Spatio-temporal Change-point Detection

Joshua Inoue

July 2022

supervised by *Matt Nunes* & *Sandipan Roy*

The University of Bath

Department of Mathematical Sciences

SAMBa

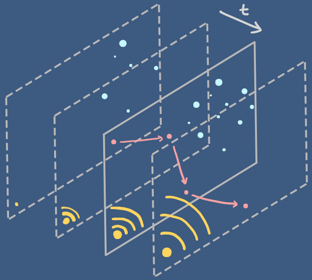
Outline

I. Motivation

II. Methods

III. Problems

IV. Conclusions



Spatio-temporal Data

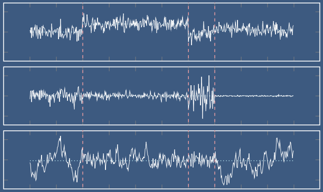
Data collected over time across different points in space

- Geospatial Data
- Point Processes
- Trajectories

Change-points

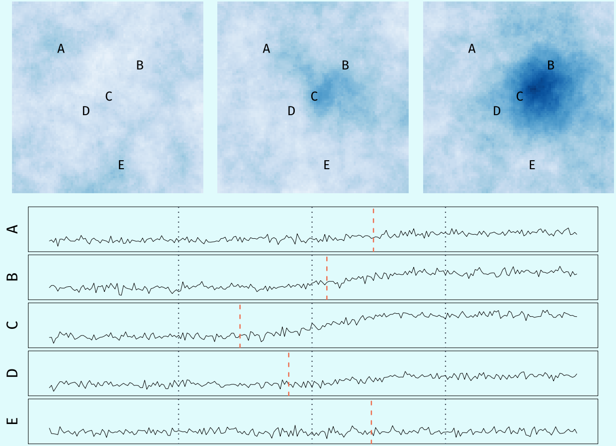
A change over time in the model parameters

- Piecewise stationary models
- High dimensional parameter estimation



Illustrative Problem

Simulated spatio-temporal data with a lagged and dampened change in mean



The Two Perspectives

Descriptive Statistical Model

$$\mathrm{Y}_t(\mathbf{s}) = \mu_t(\mathbf{s}) + \beta(\mathbf{s}) + \gamma_t + \kappa_t(\mathbf{s}) + \varepsilon_t(\mathbf{s})$$

Dynamic Statistical Model

$$\mathbf{Y}_t = \sum_{\ell=1}^{p} A_\ell \mathbf{Y}_{t-\ell} + \boldsymbol{\varepsilon}_t$$

Defining the Problem

- Time has a natural ordering, space does not
- Data locations fixed in space over time
- Time and Space discretised
- Space can be represented by some network
- Change-point defined as a change in model parameters

Capture characteristics of space and time

Simplify model via dimension reduction

Use spatial contextual information

Apply change-point detection methods

Vector Autoregressive Models: Discrete and dynamically derived

$$\boldsymbol{Y}_t = \boldsymbol{b} + A_1 \boldsymbol{Y}_{t-1} + \cdots + A_p \boldsymbol{Y}_{t-p} + \boldsymbol{\varepsilon}_t$$

Extension of Autoregressive models to multiple variables:

$$AR(1) : X_t = \alpha X_{t-1} + \varepsilon_t$$

$$Y_{1,t} = a_{11} Y_{1,t-1} + a_{12} Y_{2,t-1} + \varepsilon_{1,t}$$

$$Y_{2,t} = a_{21} Y_{1,t-1} + a_{22} Y_{2,t-1} + \varepsilon_{2,t}$$

$$\begin{pmatrix} Y_{1,t} \\ Y_{2,t} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$

II. Methods

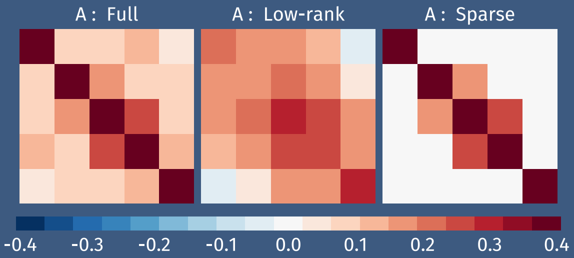
Low-rank and Sparse Decomposition

$$\mathbf{Y}_t = \sum_{\ell=1}^{p} (L_\ell + S_{j,\ell}) \mathbf{Y}_{t-\ell} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}_d(\mathbf{0}, \Sigma_\varepsilon)$$

Low-rank: *Invariant cross-autocorrelation*

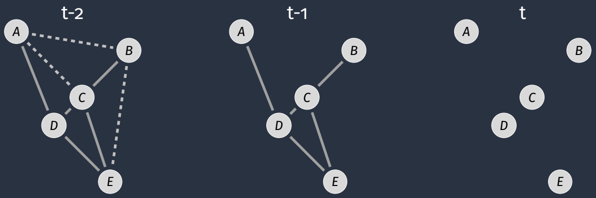
Sparse: *Time-evolving additional cross-autocorrelations*

Coefficient matrices for full and approximated processes



Spatially Structured Vector Autoregressive Models

Network representation of spatial locations in simulation with neighbourhood sets at each lag



$$(A_\ell)_{ij} = 0 \iff \mathbf{s}_j \notin \mathcal{N}_r(\mathbf{s}_i)\ \forall r \leq \Pi(\ell)$$

$$\mathcal{N}_r(\mathbf{s}) \cap \mathcal{N}_s(\mathbf{s}) = \varnothing, \text{ for } r \neq s \qquad \mathbf{s}' \in \mathcal{N}_r(\mathbf{s}) \iff \mathbf{s} \in \mathcal{N}_r(\mathbf{s}')$$

$\mathcal{N}_0(\mathbf{s}) = \{\mathbf{s}\}$ and $\Pi(\ell)$ says how many orders of the neighbourhoods are included at lag ℓ, (Knoblauch and Damoulas, 2018)

Change-point Model Formulation

For $n+1$ time points and m_0 change-points with

$$0 = t_0 < t_1 < \cdots < t_{m_0} < t_{m_0+1} = n$$

for each piecewise stationary component $t_{j-1} \leq t < t_j$, $j = 1, \ldots, m_0 + 1$,

$$\mathbf{Y}_t = \sum_{\ell=1}^{p} (L_\ell + S_{j,\ell}) \mathbf{Y}_{t-\ell} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}_d(\mathbf{0}, \Sigma_\varepsilon)$$

where $\mathbf{Y}_t, \boldsymbol{\varepsilon}_t \in \mathbb{R}^d$, $L_\ell, S_{j,\ell} \in \mathbb{R}^{d \times d}$.

Three Step Procedure (Bai, Safikhani, and Michailidis, 2020)

1. Over-select candidate change-points
2. Use an information criterion to screen out redundant ones
3. Estimate the model parameters

1. Block fused LASSO

$$(\widehat{\mathbf{\Theta}}, \widehat{L}) = \underset{\mathbf{\Theta}, L \in \Omega}{\arg\min} \frac{1}{n} \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{X}} L - \boldsymbol{\mathcal{Z}} \mathbf{\Theta}\|_2^2 + \lambda_{1,n} \|L\|_* + \lambda_{2,n} \|\mathbf{\Theta}\|_1 + \lambda_{3,n} \sum_{l=1}^{k_n} \left\| \sum_{j=1}^{l} \Theta_j \right\|_1$$

2. Screening via Information Criterion

$$\mathrm{IC}(\mathbf{s}; \eta_n) = \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{X}} \widehat{L} - \boldsymbol{\mathcal{Z}}_{s_1,\ldots,s_m} \widehat{\mathbf{\Theta}}_{s_1,\ldots,s_m}\|_F^2 + \eta_L \|\widehat{L}\|_* + \sum_{i=1}^{m+1} \eta_{(s_{i-1}, s_i)} \|\widehat{\Theta}_{(s_{i-1}, s_i)}\|_1 + m\omega_n$$

3. Consistent Parameter Estimation

$$(\widehat{L}, \widehat{S}_j) = \underset{L \in \Omega, S_j}{\arg\min} \frac{1}{N_j} \|\boldsymbol{\mathcal{Y}}_j - \boldsymbol{\mathcal{X}}_j (L + S_j)\|_F^2 + \rho_L \|L\|_* + \rho_j \|S_j\|_1$$

Capture characteristics of space and time

VAR Models

Simplify model via dimension reduction

Low-rank plus Sparse approximation

Use spatial contextual information

Neighbourhoods of spatial dependencies

Apply change-point detection methods

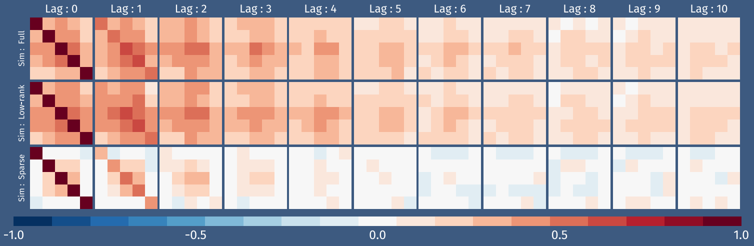
Three step procedure

Merging statistical with dynamical models

Add spatial information to low-rank plus sparse VAR models

- Use descriptive models to constrain dynamic models
- Inform networks of dependencies based on spatial proximity and on contextual spatial covariates

Comparison of cross-correlation matrices of full and approximate models at different lags

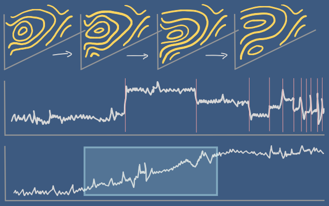


III. Problems

What change-points can be detected?

Characterise types of change-points:

- Localised change-points
- Consistency of estimation
- Non-instantaneous change-points



Other Issues

- How to identify and attribute the source of variation?
- How well do low-rank and sparse methods fit with statistical models?
- Do the models match up with real data?

SAMBa

Capture characteristics of space and time

- *VAR Models*

Simplify model via dimension reduction

- *Discretisation and Low-rank plus Sparse approximation*
- *Spatio-temporal Constraints*

Use spatial contextual information

- *Neighbourhoods of spatial dependencies*
- *Covariates and Latent Factors*

Apply change-point detection methods

- *Interpreting localised changes*

Questions?