Identification of fuzzy regression models

A. Bisserier, S. Galichet, R. Boukezzoula

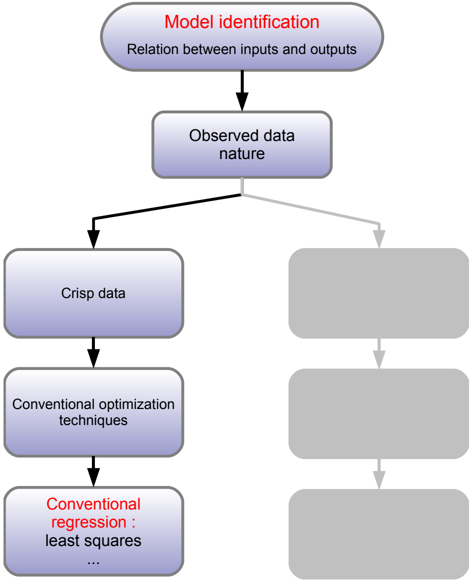
LISTIC

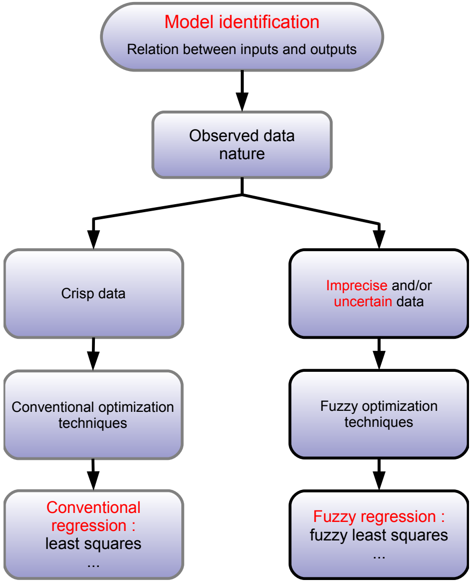
13 Novembre 2008

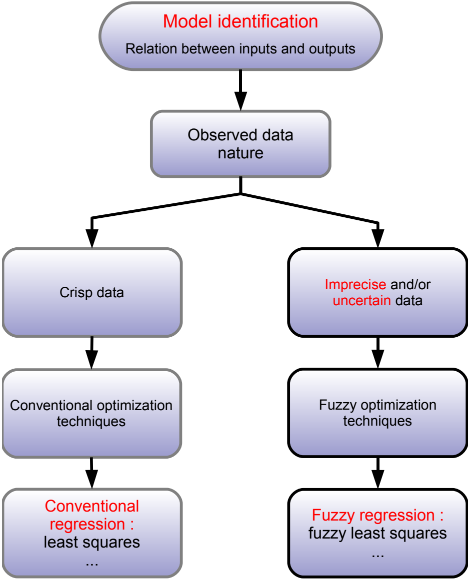
Outline

1. Introduction
2. Notations and concepts
3. Fuzzy linear regression
4. Proposed approach
5. Illustrative examples
6. Conclusion

Introduction



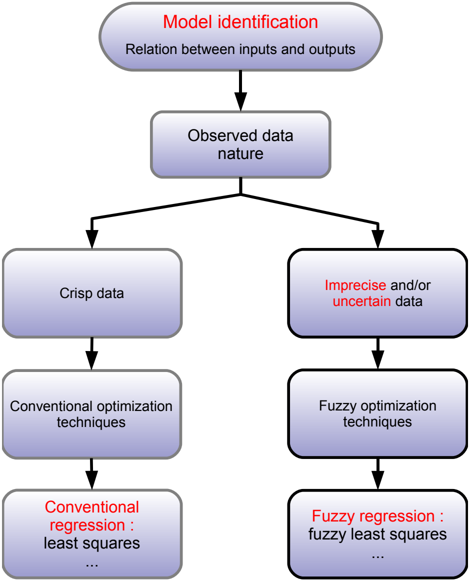

Model identification
Relation between inputs and outputs
Observed data nature
Crisp data
Imprecise and/or uncertain data
Conventional optimization techniques
Fuzzy optimization techniques
Conventional regression : least squares ...
Fuzzy regression : fuzzy least squares ...



General problem

- What kind of fuzzy models have to be used ?
- What kind of data have to be manipulated for the identification process ?

Introduction

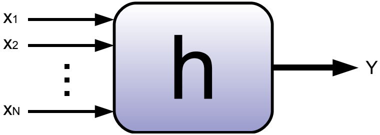


General problem

- What kind of fuzzy models have to be used ?
- What kind of data have to be manipulated for the identification process ?

⇒ How can we identify a fuzzy model on this kind of data ?

Notations and concepts



Considered model : multi-inputs, single-ouput

- Objective : to determine a relationship $Y = h(\mathbf{x})$ between inputs and output
- h considered as linear $\longrightarrow$ model of the form :

$$Y = \sum_{i=1}^{N} A_i \cdot x_i$$

- fuzzy model : fuzzy coefficients A_i

Notations and concepts

Observed data for the identification

Set of M observed data :

The j^{th} data :

- a crisp inputs vector $x_j = (x_{0j}, x_{1j}, ..., x_{Nj})$
- the corresponding fuzzy output Y_j

  $\longrightarrow$ a symmetrical triangular fuzzy number

Remark : choice made for the sake of simplicity

Notations and concepts

Conventional intervals

- Set of elements in $\mathbb{R}$ between a lower and an upper bound

$$a = \{x | a^- \leq x \leq a^+, x \in \mathbb{R}\}$$

- Midpoint and Radius

$$M(a) = M_a = (a^- + a^+)/2$$

$$R(a) = R_a = (a^+ - a^-)/2$$

- calculus on sets

Notations and concepts

Conventional intervals

- Set of elements in $\mathbb{R}$ between a lower and an upper bound

$$a = \{x | a^- \leq x \leq a^+, x \in \mathbb{R}\}$$

- Midpoint and Radius

$$M(a) = M_a = (a^- + a^+)/2$$

$$R(a) = R_a = (a^+ - a^-)/2$$

- calculus on sets

Conventional intervals : inclusion

- general relationship :

$$a \subseteq b \Leftrightarrow |M(b) - M(a)| \leq R(b) - R(a)$$

Notations and concepts

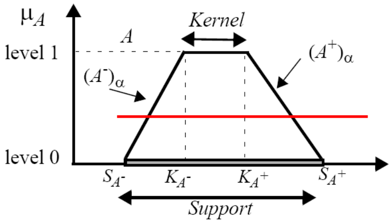
Fuzzy intervals

- Horizontal AND vertical dimension

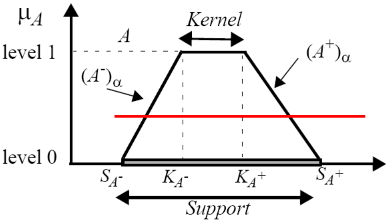
$\Rightarrow$ intervals at two levels :

  - kernel : $K_A = [K_A^-, K_A^+]$
  - support : $S_A = [S_A^-, S_A^+]$

- Kernel included in the Support
- function linking the two levels : profiles
- α-cut $\Rightarrow$ conventional interval



Notations and concepts



Fuzzy intervals : linear profiles

- general case : trapezoidal fuzzy interval

$$A = (K_A, S_A) = ([K_A^-, K_A^+], [S_A^-, S_A^+])$$

- a particular case : symmetrical triangular fuzzy interval

$$A = (K_A, R_A)$$

Notations and concepts

Notations and concepts

- inclusion of two fuzzy intervals : general definition

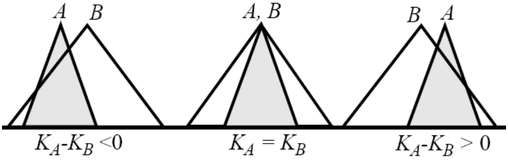
$$A \subseteq B \Leftrightarrow \forall x, \mu_A(x) \leq \mu_B(x)$$

- in the case of trapezoidal fuzzy intervals :

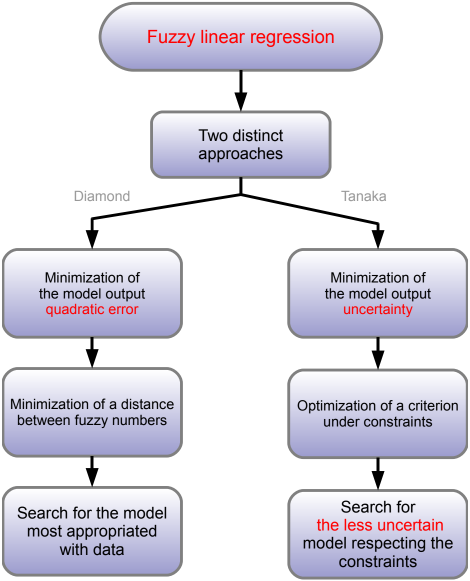
$$A \subseteq B \Leftrightarrow K_A \subseteq K_B \text{ et } S_A \subseteq S_B$$

- in the case of symmetrical triangular fuzzy intervals :

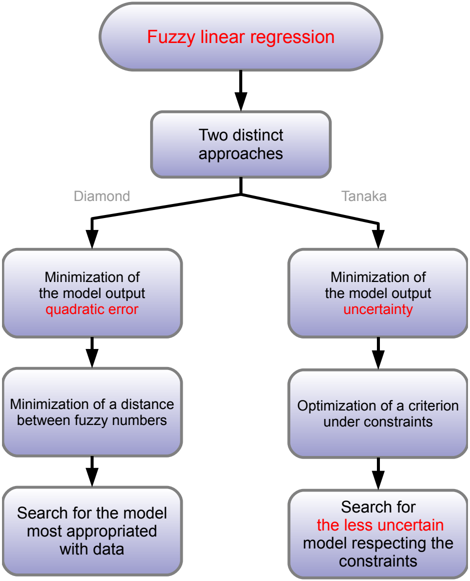
$\Rightarrow$ equality of the Kernel values



Fuzzy linear regression



Fuzzy linear regression

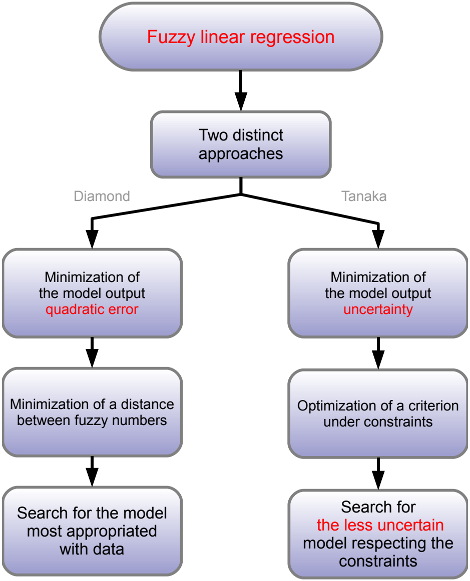


Which constraints ?

Relationship between observed and predicted outputs :

- *necessity* model : predicted outputs included in observed ones
- *conjunction* model : no empty intersection between predicted and observed outputs

Fuzzy linear regression



Which constraints ?

Relationship between observed and predicted outputs :

- *necessity* model : predicted outputs included in observed ones
- *conjunction* model : no empty intersection between predicted and observed outputs
- *possibility* model : observed outputs included in predicted ones

$\Rightarrow$ data total uncertainty is taken in care

Context of the study

- observed outputs : symmetrical triangular fuzzy intervals

  $\Rightarrow$ identified model with triangular parameters
- study leaded for the identification of single input models :

$$\hat{Y}(x) = A_0 \oplus A_1 x$$

with $\oplus$ sum of fuzzy intervals

  $\surd$ *concepts can be extended to the multi inputs case*
- possibility model : observed inputs included in predicted ones

Fuzzy linear regression : possibilistic approach

Basic concepts

- use of α-cuts

  $\Rightarrow$ conventionnal intervals are handled
- minimization of a linear criterion under constraints :

  $\Rightarrow$ criterion : a representation of the model uncertainty

  $\Rightarrow$ constraints : inclusion of conventional intervals at the level α

$\hookrightarrow$ What are the limits of this approach ?

Fuzzy linear regression : possibilistic approach

The use of α-cuts

- inclusion constraints defined for this level α :

$$[Y_j]_\alpha \subseteq [\hat{Y}_j]_\alpha$$

$$\Leftrightarrow \begin{cases} K_{\hat{Y}_j} + (1-\alpha)R_{\hat{Y}_j} \geq K_{Y_j} + (1-\alpha)R_{Y_j} \\ K_{\hat{Y}_j} - (1-\alpha)R_{\hat{Y}_j} \leq K_{Y_j} - (1-\alpha)R_{Y_j} \end{cases}$$

- after the identification : parameters considered as valid $\forall\alpha \in [0,1]$
- BUT : triangular identified model, *equality of the Kernels necessary* for the total inclusion
- Total inclusion $\forall\alpha \in [0,1]$ not guaranteed !

Fuzzy linear regression : possibilistic approach

Conventionnal criteria

- Minimization of the model uncertainty

  $\Rightarrow$ How can it be quantified ?

$\hookrightarrow$ Most used criterion (Tanaka) : sum of the predicted intervals radius

$$Somme = M \cdot R_{A_0} + R_{A_1} \cdot \sum_{j=1}^{M} | x_j |$$

- BUT : minimization at the observed points

$\Rightarrow$ strongly dependant on the learning points : weak robustness

Model representativity

Conventionnal linear model : study of the output variation on the domain

- kernel : $M(\hat{Y}(x)) = K_{\hat{Y}(x)} = K_{A_0} + K_{A_1} \cdot x$

$\Rightarrow$ variation according to the sign of K_{A_1} $\rightarrow$ **any variation**

- radius : $R(\hat{Y}(x)) = R_{\hat{Y}(x)} = R_{A_0} + R_{A_1} \cdot |x|$

$\Rightarrow$ Radius always positive !

$\Rightarrow$ **Variation of the output radius limited by the input sign**

Propositions

Points to improve in the previous method : summary

- Is it possible to consider a model respecting the inclusion $\forall\alpha \in [0, 1]$?
- Is it possible to identify such a model with a more robust criterion ?
- Is it possible to impove the representativity of a fuzzy linear model ?

Propositions

Solution to the inclusion problem

- identification of a trapezoidal fuzzy model

$\longrightarrow$ inclusion constraints at two levels α : $\alpha = 0$ and $\alpha = 1$

- $\alpha = 1 : K_{Y_j} \in [K^-_{\hat{Y}_j}, K^+_{\hat{Y}_j}]$
- $\alpha = 0 : \left[K_{Y_j} - R_{Y_j}, K_{Y_j} + R_{Y_j}\right] \subseteq [S^-_{\hat{Y}_j}, S^+_{\hat{Y}_j}]$

$\longrightarrow$ linear membership function

$\Rightarrow$ inclusion guaranteed $\forall \alpha \in [0, 1]$

Remark : model output for a data j :

$$\hat{Y}_j = ([K^-_{\hat{Y}_j}, K^+_{\hat{Y}_j}], [S^-_{\hat{Y}_j}, S^+_{\hat{Y}_j}])$$

Propositions

A more robust criterion

- Objective : make the criterion independant of data

$\longrightarrow$ definition of a total uncertainty on the domain D

$\longrightarrow$ vertical dimension (trapezes) taken into account

$\Longrightarrow$ Minimization of Volume delimited by the model on D

$$Volume = R(K_{A_0}) + R(S_{A_0}) + (R(K_{A_1}) + R(S_{A_1})) \cdot |M(D)|$$

$\hookrightarrow$ independant of observed inputs x_j

$\hookrightarrow$ still linear criterion

Propositions

A model more representative of the data

- To have any kind of the output radius variation :
  - $\longrightarrow$ the input variable sign must be modified
  - $\longrightarrow$ the linear behavior must be kept

$\Rightarrow$ The origin of the model is set on a bound of $D = [x_{min}, x_{max}]$

$\Rightarrow$ New model defined on the domain D :

$$\hat{Y}(x) = A_0 \oplus A_1(x - shift)$$

Study of the variation of this new model output radius

$$R([S_{\hat{Y}}]) = R([S_{A_0}]) + R([S_{A_1}]) \cdot |x - shift|$$

$\longrightarrow$ $x - shift \geq 0 \ \forall x \in D \Rightarrow$ increasing radius

$\longrightarrow$ $x - shift \leq 0 \ \forall x \in D \Rightarrow$ decreasing radius

output spread variation	$\nearrow$	$\searrow$
Used model	$A_0 \oplus A_1(x - x_{min})$	$A_0 \oplus A_1(x - x_{max})$

Identification method : summary

Main steps

- What is the global tendency of the oberved outputs radius on D ?

$\Rightarrow$ Choice of the appropriated value of shift

- Identification of trapezoidal coefficients :

$\Rightarrow$ Minimization of the criterion **Volume** under the inclusion constraints :

  - at $\alpha = 0$ and $\alpha = 1 \rightarrow$ total inclusion guaranteed

Optimization linear program

$$Min\ R(K_{A_0}) + R(S_{A_0}) + (R(K_{A_1}) + R(S_{A_1})) \cdot |M(D)|$$

$$s.c. \begin{cases} K_{Y_j} \in [K^-_{\hat{Y}_j}, K^+_{\hat{Y}_j}] \\ \left[K_{Y_j} - R_{Y_j}, K_{Y_j} + R_{Y_j}\right] \subseteq [S^-_{\hat{Y}_j}, S^+_{\hat{Y}_j}] \end{cases}$$

Performance indicators

Comparison of two models : it is necessary to use performance indicators

Two approaches :

- indicators about the model uncertainty :
  - **Sum** of the radius
  - **Volume** delimited by the model on D
- indicators about the model fitting with data
  - **Compatibility** between observed and predicted data : inclusion degree
  - **Distance** between observed and predicted data : quadratic error

Extension to piecewise linear models

Global model :

- shifted submodels with trapezoidal coefficients
- each submodel defined on its own domain

⇒ independant submodels for identification and using

Identification method :

- data segmentation :
  - on observed kernels
  - on observed radius

  ⇒ More representative models !
- optimization of the parameters with the criterion **Volume**

⇒ identification of a gobal model with a minimal uncertainty and respecting the total inclusion

Generalization to multi-inputs models

Previous concepts applicable to multi-inputs ($x_i, i = 1, ..., N$)

- Trapezoidal coefficients
- Appropriated shift for each x_i
- Criterion **Volume** can be extended, still linear

$\Rightarrow$ identification of the optimal model for the total ouput uncertainty on the domain

Case of a piecewise model

- Data segmentation : cartesian product
- Identification of submodels

$\Rightarrow$ Independant submodels on each region

Numerical examples : inclusion problem

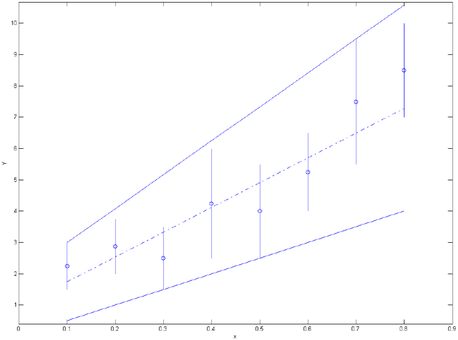
j	x_j	Y_j
1	0.1	(2.25, 0.75)
2	0.2	(2.875, 0.875)
3	0.3	(2.5, 1)
4	0.4	(4.25, 1.75)
5	0.5	(4.0, 1.5)
6	0.6	(5.25, 1.25)
7	0.7	(7.5, 2)
8	0.8	(8.5, 1.5)

- 8 observed data, $D = [0.1, 0.8]$
- crisp inputs
- observed outputs : symmetrical triangular fuzzy numbers
- form of the identified model :

$$\hat{Y}(x) = A_0 \oplus A_1 x$$

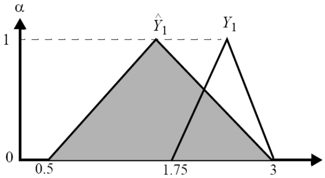
Numerical examples : inclusion problem

Triangular coefficients : identification at $\alpha = 0$



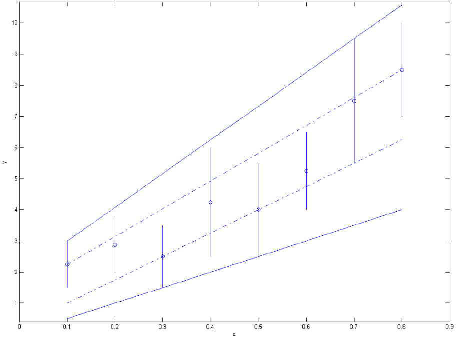
	modèle flou triangulaire
A_0	(0.96, 0.96)
A_1	(7.92, 2.92)
Compatibilité	0.83

study of the inclusion : data $j = 1$



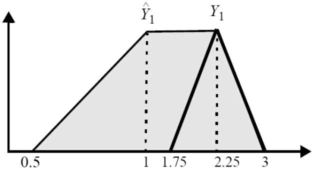
- inclusion is respected at $\alpha = 0$
- inclusion is not guaranteed $\forall \alpha$

Trapezoidal coefficients



	modèle flou trapézoidal
A_0	([0.25 1.36],[0 1.92])
A_1	([7.5 8.93],[5 10.83])
compatibilité	1

study of the inclusion : data $j = 1$



- inclusion is respected $\forall\alpha$!

Numerical examples : criterion robustness

presentation of the data set

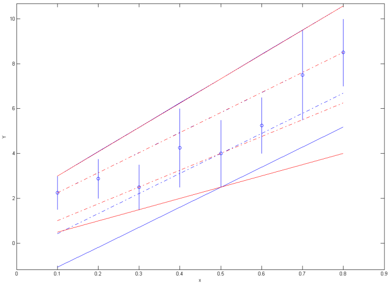
- data $j = 8$ duplicated three times, $D = [0.1, 0.8]$

$\Rightarrow$ study of the influence of data redundancy

- trapezoidal identified model
- 2 minimized criteria : *Sum* and *Volume*

	Somme	*Volume*
A_0	([-0.46 1.36],[-1.96 1.92])	([0.25 1.36],[0 1.92])
A_1	([8.93 8.93],[8.93 10.83])	([7.5 8.93],[5 10.83])
distance	71.05	91.29
somme	37.08	38.42
volume	3.28	3.15

Identified models : *Sum (blue)* and *Volume (red)*



- initial model identified again for *Volume*
- Minimal total uncertainty
- the indicators *distance* and *sum* are not robust !

Numerical examples : output representativity

presentation of the data set

- initial data shifted to have negative inputs

$\Rightarrow$ increasing outputs radius

- trapezoidal identified models, for the minimum *Volume* :
  - conventionnal model : $\hat{Y}(x) = A_0 \oplus A_1 x$
  - shifted model : $\hat{Y}(x) = A_0 \oplus A_1(x - shift)$

j	x_j	Y_j
1	-0.8	(2.25, 0.75)
2	-0.7	(2.875, 0.875)
3	-0.6	(2.5, 1)
4	-0.5	(4.25, 1.75)
5	-0.4	(4.0, 1.5)
6	-0.3	(5.25, 1.25)
7	-0.2	(7.5, 2)
8	-0.1	(8.5, 1.5)

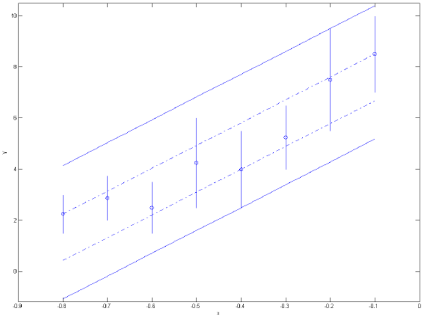
	$\hat{Y}(x) = A_0 \oplus A_1 \cdot x$	$\hat{Y}(x) = A_0 \oplus A_1(x - shift)$
A_0	([7.57 9.39],[0.07 11.29])	([1 2.25],[0.5 3])
A_1	8.93	([7.5 8.93],[5 10.83])
distance	58.83	48.08
somme	28.14	25.17
volume	3.52	3.15

All the indicators are better for the shifted model !

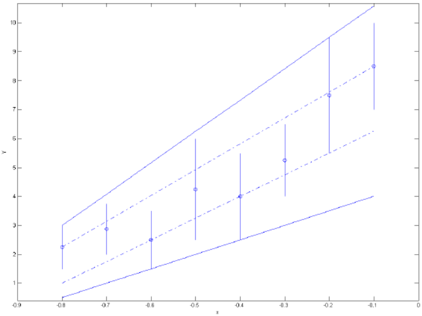
Numerical examples : output representativity

	$\hat{Y}(x) = A_0 \oplus A_1 \cdot x$	$\hat{Y}(x) = A_0 \oplus A_1(x - shift)$
A_0	([7.57 9.39],[0.07 11.29])	([1 2.25],[0.5 3])
A_1	8.93	([7.5 8.93],[5 10.83])
distance	58.83	48.08
somme	28.14	25.17
volume	3.52	3.15

Conventionnal model



Model with shifted input

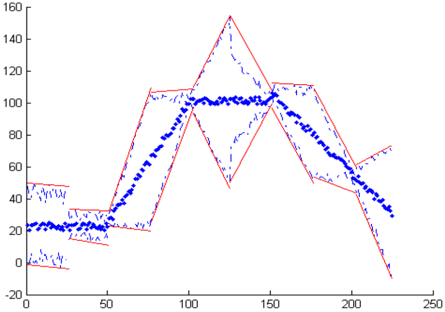


Numerical examples : piecewise model

- Great number of data
- Various tendencies of kernel and radius variations, noise : segmentation

$\Rightarrow$ Identification of shifted trapezoidal multi-inputs submodels

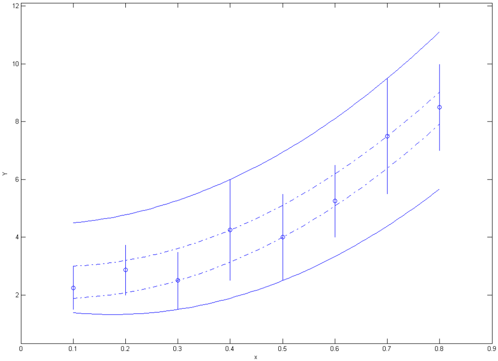
$\Rightarrow$ Identification of the optimal global model for the minimal uncertainty (**Volume**)



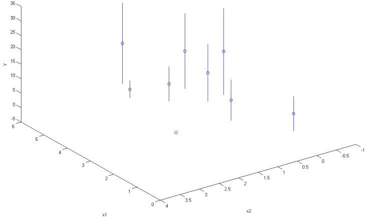
Numerical examples : higher order model

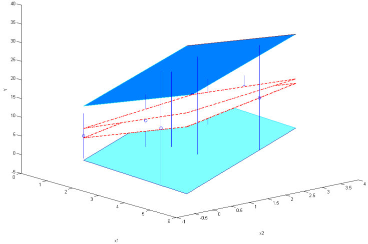
- Initial data set
- 2-order identified model :

$$\hat{Y}(x) = A_0 \oplus A_1 x \oplus A_2 x^2$$



Numerical examples : multi-inputs model





- Two inputs → two appropriated values of shifts :
  - Increasing radius on x_1 and decreasing one on x_2
- Identification of the trapezoidal model for **Volume**

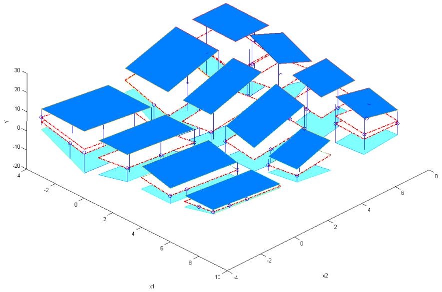
⇒ "Planes" with total inclusion of data

Numerical examples : multi-inputs piecewise model

- Various tendencies of kernel and radius variations : segmentation

⇒ Identification of shifted trapezoidal multi-inputs submodels

⇒ Identification of the optimal global model for the minimal uncertainty (**Volume**)



Conclusion

Propositions

- Identification of trapezoidal models : total inclusion is guaranteed
- Improvement of the model representativity
- Identification for the minimal total uncertainty : increased robustness
- Application of these concepts on piecewise linear regression problems

Conclusion

Future work

- Imprecise observed inputs
  - Uncertainty comes from :
    - fuzzy parameters
    - fuzzy inputs
  - Lack of representativity of conventionnal space (X, Y)

$\Rightarrow$ Representation of the model in the space $(Mid(X), Rad(X), Y)$

$\Rightarrow$ Identification of the model in this space

END

Questions ?...