Two semiparametric models arising from applied statistics

Fabrice Gamboa

Institut de Mathématiques de Toulouse

Séminaire du GREMAQ

21th of September 2010

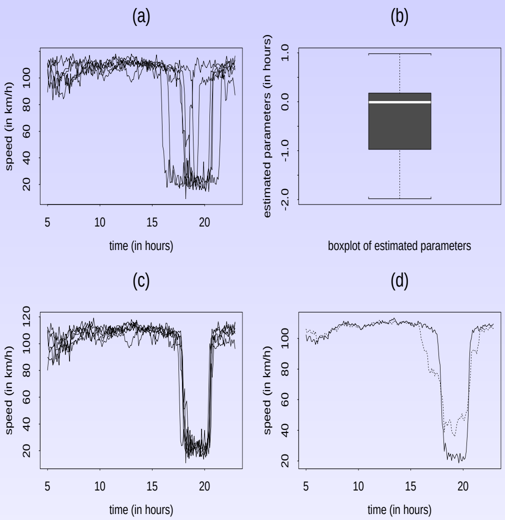
Agenda

1. Problem 1-From G, Loubes Maza EJS 2007
   - Forecasting traffic in Paris
   - The translated regression model
   - Model and estimation method
   - Numerical examples
   - Conclusion and further extensions
2. Problem 2-From Da Veiga, G submitted 2010
   - Physical modelisation and uncertainty
   - Sensitivity analysis
   - Efficient estimation of integrals of non linear functional of a density

Agenda

1. Problem 1-From G, Loubes Maza EJS 2007
   - Forecasting traffic in Paris
   - The translated regression model
   - Model and estimation method
   - Numerical examples
   - Conclusion and further extensions
2. Problem 2-From Da Veiga, G submitted 2010
   - Physical modelisation and uncertainty
   - Sensitivity analysis
   - Efficient estimation of integrals of non linear functional of a density

Forecasting traffic in Paris



The translated regression model

$$Y_{i,j} = f^*(x_i - \theta_j^*) + \varepsilon_{ij}, i = 1, \dots, N, j = 1, \dots, J. \quad (1)$$

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\dots J}$ is an unknown parameter of $\mathbb{R}^J$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

Aim: Statistical inference on $(\theta_j^*)_{j=1\dots J}$, when f^* is unknown

The translated regression model

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

- **f^* is an unknown T-periodic function**
- $(\theta_j^*)_{j=1\ldots J}$ is an unknown parameter of $\mathbb{R}^J$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

Aim: Statistical inference on $(\theta_j^*)_{j=1\ldots J}$, when f^* is unknown

The translated regression model

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

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\ldots J}$ **is an unknown parameter of** $\mathbb{R}^J$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

Aim: Statistical inference on $(\theta_j^*)_{j=1\ldots J}$, when f^* is unknown

The translated regression model

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

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\ldots J}$ is an unknown parameter of $\mathbb{R}^J$
- (ε_{ij}) **is a Gaussian white noise with variance** σ^2

Aim: Statistical inference on $(\theta_j^*)_{j=1\ldots J}$, when f^* is unknown

The translated regression model

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

- f^* is an unknown T-periodic function
- $(\theta_j^*)_{j=1\ldots J}$ is an unknown parameter of $\mathbb{R}^J$
- (ε_{ij}) is a Gaussian white noise with variance σ^2

Aim: Statistical inference on $(\theta_j^*)_{j=1\ldots J}$, when f^* is unknown

Identifiability

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

Replacing

- α^* by
$$\alpha^* + c\mathbf{1} + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J) \qquad (2)$$
- f^* by $f^*(\cdot - c)$

the observation equation remains invariant

Identifiability constraints

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

Examples

$A_1 = \{\alpha \in [-\pi, \pi[^J : \alpha_1 = 0\}$
$A_2 = \{\alpha \in [-\pi, \pi[^J : \sum \alpha_j = 0$ and $\alpha_1 \in [0, 2\pi/J]\}$

Identifiability

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

Replacing

- α^* by

$$\alpha^* + c\mathbf{1} + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J) \tag{2}$$

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

the observation equation remains invariant

Identifiability constraints

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

Examples

$A_1 = \{\alpha \in [-\pi, \pi[^J : \alpha_1 = 0\}$

$A_2 = \{\alpha \in [-\pi, \pi[^J : \sum \alpha_j = 0$ and $\alpha_1 \in [0, 2\pi/J]\}$

Identifiability

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

Replacing

- α^* by
$$\alpha^* + c\mathbf{1} + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J) \tag{2}$$
- f^* by $f^*(\cdot - c)$

the observation equation remains invariant

Identifiability constraints

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

Examples

$A_1 = \{\alpha \in [-\pi, \pi[^J : \alpha_1 = 0\}$

$A_2 = \{\alpha \in [-\pi, \pi[^J : \sum \alpha_j = 0$ and $\alpha_1 \in [0, 2\pi/J]\}$

Identifiability

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

Replacing

- α^* by

$$\alpha^* + c\mathbf{1} + 2k\pi \quad (c \in \mathbb{R}, k \in \mathbb{Z}^J) \tag{2}$$

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

the observation equation remains invariant

Identifiability constraints

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

Examples

$A_1 = \{\alpha \in [-\pi, \pi[^J\colon \alpha_1 = 0\}$

$A_2 = \{\alpha \in [-\pi, \pi[^J\colon \sum \alpha_j = 0$ and $\alpha_1 \in [0, 2\pi/J]\}$

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

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

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

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)

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

- $c_l(f^*)$ is the Fourier coefficient of f^*
- (w_{jl}) is a complex Gaussian white noise with variance σ^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 σ^2/N

Building a M-function

- Re phased Fourier coefficients

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

- Mean of Re phased Fourier coefficients

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

$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl}$ and $\hat{c}_l(\alpha^*) = c_l(f^*) + 1/J \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$

Building a M-function

- **Re phased Fourier coefficients**

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

- Mean of Re phased Fourier coefficients

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

$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl}$ and $\hat{c}_l(\alpha^*) = c_l(f^*) + 1/J \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$

Building a M-function

- Re phased Fourier coefficients

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

- **Mean of Re phased Fourier coefficients**

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

$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl}$ and $\hat{c}_l(\alpha^*) = c_l(f^*) + 1/J \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$

Building a M-function

- Re phased Fourier coefficients

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

- Mean of Re phased Fourier coefficients

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

$\tilde{c}_{jl}(\alpha^*) = c_l(f^*) + e^{il\alpha_j^*} w_{jl}$ and $\hat{c}_l(\alpha^*) = c_l(f^*) + 1/J \sum_{j=1}^{J} e^{il\alpha_j^*} w_{jl}$

The M-function

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

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

- (δ_l) is a positive sequence of weights discussed later

The M-function

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

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

- (δ_l) is a positive sequence of weights discussed later

Convergence

Theorem

Let $(\widehat{\alpha}_n)$ *be any sequence of minimizers for* M_n. *Assume further*

- *the function* f^* *lies in* $L^2([0,T])$
- *the weight sequence satisfies* $\sum_{l \in \mathbb{Z}} l^2 \delta_l^4 < +\infty$

Then $\widehat{\alpha}_n$ *converges in probability to* α^*

Proof.

Falls in the theory of M-estimation (uniform convergence+ minimum of the limit function+long and tedious calculations!!) □

Convergence

Theorem

Let $(\widehat{\alpha}_n)$ be any sequence of minimizers for M_n. Assume further

- ***the function f^* lies in $L^2([0,T])$***
- *the weight sequence satisfies $\sum_{l\in\mathbb{Z}} l^2\delta_l^4 < +\infty$*

Then $\widehat{\alpha}_n$ converges in probability to α^*

Proof.

Falls in the theory of M-estimation (uniform convergence+ minimum of the limit function+long and tedious calculations!!) □

Convergence

Theorem

Let $(\widehat{\alpha}_n)$ be any sequence of minimizers for M_n. Assume further

- *the function f^* lies in $L^2([0,T])$*
- ***the weight sequence satisfies*** $\sum_{l\in\mathbb{Z}} l^2\delta_l^4 < +\infty$

Then $\widehat{\alpha}_n$ converges in probability to α^*

Proof.

Falls in the theory of M-estimation (uniform convergence+ minimum of the limit function+long and tedious calculations!!) □

Convergence

Theorem

Let $(\widehat{\alpha}_n)$ be any sequence of minimizers for M_n. Assume further

- *the function f^* lies in $L^2([0,T])$*
- *the weight sequence satisfies $\sum_{l\in\mathbb{Z}} l^2\delta_l^4 < +\infty$*

Then $\widehat{\alpha}_n$ converges in probability to α^*

Proof.

Falls in the theory of M-estimation (uniform convergence+ minimum of the limit function+long and tedious calculations!!) □

Convergence

Theorem

Let $(\widehat{\alpha}_n)$ *be any sequence of minimizers for* M_n*. Assume further*

- *the function* f^* *lies in* $L^2([0,T])$
- *the weight sequence satisfies* $\sum_{l\in\mathbb{Z}} l^2\delta_l^4 < +\infty$

Then $\widehat{\alpha}_n$ *converges in probability to* α^*

Proof.

Falls in the theory of M-estimation (uniform convergence+ minimum of the limit function+long and tedious calculations!!) □

Asymptotic normality

Working with parameter set A_1

$$A_1 = \{\alpha \in [-\pi,\pi[^J : \alpha_1 = 0\} = [-\pi,\pi[^{J-1}$$

Theorem

Under the assumptions of last Theorem. Assume further

- *the weight sequence satisfies $\sum_{l\in\mathbb{Z}} l^4\delta_l^4 < +\infty$*

Then $\sqrt{N}(\widehat{\alpha}_n - \alpha^*)$ converges in distribution to $\mathcal{N}(0,\Gamma)$

$$\Gamma = \frac{\sigma^2 \sum_{l\in\mathbb{Z}} \delta_l^4 l^2 |c_l(f^*)|^2}{\left(\sum_{l\in\mathbb{Z}} \delta_l^2 l^2 |c_l(f^*)|^2\right)^2} (I_{J-1} + \mathbf{1}\mathbf{1}^T)$$

Asymptotic normality

Working with parameter set A_1

$$A_1 = \{\alpha \in [-\pi, \pi[^J : \alpha_1 = 0\} = [-\pi, \pi[^{J-1}$$

Theorem

Under the assumptions of last Theorem. Assume further

- *the weight sequence satisfies* $\sum_{l \in \mathbb{Z}} l^4 \delta_l^4 < +\infty$

Then $\sqrt{N}(\widehat{\alpha}_n - \alpha^*)$ *converges in distribution to* $\mathcal{N}(0, \Gamma)$

$$\Gamma = \frac{\sigma^2 \sum_{l \in \mathbb{Z}} \delta_l^4 l^2 |c_l(f^*)|^2}{\left(\sum_{l \in \mathbb{Z}} \delta_l^2 l^2 |c_l(f^*)|^2\right)^2} (I_{J-1} + \mathbf{1}\mathbf{1}^T)$$

Asymptotic normality

Working with parameter set A_1

$$A_1 = \{\alpha \in [-\pi, \pi[^J \colon \alpha_1 = 0\} = [-\pi, \pi[^{J-1}$$

Theorem

Under the assumptions of last Theorem. Assume further

- ***the weight sequence satisfies*** $\sum_{l \in \mathbb{Z}} l^4 \delta_l^4 < +\infty$

Then $\sqrt{N}(\widehat{\alpha}_n - \alpha^*)$ *converges in distribution to* $\mathcal{N}(0, \Gamma)$

$$\Gamma = \frac{\sigma^2 \sum_{l \in \mathbb{Z}} \delta_l^4 l^2 |c_l(f^*)|^2}{\left(\sum_{l \in \mathbb{Z}} \delta_l^2 l^2 |c_l(f^*)|^2\right)^2} (I_{J-1} + \mathbf{1}\mathbf{1}^T)$$

Asymptotic normality

Working with parameter set A_1

$$A_1 = \{\alpha \in [-\pi, \pi[^J \colon \alpha_1 = 0\} = [-\pi, \pi[^{J-1}$$

Theorem

Under the assumptions of last Theorem. Assume further

- *the weight sequence satisfies* $\sum_{l \in \mathbb{Z}} l^4 \delta_l^4 < +\infty$

Then $\sqrt{N}(\widehat{\alpha}_n - \alpha^*)$ ***converges in distribution to*** $\mathcal{N}(0, \Gamma)$

$$\Gamma = \frac{\sigma^2 \sum_{l \in \mathbb{Z}} \delta_l^4 l^2 |c_l(f^*)|^2}{\left(\sum_{l \in \mathbb{Z}} \delta_l^2 l^2 |c_l(f^*)|^2\right)^2} (I_{J-1} + \mathbf{1}\mathbf{1}^T)$$

Proof.

Again classical (Taylor expansion+ quite long and tedious calculations!!) □

Remarks

- On the one hand very weak assumption on f^*
- On the other hand certainly not the *least* variance

Proof.

Again classical (Taylor expansion+ quite long and tedious calculations!!) □

Remarks

- **On the one hand very weak assumption on f^***
- On the other hand certainly not the *least* variance

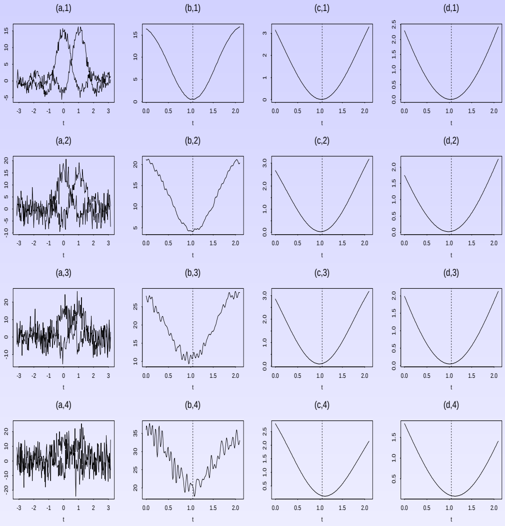
Proof.

Again classical (Taylor expansion+ quite long and tedious calculations!!) □

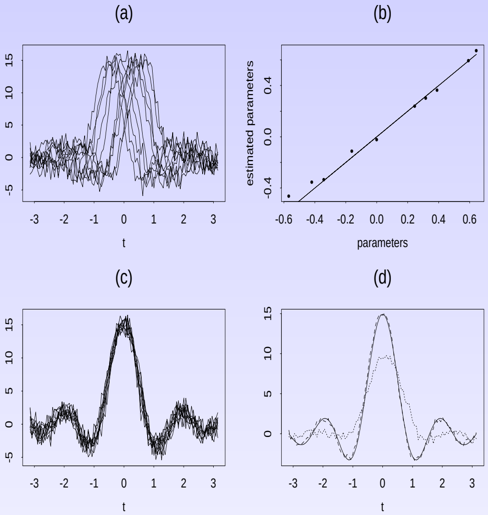
Remarks

- On the one hand very weak assumption on f^*
- **On the other hand certainly not the *least* variance**

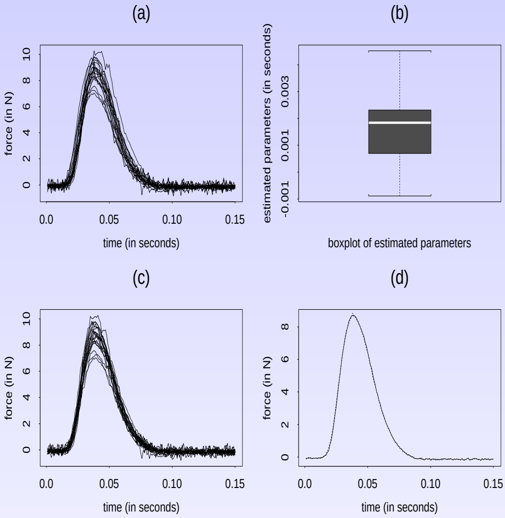
the M-function



An artificial data example



A real data example



Conclusion and further extensions

- **Simple characterization of translation in Fourier domain leads a very good statistical method: Numerically simple-Good asymptotic properties**
- Extension to non commutative group by Bigot, Loubes, Vimond (2010) in Revision in PTRF
- Efficient version for SIM (in the semi-parametric meaning) performed by Vimond (2010) in Annals of Statistics. (δ_n) replaced by a window on the Fourier coefficients
- Two dimensional regression problem including both translation, rotation and scale. Bigot, G, Vimond (2009) in SIAM Journal on Imaging
- An interesting open problem: Adaptive model selection in order to both estimate optimally f^* and the deformation parameters.

Conclusion and further extensions

- Simple characterization of translation in Fourier domain leads a very good statistical method: Numerically simple-Good asymptotic properties
- **Extension to non commutative group by Bigot, Loubes, Vimond (2010) in Revision in PTRF**
- Efficient version for SIM (in the semi-parametric meaning) performed by Vimond (2010) in Annals of Statistics. (δ_n) replaced by a window on the Fourier coefficients
- Two dimensional regression problem including both translation, rotation and scale. Bigot, G, Vimond (2009) in SIAM Journal on Imaging
- An interesting open problem: Adaptive model selection in order to both estimate optimally f^* and the deformation parameters.

Conclusion and further extensions

- Simple characterization of translation in Fourier domain leads a very good statistical method: Numerically simple-Good asymptotic properties
- Extension to non commutative group by Bigot, Loubes, Vimond (2010) in Revision in PTRF
- **Efficient version for SIM (in the semi-parametric meaning) performed by Vimond (2010) in Annals of Statistics.** (δ_n) **replaced by a window on the Fourier coefficients**
- Two dimensional regression problem including both translation, rotation and scale. Bigot, G, Vimond (2009) in SIAM Journal on Imaging
- An interesting open problem: Adaptive model selection in order to both estimate optimally f^* and the deformation parameters.

Conclusion and further extensions

- Simple characterization of translation in Fourier domain leads a very good statistical method: Numerically simple-Good asymptotic properties
- Extension to non commutative group by Bigot, Loubes, Vimond (2010) in Revision in PTRF
- Efficient version for SIM (in the semi-parametric meaning) performed by Vimond (2010) in Annals of Statistics. (δ_n) replaced by a window on the Fourier coefficients
- **Two dimensional regression problem including both translation, rotation and scale. Bigot, G, Vimond (2009) in SIAM Journal on Imaging**
- An interesting open problem: Adaptive model selection in order to both estimate optimally f^* and the deformation parameters.

Conclusion and further extensions

- Simple characterization of translation in Fourier domain leads a very good statistical method: Numerically simple-Good asymptotic properties
- Extension to non commutative group by Bigot, Loubes, Vimond (2010) in Revision in PTRF
- Efficient version for SIM (in the semi-parametric meaning) performed by Vimond (2010) in Annals of Statistics. (δ_n) replaced by a window on the Fourier coefficients
- Two dimensional regression problem including both translation, rotation and scale. Bigot, G, Vimond (2009) in SIAM Journal on Imaging
- **An interesting open problem: Adaptive model selection in order to both estimate optimally f^* and the deformation parameters.**

Agenda

1. Problem 1-From G, Loubes Maza EJS 2007
   - Forecasting traffic in Paris
   - The translated regression model
   - Model and estimation method
   - Numerical examples
   - Conclusion and further extensions
2. Problem 2-From Da Veiga, G submitted 2010
   - Physical modelisation and uncertainty
   - Sensitivity analysis
   - Efficient estimation of integrals of non linear functional of a density

Modelling

- Physical, chemical or econometrical phenomena: modelling
- Exits, entering variables, parameters, etc.
- Confidence in the du model?
  - Uncertainty on entering variables,
  - Uncertainty on parameters.

Consequence: Uncertainty on exit variables

Model:

$$Y = f(X_1, \ldots, X_d)$$

with Y model exits and $\mathbf{X} = (X_1, \ldots, X_d)$ vector of d factors (entering variables + badly known parameters).

Aim

Which entering variables have most impact on model exit?

Probabilistic frame

Sensitivity indexes

$$S_i = \frac{\mathrm{Var}(\mathbb{E}(Y|X_i))}{\mathrm{Var}(Y)} \qquad \text{(1th order)}$$

- Interpretation:
  - $\mathbb{E}(Y|X_i)$: function of X_i only that best *approximates* Y
  - $\mathrm{Var}(\mathbb{E}(Y|X_i))$: fluctuation of the exit variable if it would be of X_i only
  - then normalisation by the global fluctuation of Y, $\mathrm{Var}(Y)$

Probabilistic frame

Sensitivity indexes

$$S_i = \frac{\mathrm{Var}(\mathbb{E}(Y|X_i))}{\mathrm{Var}(Y)} \qquad \text{(1th order)}$$

- Interpretation:
  - $\mathbb{E}(Y|X_i)$: function of X_i only that best *approximates* Y
  - $\mathrm{Var}(\mathbb{E}(Y|X_i))$: fluctuation of the exit variable if it would be of X_i only
  - then normalisation by the global fluctuation of Y, $\mathrm{Var}(Y)$

Probabilistic frame

- Usual techniques to estimate S_i:
  - Sobol (functional variance analysis- Hoefding decomposition)
  - FAST (Biased method based on FFT)

Limitations

Run number
Independence assumption on entering variables

- Examples of problematic cases
  - Oil reservoir (CPU time)
  - Chemical cinetic (no independence)

Probabilistic frame

- Usual techniques to estimate S_i:
  - Sobol (functional variance analysis- Hoefding decomposition)
  - FAST (Biased method based on FFT)

Limitations

Run number
Independence assumption on entering variables

- Examples of problematic cases
  - Oil reservoir (CPU time)
  - Chemical cinetic (no independence)

Probabilistic frame

- Usual techniques to estimate S_i:
  - Sobol (functional variance analysis- Hoefding decomposition)
  - FAST (Biased method based on FFT)

Limitations

Run number

Independence assumption on entering variables

- Examples of problematic cases
  - Oil reservoir (CPU time)
  - Chemical cinetic (no independence)

Probabilistic frame

Without independence asumption

- Oakley 0'Hagan method:
  - $f(\mathbf{X})$ approximated by a Gaussian process
  - Further integration / distribution of $\mathbf{X}$

Involved numerical evaluation of multiple integrals

- Da-Veiga, Wahl and G method :
  - $\mathbb{E}(Y|X_i)$ approximated by local polynomials
  - Plug-in for $\mathrm{Var}(\mathbb{E}(Y|X_i))$

Non parametrical rates of convergence

Probabilistic frame

Without independence asumption

- Oakley 0'Hagan method:
  - $f(\mathbf{X})$ approximated by a Gaussian process
  - Further integration / distribution of $\mathbf{X}$

Involved numerical evaluation of multiple integrals

- Da-Veiga, Wahl and G method :
  - $\mathbb{E}(Y|X_i)$ approximated by local polynomials
  - Plug-in for $\mathrm{Var}(\mathbb{E}(Y|X_i))$

Non parametrical rates of convergence

Probabilistic frame

Without independence asumption

- Oakley 0'Hagan method:
  - $f(\mathbf{X})$ approximated by a Gaussian process
  - Further integration / distribution of $\mathbf{X}$

Involved numerical evaluation of multiple integrals

- Da-Veiga, Wahl and G method :
  - $\mathbb{E}(Y|X_i)$ approximated by local polynomials
  - Plug-in for $\mathrm{Var}(\mathbb{E}(Y|X_i))$

Non parametrical rates of convergence

Problem rewriting

(X_i, Y_i), $i = 1, \dots, n$ sample of (X, Y), density f
We wish to estimate

$$\mathrm{Var}(\mathbb{E}(Y|X)) = \mathbb{E}(\mathbb{E}(Y|X)^2) - (\mathbb{E}(Y))^2.$$

We define the operator T with

$$T(f) = \mathbb{E}(\mathbb{E}(Y|X)^2) = \iint \left(\frac{\int y f(x,y)dy}{\int f(x,y)dy} \right)^2 f(x,y)dxdy.$$

Existing works

B. Laurent (1996) studied efficient estimation of non linear functional as

$$T(f) = \int \phi(f(\mathbf{x}), \mathbf{x}) d\mathbf{x}$$

where ϕ is a regular function and $X_1, \ldots, X_n$ sample of density f.

Idea to build an estimate

Expand $u \to \phi(u, x) \in C^3$ around a preliminary estimate $\hat{f}$ of f

Existing works

B. Laurent (1996) studied efficient estimation of non linear functional as

$$T(f) = \int \phi(f(\mathbf{x}), \mathbf{x}) d\mathbf{x}$$

where ϕ is a regular function and $X_1, \ldots, X_n$ sample of density f.

Idea to build an estimate

Expand $u \to \phi(u, x) \in C^3$ around a preliminary estimate $\hat{f}$ of f

Existing works

B. Laurent (1996) studied efficient estimation of non linear functional as

$$T(f) = \int \phi(f(\mathbf{x}), \mathbf{x}) d\mathbf{x}$$

where ϕ is a regular function and $X_1, \ldots, X_n$ sample of density f.

Idea to build an estimate

Expand $u \to \phi(u, x) \in C^3$ around a preliminary estimate $\hat{f}$ of f

$$\begin{aligned}\phi(f(\mathbf{x}), \mathbf{x}) &= \phi(\hat{f}(\mathbf{x}), \mathbf{x}) + \frac{\partial \phi}{\partial u}(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x})\right) \\ &\quad + \frac{\partial^2 \phi}{\partial u^2}(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x})\right)^2 + \gamma.\end{aligned}$$

Existing works

B. Laurent (1996) studied efficient estimation of non linear functional as

$$T(f) = \int \phi(f(\mathbf{x}), \mathbf{x}) d\mathbf{x}$$

where ϕ is a regular function and $X_1, \ldots, X_n$ sample of density f.

Idea to build an estimate

Expand $u \to \phi(u, x) \in C^3$ around a preliminary estimate $\hat{f}$ of f

$$\begin{aligned} \phi(f(\mathbf{x}), \mathbf{x}) &= \phi(\hat{f}(\mathbf{x}), \mathbf{x}) + \phi_1'(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right) \\ &+ \phi_1''(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right)^2 + \gamma. \end{aligned}$$

Work of B. Laurent

$$T(f) = \int \phi(f(\mathbf{x}), \mathbf{x}) d\mathbf{x}$$

- Γ_n neglectable
- $\int H(\hat{f}, \cdot)\, f$ linear functional of f: no problem
- $\int K(\hat{f}, \cdot)\, f^2$ quadratic functional of f : optimal estimation (efficiency)?

Work of B. Laurent

$$\begin{aligned} T(f) &= \int \phi(f(\mathbf{x}),\mathbf{x})d\mathbf{x} \\ &= \int \phi(\hat{f}(\mathbf{x}),\mathbf{x})d\mathbf{x} + \int \phi_1'(\hat{f}(\mathbf{x}),\mathbf{x})\left(f(\mathbf{x}) - \hat{f}(\mathbf{x})\right)d\mathbf{x} \\ &+ \int \phi_1''(\hat{f}(\mathbf{x}),\mathbf{x})\left(f(\mathbf{x}) - \hat{f}(\mathbf{x})\right)^2 d\mathbf{x} + \Gamma_n \end{aligned}$$

- Γ_n neglectable
- $\int H(\hat{f},\cdot)\, f$ linear functional of f: no problem
- $\int K(\hat{f},\cdot)\, f^2$ quadratic functional of f : optimal estimation (efficiency)?

Work of B. Laurent

$$
\begin{aligned}
T(f) &= \int \phi(f(\mathbf{x}),\mathbf{x})d\mathbf{x} \\
&= \int \phi(\hat{f}(\mathbf{x}),\mathbf{x})d\mathbf{x} + \int \phi_1'(\hat{f}(\mathbf{x}),\mathbf{x})\left(f(\mathbf{x}) - \hat{f}(\mathbf{x})\right) d\mathbf{x} \\
&\quad + \int \phi_1''(\hat{f}(\mathbf{x}),\mathbf{x})\left(f(\mathbf{x}) - \hat{f}(\mathbf{x})\right)^2 d\mathbf{x} + \Gamma_n \\
&= \int G(\hat{f},\cdot) + \int H(\hat{f},\cdot)\, f + \int K(\hat{f},\cdot)\, f^2 + \Gamma_n
\end{aligned}
$$

- Γ_n neglectable
- $\int H(\hat{f},\cdot)\, f$ linear functional of f: no problem
- $\int K(\hat{f},\cdot)\, f^2$ quadratic functional of f : optimal estimation (efficiency)?

Work of B. Laurent

$$
\begin{aligned}
T(f) &= \int \phi(f(\mathbf{x}), \mathbf{x}) d\mathbf{x} \\
&= \int \phi(\hat{f}(\mathbf{x}), \mathbf{x}) d\mathbf{x} + \int \phi_1'(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right) d\mathbf{x} \\
&+ \int \phi_1''(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right)^2 d\mathbf{x} + \Gamma_n \\
&= \int G(\hat{f}, \cdot) + \int H(\hat{f}, \cdot) f + \int K(\hat{f}, \cdot) f^2 + \Gamma_n
\end{aligned}
$$

- Γ_n neglectable
- $\int H(\hat{f}, \cdot) f$ linear functional of f: no problem
- $\int K(\hat{f}, \cdot) f^2$ quadratic functional of f : optimal estimation (efficiency)?

Work of B. Laurent

$$
\begin{aligned}
T(f) &= \int \phi(f(\mathbf{x}), \mathbf{x}) d\mathbf{x} \\
&= \int \phi(\hat{f}(\mathbf{x}), \mathbf{x}) d\mathbf{x} + \int \phi_1'(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right) d\mathbf{x} \\
&+ \int \phi_1''(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right)^2 d\mathbf{x} + \Gamma_n \\
&= \int G(\hat{f}, \cdot) + \int H(\hat{f}, \cdot) f + \int K(\hat{f}, \cdot) f^2 + \Gamma_n
\end{aligned}
$$

- Γ_n neglectable
- $\int H(\hat{f}, \cdot) f$ linear functional of f: no problem
- $\int K(\hat{f}, \cdot) f^2$ quadratic functional of f : optimal estimation (efficiency)?

Work of B. Laurent

$$
\begin{aligned}
T(f) &= \int \phi(f(\mathbf{x}), \mathbf{x}) d\mathbf{x} \\
&= \int \phi(\hat{f}(\mathbf{x}), \mathbf{x}) d\mathbf{x} + \int \phi_1'(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right) d\mathbf{x} \\
&+ \int \phi_1''(\hat{f}(\mathbf{x}), \mathbf{x}) \left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right)^2 d\mathbf{x} + \Gamma_n \\
&= \int G(\hat{f}, \cdot) + \int H(\hat{f}, \cdot) f + \int K(\hat{f}, \cdot) f^2 + \Gamma_n
\end{aligned}
$$

- Γ_n neglectable
- $\int H(\hat{f}, \cdot) f$ linear functional of f: no problem
- $\int K(\hat{f}, \cdot) f^2$ quadratic functional of f : optimal estimation (efficiency)?

Efficient estimation of quadratic functional

We wish to estimate

$$\int K(\hat{f},\cdot)\, f^2 \equiv \int \psi\, f^2$$

where $f \in \mathbb{L}^2(dx)$ density of X, sample $(X_1,\ldots,X_n)$.

Efficient estimation of quadratic functional

We wish to estimate

$$\int K(\hat{f},\cdot)\, f^2 \equiv \int \psi\, f^2$$

where $f \in \mathbb{L}^2(dx)$ density of X, sample $(X_1,\ldots,X_n)$.

Let $(p_i)_{i\in D}$ orthonormal basis of $\mathbb{L}^2(dx)$, D numerable, $a_i = \int f p_i$.

Projection estimator

Efficient estimation of quadratic functional

We wish to estimate

$$\int K(\hat{f},\cdot)\, f^2 \equiv \int \psi f^2$$

where $f \in \mathbb{L}^2(dx)$ density of X, sample $(X_1,\dots,X_n)$.

Let $(p_i)_{i\in D}$ orthonormal basis of $\mathbb{L}^2(dx)$, D numerable, $a_i = \int f p_i$.

Projection estimator

Particular case: $\psi = 1$

- $\displaystyle\int f^2 = \sum_{i\in D} a_i^2 \Rightarrow \tilde{\theta} = \sum_{i\in M} \hat{a}_i^2,\ \hat{a}_i = \frac{1}{n}\sum_{j=1}^{n} p_i(X_j),\ M \subset D$
- Bias reduction: $\displaystyle\tilde{\theta} = \frac{1}{n(n-1)} \sum_{i\in M} \sum_{j\neq k=1}^{n} p_i(X_j)p_i(X_k)$

Efficient estimation of a quadratic functional

$$\psi = 1 \qquad \tilde{\theta} = \frac{1}{n(n-1)} \sum_{i \in M} \sum_{j \neq k = 1}^{n} p_i(X_j) p_i(X_k)$$

Efficient estimation of a quadratic functional

$$\psi = 1 \qquad \tilde{\theta} = \frac{1}{n(n-1)} \sum_{i \in M} \sum_{j \neq k = 1}^{n} p_i(X_j) p_i(X_k)$$

$$\text{Bias}(\tilde{\theta}) = -\int (S_M f - f)^2 \quad \text{avec} \quad S_M f = \sum_{i \in M} a_i p_i$$

Only projection bias

Efficient estimation of a quadratic functional

$$\psi = 1 \qquad \tilde{\theta} = \frac{1}{n(n-1)} \sum_{i \in M} \sum_{j \neq k = 1}^{n} p_i(X_j) p_i(X_k)$$

$$\mathrm{Bias}(\tilde{\theta}) = -\int (S_M f - f)^2 \quad \text{avec} \quad S_M f = \sum_{i \in M} a_i p_i$$

Only projection bias

any ψ: $-\int (S_M f - f)^2 \psi = 2\int (S_M f) f \psi - \int (S_M f)^2 \psi - \int \psi f^2$

Efficient estimation of a quadratic functional

$$\psi = 1 \qquad \tilde{\theta} = \frac{1}{n(n-1)} \sum_{i \in M} \sum_{j \neq k = 1}^{n} p_i(X_j) p_i(X_k)$$

$$\mathrm{Bias}(\tilde{\theta}) = -\int (S_M f - f)^2 \quad \text{avec} \quad S_M f = \sum_{i \in M} a_i p_i$$

Only projection bias

any ψ: $-\int (S_M f - f)^2 \psi = 2\int (S_M f) f \psi - \int (S_M f)^2 \psi - \int \psi f^2$

Proposed estimator:

$$\begin{aligned} \hat{\theta} \quad &= \quad \frac{2}{n(n-1)} \sum_{i \in M} \sum_{j \neq k = 1}^{n} p_i(X_j)(p_i \psi)(X_k) \\ & \quad - \frac{1}{n(n-1)} \sum_{i, i' \in M} \sum_{j \neq k = 1}^{n} p_i(X_j) p_{i'}(X_k) \int p_i p_{i'} \psi(x) dx \end{aligned}$$

Efficient estimation of a quadratic functional

Theorem (Efficient estimation of $\theta = \int \psi f^2$)

(i) If $|M|/n \to 0$

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \to \mathcal{N}\left(0, \Lambda(f,\psi)\right),$$

$$\left|\mathbb{E}\left(\hat{\theta} - \theta\right)^2 - \Lambda(f,\psi)\right| \leq \gamma_1 \left[\frac{|M|}{n} + \|S_M f - f\|_2 + \|S_M(f\psi) - f\psi\|_2\right]$$

$$\text{where} \quad \Lambda(f,\psi) = 4\left[\int f^3\psi^2 - \left(\int f^2\psi\right)^2\right],$$

(ii) Further

$$\mathbb{E}\left(\hat{\theta} - \theta\right)^2 \leq \gamma_2 \frac{|M|}{n}.$$

Efficient estimation of a quadratic functional

Theorem (Efficient estimation of $\theta = \int \psi f^2$)

(i) If $|M|/n \to 0$

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \to \mathcal{N}\left(0, \Lambda(f,\psi)\right),$$

$$\left|\mathbb{E}\left(\hat{\theta} - \theta\right)^2 - \Lambda(f,\psi)\right| \le \gamma_1 \left[\frac{|M|}{n} + \|S_M f - f\|_2 + \|S_M(f\psi) - f\psi\|_2\right]$$

$$\text{where} \quad \Lambda(f,\psi) = 4\left[\int f^3\psi^2 - \left(\int f^2\psi\right)^2\right],$$

(ii) Further

$$\mathbb{E}\left(\hat{\theta} - \theta\right)^2 \le \gamma_2 \frac{|M|}{n}.$$

$$\lim_{n\to\infty} n\mathbb{E}\left(\hat{\theta}_n - \theta\right)^2 = \Lambda(f,\psi)$$

Efficient estimation of a quadratic functional

Theorem (Efficient estimation of $\theta = \int \psi f^2$)

(i) If $|M|/n \to 0$

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \to \mathcal{N}\left(0, \Lambda(f,\psi)\right),$$

$$\left|\mathbb{E}\left(\hat{\theta} - \theta\right)^2 - \Lambda(f,\psi)\right| \le \gamma_1 \left[\frac{|M|}{n} + \|S_M f - f\|_2 + \|S_M(f\psi) - f\psi\|_2\right]$$

$$\text{where} \quad \Lambda(f,\psi) = 4\left[\int f^3\psi^2 - \left(\int f^2\psi\right)^2\right],$$

(ii) Further

$$\mathbb{E}\left(\hat{\theta} - \theta\right)^2 \le \gamma_2 \frac{|M|}{n}.$$

$\Lambda(f,\psi)$ optimal (Cramér-Rao semiparametric bound)

Efficient estimation of a quadratic functional

Theorem (Efficient estimation of $\theta = \int \psi f^2$)

(i) If $|M|/n \to 0$

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \to \mathcal{N}\left(0, \Lambda(f,\psi)\right),$$

$$\left|\mathbb{E}\left(\hat{\theta} - \theta\right)^2 - \Lambda(f,\psi)\right| \leq \gamma_1 \left[\frac{|M|}{n} + \|S_M f - f\|_2 + \|S_M(f\psi) - f\psi\|_2\right]$$

$$\text{where} \quad \Lambda(f,\psi) = 4\left[\int f^3\psi^2 - \left(\int f^2\psi\right)^2\right],$$

(ii) Further

$$\mathbb{E}\left(\hat{\theta} - \theta\right)^2 \leq \gamma_2 \frac{|M|}{n}.$$

$\Lambda(f,\psi)$ optimal (Cramér-Rao semiparametric bound)

Coming back to Taylor expansion

$$T(f) = \int G(\hat{f},\cdot) + \int \left(\phi_1'(\hat{f},\cdot) - \hat{f}\phi_1''(\hat{f},\cdot)\right) f + \int \frac{1}{2}\phi_1''(\hat{f},\cdot)\, f^2 + \Gamma_n$$

estimated by

$$\begin{aligned}
\hat{T}_n &= \int G(\hat{f},\cdot) + \frac{1}{n_2}\sum_{j=1}^{n_2}\left(\phi_1'(\hat{f},\cdot) - \hat{f}\phi_1''(\hat{f},\cdot)\right)(X_j) \\
&+ \frac{2}{n_2(n_2-1)}\sum_{i\in M}\sum_{j\neq k=1}^{n_2} p_i(X_j)(p_i\frac{1}{2}\phi_1''(\hat{f},\cdot))(X_k) \\
&- \frac{1}{n_2(n_2-1)}\sum_{i,i'\in M}\sum_{j\neq k=1}^{n_2} p_i(X_j)p_{i'}(X_k)\int p_i p_{i'}\frac{1}{2}\phi_1''(\hat{f},\cdot)
\end{aligned}$$

Conclusion of Laurent work

Theorem (efficient estimation of $T(f)$)

(i) If $|M|/n \to 0$

$$\sqrt{n}\left(\hat{T}_n - T(f)\right) \to \mathcal{N}(0, C(f,\phi))$$

(ii)

$$\lim_{n\to\infty} n\mathbb{E}(\hat{T}_n - T(f))^2 = C(f,\phi)$$

where

$$C(f,\phi) = \int (\phi_1'(f,\cdot))^2 f - \left(\int \phi_1'(f,\cdot) f\right)^2$$

$C(f,\phi)$ optimal variance

Coming back to sensitivity indexe estimation

We wish to estimate

$$T(f) = \mathbb{E}(\mathbb{E}(Y|X)^2) = \iint \left(\frac{\int y f(x,y)dy}{\int f(x,y)dy} \right)^2 f(x,y)dxdy.$$

We follow the previous method : expansion of $T(f)$ around of preliminary estimator $\hat{f}$ of f.

Expansion of $T(f)$

$$
\begin{aligned}
T(f) &= \iint \left[2y\hat{m}(x) - \hat{m}(x)^2\right] f(x,y)dxdy \\
&+ \iiint \frac{1}{\left(\int \hat{f}(x,y)dy\right)} \left[yz + \hat{m}(x)^2 - (y+z)\hat{m}(x)\right] f(x,y)f(x,z)dxdydz \\
&+ \Gamma_n \\
&= \iint H(\hat{f},x,y)f(x,y)dxdy + \iiint K(\hat{f},x,y,z)f(x,y)f(x,z)dxdydz \\
&+ \Gamma_n
\end{aligned}
$$

where

$$
\begin{aligned}
H(\hat{f},x,y) &= 2y\hat{m}(x) - \hat{m}(x)^2 \\
K(\hat{f},x,y,z) &= \frac{1}{\left(\int \hat{f}(x,y)dy\right)} \left[yz + \hat{m}(x)^2 - (y+z)\hat{m}(x)\right].
\end{aligned}
$$

Quadratic functional estimation

we wish to estimate

$$\iiint K(\hat{f}, x, y, z) f(x, y) f(x, z) = \iiint \psi(x, y, z) f(x, y) f(x, z)$$

- Idea : same procedure as Laurent
- looking an estimator having bias

$$\begin{aligned}
- & \iiint [S_M f(x, y) - f(x, y)][S_M f(x, z) - f(x, z)] \psi(x, y, z) dx dy dz \\
= & \ 2 \iiint S_M f(x, y) f(x, z) \psi(x, y, z) dx dy dz \\
& - \iiint S_M f(x, y) S_M f(x, z) \psi(x, y, z) dx dy dz \\
& - \iiint f(x, y) f(x, z) \psi(x, y, z) dx dy dz
\end{aligned}$$

Our estimator:

$$
\begin{aligned}
\hat{\theta}_n &= \frac{2}{n(n-1)} \sum_{i \in M} \sum_{j \neq k = 1}^{n} p_i(X_j, Y_j) \int p_i(X_k, u) \psi(X_k, u, Y_k) du \\
&\quad - \frac{1}{n(n-1)} \sum_{i, i' \in M} \sum_{j \neq k = 1}^{n} p_i(X_j, Y_j) p_{i'}(X_k, Y_k) \\
&\quad \iiint p_i(x, y) p_{i'}(x, z) \psi(x, y, z) dx dy dz.
\end{aligned}
$$

Quadratic functional estimation

Theorem (Efficient estimation of $\theta = \int \psi(x,y,z) f(x,y) f(x,z)$)

(i) If $|M|/n \to 0$

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \to \mathcal{N}\left(0, \Lambda(f,\psi)\right),$$

$$\left|\mathbb{E}\left(\hat{\theta} - \theta\right)^2 - \Lambda(f,\psi)\right| \leq \gamma_1 \left[\frac{|M|}{n} + \|S_M f - f\|_2 + \|S_M(g) - g\|_2\right]$$

$$\text{où} \quad \Lambda(f,\psi) = 4\left[\iint g(x,y)^2 f(x,y)\,dxdy - \left(\iint g(x,y) f(x,y)\,dxdy\right)^2\right]$$

where

$$g(x,y) = \int f(x,u)\psi(x,y,u)\,du$$

Coming back to Taylor expansion

$$
\begin{aligned}
T(f) &= \iint H(\hat{f}, x, y) f(x,y)\,dxdy + \iiint K(\hat{f}, x, y, z) f(x,y) f(x,z)\,dxdydz \\
&+ \Gamma_n
\end{aligned}
$$

Our estimator:

$$
\begin{aligned}
\hat{\theta}_n &= \frac{1}{n_2} \sum_{j=1}^{n_2} H(\hat{f}, X_j, Y_j) \\
&+ \frac{2}{n_2(n_2-1)} \sum_{i \in M} \sum_{j \neq k = 1}^{n_2} p_i(X_j, Y_j) \int p_i(X_k, u) K(\hat{f}, X_k, u, Y_k)\,du \\
&- \frac{1}{n_2(n_2-1)} \sum_{i, i' \in M} \sum_{j \neq k = 1}^{n_2} p_i(X_j, Y_j) p_{i'}(X_k, Y_k) \\
&\iiint p_i(x,y) p_{i'}(x,z) K(\hat{f}, x, y, z)\,dxdydz.
\end{aligned}
$$

Properties of $\hat{T}_n$

Theorem (Efficient estimation of $T(f)$)

(i) If $|M|/n \to 0$

$$\sqrt{n}\left(\hat{T}_n - T(f)\right) \to \mathcal{N}(0, C(f))$$

(ii)

$$\lim_{n\to\infty} n\mathbb{E}(\hat{T}_n - T(f))^2 = C(f)$$

where

$$C(f) = \iint H(f,x,y)^2 f(x,y)d\mu(x,y) - \left(\iint H(f,x,y)f(x,y)d\mu(x,y)\right)^2$$

$C(f) = 4\mathbb{E}\left(\mathrm{Var}(Y|X)\mathbb{E}(Y|X)^2\right) + \mathrm{Var}\left(\mathbb{E}(Y|X)^2\right)$ optimal variance

Extensions

- More general functional

$$\mathbb{E}\Big(\psi\big(\mathbb{E}(\phi(Y)|X)\big)\Big)$$

si $f \in \mathbb{L}^2(dxdy)$, $\phi \in C^0$ et $\psi \in C^3$.

- Optimal variance:

$$C(f) = \mathbb{E}\left(\mathrm{Var}(\phi(Y)|X)\left[\dot{\psi}\left(\mathbb{E}(Y|X)\right)\right]^2\right) + \mathrm{Var}\left(\psi\left(\mathbb{E}(\phi(Y)|X)\right)\right)$$

Best asymptotic estimation procedure

- To estimate $\mathbb{E}(\mathbb{E}(Y|X^i)^2)$, one use a sample of the pair (X^i, Y)
- The other variables $X^j, j \neq i$ are not used

Do we loose information?

Asymptotically, the others components do not give more information to estimate S_i

Best asymptotic estimation procedure

- To estimate $\mathbb{E}(\mathbb{E}(Y|X^i)^2)$, one use a sample of the pair (X^i, Y)
- The other variables $X^j, j \neq i$ are not used

Do we loose information?

Asymptotically, the others components do not give more information to estimate S_i

Best asymptotic estimation procedure

- To estimate $\mathbb{E}(\mathbb{E}(Y|X^i)^2)$, one use a sample of the pair (X^i, Y)
- The other variables X^j, $j \neq i$ are not used

Do we loose information?

Answer:

- Optimal variance taking into account all components $(X^1, \dots, X^d)$ is still $C(f)$
- Or notre estimateur $\widehat{T}_n$ atteint cette vitesse en travaillant uniquement avec (X^i, Y)

Asymptotically, the others components do not give more information to estimate S_i

Best asymptotic estimation procedure

- To estimate $\mathbb{E}(\mathbb{E}(Y|X^i)^2)$, one use a sample of the pair (X^i, Y)
- The other variables X^j, $j \neq i$ are not used

Do we loose information?

Answer:

- Optimal variance taking into account all components $(X^1, \ldots, X^d)$ is still $C(f)$
- Or notre estimateur $\widehat{T}_n$ atteint cette vitesse en travaillant uniquement avec (X^i, Y)

Asymptotically, the others components do not give more information to estimate S_i

Best asymptotic estimation procedure

- To estimate $\mathbb{E}(\mathbb{E}(Y|X^i)^2)$, one use a sample of the pair (X^i, Y)
- The other variables X^j, $j \neq i$ are not used

Do we loose information?

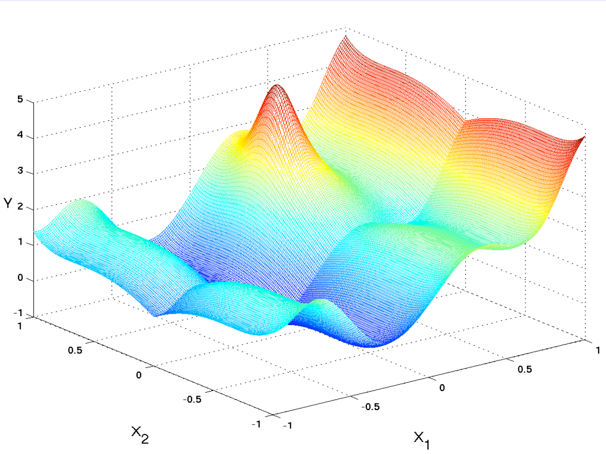
Answer:

- Optimal variance taking into account all components $(X^1, \ldots, X^d)$ is still $C(f)$
- Or notre estimateur $\widehat{T}_n$ atteint cette vitesse en travaillant uniquement avec (X^i, Y)

Asymptotically, the others components do not give more information to estimate S_i

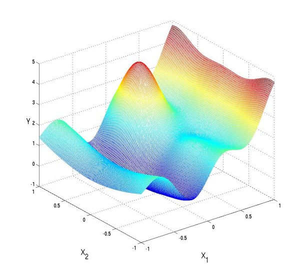
An analytical example

$$Y = 0.2\exp(X_1 - 3) + 2.2|X_2| + 1.3X_2^6 - 2X_2^2 - 0.5X_2^4 - 0.5X_1^4 + 2.5X_1^2 + 0.7X_1^3 + \frac{3}{(8X_1 - 2)^2 + (5X_2 - 3)^2 + 1} + \sin(5X_1)\cos(3X_1^2)$$

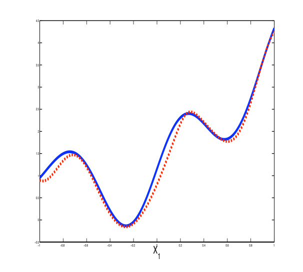


Kriging (true function, approximation)

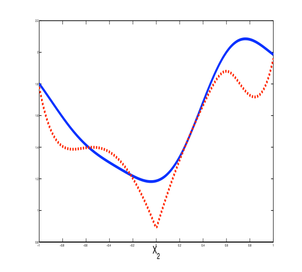
$f(X_1, X_2)$



$\mathbb{E}(Y|X_1)$

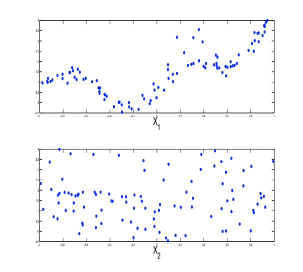


$\mathbb{E}(Y|X_2)$

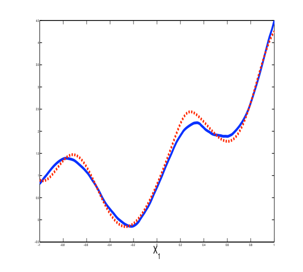


Local polynomial (true function, approximation)

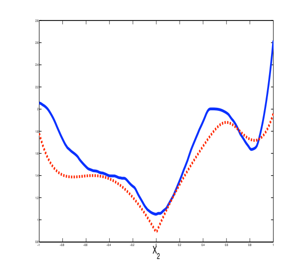
Marginal samples



$\mathbb{E}(Y|X_1)$



$\mathbb{E}(Y|X_2)$



An analytical example

		Kriging	Loc poly	Our estim
		100 pts	100 pts	100 pts
$\mathrm{Var}(\mathbb{E}(Y\|X_1))$	1.0932	1.0539	1.0643	1.1701
$\mathrm{Var}(\mathbb{E}(Y\|X_2))$	0.0729	0.1121	0.0527	0.0939

X_1 more or less same accuracy

X_2 : marginal approximation gives better results

Thanks a lot for you attention