Machine Learning and Optimization for Understanding Spatiotemporal Systems

Time Series Imputation & Periodicity Quantification

Xinyu Chen

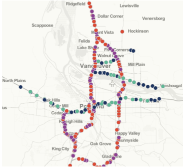
Postdoctoral Associate, MIT

May 22, 2025

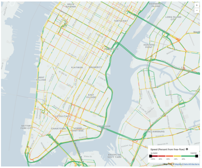
Orlando, USA

Spatiotemporal Data

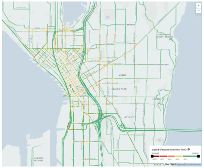
- Transport & mobility & climate application scenarios



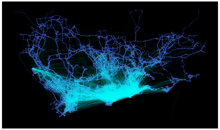
Highway (Portland)



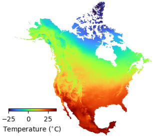
Uber movement (NYC)



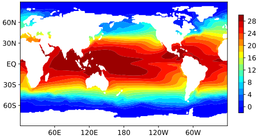
Uber movement (Seattle)



Taxi trajectory (Shenzhen)



Temperature (NA)



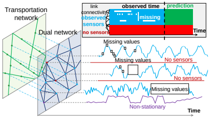
Temperature (sea surface)

- Challenges: Sparsity, high-dimensionality (network-scale), and multi-dimensionality (complicated data structure), time-varying systems

Research Contributions

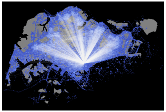
- (Practical) Formulating challenging spatiotemporal problems

Traffic Imputation — Mobility Pattern Discovery — Climate Seasonality Qualification



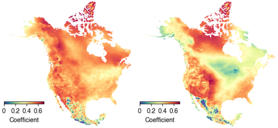
Missing data in transport

Matrix/tensor decomposition
Temporal dynamics



Singapore movement

Dynamical system modeling
Tensor decomposition

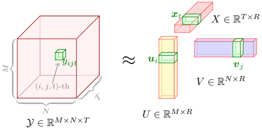


2000s 2010s

Interpretable ML
Mixed-integer programming

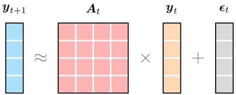
- (Methodological) Advancing ML development

Tensor decomposition



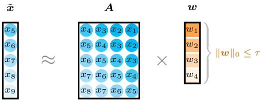
Integrate traffic flow patterns

Dynamic mode decomposition



Formulate dynamic patterns

Sparse autoregression



Interpret sparse autocorrelations

Reproducible Research

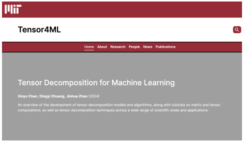
- The last mile of AI for spatiotemporal data computing

Human mobility & smart cities
Data-driven transport analytics
Spatiotemporal data modeling
Tensor decomposition for ML
Optimization for interpretable ML
. . .

Directions & Topics **Reproducible Research**

- Advancing ML development with open-source research

transdim
(1,200+ GitHub stars)
ML for Transport Data Imputation
https://github.com/xinychen/transdim

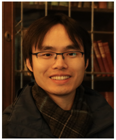


Tensor Decomposition for ML
(ML project)
Math & ML Tutorials
https://sites.mit.edu/tensor4ml

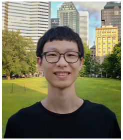
Spatiotemporal Data Computing
(Data valorization project)
Model Development of ML & Data Science
https://spatiotemporal-data.github.io

Spatiotemporal Data Imputation

❒ Convolution ❒ Fast Fourier transform ❒ Optimization w/ ℓ_1-norm
❒ Time series imputation ❒ Speed field reconstruction

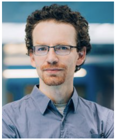


Xinyu Chen
UdeM → MIT



Zhanhong Cheng
McGill → UF

HanQin Cai
UCF



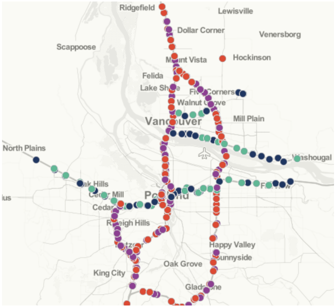
Nicolas Saunier
PolyMtl

Lijun Sun
McGill

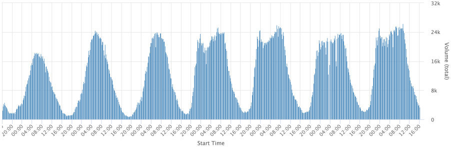
- **Xinyu Chen**, Zhanhong Chen, HanQin Cai, Nicolas Saunier, Lijun Sun (2024). "Laplacian Convolutional Representation for Traffic Time Series Imputation". *IEEE Transactions on Knowledge and Data Engineering*, 36 (11): 6490–6502.
- Blog post: Understanding time series convolution. https://spatiotemporal-data.github.io/posts/ts_conv

Motivation

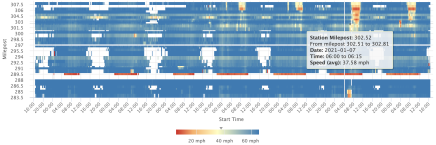
- Portland highway traffic data[1]



Highway network & sensor locations



Traffic volume

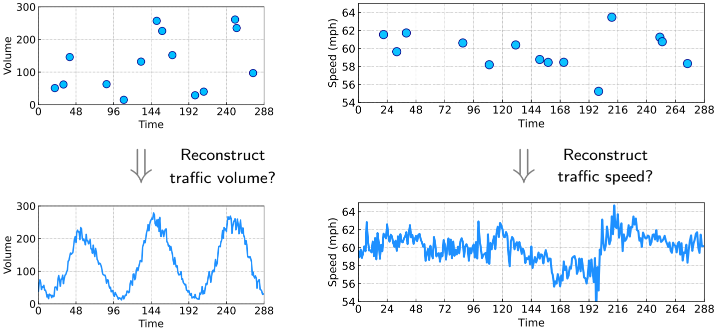


Traffic speed

- $\boldsymbol{X} \in \mathbb{R}^{N \times T}$ with N spatial locations $\times$ T time steps
- Traffic volume/speed shows strong spatial/temporal dependencies
- Missing data are there, how to improve data quality?

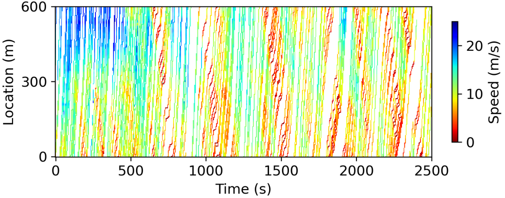
[1] https://portal.its.pdx.edu/home

Motivation

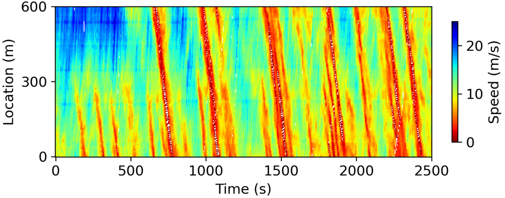


- How to utilize the global trends of traffic time series?
- How to produce local consistency of traffic data?

Motivation



200-by-500 matrix (NGSIM) $\Downarrow$ Reconstruct speed field from 20% sparse trajectories?

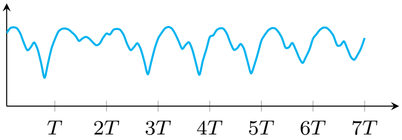


- How to learn from sparse spatiotemporal data?
- How to characterize spatial/temporal local dependencies?

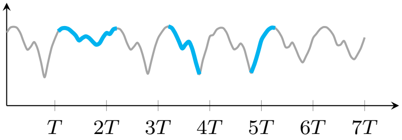
Time Series Imputation

Global/local trends in sparse data?

- Global trends (e.g., daily/weekly periodicity):

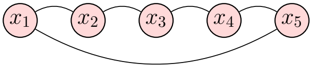


- Local trends (e.g., short-term time series trends):



Local Trend Modeling

- Intuition of Laplacian matrix



Undirected and circulant graph

$$\xrightarrow{\text{Modeling}} \quad \boldsymbol{L} = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}$$

(Circulant) Laplacian matrix

- Laplacian kernel:

$$\boldsymbol{\ell} \triangleq \underbrace{(2, -1, 0, 0, -1)^\top}_{\text{first column of } \boldsymbol{L}}$$

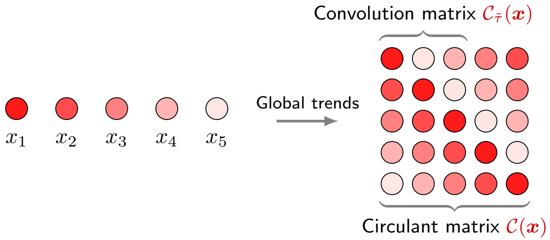
extending to the degree 2τ (i.e., graph connectivity) for $\boldsymbol{x} \in \mathbb{R}^T$.

- Temporal regularization:

$$\mathcal{R}(\boldsymbol{x}) = \underbrace{\frac{1}{2}\|\boldsymbol{L}\boldsymbol{x}\|_2^2}_{\text{mat-vec mul.}} = \underbrace{\frac{1}{2}\|\boldsymbol{\ell} \star \boldsymbol{x}\|_2^2}_{\text{convolution } \star}$$

Global Trend Modeling

Circulant matrix $\mathcal{C}(\boldsymbol{x})$ vs. convolution matrix $\mathcal{C}_{\tilde{\tau}}(\boldsymbol{x})$



- Circulant/Convolution nuclear norm minimization
  - A balance between global and local trends modeling?

CircNNM (Liu'22, Liu & Zhang'23)

$$\min_{\boldsymbol{x}} \;\|\mathcal{C}(\boldsymbol{x})\|_*$$
$$\text{s.t. } \|\mathcal{P}_\Omega(\boldsymbol{x}-\boldsymbol{y})\|_2 \leq \epsilon$$

on data $\boldsymbol{y}$ w/ observed index set Ω.

ConvNNM (Liu'22, Liu & Zhang'23)

$$\min_{\boldsymbol{x}} \;\|\mathcal{C}_{\tilde{\tau}}(\boldsymbol{x})\|_*$$
$$\text{s.t. } \|\mathcal{P}_\Omega(\boldsymbol{x}-\boldsymbol{y})\|_2 \leq \epsilon$$

on data $\boldsymbol{y}$ w/ observed index set Ω.

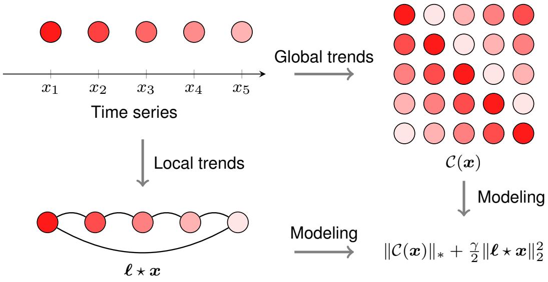
Global + Local Trends?

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/local time series trends:

$$\min_{\boldsymbol{x}} \; \underbrace{\|\mathcal{C}(\boldsymbol{x})\|_*}_{\text{global}} + \frac{\gamma}{2} \underbrace{\|\boldsymbol{\ell} \star \boldsymbol{x}\|_2^2}_{\text{local}}$$

$$\text{s.t. } \|\mathcal{P}_\Omega(\boldsymbol{x} - \boldsymbol{y})\|_2 \leq \epsilon$$



Laplacian Convolutional Representation

- LCR model:

$$\begin{aligned}\min_{\boldsymbol{x}}\ & \|\mathcal{C}(\boldsymbol{x})\|_* + \frac{\gamma}{2}\|\boldsymbol{\ell}\star\boldsymbol{x}\|_2^2 \\ \text{s.t. } & \|\mathcal{P}_\Omega(\boldsymbol{x}-\boldsymbol{y})\|_2 \leq \epsilon\end{aligned}$$

$$\Longrightarrow \quad \begin{aligned}\min_{\boldsymbol{x}}\ & \underbrace{\|\mathcal{C}(\boldsymbol{x})\|_* + \frac{\gamma}{2}\|\boldsymbol{\ell}\star\boldsymbol{x}\|_2^2}_{\text{global + local}} + \underbrace{\frac{\eta}{2}\|\mathcal{P}_\Omega(\boldsymbol{z}-\boldsymbol{y})\|_2^2}_{\text{regularization}} \\ \text{s.t. } & \boldsymbol{z}=\boldsymbol{x}\end{aligned}$$

"The alternating direction method of multipliers (ADMM) is an algorithm that solves convex optimization problems by breaking them into smaller pieces, each of which are then easier to handle."

— Source: https://stanford.edu/~boyd/admm.html

Laplacian Convolutional Representation

- Augmented Lagrangian function:

$$\mathcal{L}=\underbrace{\|\mathcal{C}(\boldsymbol{x})\|_*+\frac{\gamma}{2}\|\boldsymbol{\ell}\star\boldsymbol{x}\|_2^2}_{\text{global + local}}+\underbrace{\frac{\lambda}{2}\|\boldsymbol{x}-\boldsymbol{z}\|_2^2+\langle\boldsymbol{w},\boldsymbol{x}-\boldsymbol{z}\rangle}_{\text{Lagrangian multiplier } \boldsymbol{w}}+\underbrace{\frac{\eta}{2}\|\mathcal{P}_{\Omega}(\boldsymbol{z}-\boldsymbol{y})\|_2^2}_{\text{observations } \boldsymbol{y}}$$

- Optimize $\boldsymbol{x}$ w/ FFT in $\mathcal{O}(T\log T)$ time: :

$$\begin{cases}\|\mathcal{C}(\boldsymbol{x})\|_*=\|\mathcal{F}(\boldsymbol{x})\|_1=\|\hat{\boldsymbol{x}}\|_1 & \text{(circulant matrix)}\\ \frac{1}{2}\|\boldsymbol{\ell}\star\boldsymbol{x}\|_2^2=\frac{1}{2T}\|\mathcal{F}(\boldsymbol{\ell})\circ\mathcal{F}(\boldsymbol{x})\|_2^2=\frac{1}{2T}\|\hat{\boldsymbol{\ell}}\circ\hat{\boldsymbol{x}}\|_2^2 & \text{(circular convolution)}\end{cases}$$

- Reformulate the optimization as ℓ_1-norm minimization:

$$\boldsymbol{x}:=\arg\min_{\boldsymbol{x}}\;\|\mathcal{C}(\boldsymbol{x})\|_*+\frac{\gamma}{2}\|\boldsymbol{\ell}\star\boldsymbol{x}\|_2^2+\frac{\lambda}{2}\|\boldsymbol{x}-\boldsymbol{z}+\boldsymbol{w}/\lambda\|_2^2$$

$$\Longrightarrow\hat{\boldsymbol{x}}:=\arg\min_{\hat{\boldsymbol{x}}}\;\underbrace{\|\hat{\boldsymbol{x}}\|_1}_{\ell_1\text{-norm}}+\frac{\gamma}{2T}\|\hat{\boldsymbol{\ell}}\circ\hat{\boldsymbol{x}}\|_2^2+\frac{\lambda}{2T}\|\hat{\boldsymbol{x}}-\hat{\boldsymbol{z}}+\hat{\boldsymbol{w}}/\lambda\|_2^2$$

Laplacian Convolutional Representation

ℓ_1-norm Minimization (Liu & Zhang'23)

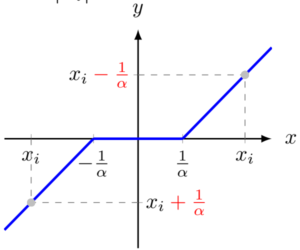
For any $\hat{\boldsymbol{h}} \in \mathbb{C}^T$ and $\delta \in \mathbb{R}$:

$$\min_{\hat{\boldsymbol{x}}} \;\|\hat{\boldsymbol{x}}\|_1 + \frac{\delta}{2}\|\hat{\boldsymbol{x}} - \hat{\boldsymbol{h}}\|_2^2$$

The solution to $\hat{\boldsymbol{x}}$:

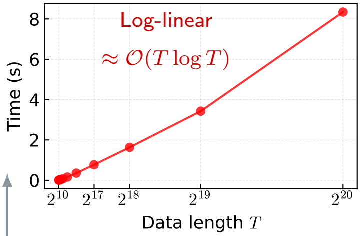
$$\hat{x}_t := \frac{\hat{h}_t}{|\hat{h}_t|} \cdot \underbrace{\max\{0, |\hat{h}_t| - 1/\delta\}}_{\text{shrinkage (e.g., ReLU)}}, t \in [T]$$

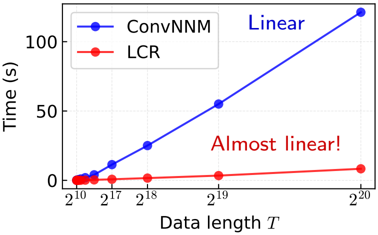
$$y_i = \frac{x_i}{|x_i|} \cdot \max\{|x_i| - 1/\alpha, 0\}$$



Laplacian Convolutional Representation

Time complexity & scalability & efficiency?



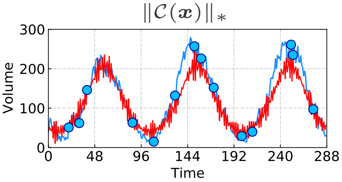


Empirical time complexity

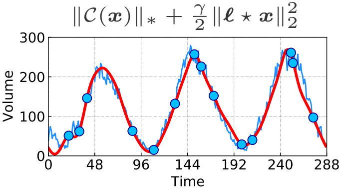
On the synthetic data $\boldsymbol{y} \in \mathbb{R}^T$ with $T \in \{2^{10}, 2^{11}, \ldots, 2^{20}\}$

Experiments

- Traffic speed imputation[2] (95% missing rate)



$\Downarrow$ Plus **local** time series trends



Highlights:

- Rethink the importance of local trend modeling in traffic data imputation tasks.
- Find a unified global and local trend modeling framework whose optimization can be efficiently solved by FFT:

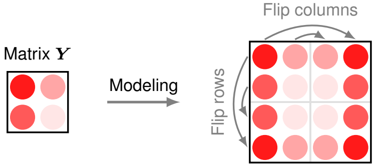
$$\min_{\boldsymbol{x}} \ \underbrace{\|\mathcal{C}(\boldsymbol{x})\|_*}_{\text{global}} + \frac{\gamma}{2} \underbrace{\|\boldsymbol{\ell} \star \boldsymbol{x}\|_2^2}_{\text{local}}$$
$$\text{s.t.} \ \ \|\mathcal{P}_\Omega(\boldsymbol{x} - \boldsymbol{y})\|_2 \leq \epsilon$$

[2] Blue dot: partial observation; red line: imputation.

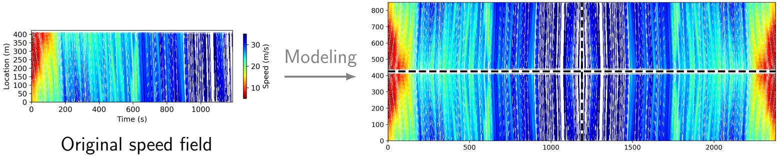
Experiments

Speed field reconstruction[3]

- Flipping operation on a matrix:



- Flipping operation on a speed field of vehicular traffic flow:



Original speed field

Flipped speed field

[3]Highway Drone (HighD) dataset at https://www.highd-dataset.com/

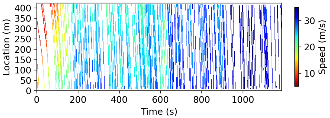
Experiments

Speed field reconstruction in German highways[4]

- Scenario: Mask trajectories of 70% vehicles
- LCR-2D on partially observed $\boldsymbol{Y} \in \mathbb{R}^{N \times T}$:

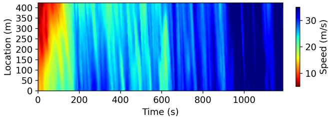
$$\min_{\boldsymbol{X}} \quad \underbrace{\|\mathcal{C}(\boldsymbol{X})\|_*}_{\text{global trend}} + \frac{\gamma}{2} \underbrace{\|(\boldsymbol{\ell}_s \boldsymbol{\ell}^\top) \star \boldsymbol{X}\|_F^2}_{\text{local trend}}$$

$$\text{s.t. } \|\mathcal{P}_\Omega(\boldsymbol{X} - \boldsymbol{Y})\|_F \leq \epsilon$$

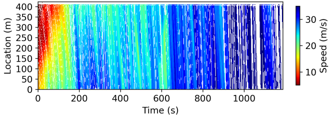


Sparse speed field

LCR-2D



Reconstructed speed field



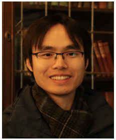
Ground-truth speed field

[4]Highway Drone (HighD) dataset at https://www.highd-dataset.com/

Quantifying Time Series Periodicity

(Ongoing Research)

❐ Interpretable ML ❐ Optimization w/ ℓ_0-norm ❐ Mixed-integer programming
❐ Human mobility regularity ❐ Climate system seasonality



Xinyu Chen
MIT

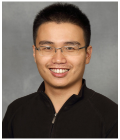
Dingyi Zhuang
MIT

Yunhan Zheng
MIT

Jinhua Zhao
MIT

HanQin Cai
UCF

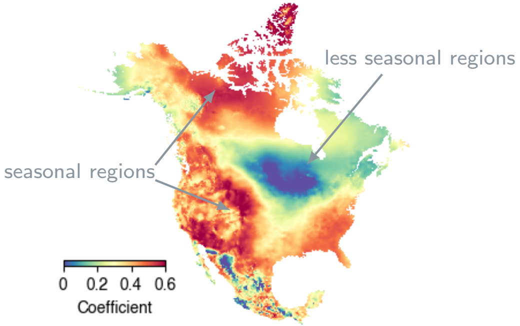
Ryan Qi Wang
Northeastern



Lijun Ding
UCSD

Vassilis Digalakis Jr
BU

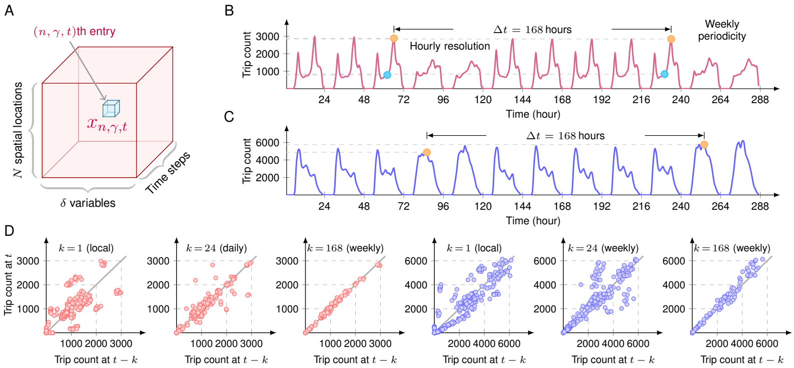
Motivation



Yearly temperature **seasonality** pattern in 2010s

Motivation

Human mobility data show daily/weekly regularity and periodicity?

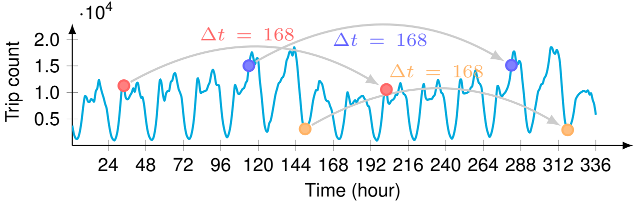


"**Closeness**" to the anti-diagonal $y = x$

$x_t \approx x_{t-168}$ (weekly periodicity)

Motivation

Weekly periodicity of ridesharing trip time series in Chicago



What motivate us most about **periodicity**?

❶ **Resilience and stability of systems**: Empirically measure the periodicity and predictability of urban systems.

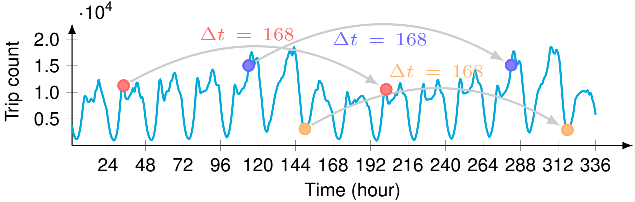
❷ **Optimization of transport systems**: Optimize resources (e.g., public transit, taxi, rideshare, and micromobility) to meet transport demand efficiently.

❸ **Design of sustainable transport & infrastructure**: Implement energy-efficient solutions tailored to peak hours.

Motivation

- Time series autoregression on $\boldsymbol{x} \in \mathbb{R}^T$

$$\boldsymbol{w} := \arg\min_{\{w_k\}_{k\in[d]}} \sum_{t\in[d+1,T]} \Big(x_t - \sum_{k\in[d]} w_k x_{t-k}\Big)^2$$



Periodicity of rideshare trip time series

- Sparse coefficient vector $\mapsto$ **Interpretability**?

$$\boldsymbol{w} = (\underbrace{0.33}_{k=1}, 0, \cdots, 0, \underbrace{0.20}_{k=167}, \underbrace{0.46}_{k=168})^\top \in \mathbb{R}^{168}$$

Valorizing Autoregression

- Time series autoregression

$$\begin{aligned} \boldsymbol{w} &:= \arg\min_{\{w_k\}_{k\in[d]}} \sum_{t\in[d+1,T]} \Big(x_t - \sum_{k\in[d]} w_k x_{t-k}\Big)^2 \\ &= \arg\min_{\boldsymbol{w}} \|\tilde{\boldsymbol{x}} - \boldsymbol{A}\boldsymbol{w}\|_2^2 \end{aligned}$$

$$\overset{\tilde{\boldsymbol{x}}}{\begin{bmatrix} x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{bmatrix}} \approx \overset{\boldsymbol{A}}{\begin{bmatrix} x_4 & x_3 & x_2 & x_1 \\ x_5 & x_4 & x_3 & x_2 \\ x_6 & x_5 & x_4 & x_3 \\ x_7 & x_6 & x_5 & x_4 \\ x_8 & x_7 & x_6 & x_5 \end{bmatrix}} \times \overset{\boldsymbol{w}}{\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix}} \Big\} \|\boldsymbol{w}\|_0 \leq \tau$$

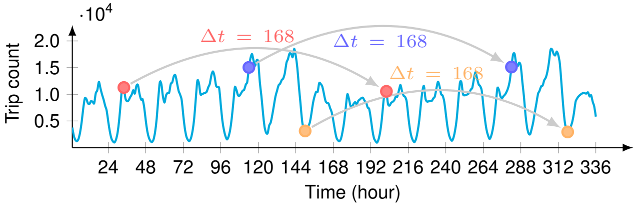
Autoregression on time series $\boldsymbol{x} = (x_1, x_2, \ldots, x_9)^\top$ w/ sparsity $\tau \in \mathbb{Z}^+$

- Sparse autoregression

$$\begin{aligned} &\min_{\boldsymbol{w}\geq 0} \ \|\tilde{\boldsymbol{x}} - \boldsymbol{A}\boldsymbol{w}\|_2^2 \\ &\text{s.t.} \ \underbrace{\|\boldsymbol{w}\|_0 \leq \tau}_{\text{sparsity w/ } \ell_0\text{-norm}} \end{aligned} \iff \begin{aligned} &\min_{\boldsymbol{w},\boldsymbol{\beta}} \ \|\tilde{\boldsymbol{x}} - \boldsymbol{A}\boldsymbol{w}\|_2^2 \\ &\text{s.t.} \ \begin{cases} 0 \leq \boldsymbol{w} \leq \boldsymbol{\beta}, \ \boldsymbol{\beta} \in \{0,1\}^d \\ \|\boldsymbol{\beta}\|_1 \leq \tau \end{cases} \end{aligned}$$

Solution Quality

- Subspace pursuit (SP) sometimes fails



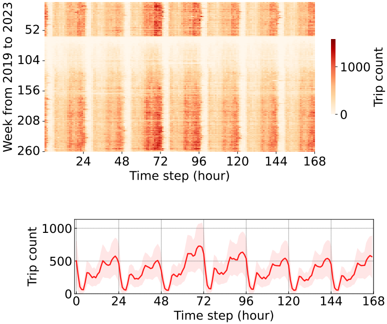
Periodicity of ridesharing trip time series

- Exact solution w/ mixed-integer programming (MIP)
- An intuitive example (sparsity $\tau = 2$):

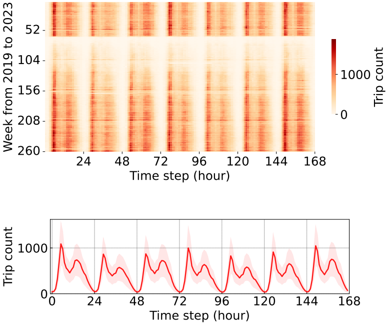
$$\underbrace{\boldsymbol{w} = (\cdots, \underbrace{0.02}_{k=53}, \cdots, \underbrace{0.96}_{k=168})^\top}_{\text{loss func.} = 8.32 \times 10^7 \text{ (SP)}} \quad \text{vs.} \quad \underbrace{\boldsymbol{w} = (\underbrace{0.22}_{k=1}, \cdots, \underbrace{0.77}_{k=168})^\top}_{\text{loss func.} = 6.25 \times 10^7 \text{ (MIP)}}$$

John F. Kennedy International Airport

- Pickup/Dropoff trips in airport
  - Pickup trips are relevant to flight delay, baggage claim, and other factors.
  - Dropoff trips to airport are highly related to flight schedules.



Pickup trips from airport



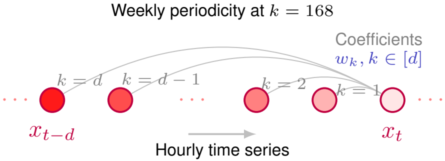
Dropoff trips to airport

- Sparse coefficient vectors (sparsity $\tau = 3$):

$$\boldsymbol{w} = (\underbrace{0.31}_{k=1}, \cdots, \underbrace{0.28}_{k=24}, \cdots, \underbrace{0.41}_{k=168})^\top \quad \text{vs.} \quad \boldsymbol{w} = (\underbrace{0.18}_{k=1}, \cdots, \underbrace{0.35}_{k=24}, \cdots, \underbrace{0.47}_{k=168})^\top$$

Spatially- and Time-Varying Autoregression

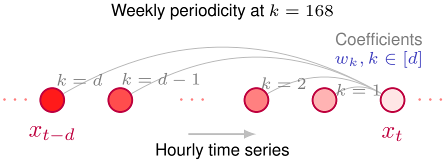
Univariate autoregression



$$\min \sum_{t} \Big(x_t - \sum_{k \in [d]} w_k x_{t-k} \Big)^2$$

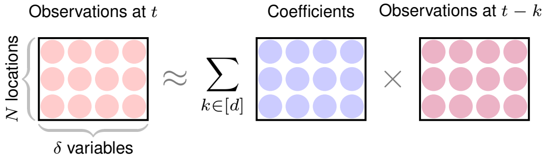
Spatially- and Time-Varying Autoregression

Univariate autoregression



$$\min \sum_{t} \Big(x_t - \sum_{k \in [d]} w_k x_{t-k} \Big)^2$$

Multidimensional autoregression



$$\min \sum_{n \in [N]} \sum_{\gamma \in [\delta]} \sum_{t} \Big(x_{n,\gamma,t} - \sum_{k \in [d]} w_{n,\gamma,k} x_{n,\gamma,t-k} \Big)^2$$

Envisioning Human Mobility

- Ridesharing trip data $\{\boldsymbol{x}_{n,\gamma}\}$ across $\gamma \in [\delta]$ years
- Reformulate sparse autoregression:

year

spatial location

hourly time step

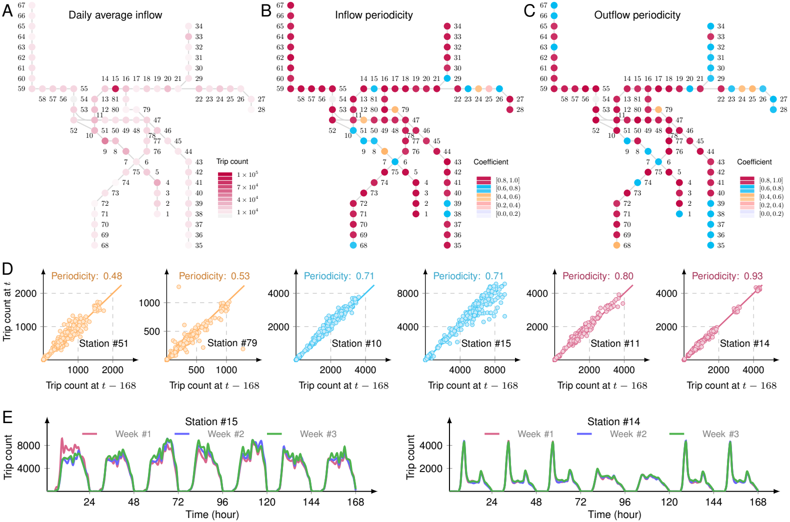
$$\min_{\{\boldsymbol{w}_{n,\gamma}\},\boldsymbol{\beta}} \sum_{n\in[N]} \sum_{\gamma\in[\delta]} \sum_{t\in[d+1,T_\gamma]} \Bigl(x_{n,\gamma,t} - \sum_{k\in[d]} w_{n,\gamma,k} x_{n,\gamma,t-k}\Bigr)^2$$

$$\text{s.t.} \quad \underbrace{\boldsymbol{\beta} \in \{0,1\}^d}_{\text{binary var.}} \quad \underbrace{0 \leq \boldsymbol{w}_{n,\gamma} \leq \boldsymbol{\beta}}_{\text{upper bound in } \{0,1\}} \quad \underbrace{\|\boldsymbol{\beta}\|_1 \leq \tau}_{\text{sum of binary var.}}$$

- MIP problem w/ $(N\delta + 1)d$ variables!
- How to handle thousands or millions of (e.g., $N\delta = 10^6$) time series?

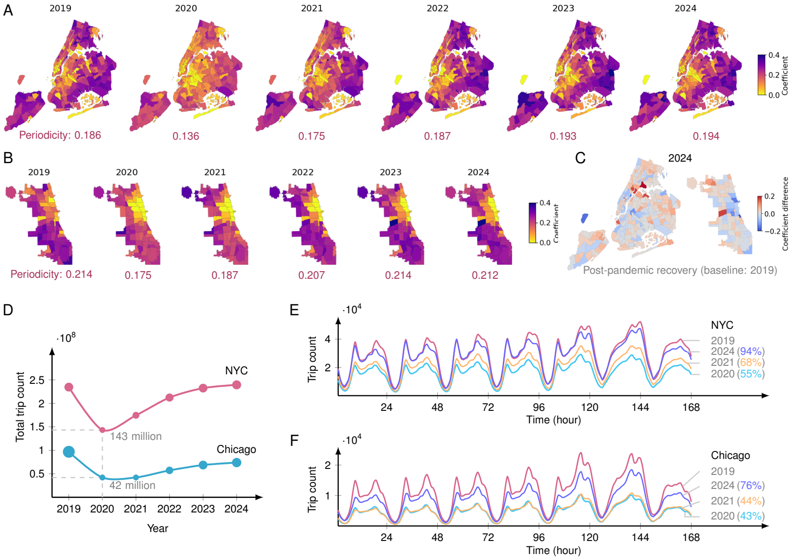
Quantify weekly periodicity by $\{w_{n,\gamma,k}\}$ at index $k = 168$

Envisioning Human Mobility



Hangzhou metro passenger flow in January 2019

Envisioning Human Mobility



Weekly periodicity reveals spatial patterns of ridesharing systems

Understanding Climate Systems

Quantify yearly seasonality by $\{w_{m,n,\gamma,k}\}$ at index $k = 12$

longitude

latitude　decade　monthly time step　temperature

$$\min_{\{\boldsymbol{w}_{m,n,\gamma}\},\boldsymbol{\beta}} \sum_{m\in[M]} \sum_{n\in[N]} \sum_{\gamma\in[\delta]} \sum_{t\in[d+1,T_\gamma]} \Big(x_{m,n,\gamma,t} - \sum_{k\in[d]} w_{m,n,\gamma,k} x_{m,n,\gamma,t-k} \Big)^2$$

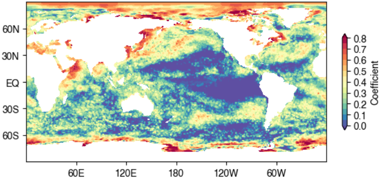
$$\text{s.t.} \begin{cases} \boldsymbol{\beta} \in \{0,1\}^d \\ 0 \leq \boldsymbol{w}_{m,n,\gamma} \leq \boldsymbol{\beta},\ \forall m,n,\gamma \\ \|\boldsymbol{\beta}\|_1 \leq \tau \\ \|\boldsymbol{w}_{m,n,\gamma}\|_1 = 1,\ \forall m,n,\gamma \end{cases}$$

binary decision var.

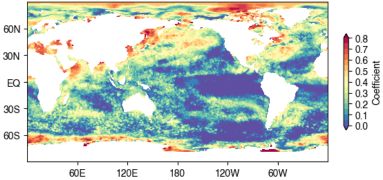
sparsity constraints

ℓ_1-normalization

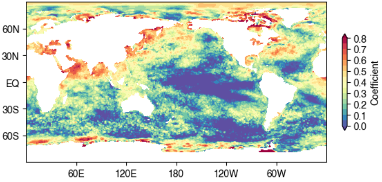
Sea Surface Temperature



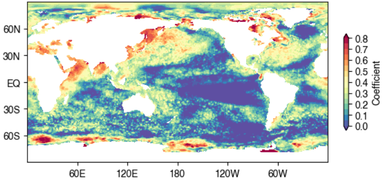
1980s



1990s



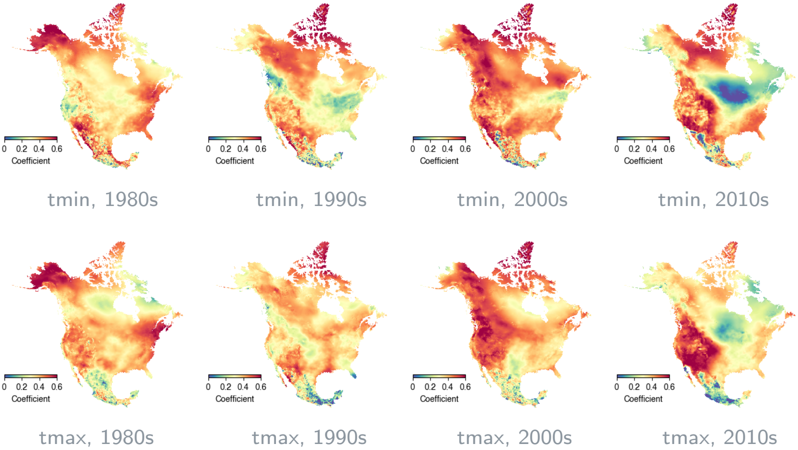
2000s



2010s

- Identify yearly periodicity at $k = 12$ from SST data ($\tau = 3$)
  1. The areas of El Nino events are less seasonal/predictable
  2. Arctic becomes less seasonal/predictable in the past 20 years
- Insights into climate system monitoring

North America Temperature



- Identify yearly periodicity at $k = 12$ from temperature data ($\tau = 3$)
  1. Stronger yearly seasonality in high-latitude areas
  2. Less seasonal temperature in south areas (e.g., Mexico)
  3. Seasonality patterns in 2000s & 2010s are different from 1980s & 1990s

Thanks for your attention!

Any Questions?

Slides: https://xinychen.github.io/slides/ml_opt_stsystem.pdf

Essential AR: https://xinychen.github.io/slides/essential_ar.pdf

About me:

- Homepage: https://xinychen.github.io
- MIT sites: https://sites.mit.edu/xinychen
- How to reach me: xinychen@mit.edu