Ensuring monotonic gain characteristics in estimated models by fuzzy model structures

Peter Lindskog, Lennart Ljung
Division of Automatic Control
E-mail: peterl@isy.liu.se, ljung@isy.liu.se
25th June 2007

Report no.: LiTH-ISY-R-2802
Accepted for publication in Automatica, 2000

Address:
Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

WWW: http://www.control.isy.liu.se
Abstract
We consider the situation where a non-linear physical system is identified from input-output data. In case no specific physical structural knowledge about the system is available, parameterized grey-box models cannot be used. Identification in black box type of model structures is then the only alternative, and general approaches like neural nets, neuro-fuzzy models, etc., have to be applied. However, certain non-structural knowledge is sometimes available. It could be known, e.g., that the step response is monotonic, or that the steady-state gain curve is monotonic. The main question is then how to utilize and maintain such information in an otherwise black-box framework. In this paper we show how this can be done, by applying a specific fuzzy model structure, with strict parametric constraints. The usefulness of the approach is illustrated by experiments on real-world data.

Keywords: identification, model reduction, variance reduction
Ensuring monotonic gain characteristics in estimated models by fuzzy model structures

Peter Lindskog, Lennart Ljung

Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden

Received 25 April 1997; revised 29 March 1999; received in final form 4 June 1999

Abstract

We consider the situation where a non-linear physical system is identified from input-output data. In case no specific physical structural knowledge about the system is available, parameterized grey-box models cannot be used. Identification in black-box type of model structures is then the only alternative, and general approaches like neural nets, neuro-fuzzy models, etc., have to be applied. However, certain non-structural knowledge about the system is sometimes available. It could be known, e.g., that the step response is monotonic, or that the steady-state gain curve is monotonic. The main question is then how to utilize and maintain such information in an otherwise black-box framework. In this paper we show how this can be done, by applying a specific fuzzy model structure, with strict parametric constraints. The usefulness of the approach is illustrated by experiments on real-world data.

Keywords: Non-linear systems; Fuzzy modeling; Monotonicity; Model structures; Parameter estimation

1. Introduction

Don’t estimate what you already know!

This is a basic principle in estimation and identification and is also a pragmatic variant of the principle of parsimony — to be parsimonious with parameters to estimate (Ljung, 1999). In an identification context, the concept of grey boxes has been introduced to denote model structures that use some kind of prior information about the system. See, e.g., Bohlin (1991). The term tailor-made model structure has also been used. This is, of course, in contrast to black boxes or ready-made model structures, which just use “size” as the basic structure option.

Now, there are several shades of grey. Often grey boxes employ rather specific knowledge of the system; as an extreme it may correspond to a complete physical parameterization having some unknown parameters. These parameters are typically estimated by maximum likelihood/prediction error techniques.

More often, in particular in the process industry, the prior knowledge may be much “softer”, like knowing from physical grounds that the step response is monotonic. In such situations it is desirable to combine the richness and flexibility of non-linear black-box models, see, e.g., Sjöberg et al. (1995), with prior physical knowledge that is not of precise, analytical character. We would thus like to work with “boxes” that are just a shade lighter than black.

In this contribution we shall illustrate how this can be done using fuzzy model structures, see, e.g., Driankov, Hellendoorn and Reinfrank (1993), Lee (1990), Roger Jang and Sun (1995), Takagi and Sugeno (1985) and Wang (1994). We shall concentrate on one particular, but common, type of prior knowledge for a non-linear dynamic system, namely that the static gain curve is monotonic. If \( u \) denotes a constant input level, and if the system is stable, the output will converge to a constant level \( y \), and the function

\[
y = f(u)
\]

is the static gain curve. For many systems it follows from simple considerations that this curve must be monotonic.
(e.g., $u$ is heating power and $y$ is temperature, or $u$ is inflow in some system and $y$ is level, etc.). For a linear model, the gain curve is linear too, but if a flexible non-linear black-box structure is estimated from data, there is no guarantee that the obtained model will have a physically feasible gain curve. As illustrated in Section 4, this could lead to serious problems when applying the model.

2. The model structure

Let $\phi(t)$ denote the regressors of the model. Typically, it will consist of lagged, measured inputs and outputs from the system:

$$\phi(t) = y(t - 1), \ldots, y(t - n_y), u(t - 1), \ldots, u(t - n_u).$$

For the moment we do not assume any specific structure of $\phi(t)$, other than being an $r$-dimensional vector, measured and known at time $t - 1$. The $k$th component is denoted by $\phi_k(t)$. The model structure for the prediction of the output $y(t)$ is defined as

$$\hat{y}(t | \theta) = g(\phi(t), \theta) = \sum_{j_1=1}^{n_1} \cdots \sum_{j_r=1}^{n_r} \varphi_{j_1} \cdots \varphi_{j_r} \prod_{k=1}^{n_k} \mu_{j_k, k}(\phi_k(t), \beta_{j_k,k}, \gamma_{j_k,k}).$$

(3)

In a moment, we shall interpret this model parameterization as a fuzzy model structure, but first we just note that $\mu_{j_k,k}(\cdot)$ are functions of the regressor component parameters, parameterized by parameters $\gamma$ and $\beta$. The overall parameter vector is

$$\theta^T = [z^T \beta^T \gamma^T].$$

(4)

With a fuzzy model interpretation, the $\mu_{j_k,k}$ is the membership function (MF) for the (linguistic) variable $\phi_k$ belonging to the fuzzy set $j_k$ (denoting some characteristics of the variable, like “hot”, “cold”, etc.). In fuzzy model language, the parameterization corresponds to a singleton fuzzifier, a sup-star-based inference mechanism with Mamdani implication and algebraic product for AND, along with a center-of-sums defuzzifier; consult Driankov et al. (1993) for the definitions, and Ljung (1999) for the predictor structure interpretation. The membership functions at the output side (associated with the fuzzy sets $B_1, \ldots, B_{n_y}$) are chosen to be fuzzy singletons or centers, i.e., $\varphi_{j_k} = \alpha_{j_k,k}$, whereas the membership functions at the input side (associated with the fuzzy sets $A_1, \ldots, A_{n_u}$) generally involve both scale $\beta$ and position $\gamma$ parameters. This is the same terminology as is suggested in Sjöberg et al. (1995) and Lee (1990), where the series expansion (3) is referred to as a tensor product construction. However, contrary to many other similar approaches, it should here be noted that the parameters can be linguistically interpreted. As we shall see in the following section, this is an important property to preserve in the parameter estimation step.

Furthermore, we shall assume that the parameterization corresponds to a fuzzy partition, i.e., that

$$\sum_{j_k=1}^{n_k} \mu_{j_k,k}(\phi_k(t), \beta_{j_k,k}, \gamma_{j_k,k}) = 1$$

(5)

holds for all of the $r$ linguistic variables. If, in addition, we assume that the rule base is complete in the sense that it covers the whole input domain $\mathbb{U}^n$, it immediately follows that the model structure (3) simplifies to

$$\hat{y}(t | \theta) = g(\phi(t), \theta) = \sum_{j_1=1}^{n_1} \cdots \sum_{j_r=1}^{n_r} \varphi_{j_1} \cdots \varphi_{j_r} \prod_{k=1}^{n_k} \mu_{j_k, k}(\phi_k(t), \beta_{j_k,k}, \gamma_{j_k,k}).$$

(6)

At this point, note that a fuzzy partition puts certain demands on the MFs and their parameters. For example, we cannot in general use sigmoidal or Gaussian MFs because of their spreading and curvature. Piecewise linear MFs, on the other hand, can easily be parameterized so that a fuzzy partition is obtained. Within this category we will restrict the discussion to the open left mfi$(u, \gamma_1, \gamma_2)$, the open right mfr$(u, \gamma_1, \gamma_2)$ and the triangular mftri$(u, \gamma_1, \gamma_2, \gamma_3)$ MFs, which in order are given by

$$\mu(u, \gamma_1, \gamma_2) = \text{mfi}() = \max\left(0, \frac{\gamma_2 - u}{\gamma_2 - \gamma_1}\right).$$

(7)

$$\mu(u, \gamma_1, \gamma_2) = \text{mfr}() = \max\left(0, \frac{u - \gamma_1}{\gamma_2 - \gamma_1}\right).$$

(8)

$$\mu(u, \gamma_1, \gamma_2, \gamma_3) = \text{mftri}() = \max\left(0, \frac{u - \gamma_1}{\gamma_2 - \gamma_1}, \frac{\gamma_3 - u}{\gamma_3 - \gamma_2}\right).$$

(9)

3. Ensuring a monotonic gain curve

Consider now the case with a single-input linguistic variable ($r = 1$), so that (6) simplifies to

$$\hat{y}(t | \theta) = \sum_{j=1}^{n} \varphi_j \mu_j(\phi(t), \beta_j, \gamma_j).$$

(10)

We are here searching for parameter values and MFs guaranteeing that the predictor is monotonically increasing in $\phi(t)$. This can be achieved in many different ways (especially if the origin of (10) is neglected), but then it can be hard to express restrictions on the parameters that guarantee that the monotonic behavior is preserved in
the estimation step. However, this turns out to be a straightforward task when the input MFs form a fuzzy partition.

To see this, assume that all input MFs are ordered on the universe $\mathbb{U}$ in such a way that $\mu_j(\cdot)$ reaches a full degree of membership for a value of $\varphi(\cdot)$ that is lower than what is the case for $\mu_{j+1}(\cdot)$. If the ordered MFs at the input side form a fuzzy partition and the corresponding centers $x_j$ reflecting the output MFs are such that

$$x_1 < x_2 < \cdots < x_n,$$

then $\hat{y}(\theta)$ will be monotonically increasing in $\varphi(t)$. In verifying this, we first notice that at intervals where the $j$th input MF is fully active then the corresponding output becomes $x_j$. With fuzzy partitions constructed by piecewise linear MFs we also have that

$$\hat{y}(\theta) = x_j \mu_j(\cdot) + x_{j+1} \mu_{j+1}(\cdot) = (x_j - x_{j+1}) \mu_{j+1}(\cdot) + x_j$$

for all intervals $[\gamma_j, \gamma_{j+1}] \subset \mathbb{U}$ such that $\mu_j(\cdot)$ and $\mu_{j+1}(\cdot)$ are not always zero. Since $x_{j+1} > x_j$ (equality gives a constant output on the current interval) and $\mu_{j+1}(\cdot)$ is an increasing function on $[\gamma_j, \gamma_{j+1}]$, it follows that also $\hat{y}(\theta)$ is an increasing function on that interval, with values ranging from $x_j$ to $x_{j+1}$. These facts give that the overall predictor is a non-decreasing function. To get a strictly increasing mapping it additionally must be required that the input MFs lack intervals with a full degree of membership.

The next step is to generalize this result to predictors having $r$ regressors. To do this we start by formally defining what is meant by a monotonically increasing predictor.

**Definition 1 (Regressor ordering).** Let $\varphi(t)$, $\tilde{\varphi}(t) \in \mathbb{R}^r$. We say that $\varphi(t) \geq \tilde{\varphi}(t)$ if $\varphi_k(t) \geq \tilde{\varphi}_k(t)$ for $k = 1, \ldots, r$.

**Definition 2 (Monotonically increasing predictor).** Let $\varphi(t)$, $\tilde{\varphi}(t) \in \mathbb{R}^r$. We say that a predictor $g(\varphi(t), \theta)$ is monotonically increasing in the regressors if whenever $\varphi(t) \geq \tilde{\varphi}(t)$ it holds that $g(\varphi(t), \theta) \geq g(\tilde{\varphi}(t), \theta)$.

We now have the following result.

**Lemma 3.** Let the model structure be complete and given by (6). If, for all $k = 1, \ldots, r$, it holds that

$$\sum_{j_k = 1}^{n_k} x_{j_1, \ldots, j_k} \mu_{j_k, k}(\varphi_k(t) \beta_{j_k, k}, \gamma_{j_k, k})$$

are increasing functions in $\varphi_k(t)$ on $\mathbb{U}_k$ for all possible combinations of fixed values of $j_1, \ldots, j_{k-1}, j_{k+1}, \ldots, j_r$, then the predictor (6) is monotonically increasing in the regressors $\varphi(t)$.

**Proof.** Let $\varphi_k(t)$ denote any regressor and fix all other regressors $\varphi_{k'}(t) = \varphi_{k'}^* \in \mathbb{U}_{k'}$, $k = 1, \ldots, l - 1, l + 1, \ldots, r$. Rearranging the terms of the predictor (6) gives

$$\hat{y}(t | \theta) = \sum_{j_1 = 1}^{n_1} \cdots \sum_{j_{r-1} = 1}^{n_{r-1}} \sum_{j_r = 1}^{n_r} \prod_{k=1}^{r} \mu_{j_k, k}(\varphi_k^*) \sum_{\gamma_{j_k, k}} x_{j_k, \ldots, j_r} \mu_{j_{r+1}, k}(\varphi_r(t)),$$

where for simplicity the $\beta_{j_k, k}$ and the $\gamma_{j_k, k}$ parameters have been dropped. The first part of the expression (including the product) returns weights formed by taking the product of $r - 1$ MFs, i.e., all the weights lie in $[0, 1]$. By the assumption, the last sum returns functions that are increasing in $\varphi_r(t)$ on $\mathbb{U}_r$, which means that the predictor is a weighted (positive weights) sum of increasing functions. This gives that the overall predictor is monotonically increasing in the regressors $\varphi(t)$.

The main point with Theorem 3 is that it is sufficient to work with one-dimensional functions. A simple way to ensure increasing functions in all $\varphi_k(t)$ is now to restrict the input MFs to fuzzy partitions and order the corresponding centers as was done in the one-dimensional case.

**Lemma 4.** Let the model structure be (6) and let $\varphi_k(t)$ denote one of its regressors. Assume that the ordered (on $\mathbb{U}_k$) MFs associated with $\varphi_k(t)$ are piecewise linear and such that (5) holds. If, for all combinations of $j_1, \ldots, j_{k-1}, j_{k+1}, \ldots, j_r$, it holds that

$$x_{j_1, \ldots, j_{k-1}, j_k} \leq x_{j_1, \ldots, j_{k+1}}$$

then every

$$\sum_{j_k = 1}^{n_k} x_{j_1, \ldots, j_k} \mu_{j_k, k}(\varphi_k(t), \gamma_{j_k, k})$$

is a monotonically increasing function in $\varphi_k(t)$.

This lemma follows directly from the one-dimensional case discussed above. The requirements for Lemma 3 to hold are fulfilled if all MFs are chosen according to Lemma 4. This is the case for the rule base in Fig. 1, from which it is clear that the resulting predictor returns a larger (or unchanged) output if one or more of the regressors become larger. Moreover, if the orders among the parameters $\alpha$ and $\gamma$ are maintained in the estimation step, then this is a fact that cannot be altered by the estimation procedure. Using a squared prediction error optimization criterion, the obtained constrained minimization problem can be solved, e.g., by a barrier function method; see Fletcher (1987) and Scales (1985) for algorithmic details.
At this point, assume that the regressors include dynamics
\[ \varphi(t) = [y(t-1) \ y(t-2) \ \ldots \ u(t) \ u(t-1) \ \ldots]^T, \quad (14) \]
where, without loss of generality, only one input signal is present. A globally asymptotically stable predictor in \( \varphi(t) \) implies that a constant input \( u^* = u(t) = u(t-1) = \ldots \)
leads to a constant output \( y^* \) as \( t \to \infty \). Plotting \( y^* \) for each value of \( u^* \) gives the steady-state gain curve.

**Lemma 5.** Let \( u^*, y^* \) and \( \bar{u}^*, \bar{y}^* \) be two steady-state solutions to a globally asymptotically stable predictor \( g(\varphi(t), \theta) \), i.e.,
\[ y^* = g([y^* \ y^* \ \ldots \ u^* \ u^* \ \ldots]^T, \theta), \]
\[ \bar{y}^* = g([\bar{y}^* \ \bar{y}^* \ \ldots \ \bar{u}^* \ \bar{u}^* \ \ldots]^T, \theta). \]
If \( g(\varphi(t), \theta) \) is monotonically increasing in \( \varphi(t) \) and \( u^* \geq \bar{u}^* \), then \( y^* \geq \bar{y}^* \).

**Proof.** Suppose that \( y^* < \bar{y}^* \). Let \( u(t) = \bar{u}^* \) for \( t \leq 0 \), whereupon \( u(t) = u^* \) for \( t > 0 \). This input sequence results in an output sequence \( \{y(t)\} \) and \( r \) regressor sequences \( \{\varphi_k(t)\} \). If \( y^* < \bar{y}^* \), then there exists a \( t^* \) such that \( y(t^*) < \bar{y}^* \) occurs for the first time:
\[ y(t^*) = g([y(t^*-1) \ y(t^*-2) \ \ldots \ u^* \ u^* \ \ldots]^T, \theta) = g(\varphi(t^*), \theta). \]
Since this happens for the first time we have that \( \varphi(t^*) \geq \bar{\varphi}^* \), which by the monotonicity assumption implies that \( y(t^*) \geq \bar{y}^* \). This is a contradiction! \( \square \)

Thus, combining the results of the lemmas shows that maintaining (13) is sufficient to guarantee a monotonic gain curve for the resulting model. Furthermore, starting from a steady-state solution and increasing the input in a stepwise fashion, it follows by simple induction that \( \bar{y}(\theta) \) increases monotonically with \( t \). This, in particular, means that the predictor shows a non-oscillatory step response behavior, which is a restriction but also a property that is valid for many industrial processes (e.g., thermal systems), as will be illustrated next.

4. Example — water heating system

This section considers identification of a water heating system as depicted in Fig. 2. The process has earlier been investigated by Koivisto (1995), from which the following system description is gathered.

Water from the domestic water network flows across an uninsulated 0.4 l tank. On its way, the water is heated by a resistor element, which in turn is controlled by the voltage \( u(t) \) applied to a thyristor. At the outlet, the water temperature \( T(t) \) is measured with a Pt-100 transducer. As in Koivisto (1995), we will here restrict the discussion to a situation where \( Q_{\text{in}}(t) \) as well as the inlet water temperature do not change. The modeling problem is then to describe the outlet water temperature \( T(t) \) given the voltage \( u(t) \).

The data to be used originate from a real-time identification run (performed by Koivisto), where the process was driven by a pseudo-random type of input signal \( u(t) \) (given in percent of the maximum allowed voltage). The experiment lasted 9000 s and measurements were recorded every 3rd second, thereby resulting in a data set of 3000 samples. This record was then divided into an estimation set of 2000 samples, see Fig. 2, and a validation set consisting of the remaining 1000 samples.
The thyristor has a saturating characteristics.

Step response tests show that the time delay from a change in the input can be seen in the output is 12–15 s. Since the sampling interval is 3 s, useful regressors stemming from the input are \( u(t - 4), u(t - 5) \), and so on.

The thyristor has a saturating characteristics.

The temperature \( T(t) \) will increase if more power (a larger \( u(t) \)) is applied to the heater element. The steady-state gain curve of a physically sound model should thus be monotonically increasing in \( u(t) = u^* \).

The third item is extremely important to retain as the model is going to be used in a model predictive control (MPC) arrangement, see Koivisto (1995), where the aim of the control is to drive the temperature \( T(t) \) to a desired set-point value. Let us briefly explain this. Suppose that the current steady-state point lies in an area where the steady-state gain curve shows a decreasing characteristics. Locally this curve indicates that to decrease the temperature \( T(t) \), one should actually increase the voltage \( u(t) \). Such a decision is of course fatal as it is known from physics that an increase in the voltage leads to an increase in the outlet temperature. A controller based on this model will thus react in a qualitatively opposite manner to what is reasonable, and it is not hard to understand that such a behavior can cause severe stability problems. Koivisto (1995) has reported such problems, when using general non-linear black-box models of neural network type.

This gives us a reason to try the fuzzy model structure (6). Desiring a model of low complexity, correlation tests indicate that \( T(t - 1) \) and \( u(t - 5) \) are reasonable regressor candidates. The corresponding linguistic variables are, e.g.,

\[
\langle \text{TEMP}(t), \{\text{VL}, \text{L}, \text{M}, \text{H}, \text{VL}, \text{L}, \text{M}, \text{H} \}, \hat{T}(t | \theta) \rangle,
\]

\[
\langle \text{TEMP}(t - 1), \{1, \text{M}, \text{H} \}, \phi_1(t) = z_1(t - 1) \rangle,
\]

\[
\langle \text{VOLTAGE}(t - 5), \{1, \text{H} \}, \phi_2(t) = z_2(t - 5) \rangle,
\]

\[
\hat{T}(t | \theta) \in \mathcal{Y} = [10, 50], \quad \phi_1(t) \in \mathcal{U}_1 = [10, 50], \quad \phi_2(t) \in \mathcal{U}_2 = [0, 100],
\]

where \( z(t) = [T(t) \ u(t)] \) and \( \text{VL}, \text{L}, \text{R}, \text{M}, \text{H} \) and \( \text{VL} \) are abbreviations for \text{VERY LOW}, \text{LOW}, \text{RATHER LOW}, \text{MEDIUM}, \text{HIGH} \) and \text{VERY HIGH}. Note that this choice addresses the first system property, while the last two properties are guaranteed if the MFs

\[
\mu_{a_1}(\hat{T}(t | \theta)) = \alpha_{1,1}, \quad \mu_{a_2}(\hat{T}(t | \theta)) = \alpha_{2,1}, \quad \mu_{a_3}(\hat{T}(t | \theta)) = \alpha_{3,1}, \quad \mu_{a_4}(\hat{T}(t | \theta)) = \alpha_{3,2}
\]

are used in the predictor

\[
\hat{T}(t | \theta) = \sum_{j_1 = 1}^{3} \sum_{j_2 = 1}^{2} \sum_{j_3 = 1}^{2} \sum_{j_4 = 1}^{2} \sum_{j_5 = 1}^{2} \mu_{a_1}(\phi_1(t), \gamma_{j_1,1}) \mu_{a_2}(\phi_2(t), \gamma_{j_2,1}) \mu_{a_3}(\phi_3(t), \gamma_{j_3,1}) \mu_{a_4}(\phi_4(t), \gamma_{j_4,1}) \mu_{a_5}(\phi_5(t), \gamma_{j_5,1}),
\]

which contains 11 parameters

\[
\theta = \begin{bmatrix} \alpha_{1,1} \alpha_{1,2} \alpha_{2,1} \alpha_{2,2} \alpha_{3,1} \alpha_{3,2} \gamma_{1,1} \gamma_{2,1} \gamma_{2,2} \gamma_{3,1} \gamma_{3,1} \gamma_{3,2} \gamma_{3,2} \end{bmatrix}^T
\]

chosen so that

\[
10 < \alpha_{1,1} < \alpha_{1,2} < \alpha_{2,1} < \alpha_{2,2} < \alpha_{3,1} < \alpha_{3,2} < 50,
\]

\[
10 < \gamma_{1,1} < \gamma_{2,1} < \gamma_{3,1} < 50,
\]

\[
0 < \gamma_{1,2} < \gamma_{2,2} < 100.
\]

A graphical representation of the corresponding fuzzy rule base is shown in Fig. 3, where dotted lines represent
the initial positions of the MFs. Constrained estimation of \( \theta \) subject to constraints (17) results in a model (with MFs according to Fig. 3), whose simulation behavior is reproduced in Fig. 4.

A linear model will of course always have a monotonic gain curve. The dynamic properties of the best linear model found (an ARX model with four parameters) are shown in Fig. 5. Its root mean square (RMS) error is 2.08 compared to the fuzzy model’s 1.02, i.e., it is more than twice as big. The improvements in the non-linear model are significant at low and high temperatures, since the linear model cannot capture the saturating characteristics of the thyristor. There are consequently important reasons to use a non-linear model. The built-in increasing nature of the steady-state gain curve of the estimated fuzzy model is shown in Fig. 6.

From these experiments we conclude that the derived fuzzy model is able to accurately describe the water
heating system, at the same time as the important monotonicity property is ensured. The obtained RMS error 1.02 should be compared to 0.92, which is obtained in Koivisto (1995) using a neural net with many more parameters (31 compared to 11 in the fuzzy case). This along with the ensured monotonic behavior suggest that the above fuzzy approach might be a good alternative to neural nets when applied to predictive control.

5. Conclusions

We have in this paper addressed the problem of “dark grey boxes”: How to include, ensure and maintain a known qualitative behavior of a process in a model structure that is otherwise quite flexible. This is difficult to achieve in traditional grey-box structures, unless rather precise physical knowledge in analytic form is at hand.

To deal with the problem, we have turned to black-box structures of kind (3), using basis functions that have some useful interpretation related to the system’s behavior. This is the basic property of fuzzy models (“neuro-fuzzy models”). We have shown how to parameterize these basis functions (“membership functions”) and also how to constrain their parameters so that certain monotonicity properties in the model’s response are guaranteed.

Experiments on real-world data — in this case Koivisto’s heating system — as well as other applications (Lindskog, 1996), have demonstrated the usefulness of the approach.

Acknowledgements

This work was supported by the Swedish Research Council for Engineering Sciences (TFR), which is gratefully acknowledged.

References


Ljun, L. (1999). System identification — parts I and II. Ph.D. Thesis 170, Tampere University of Technology. His current research interests include system identification and modeling. In 1997, he joined Ericsson Radio Systems AB and started to work with radio network modeling and control. He is now with Atero AB, focusing mainly on design, modeling and control of complex systems.

Peter Lindskog was born in Nyköping, Sweden, in 1965. He received the M.Sc. in computer science and engineering in 1990 and the Ph.D. degree in automatic control in 1996, both from Linköping Institute of Technology. His current research interests include system identification and modeling. In 1997, he joined Ericsson Radio Systems AB and started to work with radio network modeling and control. He is now with Atero AB, focusing mainly on design, modeling and control of complex systems.

Lennart Ljung received his Ph.D. in Automatic control from Lund Institute of Technology in 1974. Since 1976 he is Professor of the chair of Automatic Control in Linköping, Sweden, and is currently Director of the Competence Center “Information Systems for Industrial Control and Supervision” (ISIS). He has held visiting positions at Stanford and MIT and has written several books on System Identification and Estimation. He is an IEEE Fellow, an IFAC Advisor, a member of the Royal Swedish Academy of Sciences (KVA) and of the Royal Swedish Academy of Engineering Sciences (IVA), and has received honorary doctorates from the Baltic State Technical University in St. Petersburg, and from Uppsala University.
Ensuring monotonic gain characteristics in estimated models by fuzzy model structures

We consider the situation where a non-linear physical system is identified from input-output data. In case no specific physical structural knowledge about the system is available, parameterized grey-box models cannot be used. Identification in black-box type of model structures is then the only alternative, and general approaches like neural nets, neuro-fuzzy models, etc., have to be applied. However, certain non-structural knowledge is sometimes available. It could be known, e.g., that the step response is monotonic, or that the steady-state gain curve is monotonic. The main question is then how to utilize and maintain such information in an otherwise black-box framework. In this paper we show how this can be done, by applying a specific fuzzy model structure, with strict parametric constraints. The usefulness of the approach is illustrated by experiments on real-world data.

Keywords identification, model reduction, variance reduction