Fuzzy System Identification

Takagi-Sugeno Modelling and Identification

Olaf Wolkenhauer

Control Systems Centre

UMIST

o.wolkenhauer@umist.ac.uk

www.csc.umist.ac.uk/people/wolkenhauer.htm

Contents

1. Model Structures

- NARX model structure :

$$\begin{aligned} R_i: \quad & \text{IF } y(k) \text{ is } A_{i1} \text{ AND} \ldots \text{AND } y(k-n_y+1) \text{ is } A_{in_y} \\ & \text{AND } u(k) \text{ is } B_{i1} \text{ AND} \ldots \text{AND } u(k-n_u+1) \text{ is } B_{in_u}, \\ & \text{THEN } y(k+1) \text{ is } C_i \end{aligned}$$

- Linguistic Model
- Takagi-Sugeno Model

1.1. Linguistic Model

▷ *Linguistic Model :*

$$R_i : \quad \text{IF } \mathbf{x} \text{ is } A_i, \text{ THEN } y \text{ is } B_i, \qquad i = 1, 2, \ldots, n_R \tag{1}$$

where $\mathbf{x}$ is the *antecedent variable* and y the *consequent variable* and *fuzzy restriction* $\mu_{A_i}(\mathbf{x})\colon X_1 \times \cdots \times X_r \rightarrow [0, 1]$

▷ *conjunctive form* :

$$R_i : \quad \text{IF } x_1 \text{ is } A_{i1} \text{ AND } x_2 \text{ is } A_{i2} \ldots \text{AND } x_r \text{ is } A_{ir}, \\ \text{THEN } y \text{ is } B_i \tag{2}$$

▷ *Degree of fullfillment* :

$$\beta_i(\mathbf{x}) \doteq \mu_{A_{i1}}(x_1) \wedge \mu_{A_{i2}}(x_2) \wedge \cdots \wedge \mu_{A_{ir}}(x_r) \tag{3}$$

1.2. Fuzzy System as Basis Function Approximator

Singleton model :

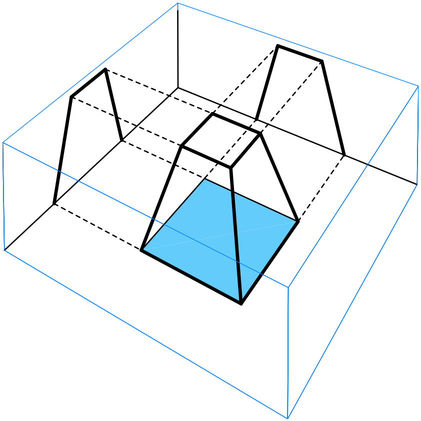
$$R_i : \quad \text{IF } \mathbf{x} \text{ is } A_i, \text{ THEN } y \text{ is } b_i, \qquad i = 1, 2, \ldots, n_R \tag{4}$$

Normalised degree of fullfillment :

$$\phi_i(\mathbf{x}) \doteq \frac{\beta_i(\mathbf{x})}{\sum_{k=1}^{n_R} \beta_k(\mathbf{x})} \ ,$$

..showing that a fuzzy system is a *basis function approximator* :

$$y = \sum_{i=1}^{n_R} \phi_i(\mathbf{x}) \cdot b_i \tag{5}$$



1.3. Takagi-Sugeno Model

▷ *Takagi-Sugeno Model :*

$$R_i : \quad \text{IF } \mathbf{x} \text{ is } A_i, \ \text{THEN } y_i = f_i(\mathbf{x}), \qquad i = 1, 2, \ldots, n_R \tag{6}$$

▷ Conjunctive form :

$$\begin{aligned} R_i : \quad \text{IF } x_1 \text{ is } A_{i1} \text{ AND } x_2 \text{ is } A_{i2} \ldots \text{AND } x_r \text{ is } A_{ir}, \\ \text{THEN } y_i = f_i(\mathbf{x}) \ . \end{aligned} \tag{7}$$

▷ *Affine linear TS model :*

$$y_i = \mathbf{a}_i^T \mathbf{x} + b_i \tag{8}$$

1.4. Identification of Antecedent Fuzzy Sets

Linguistic Model :

$$R_i: \quad \text{IF } \mathbf{x} \text{ is } A_i, \text{ THEN } \dots , \qquad i = 1, 2, \dots , n_R$$

r-dimensional fuzzy set A_i defined by the ith row in $\mathbf{U}$:

$$\forall\,(i, j) \qquad \mu_{A_i}(\mathbf{x}_j) = \max_{j'=1,\dots,d} \{u_{ij'} \in \mathbf{U} : \ \mathbf{x}_{j'} = \mathbf{x}_j\} \tag{9}$$

where

$$u_{ij} \in \mathbf{U}, \qquad i = 1, 2, \dots , c\ , \qquad j = 1, 2, \dots , d\ .$$

The regressor vector (antecedent variables) is denoted

$$\mathbf{x} \doteq [x_1, \dots , x_k, \dots , x_r]^T, \qquad k = 1, 2, \dots , r$$

Vectors of sampled values for antecedent variables $x_1, \dots , x_r$, $m_{kj} \in \mathbf{M}$ and $\mathbf{M}$ is a $(r + 1) \times d$ matrix :

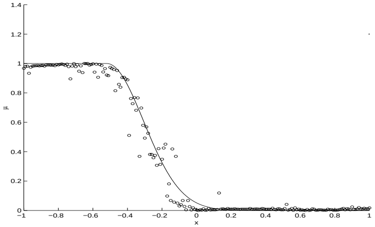
$$\mathbf{x}_j \doteq [m_{1j}, \dots , m_{kj}, \dots , m_{rj}]^T\ . \tag{10}$$

Conjunctive form :

$$R_i : \quad \text{IF } x_1 \text{ is } A_{i1} \text{ AND } \ldots \text{ AND } x_k \text{ is } A_{ik} \text{ AND } \ldots \text{ AND } x_r \text{ is } A_{ir}, \text{ THEN } \ldots$$

For all (i, j, k)

$$\mu_{A_{ik}}(x_{kj}) = \max_{m_{k'j'} \in \mathbf{M}} \{u_{ij'} \in \mathbf{U} : m_{kj} = m_{k'j'}, j' = 1, \ldots, d, \; k' = 1, \ldots, r\} . \tag{11}$$



1.5. Consequent Parameter Identification

Inference in the Takagi-Sugeno model:

$$y = \frac{\sum_{i=1}^{n_R} \beta_i(\mathbf{x}) \cdot \left(\mathbf{a}_i^T \mathbf{x} + b_i\right)}{\sum_{i=1}^{n_R} \beta_i(\mathbf{x})} \, . \tag{12}$$

Normalised degree of fullfillment :

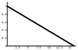
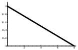
$$\phi_i(\mathbf{x}) \doteq \frac{\beta_i(\mathbf{x})}{\sum_{k=1}^{n_R} \beta_k(\mathbf{x})} \, , \tag{13}$$

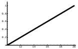
Quasilinear model :

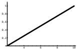
$$\begin{aligned} y &= \left(\sum_{i=1}^{c} \phi_i(\mathbf{x}) \cdot \mathbf{a}_i^T\right) \mathbf{x} + \sum_{i=1}^{c} \phi_i(\mathbf{x}) \cdot b_i \\ &\doteq \mathbf{a}^T(\mathbf{x}) + b(\mathbf{x}) \, . \end{aligned} \tag{14}$$

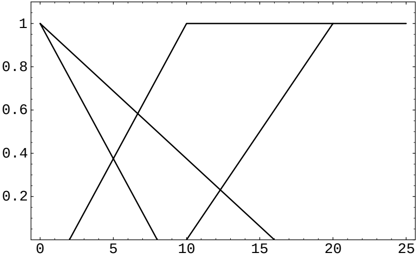
2. Examples: TS-Modelling and Identification

Suppose that we have the following three rules (implications) :

R_1 : IF x_1 is  AND x_2 is , THEN $y = x_1 + x_2$

R_2 : IF x_1 is , THEN $y = 2x_1$

R_3 : IF x_2 is , THEN $y = 3x_2$.

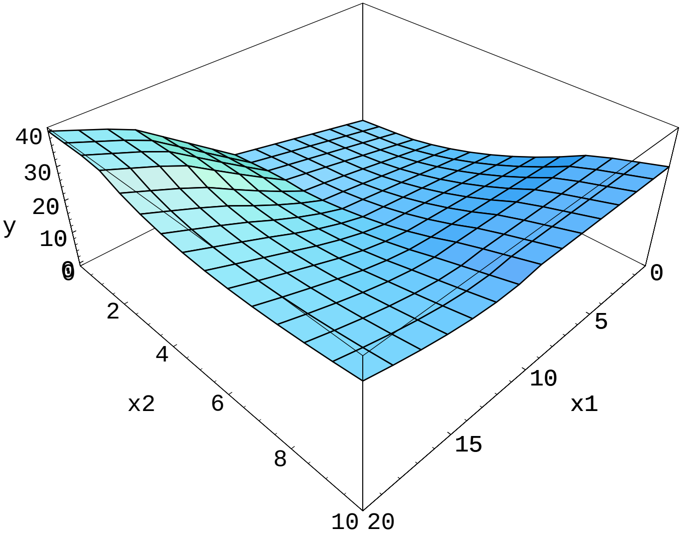


The output for the conjunctive TS model structure is given by

$$y = \frac{\sum_{i=1}^{n_R} \left(\mu_{A_{i1}}(x_1) \wedge \cdots \wedge \mu_{A_{ir}}(x_r)\right) \cdot y_i}{\sum_{i=1}^{n_R} \left(\mu_{A_{i1}}(x_1) \wedge \cdots \wedge \mu_{A_{ir}}(x_r)\right)} \tag{15}$$

Example: We are given $x_1 = 12$, $x_2 = 5$, then

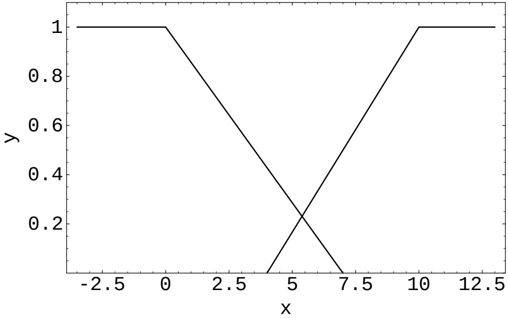
$$y = \frac{0.25 \cdot 17 + 0.2 \cdot 24 + 0.375 \cdot 15}{0.25 + 0.2 + 0.375} \simeq 17.8$$


y
40
30
20
10
0
x2
0
2
4
6
8
10
x1
0
5
10
15
20

2.1. Example: Consequent Parameter Estimation

Suppose that we have the following two implications (rules):

$$R_1 : \quad \text{IF } x \text{ is } A_{11}, \text{ THEN } y = 2 + 0.6x$$
$$R_2 : \quad \text{IF } x \text{ is } A_{21}, \text{ THEN } y = 9 + 0.2x \, .$$



The output y for the input x is obtained from (15),

$$y = \frac{\sum_{i=1}^{n_R} \big(\mu_{A_{i1}}(x_1) \wedge \cdots \wedge \mu_{A_{ir}}(x_r)\big) \cdot (b_i + a_{i1}x_1 + \cdots + a_{ir}x_l)}{\sum_{i=1}^{n_R} \big(\mu_{A_{i1}}(x_1) \wedge \cdots \wedge \mu_{A_{ir}}(x_r)\big)}$$

$$= \frac{\sum_{i=1}^{2} \mu_{A_{i1}}(x) \cdot (b_i + a_{i1}x)}{\sum_{i=1}^{2} \mu_{A_{i1}}(x)}$$

Define ϕ_i :

$$\phi_i = \frac{\mu_{A_{i1}}(x_1) \wedge \cdots \wedge \mu_{A_{ir}}(x_r)}{\sum_{k=1}^{n_R} \left(\mu_{A_{k1}}(x_1) \wedge \cdots \wedge \mu_{A_{kr}}(x_r)\right)}$$
$$= \frac{\mu_{A_{i1}}(x)}{\sum_{k=1}^{2} \mu_{A_{k1}}(x)}$$

then

$$y = \sum_{i=1}^{n_R} \phi_i \cdot (b_i + a_{i1}x_1 + \cdots + a_{ir}x_r)$$
$$= \sum_{i=1}^{2} \phi_i \cdot (b_i + a_{i1}x) \ .$$

We obtain the *consequence parameters* $b_i, a_{i1}, \ldots, a_{ir}$, $i = 1, \ldots, n_R$, using *least squares estimation*:

$$\mathbf{Y} = [y_1, \cdots, y_d]^T$$

$$\boldsymbol{\theta} = [b_1, \cdots, b_{n_R}, a_{11}, \cdots, a_{n_R 1}, \cdots, a_{1r}, \cdots, a_{n_R r}]^T$$

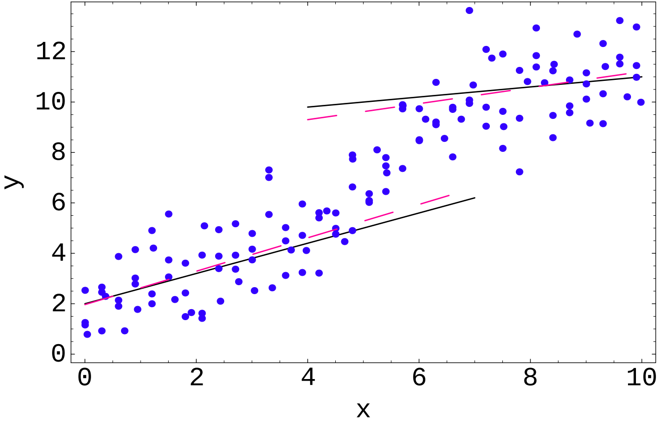
$$\mathbf{X} = \begin{bmatrix} \phi_{11} & \phi_{21} \cdots \phi_{n_R 1} & m_{11}\phi_{11} \cdots m_{11}\phi_{11} \cdots m_{r1}\phi_{11} \cdots m_{r1}\phi_{n_R 1} \\ \phi_{12} & \phi_{22} \cdots \phi_{n_R 2} & m_{12}\phi_{12} \cdots m_{12}\phi_{22} \cdots m_{r2}\phi_{12} \cdots m_{r2}\phi_{n_R 2} \\ \vdots & \vdots & \vdots \\ \phi_{1d} & \phi_{2d} \cdots \phi_{n_R d} & m_{1d}\phi_{1d} \cdots m_{1d}\phi_{2d} \cdots m_{rd}\phi_{1d} \cdots m_{rd}\phi_{n_R d} \end{bmatrix}$$

where

$$\phi_{ij} \doteq \frac{\mu_{A_{i1}}(x_{1j}) \wedge \cdots \wedge \mu_{A_{ir}}(x_{rj})}{\sum_k \mu_{A_{k1}}(x_{kj}) \wedge \cdots \wedge \mu_{A_{kr}}(x_{rj})} \; .$$

Then the *parameter vector* $\boldsymbol{\theta}$ is calculated as

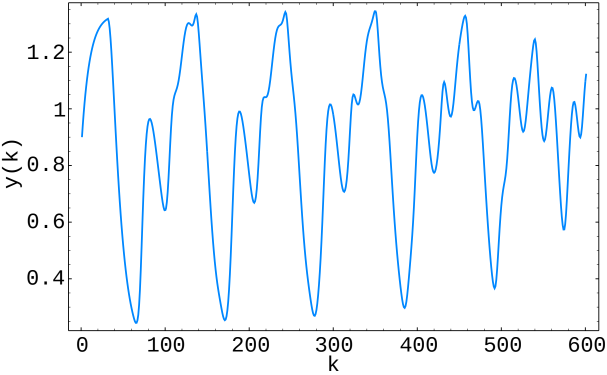
$$\hat{\boldsymbol{\theta}} = \left[\mathbf{X}^T\mathbf{X}\right]^{-1}\mathbf{X}^T\mathbf{Y} \ .$$



2.2. Example: Forecasting, Multidimensional Antecedent Set

Chaotic Mackey-Glass time-series :

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \frac{0.2 \cdot x(t-\tau)}{1 + x^{10}(t-\tau)} \ .$$



In forecasting, $\mu_{A_i}(\mathbf{x}) \doteq \beta_i(\mathbf{x})$ is calculated directly. Since we consider only subspace X of $X \times Y$,

$$\mathbf{A}_{\mathbf{x}}^{(i)} \doteq \left(|\mathbf{F}^{(i)}|\right)^{1/r} \cdot \left(\mathbf{F}^{(i)}\right)^{-1} .$$

Similar, let $\mathbf{c}_{\mathbf{x}}^{(i)}$ denote the projection of cluster prototype $\mathbf{c}^{(i)}$, onto X such that

$$d^2_{\mathbf{A}_{\mathbf{x}}^{(i)}}\left(\mathbf{x}, \mathbf{c}_{\mathbf{x}}^{(i)}\right) = \left(\mathbf{c}_{\mathbf{x}}^{(i)} - \mathbf{m}_j\right)^T \mathbf{A}_{\mathbf{x}}^{(i)} \left(\mathbf{c}_{\mathbf{x}}^{(i)} - \mathbf{m}_j\right) .$$

Probabilistic Method:

$$\beta_i(\mathbf{x}) = \frac{1}{\sum_{k=1}^{c} \left(d^2_{\mathbf{A}_{\mathbf{x}}^{(i)}}\left(\mathbf{x}, \mathbf{c}_{\mathbf{x}}^{(i)}\right) / d^2_{\mathbf{A}_{\mathbf{x}}^{(i)}}\left(\mathbf{x}, \mathbf{c}_{\mathbf{x}}^{(i)}\right) \right)^{1/(w-1)}} .$$

Possibilistic Method:

$$\beta_i(\mathbf{x}) = \frac{1}{1 + d^2_{\mathbf{A}_{\mathbf{x}}^{(i)}}\left(\mathbf{x}, \mathbf{c}_{\mathbf{x}}^{(i)}\right)} .$$

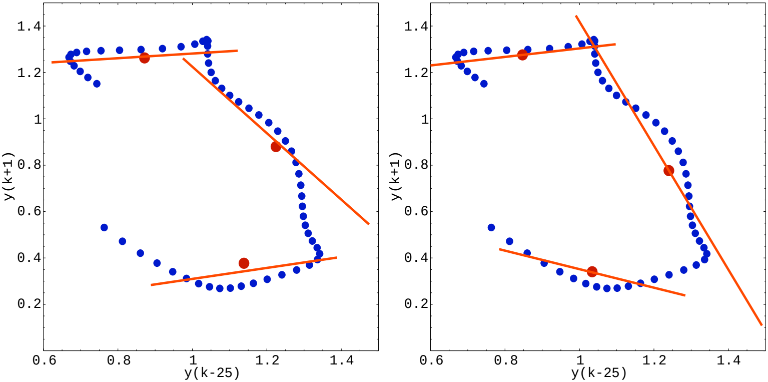


Figure 1: *Result of fuzzy-c-means (left) and GK-clustering (right).*

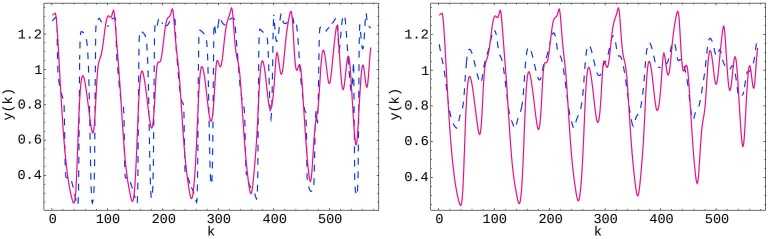


Figure 2: *Fuzzy-c-means: probabilistic method (left) vs. possibilistic method (right).*

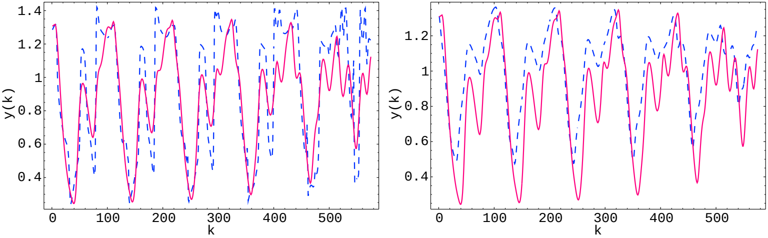


Figure 3: *GK-algorithm: probabilistic method (left) vs. possibilistic method (right).*

3. Discussion and Summary

- ✘ Fuzzy systems are nonlinear, universal approximator.
- ✘ Antecedent fuzzy sets partition the input space, while each rule defines a linear (sub)model in the consequent part of the if-then rules (facilitating least-squares estimation).
- ✘ Interpretability?
- ✘ Structure selection?
- ✘ GK-algorithm more accurate but computationally more intensive than FCM.

References

[1] Babuska, R. : *Fuzzy Modelling for Control*. Kluwer, 1998.
See http://lcewww.et.tudelft.nl/

[2] Takagi, T. and Sugeno, M. : *Fuzzy Identification of Systems and Its Application to Modelling and Control.*
IEEE Transactions on Systems, Men, and Cybernetics, Vol. 15, No. 1, January-February 1985, pp. 116–132.

[3] Wang, L.-X. : *A Course in Fuzzy Systems and Control.*
Prentice Hall, 1997.

[4] Wolkenhauer, O. : *Data Engineering*.
http://www.csc.umist.ac.uk/people/wolkenhauer.htm.