Statistical Models for Road Traffic Forecasting

J-M. Loubes
Collaboration with: Guillaume Allain, Thibault Espinasse, Fabrice Gamboa, Jean-Noël Kien

INRIA-12th of June
Overview

1. Collaboration overview
2. Industrial context
3. Road traffic models: examples
4. Shape Invariant Models
5. Speed Models with Gaussian Field on a Graph
6. Applications
1 Collaboration overview

2 Industrial context

3 Road traffic models: examples

4 Shape Invariant Models

5 Speed Models with Gaussian Field on a Graph

6 Applications
The actors of the collaboration

- **Mediamobile Vtraffic**
  Mediamobile ensures the production and broadcasting of reliable and pertinent real-time traffic information. Founded in 1996, Mediamobile originated from a partnership between TDF Group (European leader in media content broadcast) and the automotive manufacturer Renault in the framework of a European Program for Research and Development of Intelligent Transportation.

- **Institut de Mathématiques de Toulouse**
  The Toulouse Mathematics Institute, CNRS Research Laboratory, federates the mathematics community of the Toulouse area. One of the biggest mathematical team in France (around 400 people)
People involved in the collaboration

Six years collaboration leading to three patents. Actual people involved

- **Mediamobile Vtrafic**
  - Philippe Goudal head of the prediction department
  - Guillaume Allain Engineer has been Engineer/CIFRE Ph. D Student of the project
  - Jean-Noël Kien Engineer/CIFRE Ph. D Student

- **Institut de Mathématiques de Toulouse**
  - Fabrice Gamboa Professor
  - Jean-Michel Loubes Professor
  - Elie Maza Assistant Professor
  - Thibault Espinasse Assistant Professor
  - Jean-Noël Kien Engineer/CIFRE Ph. D Student
Overview

1. Collaboration overview
2. Industrial context
3. Road traffic models: examples
4. Shape Invariant Models
5. Speed Models with Gaussian Field on a Graph
6. Applications
Mediamobile’s task

- **Gathering** raw traffic information
- **Processing** and aggregating
- **Broadcasting** (radio, www, mobile device...)

⇒ Fancy new services: **forecasting** and **dynamic routine**

**Industrial constraints:**

- **coverage**
  - each road of the network
  - from real time to long run

- **quality/accuracy**
  - controlled speed prediction error
  - controlled jam prediction error

- **user friendly**
  - automatable
  - adaptative
  - easy to update
Road traffic data - Road network

What is a road network?

- Graph composed of a set of pair (edges, vertices)
- **Complexity** of the graph → *Functional Road Classes (FRC)*
- **FRC** → road type classification (arterial, collector, local road...)

<table>
<thead>
<tr>
<th>FRC \ Number of edges</th>
<th>( \sum L ) [km]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 46 175</td>
<td>22 580</td>
</tr>
<tr>
<td>1 232 572</td>
<td>42 793</td>
</tr>
<tr>
<td>2 462 907</td>
<td>75 453</td>
</tr>
<tr>
<td>3 998 808</td>
<td>175 790</td>
</tr>
<tr>
<td>{0,1,2,3} 1 740 462</td>
<td>316 616</td>
</tr>
</tbody>
</table>

*Tab: Number of edges by FRC*
Network coverage depends on the FRC

Fig: Network coverage by all FRC \{0,1,2,3\} from 03/01/2009 to 05/31/2009
What is a speed data?

Loop sensor

- speed calculated from flow and density (conservation law)

Pros

- More accurate
- 3min constant frequency

Cons

- Located only in main roads
- Thresholded at national speed limits
Speed data

GPS sensor: Floating Car Data

- positions are mapped on a graph → building speeds

Pros

- Can potentially cover all the graph
- Raw source of data

Cons

- Less accurate → GPS and map-matching error
- More variable → outlier emergence
- Random frequency → user feedback
Overview

1. Collaboration overview

2. Industrial context

3. Road traffic models: examples
   - Sparse model to forecast
   - Punctual model of road traffic
     - Calendar influence
     - Model by classification
     - How to use the observed speed of the day?
     - Aggregation: statistical learning

4. Shape Invariant Models

5. Speed Models with Gaussian Field on a Graph
Local road trafficking forecasting with $\ell^1$

**Our Goal**

- Approach the road traffic dynamic with local statistical models

\[
V(s_q, t_{p+h}) = F(Q(s, t), \rho(s, t)...) \rightarrow V_{q,p+h} = g_{q,p,h}(\{V_{i,k}; i \in G, k \in T\})
\]

**Problems**

- High dimension of $X$
- All $V_{i,k}$ not influential

**Solution**

- Regularization
- Selection
Modelizing traffic dynamic with significative effects only

\[
V_{q,p+h} = g_{q,p,h}(V_{i,k}) \rightarrow V_{q,p+h} = \sum_{i \in G, k \in T} \beta_{i,k} \cdot V_{i,k}
\]

where

\[
\beta_{i,k} = K((i,k),(q,p+h))
\]

Kernel selection: fit road traffic dynamic

- learning a sparse set of influence parameters

\[
\hat{\beta} = \arg \min_{\beta} \left( \| V_{q,p+h} - \sum_{i \in G, k \in T} \beta_{i,k} \cdot V_{i,k} \|^2 + \lambda \sum |\beta_{i,k}| \right)
\]
Conclusion

- Short run local model
- Forecast and complete missing data
- Time and spatial road traffic dynamic used
- Exists block version to privilege certain axis
**Improve** accuracy of short/long run predictions with weather data

Partnership between Mediamobile and Meteo-France

Rupture model

\[ V(x, t_1) = V(x, t_0) + C(.) \times 1_{M(x, t_1) \neq M(x, t_0)} \quad \text{with} \quad t_1 - t_0 < \tau_{sta} \]

- \( \tau_{sta} \): timespan for a stationary traffic flow

\( C(.) \) correction term can depend on :
- edge \( x \): road specifications, geographical areas
- nature and intensity of the weather evolution
- traffic state at \( t_0 \): \( V(x, t_0) \)
Model selection based on $C(.)$ structure

Linear thresholded bias model

If $V(x, t_0) \geq \alpha$,

$$V(x, t_1) = V(x, t_0) - \beta \cdot (V(x, t_0) - \alpha)$$

correction term

break parameter

Or else,

$$V(x, t_1) = V(x, t_0)$$

Advantages

- takes traffic state into consideration
- thresholded model yields interpretable model

Drawback

- edge by edge model
Network generalization

Repartition of \((\alpha, \beta)\) parameters

Results

- \(\beta\) can be generalized
- Repartition of \(\alpha\) depends on the FRC
<table>
<thead>
<tr>
<th>Weather condition</th>
<th>$\alpha$ (% of FreeFlowSpeed)</th>
<th>$\beta$</th>
<th># obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low rain</td>
<td>93%</td>
<td>0.95</td>
<td>7236</td>
</tr>
<tr>
<td>Medium and Strong rain</td>
<td>90%</td>
<td>0.95</td>
<td>3316</td>
</tr>
<tr>
<td>Freezing rain</td>
<td>NA</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>Rain and Snow</td>
<td>94%</td>
<td>0.97</td>
<td>1011</td>
</tr>
<tr>
<td>Snow</td>
<td>83%</td>
<td>0.96</td>
<td>2621</td>
</tr>
<tr>
<td>Hail</td>
<td>NA</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>Drizzle</td>
<td>89%</td>
<td>0.90</td>
<td>615</td>
</tr>
</tbody>
</table>

For instance, let the free flow speed equals 100 km/h: a car travels at 130 km/h on a freeway and strong rain appears. Since $130 > 90\% \times 100$, car speed decreases to $130 - 95\% \times (130 - 90) = 92$ km/h.
Model the relationship between speeds and calendar

How it is used:

→ $D \neq \text{day of the prediction}$
→ the speed curve is not observed
→ « Inboard configuration » $\Rightarrow$ low complexity

Mathematical model: linear model with $k$ fixed

$$g(t_k, x) = \beta_0 + \left\{ \begin{array}{l}
\beta_1 \mathbb{1}_{\{c=\text{Monday}\}} + \beta_2 \ldots \\
\beta_8 \mathbb{1}_{\{c=\text{January}\}} + \beta_{19} \ldots \\
\beta_{20} \mathbb{1}_{\{c=\text{Hollidays}\}} + \ldots \\
\beta_{1,8} \mathbb{1}_{\{c=\text{Monday} \cap \text{January}\}} + \ldots \\
+ \beta_{1,20} \mathbb{1}_{\{c=\text{Monday} \cap \text{Hollidays}\}} + \ldots \\
\ldots 
\end{array} \right\}$$

-one order effects

Second order effects

Drawbacks:

→ Functionnal aspects are lost
→ $(N + 1) \times K$ effects
Model by classification

- **Functional mixture model**
- speed curve $V$ is represented as a **finite number of patterns**:
  \[
  V = \sum_{i=1}^{m} \mathbb{1}_{E=i} f_i + \epsilon_i \quad \text{et} \quad f^* = f_E
  \]
  \[
  E \in \{1, \ldots, m\} \text{ i.i.d. hidden R.V.}
  \epsilon_i \in \mathbb{R}^K, \quad \epsilon_i \sim \mathcal{N}(0, \Sigma_i \in \mathcal{M}_{K,K})
  \]
  \[
  \mathbb{E}[V|E = i] = f_i, \quad \text{Var}[V|E = i] = \Sigma_i
  \]

- **Classification** of $E$ then prediction of $V$ by $f^*$:
  \[
  \begin{array}{c}
  \chi \xrightarrow{\text{regression}} \hat{V} \in \mathbb{R}^K \\
  \end{array}
  \]
  \[
  \begin{array}{c}
  \chi \xrightarrow{\text{classification}} \hat{E} \\
  \end{array}
  \]
The classification model

\[ X = \{ \text{Day of the week, Holidays} \} \text{ and } m = 4 \]
Model the information contained in the speed of the day

Frame:

- Prediction **in the day** \(D\)
- Speeds \(V^p\) are **known**

\[\begin{align*}
\text{p fixed, } X &= (V^p, C) \\
\end{align*}\]

**X** Time series
- How many patterns?
  - \(h\) big et \(p\) small:
    - \(\Rightarrow m\) small
  - \(h\) small and \(p\) big:
    - \(\Rightarrow m\) **big**
Overcoming non stationarity

**Restriction** of the forecast profile:

- **STA**: \( g(V_p) = V_p \)
- **KMC10**: \( X = V_p, m = 10 \) et \( \tau = 1\text{h} \)
- **KML4**: \( m = 4 \) et \( \tau = \infty \)
- **CAL4**: \( X = C, m = 4 \)

**Avantages**:

- High stability
- Small processing time
## Prediction of the travel time

Example for a travel with $h = 1$ (forecast at one hour)

<table>
<thead>
<tr>
<th></th>
<th>REF</th>
<th>STA</th>
<th>C10</th>
<th>L4</th>
<th>CAL4</th>
<th>BP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BPI</strong> (14km)</td>
<td>32.3</td>
<td>21.2</td>
<td><strong>14.2</strong></td>
<td>15.3</td>
<td>17.5</td>
<td>14.4</td>
</tr>
<tr>
<td><strong>BPE</strong> (21km)</td>
<td>41.8</td>
<td>24.6</td>
<td>17.6</td>
<td>18.9</td>
<td>21.6</td>
<td><strong>17.1</strong></td>
</tr>
<tr>
<td><strong>A86ES</strong> (22km)</td>
<td>20.4</td>
<td>14.7</td>
<td>15.4</td>
<td>13.2</td>
<td>12.6</td>
<td><strong>10.1</strong></td>
</tr>
<tr>
<td><strong>N118W</strong> (26km)</td>
<td>25.4</td>
<td>16.7</td>
<td>9.6</td>
<td>9.8</td>
<td>14.3</td>
<td>9</td>
</tr>
<tr>
<td><strong>A4W</strong> (35km)</td>
<td>21.3</td>
<td>17.5</td>
<td>12.8</td>
<td>13</td>
<td>15</td>
<td><strong>11.8</strong></td>
</tr>
</tbody>
</table>
Overview

1. Collaboration overview
2. Industrial context
3. Road traffic models : examples
4. Shape Invariant Models
   - An artificial data example
5. Speed Models with Gaussian Field on a Graph
6. Applications
Shape Invariant Models

Shift on traffic jams

(a)

(b)

(c)

(d)

-2.0 -1.0 0.0 1.0

boxplot of estimated parameters

estimated parameters (in hours)

Shift on traffic jams

(a)

(b)

(c)

(d)

-2.0 -1.0 0.0 1.0

boxplot of estimated parameters

estimated parameters (in hours)
Shift on traffic jams

(a)

(b)

(c)

(d)
A more general model: Shape invariant model (SIM)

\[ Y_{ij} = f^*_j(x_i) + \varepsilon_{ij} \quad i = 1 \ldots n_j, \ j = 1 \ldots J. \]

- \( \varepsilon \) is as before a Gaussian white noise with variance \( \sigma^2 \)
- \( \exists f^*: \mathbb{R} \rightarrow \mathbb{R} \) with

\[ f^*_j(\cdot) = a_j^* f^*(\cdot - \theta^*_j) + \nu^*_j \quad (\theta^*_j, a^*_j, \nu^*_j) \in \mathbb{R}^3, \ \forall j = 1 \ldots J. \]
A more general model: Shape invariant model (SIM)

\[ Y_{ij} = f_j^*(x_i) + \varepsilon_{ij} \quad i = 1 \ldots n_j, \ j = 1 \ldots J. \]

- \( \varepsilon \) is as before a Gaussian white noise with variance \( \sigma^2 \)
- \( \exists f^* : \mathbb{R} \rightarrow \mathbb{R} \) with

\[ f_j^*(\cdot) = a_j^* f^*(\cdot - \theta_j^*) + v_j^* \quad (\theta_j^*, a_j^*, v_j^*) \in \mathbb{R}^3, \ \forall j = 1 \ldots J. \]
A more general model: Shape invariant model (SIM)

\[ Y_{ij} = f^*_j(x_i) + \varepsilon_{ij} \quad i = 1 \ldots n_j, \ j = 1 \ldots J. \]

- \( \varepsilon \) is as before a Gaussian white noise with variance \( \sigma^2 \)
- \( \exists f^* : \mathbb{R} \rightarrow \mathbb{R} \) with

\[ f^*_j(\cdot) = a^*_j f^*(\cdot - \theta^*_j) + v^*_j \quad (\theta^*_j, a^*_j, v^*_j) \in \mathbb{R}^3, \ \forall j = 1 \ldots J. \]
A more general model: Shape invariant model (SIM)

\[ Y_{ij} = f_j^*(x_i) + \varepsilon_{ij} \quad i = 1 \ldots n_j, \ j = 1 \ldots J. \]

- \( \varepsilon \) is as before a Gaussian white noise with variance \( \sigma^2 \)
- \( \exists f^* : \mathbb{R} \rightarrow \mathbb{R} \) with

\[ f_j^*(\cdot) = a_j^* f^*(\cdot - \theta_j^*) + v_j^* \quad (\theta_j^*, a_j^*, v_j^*) \in \mathbb{R}^3, \ \forall j = 1 \ldots J. \]
Some references on SIM

Some references on SIM

Some references on SIM

Some references on SIM

Some references on SIM

Some references on SIM

Some references on SIM

Recall the model

\[ Y_{i,j} = f^*(x_i - \theta_j^*) + \epsilon_{ij}, \ i = 1, \ldots, N, \ j = 1, \ldots, J. \]  

(1)

- \( f^* \) is an unknown \( T \)-periodic function
- \( (\theta_j^*)_{j=1}^J \) is an unknown parameter of \( \mathbb{R}^J \)
- The design is uniform: \( x_i = 2i\pi/T, \ i = 1, \ldots, N \)
- \( (\epsilon_{ij}) \) is a Gaussian white noise with variance \( \sigma^2 \)

The model is not well posed. Identifiability problem
Recall the model

\[ Y_{i,j} = f^*(x_i - \theta_j^*) + \epsilon_{ij}, \quad i = 1, \ldots, N, \quad j = 1, \ldots, J. \]  

(1)

- \( f^* \) is an unknown \( T \)-periodic function
- \( (\theta_j^*)_{j=1}^J \) is an unknown parameter of \( \mathbb{R}^J \)
- The design is uniform: \( x_i = 2i\pi/T, \quad i = 1, \ldots, N \)
- \( (\epsilon_{ij}) \) is a Gaussian white noise with variance \( \sigma^2 \)

The model is not well posed. Identifiability problem
Recall the model

\[ Y_{i,j} = f^*(x_i - \theta^*_j) + \epsilon_{ij}, \ i = 1, \ldots, N, \ j = 1, \ldots, J. \]  \hspace{1cm} (1)

- \( f^* \) is an unknown T-periodic function
- \((\theta^*_j)_{j=1}^J \) is an unknown parameter of \( \mathbb{R}^J \)
- The design is uniform: \( x_i = 2i\pi/T, \ i = 1, \ldots, N \)
- \((\epsilon_{ij}) \) is a Gaussian white noise with variance \( \sigma^2 \)

The model is not well posed. Identifiability problem.
Recall the model

\[ Y_{i,j} = f^*(x_i - \theta_j^*) + \epsilon_{ij}, \quad i = 1, \ldots, N, \quad j = 1, \ldots, J. \] (1)

- \( f^* \) is an unknown \( T \)-periodic function
- \( (\theta_j^*)_{j=1}^J \) is an unknown parameter of \( \mathbb{R}^J \)
- The design is uniform: \( x_i = 2i\pi/T, \quad i = 1, \ldots, N \)
- \( (\epsilon_{ij}) \) is a Gaussian white noise with variance \( \sigma^2 \)

The model is not well posed. Identifiability problem.
Recall the model

\[ Y_{i,j} = f^*(x_i - \theta_j) + \epsilon_{ij}, \quad i = 1, \ldots, N, \quad j = 1, \ldots, J. \]  

(1)

- \( f^* \) is an unknown \( T \)-periodic function
- \((\theta_j^*)_{j=1}^J\) is an unknown parameter of \( \mathbb{R}^J \)
- The design is uniform: \( x_i = 2i\pi/T, \quad i = 1, \ldots, N \)
- \((\epsilon_{ij})\) is a Gaussian white noise with variance \( \sigma^2 \)

The model is not well posed. Identifiability problem
Recall the model

\[ Y_{i,j} = f^*(x_i - \theta_j^*) + \varepsilon_{ij}, \ i = 1, \ldots, N, \ j = 1, \ldots, J. \]  (1)

- $f^*$ is an unknown $T$-periodic function
- $(\theta_j^*)_{j=1}^J$ is an unknown parameter of $\mathbb{R}^J$
- The design is uniform: $x_i = 2i\pi/T, \ i = 1, \ldots, N$
- $(\varepsilon_{ij})$ is a Gaussian white noise with variance $\sigma^2$

The model is not well posed. Identifiability problem
Set $\alpha^*_j = \frac{2\pi}{T} \theta^*_j$.

Replacing

- $\alpha^*$ by $\alpha^* + c1 + 2k\pi$ ($c \in \mathbb{R}$, $k \in \mathbb{Z}^J$) (2)

- $f^*$ by $f^*(\cdot - c)$

the observation equation remains invariant

**Identifiability constraints**

- Parameter set $\Lambda$ is compact
- $\alpha^* \in \Lambda$
- If $\alpha \in \Lambda$ and $\alpha = (2)$ holds then $\alpha = \alpha^*$

**Examples**

$\Lambda_1 = \{\alpha \in [-\pi, \pi]^J : \alpha_1 = 0\}$

$\Lambda_2 = \{\alpha \in [-\pi, \pi]^J : \sum \alpha_j = 0$ and $\alpha_1 \in [0, 2\pi/J]\}$
Identifiability

Set $\alpha_j^* = \frac{2\pi}{T} \theta_j^*$. 

Replacing

- $\alpha^*$ by $\alpha^* + c1 + 2k\pi$ ($c \in \mathbb{R}$, $k \in \mathbb{Z}^J$) (2)
- $f^*$ by $f^*(\cdot - c)$

the observation equation remains invariant

Identifiability constraints

- Parameter set $\Lambda$ is compact
- $\alpha^* \in \Lambda$
- If $\alpha \in \Lambda$ and $\alpha =$ (2) holds then $\alpha = \alpha^*$

Examples

$\Lambda_1 = \{\alpha \in [-\pi, \pi[^J : \alpha_1 = 0\}$

$\Lambda_2 = \{\alpha \in [-\pi, \pi[^J : \sum \alpha_j = 0$ and $\alpha_1 \in [0, 2\pi/J]\}$
Shape Invariant Models

Identifiability

Set $\alpha_j^* = \frac{2\pi}{T} \theta_j^*$.

Replacing

- $\alpha^*$ by $\alpha^* + c1 + 2k\pi$ \((c \in \mathbb{R}, k \in \mathbb{Z}^J)\) (2)
- $f^*$ by $f^*(\cdot - c)$

the observation equation remains invariant

Identifiability constraints

- Parameter set $\Lambda$ is compact
- $\alpha^* \in \Lambda$
- If $\alpha \in \Lambda$ and $\alpha = (2)$ holds then $\alpha = \alpha^*$

Examples

$\Lambda_1 = \{\alpha \in [-\pi, \pi]^J: \alpha_1 = 0\}$
$\Lambda_2 = \{\alpha \in [-\pi, \pi]^J: \sum \alpha_j = 0$ and $\alpha_1 \in [0, 2\pi/J]\}$
Identifiability

Set \( \alpha_j^* = \frac{2\pi}{T} \theta_j^* \).

Replacing

- \( \alpha^* \) by
  \[
  \alpha^* + c1 + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J)
  \]

- \( f^* \) by \( f^*(\cdot - c) \)

the observation equation remains invariant

Identifiability constraints

- Parameter set \( \Lambda \) is compact
- \( \alpha^* \in \Lambda \)
- If \( \alpha \in \Lambda \) and \( \alpha = (2) \) holds then \( \alpha = \alpha^* \)

Examples

\( \Lambda_1 = \{ \alpha \in [-\pi, \pi]^J: \alpha_1 = 0 \} \)
\( \Lambda_2 = \{ \alpha \in [-\pi, \pi]^J: \sum \alpha_j = 0 \text{ and } \alpha_1 \in [0, 2\pi/J] \} \)
Main simple idea

For any $c \in \mathbb{R}$ the shift operator $T_c$ defined on $T$-periodic functions

$$T_c(f) = f(\cdot - c)$$

has common eigenvectors

$$T_c[\exp(2i\pi/T \cdot)] = \exp(-2i\pi c/T) \exp(2i\pi/T \cdot)$$

More generally on a general group (here the torus), Fourier transform diagonalizes any translation operators acting on functions on the group (forward to extensions)

Rewrite the regression model using the eigenvectors
Estimation procedure

Main simple idea
For any \( c \in \mathbb{R} \) the shift operator \( T_c \) defined on \( T \)-periodic functions

\[
T_c(f) = f(\cdot - c)
\]

has common eigenvectors

\[
T_c[\exp(2i\pi/T\cdot)] = \exp(-2i\pi c/T) \exp(2i\pi/T\cdot)
\]

More generally on a general group (here the torus), Fourier transform diagonalizes any translation operators acting on functions on the group (forward to extensions)

Rewrite the regression model using the eigenvectors
Main simple idea
For any $c \in \mathbb{R}$ the shift operator $T_c$ defined on $T$-periodic functions

$$T_c(f) = f(\cdot - c)$$

has common eigenvectors

$$T_c[\exp(2i\pi/T \cdot)] = \exp(-2i\pi c/T) \exp(2i\pi/T \cdot)$$

More generally on a general group (here the torus), Fourier transform diagonalizes any translation operators acting on functions on the group (forward to extensions)

Rewrite the regression model using the eigenvectors
Shape Invariant Models

Estimation procedure

**Main simple idea**

For any \( c \in \mathbb{R} \) the shift operator \( T_c \) defined on \( T \)-periodic functions

\[
T_c(f) = f(\cdot - c)
\]

has common **eigenvectors**

\[
T_c[\exp(2i\pi/T\cdot)] = \exp(-2i\pi c/T) \exp(2i\pi/T\cdot)
\]

More generally on a general group (here the torus), Fourier transform diagonalizes any translation operators acting on functions on the group (forward to extensions)

**Rewrite the regression model using the eigenvectors**
Rewriting the model in terms of the Fourier transform

Taking the DFT and neglecting the (deterministic) error between the DFT and the Fourier transform. The model may be rewritten as (N is odd)

\[ d_{jl} = e^{-il\alpha_j^*} c_l(f^*) + w_{jl}, \quad l = -(N - 1)/2, \ldots, (N - 1)/2, \quad j = 1, \ldots, J \]

- \( c_l(f^*) \) is the Fourier coefficient of \( f^* \)
- \( (w_{jl}) \) is a complex Gaussian white noise with variance \( \sigma^2/N \)
Rewriting the model in terms of the Fourier transform

Taking the DFT and neglecting the (deterministic) error between the DFT and the Fourier transform. The model may be rewritten as (N is odd)

\[ d_{jl} = e^{-il\alpha_j^*} c_l(f^*) + w_{jl}, l = -(N-1)/2, \ldots, (N-1)/2, j = 1, \ldots, J \]

- \(c_l(f^*)\) is the Fourier coefficient of \(f^*\)
- \((w_{jl})\) is a complex Gaussian white noise with variance \(\sigma^2/N\)
Building a $\mathcal{M}$-function

- Re phased Fourier coefficients
  \[ \tilde{c}_{jl}(\alpha) = e^{il\alpha_j}d_{jl} \quad (\alpha \in \mathcal{A}) \]

- Mean of Re phased Fourier coefficients
  \[ \hat{c}_l(\alpha) = \frac{1}{J} \sum_{j=1}^{J} \tilde{c}_{jl}(\alpha) \]

\[ \tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*}w_{jl} \quad \text{and} \quad \hat{c}_l(\alpha^*) = c_l(f^*) + \frac{1}{J} \sum_{j=1}^{J} e^{il\alpha_j^*}w_{jl} \]
Shape Invariant Models

Building a $M$-function

- **Re phased Fourier coefficients**

$$\tilde{c}_{jl}(\alpha) = e^{il\alpha_j} d_{jl} \ (\alpha \in A)$$

- **Mean of Re phased Fourier coefficients**

$$\hat{c}_l(\alpha) = \frac{1}{J} \sum_{j=1}^{J} \tilde{c}_{jl}(\alpha)$$

$$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl} \text{ and } \hat{c}_l(\alpha^*) = c_l(f^*) + \frac{1}{J} \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$$
Building a $M$-function

- Re phased Fourier coefficients

$$\tilde{c}_{j l}(\alpha) = e^{i l \alpha_j} d_{j l} \quad (\alpha \in \mathbb{A})$$

- Mean of Re phased Fourier coefficients

$$\hat{c}_{l}(\alpha) = \frac{1}{J} \sum_{j=1}^{J} \tilde{c}_{j l}(\alpha)$$

$$\tilde{c}_{j l}(\alpha^*) = c_{l}(f^*) + e^{i l \alpha_j^*} w_{j l} \text{ and } \hat{c}_{l}(\alpha^*) = c_{l}(f^*) + \frac{1}{J} \sum_{j=1}^{J} e^{i l \alpha_j^*} w_{j l}$$
Building a $M$-function

- Re phased Fourier coefficients
  \[ \tilde{c}_{jl}(\alpha) = e^{il\alpha_j}d_{jl} \quad (\alpha \in \Lambda) \]

- Mean of Re phased Fourier coefficients
  \[ \hat{c}_l(\alpha) = \frac{1}{J} \sum_{j=1}^{J} \tilde{c}_{jl}(\alpha) \]

\[ \tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*}w_{jl} \text{ and } \hat{c}_l(\alpha^*) = c_l(f^*) + \frac{1}{J} \sum_{j=1}^{J} e^{il\alpha_j^*}w_{jl} \]
The $M$-function

**Idea:** The deviation $\tilde{c}_{j\ell}(\alpha) - \hat{c}_\ell(\alpha)$ should be small for $\alpha = \alpha^*$

$$M_n(\alpha) := \frac{1}{J} \sum_{j=1}^{J} \sum_{l=\left(\frac{N-1}{2}\right)}^{\left(\frac{N-1}{2}\right)/2} \delta_l^2 |\tilde{c}_{j\ell}(\alpha) - \hat{c}_\ell(\alpha)|^2$$

- $(\delta_l)$ is $l^2$ sequence of weights discussed later
The $M$-function

**Idea:** The deviation $\tilde{c}_{jl}(\alpha) - \hat{c}_l(\alpha)$ should be small for $\alpha = \alpha^*$

$$M_n(\alpha) := \frac{1}{J} \sum_{j=1}^{J} \sum_{-(N-1)/2}^{(N-1)/2} \delta_l^2 |\tilde{c}_{jl}(\alpha) - \hat{c}_l(\alpha)|^2$$

- $(\delta_l)$ is $l^2$ sequence of weights discussed later
An artificial data example

(a) Simulated data with shape invariant models.
(b) Estimated parameters over time.
(c) Model output data.
(d) Actual and predicted data comparison.
Overview

1. Collaboration overview
2. Industrial context
3. Road traffic models: examples
4. Shape Invariant Models
5. Speed Models with Gaussian Field on a Graph
   - General frame
   - Maximum likelihood
6. Applications
Graph of roads network

Modeling: Random process \( (X_i^{(n)})_{n \in \mathbb{Z}, i \in G} \)
Graph of roads network

Modeling: Random process $(X_i^{(n)})_{n \in \mathbb{Z}, i \in G}$

- Indexed by (discrete) time $\mathbb{Z}$ and the graph $G$ of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to graphs
Graph of roads network

Modeling: Random process \( \left( X_i^{(n)} \right)_{n \in \mathbb{Z}, i \in G} \)

- Indexed by (discrete) time \( \mathbb{Z} \) and the graph \( G \) of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to graphs
Graph of roads network

Modeling: Random process \( (X_{i}^{(n)})_{n \in \mathbb{Z}, i \in G} \)

- Indexed by (discrete) time \( \mathbb{Z} \) and the graph \( G \) of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to graphs
Graph of roads network

**Modeling**: Random process \((X_i^{(n)})_{n \in \mathbb{Z}, i \in G}\)

- Indexed by (discrete) time \(\mathbb{Z}\) and the graph \(G\) of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to graphs
Graph of roads network

Modeling: Random process \((X_i^{(n)})_{n \in \mathbb{Z}, i \in G}\)

- Indexed by (discrete) time \(\mathbb{Z}\) and the graph \(G\) of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to **graphs**
Graph of roads network

Modeling: Random process \((X^{(n)}_i)_{n \in \mathbb{Z}, i \in G}\)

- Indexed by (discrete) time \(\mathbb{Z}\) and the graph \(G\) of the road traffic network
- Gaussian
- Centered
- “Stationary“
- Extension of classical tools from time series to graphs

Objective: Yield a parametric model \((\mathcal{K}_\theta)_{\theta \in \Theta}\) for covariance operators of \(X\)
Traffic: Predict the speed of the vehicles with missing values

For now: Spatial dependency is not exploited

Aims
- Give a model that uses spatial dependency
- Estimate the spatial correlation
- Spatial filtering
Gaussian Process on Graph : Origin of the Problem

Traffic: Predict the speed of the vehicles with missing values

For now: Spatial dependency is not exploited

Aims

- Give a model that uses spatial dependency
- Estimate the spatial correlation
- Spatial filtering
Gaussian Process on Graph: Origin of the Problem

Traffic: Predict the speed of the vehicles with missing values

For now: Spatial dependency is not exploited

Aims

- Give a model that uses spatial dependency
- Estimate the spatial correlation
- Spatial filtering
Model: Speed process \((X_i)_{i \in G}\) indexed by the vertices \(G\) of a graph \(G\).

Definition (Unoriented weighted graph)

\(G = (G, W) :\)

- \(G\) set of vertices (infinite countable)
- \(W \in [-1, 1]^{G \times G}\) Weighted adjacency operator (symmetric)

Neighbors: \(i \sim j\) if \(W_{ij} \neq 0\)

Degree of a vertex: \(D_i = \#\{j, i \sim j\}\).

\(H_0:\)

- \(D := \sup_{i \in G} D_i < +\infty\), \(G\) has bounded degree
- \(\forall i \in G, \sum_{j \in G} |W_{ij}| \leq 1\) even renormalize
Model: Speed process \((X_i)_{i \in G}\) indexed by the vertices \(G\) of a graph \(G\).

**Definition (Unoriented weighted graph)**

\(G = (G, W) :\)

- \(G\) set of vertices (infinite countable)
- \(W \in [-1, 1]^{G \times G}\) Weighthed adjacency operator (symmetric)

**Neighbors:** \(i \sim j\) if \(W_{ij} \neq 0\)

**Degree of a vertex:** \(D_i = \#\{j, i \sim j\}\).

\(H_0\)

- \(D := \sup_{i \in G} D_i < +\infty, G\) has bounded degree
- \(\forall i \in G, \sum_{j \in G} |W_{ij}| \leq 1\) even renormalize
Graph

Model: Speed process \((X_i)_{i \in G}\) indexed by the vertices \(G\) of a graph \(G\).

**Definition (Unoriented weighted graph)**

\[ G = (G, W) : \]
- \(G\) set of vertices (infinite countable)
- \(W \in [-1, 1]^{G \times G}\) Weighted adjacency operator (symmetric)

**Neighbors:** \(i \sim j\) if \(W_{ij} \neq 0\)

**Degree of a vertex:** \(D_i = \#\{j, i \sim j\}\).

**H**

- \(D := \sup_{i \in G} D_i < +\infty\), \(G\) has bounded degree
- \(\forall i \in G, \sum_{j \in G} |W_{ij}| \leq 1\) even renormalize
Graph

**Model**: Speed process $(X_i)_{i \in G}$ indexed by the vertices $G$ of a graph $G$.

**Definition (Unoriented weighted graph)**

$G = (G, W)$:
- $G$ set of vertices (infinite countable)
- $W \in [-1, 1]^{G \times G}$ Weighed adjacency operator (symmetric)

**Neighbors**: $i \sim j$ if $W_{ij} \neq 0$

**Degree of a vertex**: $D_i = \# \{j, i \sim j\}$.

$H_0$
- $D := \sup_{i \in G} D_i < +\infty$, $G$ has bounded degree
- $\forall i \in G, \sum_{j \in G} |W_{ij}| \leq 1$ even renormalize
Problem

Remark:

- Our work robust to renormalization
- For $Z$, for instance: $W_{i,j}^{(Z)} = \frac{1}{2} 1_{|i-j|=1}$

$W$ acts on $l^2(G)$:

$$\forall u \in l^2(G), \forall i \in G, (Wu)_i := \sum_{j \in G} W_{i,j} u_j$$

Under $H_0$

$W$ bounded as operator of $B_G := l^2(G) \to l^2(G)$:

$$\|W\|_{2,op} \leq 1$$

$H'_0$: The entries of $W$ belongs to a finite set
Problem

Remark:

- Our work robust to renormalization
- For $\mathbb{Z}$, for instance: $W_{ij}^{(\mathbb{Z})} = \frac{1}{2} \mathbf{1}_{|i-j|=1}$

$W$ acts on $l^2(G)$:

$$\forall u \in l^2(G), \forall i \in G, (Wu)_i := \sum_{j \in G} W_{ij} u_j$$

Under $H_0$

$W$ bounded as operator of $B_G := l^2(G) \to l^2(G)$:

$$\|W\|_{2,\text{op}} \leq 1$$

$H'_0$ : The entries of $W$ belongs to a finite set
Problem

Remark:

• Our work robust to renormalization
• For $\mathbb{Z}$, for instance: $W_{ij}^{(\mathbb{Z})} = \frac{1}{2}1_{|i-j|=1}$

$W$ acts on $l^2(G)$:

$$\forall u \in l^2(G), \forall i \in G, (Wu)_i := \sum_{j \in G} W_{ij} u_j$$

Under $H_0$

$W$ bounded as operator of $B_G := l^2(G) \rightarrow l^2(G)$:

$$\|W\|_{2,op} \leq 1$$

$H'_0$ : The entries of $W$ belongs to a finite set
**Problem**

**Remark:**

- Our work robust to renormalization
- For $\mathbb{Z}$, for instance: $W_{ij}^{(\mathbb{Z})} = \frac{1}{2} \mathbf{1}_{|i-j|=1}$

$W$ acts on $l^2(G)$:

$$\forall u \in l^2(G), \forall i \in G, (Wu)_i := \sum_{j \in G} W_{ij} u_j$$

**Under $H_0$**

$W$ bounded as operator of $B_G := l^2(G) \rightarrow l^2(G)$:

$$\|W\|_{2,\text{op}} \leq 1$$

$H'_0$: The entries of $W$ belongs to a finite set
**Observation** : Correlations are independent of the position and the orientation

**Aim** : Propose a *stationary* and *isotropic* model for covariances $(X_i)_{i \in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G \times G}$:

$\Rightarrow$ Characterized by $K$

**Aim** : Extension of time series

$\Rightarrow$ Construction MA with adjacency operator

$+$ *isotropic* modification of the graph
**Observation**: Correlations are independent of the position and the orientation

**Aim**: Propose a stationary and isotropic model for covariances $(X_i)_{i \in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G \times G}$:

⇒ Characterized by $K$

**Aim**: Extension of time series

⇒ Construction MA with adjacency operator

+ isotropic modification of the graph
General approach

**Observation**: Correlations are independent of the position and the orientation

**Aim**: Propose a *stationary* and *isotropic* model for covariances \((X_i)_{i \in G}\) Gaussian, zero-mean, with covariance \(K \in \mathbb{R}^{G \times G}\):

\[\Rightarrow\] Characterized by \(K\)

**Aim**: Extension of time series

\[\Rightarrow\] Construction MA with adjacency operator

+ *isotropic* modification of the graph
Observation: Correlations are independent of the position and the orientation

Aim: Propose a stationary and isotropic model for covariances $(X_i)_{i \in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G \times G}$:
⇒ Characterized by $K$

Aim: Extension of time series
⇒ Construction MA with adjacency operator
+ isotropic modification of the graph
For $\mathbb{Z} : (\epsilon_n)_{n \in \mathbb{Z}}$ white noise

$$X_n = \sum_{k \in \mathbb{N}} a_k \epsilon_{n-k}$$
Observation: Correlations are independent of the position and the orientation.

Aim: Propose a stationary and isotropic model for covariances $(X_i)_{i \in G}$ Gaussian, zero-mean, with covariance $K \in \mathbb{R}^{G \times G}$:

$\Rightarrow$ Characterized by $K$

Aim: Extension of time series

$\Rightarrow$ Construction MA with adjacency operator

$+$ isotropic modification of the graph

Aim: Maximum Likelihood Estimation

$\Rightarrow$ Generalize Whittle’s approximation
A few bibliography

Spectral representation of stationary processes:
- $\mathbb{Z}^d$: X. Guyon
- Homogeneous tree: J-P. Arnaud
- Distance-transitive graphs: H. Heyer

Maximum Likelihood:
- $\mathbb{Z}$: here R. Azencott et D. Dacunha-Castelle
- $\mathbb{Z}^d$: X. Guyon, R. Dahlhaus
A few bibliography

Spectral representation of stationary processes:
- $\mathbb{Z}^d$: X. Guyon
- Homogeneous tree: J-P. Arnaud
- Distance-transitive graphs: H. Heyer

Maximum Likelihood
- $\mathbb{Z}$: here R. Azencott et D. Dacunha-Castelle
- $\mathbb{Z}^d$: X. Guyon, R. Dahlhaus
Example $G = \mathbb{Z}$

A Gaussian centered process with covariance $K$ is stationary if

$$\exists (r_k)_{k \in \mathbb{N}}, K_{ij} = r_{|i-j|}$$

**Spectral density**

If $r \in l^1$, $\exists f$, $K_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} f(t) \cos((j - i)t) \, dt := (T(f))_{ij}$

Let $g, f(t) = g(\cos(t))$, As

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(Z)}\right)^k\right)_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} \cos(t)^k \cos((j - i)t) \, dt,$$

**Operator representation**

We have $K = g(W^{(Z)})$
Example $G = \mathbb{Z}$

A Gaussian centered process with covariance $K$ is stationary if

$$\exists (r_k)_{k \in \mathbb{N}}, K_{ij} = r_{|i-j|}$$

**Spectral density**

If $r \in l^1$, $\exists f$, $K_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} f(t) \cos((j-i)t) \, dt := (T(f))_{ij}$

Let $g, f(t) = g(\cos(t))$, As

$$\forall i, j, k \in \mathbb{Z}, \left( (W^{(\mathbb{Z})})^k \right)_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} \cos(t)^k \cos((j-i)t) \, dt,$$

**Operator representation**

We have $K = g(W^{(\mathbb{Z})})$
Example $\mathbf{G} = \mathbb{Z}$

$X$ Gaussian centered process with covariance $K$ is stationary if

$$\exists (r_k)_{k \in \mathbb{N}}, K_{ij} = r_{|i-j|}$$

**Spectral density**

If $r \in l^1$, $\exists f, K_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} f(t) \cos ((j - i)t) \, dt := (T(f))_{ij}$

Let $g, f(t) = g(\cos(t))$, As

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})}\right)^k\right)_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} \cos(t)^k \cos ((j - i)t) \, dt,$$

**Operator representation**

We have $K = g(W^{(\mathbb{Z})})$
Example $G = \mathbb{Z}$

$X$ Gaussian centered process with covariance $K$ is stationary if

$$\exists (r_k)_{k \in \mathbb{N}}, K_{ij} = r_{|i-j|}$$

Spectral density

If $r \in l^1$, $\exists f$, $K_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} f(t) \cos((j-i)t) \, dt := (T(f))_{ij}$

Let $g$, $f(t) = g(\cos(t))$, As

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W^{(\mathbb{Z})}\right)^k\right)_{ij} = \frac{1}{2\pi} \int_{[0,2\pi]} \cos(t)^k \cos((j-i)t) \, dt,$$

Operator representation

We have $K = g(W^{(\mathbb{Z})})$
Example \( G = \mathbb{Z} \)

We can also write

\[
\forall i, j, k \in \mathbb{Z}, \left( \left( W^{(\mathbb{Z})} \right)^k \right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} d\lambda
\]

where \( T_k \) is the \( k \)-th Tchebychev’s polynomial

\[
\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} d\lambda
\]

- \( g \) polynomial of degree \( q \) : \( MA_q \)
- \( \frac{1}{g} \) polynomial of degree \( p \) : \( AR_p \) …

**Aim**: Generalize this kind of representation
Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left(\left(W(\mathbb{Z})\right)^k\right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} d\lambda$$

where $T_k$ is the $k$-th Tchebychev’s polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} d\lambda$$

- $g$ polynomial of degree $q$: $MA_q$
- $\frac{1}{g}$ polynomial of degree $p$: $AR_p$ · · ·

**Aim**: Generalize this kind of representation
Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left(\left(\mathcal{W}^{(\mathbb{Z})}\right)^k\right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} d\lambda$$

where $T_k$ is the $k$-th Tchebychev’s polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} d\lambda$$

- $g$ polynomial of degree $q : MA_q$
- $\frac{1}{g}$ polynomial of degree $p : AR_p$ · · ·

**Aim**: Generalize this kind of representation
Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left( (W(\mathbb{Z}))^k \right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} \, d\lambda$$

where $T_k$ is the $k$-th Tchebychev’s polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} \, d\lambda$$

- $g$ polynomial of degree $q : MA_q$
- $\frac{1}{g}$ polynomial of degree $p : AR_p$ ···

Aim : Generalize this kind of representation
Example $G = \mathbb{Z}$

We can also write

$$\forall i, j, k \in \mathbb{Z}, \left( (W(\mathbb{Z}))^k \right)_{ij} = \int_{[-1,1]} \lambda^k \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} d\lambda$$

where $T_k$ is the $k$-th Tchebychev’s polynomial

$$\forall i, j \in \mathbb{Z}, (K)_{ij} = \int_{[-1,1]} g(\lambda) \frac{T_{|j-i|}(\lambda)}{\sqrt{1 - \lambda^2}} d\lambda$$

- $g$ polynomial of degree $q : MA_q$
- $\frac{1}{g}$ polynomial of degree $p : AR_p$ ...

**Aim**: Generalize this kind of representation
Identity resolution

Spectral decomposition

\[ \exists \mathcal{E}, \mathcal{M}, \mathcal{W} = \int_{\mathcal{M}} \lambda d\mathcal{E}(\lambda) \]

Definition (Identity resolution)

\( \mathcal{M} \) Sigma-algebra \( \mathcal{E} : \mathcal{M} \to \mathcal{B}_G \) such that \( \forall \omega, \omega' \in \mathcal{M} \),

1) \( \mathcal{E}(\omega) \) self-adjoint operator.

2) \( \mathcal{E}(\emptyset) = 0, \mathcal{E}(\Omega) = I \)

3) \( \mathcal{E}(\omega \cap \omega') = \mathcal{E}(\omega)\mathcal{E}(\omega') \)

4) \( \text{Si} \ \omega \cap \omega' = , \ then \ \mathcal{E}(\omega \cup \omega') = \mathcal{E}(\omega) + \mathcal{E}(\omega') \)

\( \forall i, j \in G, \forall \omega \in \mathcal{M}, \mu_{ij}(\omega) = E_{ij}(\omega) \)
Identity resolution

Spectral decomposition

\[ \exists \mathcal{E}, \mathcal{M}, \mathcal{W} = \int_{\mathcal{M}} \lambda d\mathcal{E}(\lambda) \]

Definition (Identity resolution)

\( \mathcal{M} \) **Sigma-algebra** \( \mathcal{E} : \mathcal{M} \to B_G \) such that \( \forall \omega, \omega' \in \mathcal{M} \),

1) \( \mathcal{E}(\omega) \) **self-adjoint operator.**
2) \( \mathcal{E}() = 0, \mathcal{E}(\Omega) = I \)
3) \( \mathcal{E}(\omega \cap \omega') = \mathcal{E}(\omega)\mathcal{E}(\omega') \)
4) Si \( \omega \cap \omega' = \) , then \( \mathcal{E}(\omega \cup \omega') = \mathcal{E}(\omega) + \mathcal{E}(\omega') \)

\( \forall i, j \in G, \forall \omega \in \mathcal{M}, \mu_{ij}(\omega) = E_{ij}(\omega) \)
Identity resolution

Spectral decomposition

\[ \exists \mathcal{E}, \mathcal{M}, \mathcal{W} = \int_{\mathcal{M}} \lambda d\mathcal{E}(\lambda) \]

Definition (Identity resolution)

\( \mathcal{M} \) *Sigma-algebra* \( \mathcal{E} : \mathcal{M} \rightarrow B_G \) *such that* \( \forall \omega, \omega' \in \mathcal{M} \),

1. \( \mathcal{E}(\omega) \) *self-adjoint operator.*
2. \( \mathcal{E}(\emptyset) = 0, \mathcal{E}(\Omega) = I \)
3. \( \mathcal{E}(\omega \cap \omega') = \mathcal{E}(\omega)\mathcal{E}(\omega') \)
4. \( Si \ \omega \cap \omega' = , \ then \ \mathcal{E}(\omega \cup \omega') = \mathcal{E}(\omega) + \mathcal{E}(\omega') \)

\( \forall i, j \in G, \forall \omega \in \mathcal{M}, \mu_{ij}(\omega) = E_{ij}(\omega) \)
Extension to a graph

Definition

$\left( X_i \right)_{i \in G}$ Gaussian field with covariance $K$.

$$K = \int_{Sp(W)} g(\lambda) dE(\lambda),$$

- $g$ polynomial : $MA_q^{(W)}$
- $\frac{1}{g}$ polynomial : $AR_p^{(W)}$ ...

Remarks :
- Conditions about $g$
- Equivalence with $Z$
- $K = g(W)$, with normal convergence of the series
- Dependency on $W$
Extension to a graph

Definition

\( (X_i)_{i \in G} \) \textit{Gaussian field with covariance} \( K \).

\[
K = \int_{\text{Sp}(W)} g(\lambda) dE(\lambda),
\]

- \( g \) \textit{polynomial} : \( MA_q^{(W)} \)
- \( \frac{1}{g} \) \textit{polynomial} : \( AR_p^{(W)} \) …

Remarks :

- Conditions about \( g \)
- Equivalence with \( Z \)
- \( K = g(W) \), with normal convergence of the series
- Dependency on \( W \)
Extension to a graph

Definition

\((X_i)_{i \in G}\) \textit{Gaussian field with covariance } \(K\).

\[ K = \int_{\mathcal{S}_p(W)} g(\lambda) \, d\mathcal{E}(\lambda), \]

- \(g\) \textit{polynomial} : \(\mathcal{M}_q(W)\)
- \(\frac{1}{g}\) \textit{polynomial} : \(\mathcal{A}_p(W)\) \ldots

Remarks :

- Conditions about \(g\)
- Equivalence with \(\mathcal{Z}\)
- \(K = g(W)\), with normal convergence of the series
- Dependency on \(W\)
Extension to a graph

Definition

\((X_i)_{i \in G}\) Gaussian field with covariance \(K\).

\[
\text{If } K = \int Sp(W) g(\lambda) dE(\lambda),
\]

- \(g\) polynomial : \(MA_q^{(W)}\)
- \(\frac{1}{g}\) polynomial : \(AR_p^{(W)} \ldots\)

Remarks:

- Conditions about \(g\)
  - Equivalence with \(Z\)
  - \(K = g(W)\), with normal convergence of the series
  - Dependency on \(W\)
Extension to a graph

**Definition**

\[(X_i)_{i \in G} \text{ Gaussian field with covariance } K.\]

\[\text{If } K = \int_{S_p(W)} g(\lambda) dE(\lambda),\]

- **g polynomial** : \(MA_q(W)\)
- \(\frac{1}{g} \text{ polynomial} : AR_p(W) \ldots\)

**Remarks :**

- Conditions about \(g\)
- Equivalence with \(Z\)
- \(K = g(W),\) with normal convergence of the series
- Dependency on \(W\)
Extension to a graph

Definition

\((X_i)_{i \in G}\) Gaussian field with covariance \(K\).

If \(K = \int_{Sp(W)} g(\lambda) dE(\lambda)\),

- \(g\) polynomial : \(MA_q^{(W)}\)
- \(\frac{1}{g}\) polynomial : \(AR_p^{(W)}\) ...

Remarks :

- Conditions about \(g\)
- Equivalence with \(Z\)
- \(K = g(W)\), with normal convergence of the series
- Dependency on \(W\)
Extension to a graph

**Definition**

\[
(X_i)_{i \in G} \text{ Gaussian field with covariance } K.
\]

If 

\[
K = \int_{S_p(W)} g(\lambda) dE(\lambda),
\]

- \(g\) polynomial : \(MA_q^{(W)}\)
- \(\frac{1}{g}\) polynomial : \(AR_p^{(W)}\) …

**Remarks:**

- Conditions about \(g\)
- Equivalence with \(Z\)
- \(K = g(W)\), with normal convergence of the series
- Dependency on \(W\)
Problem:

- \( \Theta \subset \mathbb{R} \) compact
- \( (f_\theta)_{\theta \in \Theta} \) parametric family of densities associated to \( K(f_\theta) = f_\theta(W) \)
- Asymptotic on \( (G_n)_{n \in \mathbb{N}} \) sequence of finite nested subgraphs
  
  Example \( G = \mathbb{Z} : G_n = [1, n] \).
- \( \theta_0 \in \hat{\Theta}, \ X \sim \mathcal{N}(0, K(f_{\theta_0})) \)
- We observe the restriction \( X_n \) of \( X \) to \( G_n \), \( \text{cov} : K_n(f_\theta) \)
- \( m_n = \#G_n \)

Aim: Estimate \( \theta_0 \) by maximum likelihood:

\[
L_n(\theta) := -\frac{1}{2} \left( m_n \log(2\pi) + \log \det (K_n(f_\theta)) + X_n^T (K_n(f_\theta))^{-1} X_n \right)
\]
Problem:

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(G_n)_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}: G_n = [1, n]$.
- $\theta_0 \in \Theta$, $X \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction $X_n$ of $X$ to $G_n$, $\text{cov} : K_n(f_\theta)$
- $m_n = \#G_n$

Aim: Estimate $\theta_0$ by maximum likelihood:

$$L_n(\theta) := -\frac{1}{2} \left( m_n \log(2\pi) + \log \det(K_n(f_\theta)) + X_n^T (K_n(f_\theta))^{-1} X_n \right)$$
Problem:

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(G_n)_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  - Example $G = \mathbb{Z}$: $G_n = [1, n]$.
- $\theta_0 \in \Theta$, $X \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction $X_n$ of $X$ to $G_n$, $\text{cov} : K_n(f_\theta)$
- $m_n = \#G_n$

Aim: Estimate $\theta_0$ by maximum likelihood:

$$L_n(\theta) := -\frac{1}{2} \left( m_n \log(2\pi) + \log \det(K_n(f_\theta)) + X_n^T(K_n(f_\theta))^{-1} X_n \right)$$
Problem:

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(G_n)_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  
  **Example** $G = \mathbb{Z}$ : $G_n = [1, n]$.

  - $\theta_0 \in \hat{\Theta}$, $X \sim \mathcal{N}(0, K(f_{\theta_0}))$

- We observe the restriction $X_n$ of $X$ to $G_n$, $\text{cov} : K_n(f_{\theta})$

- $m_n = \#G_n$

Aim: Estimate $\theta_0$ by maximum likelihood:

$$L_n(\theta) := -\frac{1}{2} \left( m_n \log(2\pi) + \log \det (K_n(f_\theta)) + X_n^T (K_n(f_\theta))^{-1} X_n \right)$$
Problem:

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(G_n)_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z}$ : $G_n = [1, n]$.
- $\theta_0 \in \hat{\Theta}$, $X \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction $X_n$ of $X$ to $G_n$, $\text{cov} : K_n(f_\theta)$
- $m_n = \#G_n$

Aim: Estimate $\theta_0$ by maximum likelihood:

$$L_n(\theta) := -\frac{1}{2} \left( m_n \log(2\pi) + \log \det(K_n(f_\theta)) + X_n^T (K_n(f_\theta))^{-1} X_n \right)$$
Problem:

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(G_n)_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z} : G_n = [1, n]$.
- $\theta_0 \in \hat{\Theta}$, $X \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction $X_n$ of $X$ to $G_n$, $\text{cov} : K_n(f_\theta)$
- $m_n = \#G_n$

Aim: Estimate $\theta_0$ by maximum likelihood:

$$L_n(\theta) := -\frac{1}{2} \left( m_n \log(2\pi) + \log \det(K_n(f_\theta)) + X_n^T (K_n(f_\theta))^{-1} X_n \right)$$
Problem:

- $\Theta \subset \mathbb{R}$ compact
- $(f_\theta)_{\theta \in \Theta}$ parametric family of densities associated to $K(f_\theta) = f_\theta(W)$
- Asymptotic on $(G_n)_{n \in \mathbb{N}}$ sequence of finite nested subgraphs
  Example $G = \mathbb{Z} : G_n = [1, n]$.
- $\theta_0 \in \Theta$, $X \sim \mathcal{N}(0, K(f_{\theta_0}))$
- We observe the restriction $X_n$ of $X$ to $G_n$, $\text{cov} : K_n(f_\theta)$
- $m_n = \#G_n$

Aim: Estimate $\theta_0$ by maximum likelihood:

$$L_n(\theta) := -\frac{1}{2} \left( m_n \log(2\pi) + \log \det (K_n(f_\theta)) + X_n^T (K_n(f_\theta))^{-1} X_n \right)$$
The concrete problem
The concrete problem
Overview

1. Collaboration overview
2. Industrial context
3. Road traffic models: examples
4. Shape Invariant Models
5. Speed Models with Gaussian Field on a Graph
6. Applications
Fig: Graphe G
Applications

Fig: Empirical spectral measure
Fig: Empirical distribution of estimation error
Spectrum of the road network

Applications

Statistical tools for road traffic prediction

INRIA-12th of June
Real datas

**Aim**: Predict missing values on FRC 0 in Toulouse
Real data

**Aim**: Predict missing values on FRC 0 in Toulouse

**Protocol**:
- 10% of data hidden to test the quality of the prediction
- Model: $AR_1$
Real datas

**Aim**: Predict missing values on FRC 0 in Toulouse
A concrete problem
A solution ?

Select Source: Propagated

Date: 2012-02-17 From: 12:00 To: 22:00

Set Default Time Range  Get Traffic Data

Speed Map

Distance (km)

2012-02-17 Romance Standard Time
Let's compare
An alternative: prediction

Observations of the process on a growing sequence of subgraphs of $G$, with missing values.

Let $(O, M)$ be a partition of $G$. The set $O$ will denote the asymptotic for observed values index set, and $B$ the "blind" missing values index set (finite).

Let $(G_N)_{N \in \mathbb{N}}$ be a growing sequence of induced subgraphs of $G$. From now on, we assume that $N$ is large enough to ensure $B \subset G_N$. The observation index set will be denoted $O_N := O \cap G_N$. We consider the restriction $X_{O_N} := (X_i)_{i \in O_N}$, and assume from now on, that we dispose of a consistent estimation procedure $\hat{f}_N$ for $f$, such that there exists $(r_N)$ such that

\[ \mathbb{E} \left[ \left\| \hat{f}_N - f \right\|_\infty^4 \right]^{1/4} \leq r_N. \]

\[ \mathbb{E} \left[ \left\| \hat{f}_N - f \right\|_\infty^2 \right]^{1/2} \leq r_N. \]
Linear Prediction : Kriging

Recall that the best linear predictor of $Z_B$ (this is also the best predictor in the Gaussian case) can be written

$$\tilde{Z}_B = P_{[X_{ON}]}(f)Z_B := a_{BMO}^T(f)(O_N(f))^{-1}X_{ON}. $$

Then, remark that we asymptotically observe $X_O$ and introduce the best linear prediction of $Z_B$ knowing $X_O$ :

$$\tilde{Z}_B := P_{[X_O]}(f)Z_B := a_{BMO}^T(f)(O(f))^{-1}X_O. $$

The blind problem can be formulated as following :

- Estimation step : Estimate $P_{[X_{ON}]}(f)$ by $\hat{P}_{[X_{ON}]}(f) := P_{[X_{ON}]}(\hat{f})$
- Prediction step : Build $\hat{Z}_B := P_{[X_{ON}]}(\hat{f})Z_B$
Extension to graph of Kriging method

Under the assumption that there exists \( m, M > 0 \) such that

\[ \forall t \in \text{Sp}(A), m \leq f(t) \leq M. \]

Risk :

\[
\mathcal{R}_{K,N} = \sup_{Z_B \in [X_B] \text{var}(Z_B)=1} \mathbb{E}\left[ \left( Z_B - \hat{Z}_B \right)^2 \right]^{\frac{1}{2}}.
\]

\[
\mathcal{R}_N = \sup_{Z_B \in [X_B] \text{var}(Z_B)=1} \mathbb{E}\left[ \left( \tilde{Z}_B - \bar{Z}_B \right)^2 \right]^{\frac{1}{2}} + \sup_{Z_B \in [X_B] \text{var}(Z_B)=1} \mathbb{E}\left[ \left( \bar{Z}_B - \hat{Z}_B \right)^2 \right]^{\frac{1}{2}}.
\]

Result :

\[
\mathcal{R}_N \leq \frac{\sqrt{M(m + M)}}{m^2} r_N + \frac{1}{m^2} \left( \frac{M^5}{m} + M^3 \right) \sum_{k \geq d_G(B, (G \setminus G_N))} \left| \left( \frac{1}{f} \right)_k \right|.
\]
Thank you for your Attention