snapshots + 50 snapshots randomly selected from the last 100 snapshots)

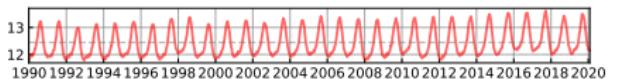


- Produce interpretable patterns
- Identify the system of different frequencies

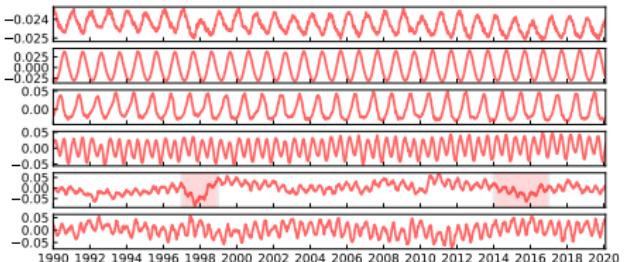
- Sea surface temperature dataset



Spatial modes in  $W$



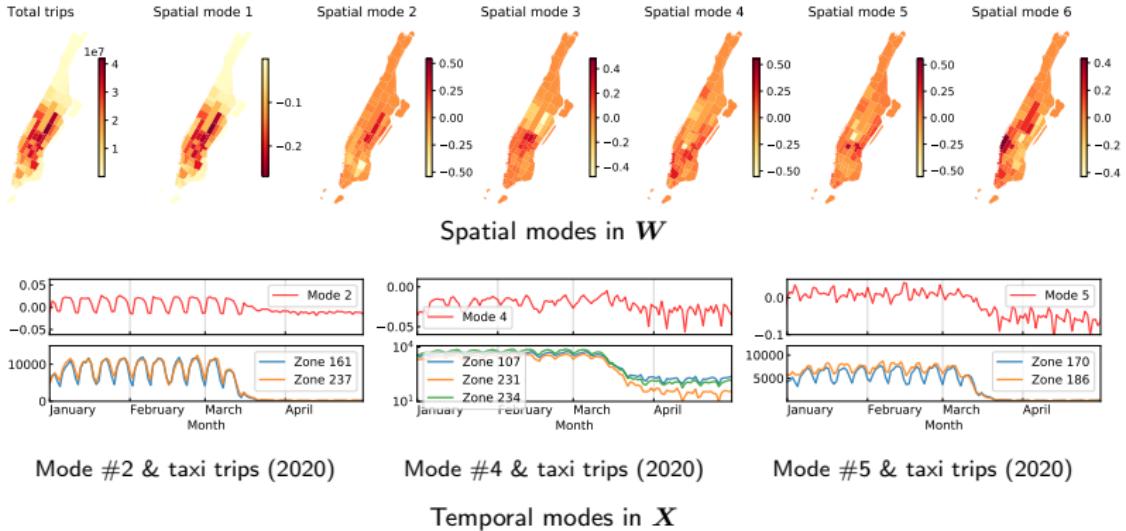
Time series of mean temperature



Temporal modes in  $X$

Identify two strongest El Nino events (on 1997-98 & 2014-16).

- NYC taxi dataset (pickup)



- Produce interpretable patterns
- Identify the changing point of the system (mainly due to COVID-19)

# Orthogonal Spatial/Temporal Patterns?

- Optimization problem with orthogonal constraints:<sup>1</sup>

$$\begin{aligned} \min_{\mathbf{W}, \mathcal{G}, \mathbf{V}, \mathbf{x}} \quad & \frac{1}{2} \sum_{t=2}^T \left\| \mathbf{y}_t - (\mathcal{G} \times_1 \mathbf{W} \times_2 \mathbf{V} \times_3 \mathbf{x}_t^\top) \mathbf{y}_{t-1} \right\|_2^2 \\ \text{s.t.} \quad & \begin{cases} \mathbf{W}^\top \mathbf{W} = \mathbf{I}_R \\ \mathbf{V}^\top \mathbf{V} = \mathbf{I}_R \\ \mathbf{X}^\top \mathbf{X} = \mathbf{I}_R \end{cases} \end{aligned}$$

## Algorithm (Chen et al.'23)

Repeat

- Estimate  $\mathcal{G}$  by LS
- Estimate  $\mathbf{W}$  by LS
- Estimate  $\mathbf{V}$  by conjugate gradient
- For  $t \in \{2, 3, \dots, T\}$  do
  - Estimate  $\mathbf{x}_t$  by LS



## Algorithm

Repeat

- Estimate  $\mathcal{G}$  by LS
- Estimate  $\mathbf{W}$  by orthogonal projection (Cai et al.'14)
- Estimate  $\mathbf{V}$  by projected conjugate gradient (Calamai & More'87)
- Estimate  $\mathbf{X}$  by projected conjugate gradient (Calamai & More'87)

<sup>1</sup><https://spatiotemporal-data.github.io/probs/orth-var>

# Spatiotemporal Mobility Networks

## State-Space Model (SSM)

- State transition equation:

$$\underbrace{x_{t+1}}_{\text{state}} = Ax_t + B \underbrace{u_t}_{\text{input}} + \underbrace{w_t}_{\mathcal{N}(0, I)}$$

- Observation equation:

$$\underbrace{y_t}_{\text{output}} = Cx_t + D \underbrace{u_t}_{\text{input}} + \underbrace{v_t}_{\mathcal{N}(0, I)}$$

VS.

## Recurrent Neural Network (RNN)

- Hidden state update:

$$h_t = \text{activation}(\mathbf{W}_{hh} h_{t-1} + \mathbf{W}_{hx} x_t + \underbrace{b_h}_{\text{bias}})$$

- Output:

$$\underbrace{y_t}_{\text{output}} = \text{activation}(\mathbf{W}_{yh} \underbrace{h_t}_{\text{state}} + \underbrace{b_y}_{\text{bias}})$$

- Long-range spatiotemporal modeling, e.g.,
  - State-space layers (Smith et al.'22)
  - Convolutional SSM (Smith et al.'23)
- Linear-time sequence modeling (e.g., selective SSM, Gu & Dao'23)

## References

---

A short list:

- [Cai et al.'14] J.-F. Cai, H. Ji, Z. Shen, and G.-B. Ye, “Data-driven tight frame construction and image denoising,” *Applied and Computational Harmonic Analysis*, vol. 37, no. 1, pp. 89–105, 2014.
- [Calamai & More'87] P. H. Calamai and J. J. More, “Projected gradient methods for linearly constrained problems,” *Mathematical programming*, vol. 39, no. 1, pp. 93–116, 1987.
- [Golub & Van Loan'13] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU press, 2013.
- [Gu & Dao'23] A. Gu and T. Dao, “Mamba: Linear-time sequence modeling with selective state spaces.” arXiv preprint arXiv:2312.00752.
- [Smith et al.'22] J. T. H. Smith, A. Warrington, and S. W. Linderman, “Simplified state space layers for sequence modeling.” arXiv preprint arXiv:2208.04933. (ICLR'23)
- [Smith et al.'23] J. T. H. Smith, S. De Mello, J. Kautz, S. W. Linderman, and W. Byeon, “Convolutional state space models for long-range spatiotemporal modeling.” arXiv preprint arXiv:2310.19694.



POLYTECHNIQUE  
MONTRÉAL

UNIVERSITÉ  
D'INGÉNIERIE



# Thanks for your attention!

Any Questions?

## About me:

- 🏡 Homepage: <https://xinychen.github.io>
- 🐙 GitHub: <https://github.com/xinychen>
- ✉️ How to reach me: [chenxy346@gmail.com](mailto:chenxy346@gmail.com)