Innovative Criteria for Input Selection

Madasu Hanmandlu
Department of Electrical Engg.,
I.I.T. Delhi, Hauz Khas,
New Delhi -110016, India.
mhmandlu@ee.iitd.ernet.in

Mohd. Fazle Azeem
Department of Electrical Engg.,
Aligarh Muslim University,
Aligarh – 202002, U.P., India.
mf_azeem@yahoo.com

Vamsi Krishna Madasu
School of ITEE,
University of Queensland,
Brisbane, QLD 4072, Australia.
madasu@itee.uq.edu.au

Abstract

In this paper an attempt is made to derive new criteria for input selection of dynamic systems using the fuzzy curve approach. The Approximate Fuzzy Data Model (AFDM), the output of which is the fuzzy curve, is shown to be a special case of the Generalized Fuzzy Model (GFM). Moreover, AFDM is proved to be an unconditional expectation of the output thus linking fuzzy rules with probability. The validity of the criteria for input selection has been studied on GFM by means of significance of inputs, which is determined from the ratio of change in the output of AFDM to the range of the actual output. The complexity of the criteria has been proved to be of the order of O(n), which is a significant achievement in comparison to the complexity of the existing criteria.

Key words: Input selection criteria, Fuzzy modeling, Expectation, Fuzzy curve

1 Introduction

To build a fuzzy model, the first step is to select significant inputs among the many input candidates [1-4]. We can identify the significant inputs to the model by using any identification algorithm [1-3]. But, these algorithms either fail to select the exact number of inputs [2] to the model or involve expensive computations [1,3]. A number of models are available for identification. The models of Xu and Lu [2] are based on two input variables, and are, therefore, not suitable for multivariable systems. Takagi and Sugeno [1] have proposed a method that requires a maximum of $n\{(2p-1)n+1\}/2$ test models and 2^p partitions for n input candidates and p outputs. Hence, the number of test models is of the order $O(n^2 p)$. Sugeno and Yasukawa [3] have proposed another method, which requires $n(n+1)/2$ test models for input selection, by decoupling the number of partitions. Hence, the number of test models is of the order of $O(n^2)$. In this paper, we seek to select the input variables that significantly affect the output by evaluating the performance of a model using some criterion.

Lin and Cunningham III [4] have proposed a fuzzy curve approach for input variable selection. This approach lists only a comparative significance of each input candidate without any further test for input selection. In this paper, we propose new criteria for input variable selection on the basis of significance evaluated from fuzzy curves, which are obtained from the Approximate Fuzzy Data Model (AFDM) over the domain of input.

2 Dynamic Fuzzy Models

Dynamic system models may be constructed by mapping from the input space to the output space. The mapping function for a fuzzy system is the result of function approximation and is called the universal function approximator. To construct fuzzy rules for a Multi-Input and Multi-Output (MIMO) system, we separate the overall system into several Multi-Input and Single-Output (MISO) systems, which can be represented by the any one of the following models with the associated rules:

Compositional Rule of Inference (CRI) Model

$$R^k : \text{if } x_1 \text{ is } \mathbf{A}_1^k \wedge x_2 \text{ is } \mathbf{A}_2^k \wedge \ . \ . \ . \ \wedge x_n \text{ is } \mathbf{A}_n^k \quad \text{then } y \text{ is } \mathbf{B}^k \qquad (1)$$

Takagi-Sugeno (TS) Model

$$R^k : \text{if } x_1 \text{ is } \mathbf{A}_1^k \wedge x_2 \text{ is } \mathbf{A}_2^k \wedge \ . \ . \ . \ \wedge x_n \text{ is } \mathbf{A}_n^k \quad \text{then } y \text{ is } f^k(\mathbf{x}) \qquad (2)$$

Generalized Fuzzy Model (GFM)

$$R^k : \text{if } x_1 \text{ is } \mathbf{A}_1^k \wedge x_2 \text{ is } \mathbf{A}_2^k \wedge \ . \ . \ . \ \wedge x_n \text{ is } \mathbf{A}_n^k$$

$$\text{then } y \text{ is } \mathbf{B}^k\left(f^k(\mathbf{x}),\ v^k\right) \quad (3)$$

where, $x_i (i = 1,2,...,n)$ are the inputs and y is the output. $\mathbf{A}_i^k$ denotes the linguistic labels of fuzzy sets describing the qualitative state of the input variable x_i in k[th] rule, i.e., R^k, $\wedge$ is a fuzzy conjunction operator and $\mathbf{B}^k\left(f^k(\mathbf{x}), v^k\right) \subset R$ denotes the linguistic labels attached to the local nonlinear regression of inputs $f^k(\mathbf{x})$ with an index of fuzziness v^k, which describes the qualitative state of the output variable y.

A linear form of $f^k(\mathbf{x})$ in Eqn. (2) is as follows:

$$f^k(\mathbf{x}) = b_0^k + b_1^k x_1 + \ldots + b_n^k x_n \quad (4)$$

The firing strength of the k[th] rule is obtained by taking T-norm of the membership functions of the premise parts of the rule as:

$$\mu^k(\mathbf{x}) = \mu_1^k(x_1) \wedge \mu_2^k(x_2) \wedge \ . \ . \ . \ \wedge \mu_n^k(x_n) \quad (5)$$

where, $\mu_i^k(x_i)$ is the membership function of fuzzy set $\mathbf{A}_i^k$. The firing strength of k[th] rule for variable y is also represented as a fuzzy set $\mathbf{A}^k \subset R^n$ in the input space. Hence, Eqn. (3) can be written in compact form as:

$$R^k : \text{if } \mathbf{x} \text{ is } \mathbf{A}^k \text{ then } y \text{ is } \mathbf{B}^k\left(f^k(\mathbf{x}), v^k\right) \quad (6)$$

For the specific class of GFM [5] where we use the multiplicative T-norm operator for mapping fuzzy subsets from the input space, $\mathbf{A}^k \subset R^n$ to fuzzy subset in the output space, $\mathbf{B}^k\left(f^k(\mathbf{x}),\ v^k\right) \subset R$, and the additive S-norm operator to join all the mapped region in the output space, the weighted average defuzzified output value y^o is:

$$y_i^o(x) = \sum_{k=1}^{m} \frac{\mu^k(x).v^k}{\sum_{k=1}^{m} \mu^k(x).v^k} . f^k(x) \quad (7)$$

This model can be reduced to either CRI-model by applying the condition $f^k(\mathbf{x}) = b^k$ or TS-model by applying the condition $v^1 = v^2 = \ . \ . \ . \ = v^m$.

3 AFDM as a special case of GFM

To prove the above, we split the MISO system, into a set of SISO systems. We intend to derive AFDM for each of the inputs x_i and then evaluate its significance. Here, we define a rule in $x_i - y$ space for each element of the input x_i so that the number of rules would be equal to the number of elements. The rule is of the form:

$$R_i^k : \text{If } x_{ik} \text{ is } X_i^k \text{ then } y \text{ is } y_k \quad (8)$$

where, X_i^k is the fuzzy set having a Gaussian membership function $\mu^k(x_i)$ corresponding to the data point (x_{ik}, y_k) given by,

$$\mu^k(x_i) = \exp\left\{-\left(\frac{x_{ik} - x_i}{w}\right)^2\right\}; k = 1,2,...m \quad (9)$$

The value of w is taken in the range of 10% to 20% of the input interval of x_i. By interpreting the terms, Eqn.(6) can be adapted to obtain the defuzzified output for the Eqn.(8). Since each membership function $\mu^k(x_i)$ of x_i is located at $x_i = x_{ik}$ with the maximum value of 1 having the same width, w, the strengths of all $\mu^k(x_i)$ are the same. This leads to $v^1 = v^2 = ... = v^m$. We replace $\mu^k(x)$ in Eqn.(6) by $\mu^k(x_i)$ and $f^k(x)$ by $f^k(x_i)$. According to the rule, $f^k(x_i) = y_k$. In view of these substitutions, Eqn. (6) becomes:

$$y_i^0(x_i) = \frac{\sum_{k=1}^{m} \mu^k(x_i) y_k}{\sum_{k=1}^{m} \mu^k(x_i)} \quad (10)$$

where, superscript '0' stands for the output of fuzzy curve and y_i^0 represents the fuzzy curve corresponding to the input x_i. Thus, GFM becomes AFDM when all elements of input have the same membership function

3.1 Proof of y as unconditional expectation

The membership function $\mu^k(x_i)$ of each point (x_{ik}, y_k), can be thought of as the distribution of x_{ik}. Let, this distribution be denoted by, $p_X^k(x_i) = \mu^k(x_i)$. The distribution of y_k is 1 since it is a singleton. In view of this interpretation, we can rewrite Eqn. (10) as,

$$y_i^0(x_i) = \frac{\sum_{k=1}^{m} p_X^k(x_i) p_Y^k(y) y_k}{\sum_{k=1}^{m} p_X^k(x_i)} \qquad (11)$$

Let, $p_{XY}^k(x_i, y) = p_X^k(x_i) p_Y{}^k(y)$ as both terms on the right side are independent, we have:

$$y_i^0(x_i) = \frac{\sum_{k=1}^{m} p_{XY}^k(x_i, y) y_k}{\sum_{k=1}^{m} p_X^k(x_i)} = \sum_{k=1}^{m} y_k p^k(y \mid x_i) = E(y \mid x_i) \qquad (12)$$

Thus, we have proved that the fuzzy curve is the unconditional expectation.

3.2 Proof of conditional expectation of the output of GFM

Replacing y_k by $f^k(x_i)$ introducing v_k in Eqn.(12) yields the defuzzified output of GFM:

$$y_i^0(x_i) = \frac{\sum_{k=1}^{m} v^k p_{XY}^k(x_i, y) f^k(x_i)}{\sum_{k=1}^{m} v^k p_X^k(x_i)}$$

$$= \sum_{k=1}^{m} f^k(x_i) p_v{}^k(y \mid x_i) = E(y \mid x_i) \qquad (13)$$

However, it is easy to show that CRI model output is the unconditional expectation contrary to TS model.

4 Input Selection Criteria

From the shape of a fuzzy curve [4], the following inferences can be drawn. Change in positive slope of the input results in an increase of the output with an increase in x_i. Based on the heuristic that fuzzy model will interpolate between maxima and minima, the minimum number of rules required to approximate the fuzzy curve can be determined from the number of maxima and minima. The importance of variable x_i is evaluated from the approximate changes in the output of AFDM, i.e., c_i over the range of input x_i. Thus, c_i is defined as in [6]:

$$c_i = \frac{\mathbf{max}\left(y_i^o(x_i)\right) - \mathbf{min}\left(y_i^o(x_i)\right)}{\mathbf{max}\left(y_k\right) - \mathbf{min}\left(y_k\right)} \times 100 \qquad (14)$$

Before applying the following criteria, we sort out all the input candidates for the model in descending order of c_i.

Criterion1*:* Select all those input candidates for which $c_i \geq h$, where $h \in [0, 100]$.

The value of "*h*" depends upon the number of input candidates as well as values of c_i. The higher the value of "h", lesser is the number of selected input candidates and vice-versa.

Criterion2*:* Among the selected input candidates, start modeling with the input candidate of highest value of c_i and go on adding input candidates one by one with the descending order of c_i. In doing so, accept only those input candidates which significantly improve the performance of model and reject those input candidates, inclusion of which deteriorate the performance of a model. Terminate the input selection when the model performance reaches the target value. Supposing that n variables are selected with criterion 1, and then we need to evaluate *n* test models by checking every input candidate to see whether it results in the best performance of model. Again, on the basis of criterion 2, we need a testing of *n*-1 models. Together, the count comes to 2*n*-1 test models, which is linear in the number of input candidates, i.e., $O(n)$.

5 Simulation Results

A normalized mean square error ***J***, is considered for the evaluation of the test models using GFM. The initial parameters for all the test models are found from the fuzzy curve method. Further these parameters are fine-tuned using Gradient Descent (GD) and Least Squares Estimation (LSE) techniques. For stopping the input selection, the target value of ***J*** is set to 2×10^{-4}. In view of a limited number of input candidates, we apply criterion 2 directly.

Example: Bilinear system

We consider a two-input and single output bilinear dynamic model [2] given by the equation:

$$y(k) = 0.8y(k-1)u_1(k) + 0.5u_1(k-1)y(k-2) + u_2(k-4) + \alpha.e(k) \qquad (15)$$

to provide the following input-output data sequence:

$\{y(k), u_1(k), u_2(k), k = 1{,}400\ \&\ \alpha = 1\}$. Note that e($t$) in Eqn. (16), is an uncorrelated random noise uniformly distributed over the interval (-0.08, 0.08). Inputs $u_1(k)$ & $u_2(k)$ are both uncorrelated random sequences uniformly distributed over the interval (0.1, 0.9).

We consider the variables $u_1(t)$, . . ., $u_1(t\text{-}4)$, $u_2(t)$, . . ., $u_2(t\text{-}6)$, $y(t\text{-}1)$, . . ., $y(t\text{-}3)$ as input candidates to the model. Using the fuzzy curves of all the input candidates, the values of c_i are obtained from Eqn. (15). The number of fuzzy rules for the model is given by the number of maxima and minima from each of the fuzzy curves, i.e., m=3 in the example.

In Table I, the values of c_i and descending order of c_i are listed. First, 80% of the data is used for model learning and the rest is used for model validation. In the figure, '+' sign shows the model performance with the each individual input candidate added, the solid line shows the performance of the models on applying criterion 2 and a circle on the solid line shows the acceptance of that input candidate. The selected variables for the data are $y(t\text{-}1)$, $y(t\text{-}2)$, $u_2(t\text{-}4)$, $u_1(t)$ and $u_1(t\text{-}1)$.

Table 1: List of input candidates with descending order of c_i for the Example.

Inputs	c_i	Order of c_i	Inputs	c_i	Order of c_i
$u_1(t)$	26.38	4	**$u_2(t\text{-}3)$**	5.66	13
$u_1(t\text{-}1)$	24.16	5	**$u_2(t\text{-}4)$**	29.00	3
$u_1(t\text{-}2)$	14.66	10	**$y(t\text{-}1)$**	42.60	1
$u_2(t)$	2.29	16	**$y(t\text{-}2)$**	33.32	2
$u_2(t\text{-}1)$	4.25	14	**$y(t\text{-}3)$**	22.30	6
$u_2(t\text{-}2)$	3.84	15	**$y(t\text{-}4)$**	16.13	9

6 Conclusions

The fuzzy curve, which is the output of AFDM, is shown to be the special case of defuzzified output of GFM in addition to proving that it is an unconditional expectation of output. This fuzzy curve is used to find the significance of inputs. The significance is evaluated from the descending order of a ratio that is the percentage change in the output of AFDM. As compared to the requirement of testing models of orders $O(n^2 p)$ and $O(n^2)$ in TS & SY-methods, the new criteria require the testing models of order $O(n)$. The validity of the proposed criteria is shown on the GFM of a dynamic system.

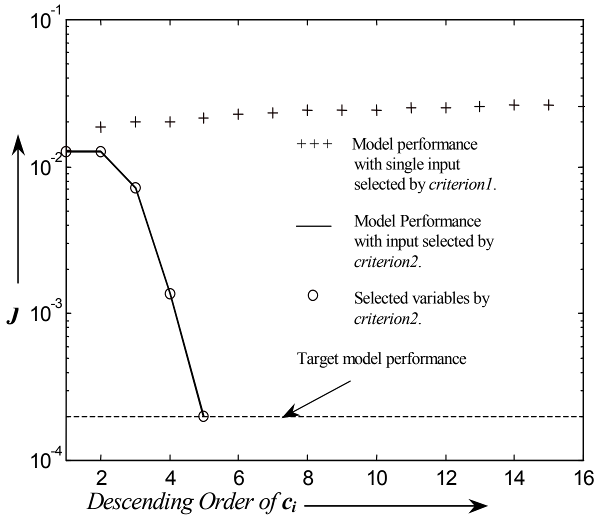

Figure 1: Model Performance vs Descending Order of c_i for the Example

References

[1] T. Takagi and M. Sugeno (1985). Fuzzy Identification of Systems and its Application to Modeling and Control. In *IEEE Transactions on Systems, Man and Cybernetics*, volume 15, pages 116-132, January-February 1985.

[2] C-W. Xu and Y-Z. Lu (1987). Fuzzy Model Identification and Self-Learning for Dynamic System. In *IEEE Transactions on Systems Man and Cybernetics*, volume 17, pages 683-689, July-August 1987.

[3] M. Sugeno and T. Yasukawa (1993). A Fuzzy logic based approach to qualitative modeling. In *IEEE Transactions on Fuzzy Systems*, volume 1, pages 7-31, 1993.

[4] Y. Lin , G.A. Cunningham III (1995). A New Approach to Fuzzy-Neural System Modeling. In *IEEE Transactions on Fuzzy Systems*, volume 3, pages 190-198, May 1995.

[5] M.F. Azeem, M. Hanmandlu and N. Ahmad, (2000). Generalization of Adaptive Neuro-Fuzzy Inference Systems. In *IEEE Transactions on Neural Networks*, volume 11, pages 1332-1346, November 2000.

[6] M.F. Azeem, M. Hanmandlu and N. Ahmad (1998). A new criteria for input variable identification of dynamical systems. In *Proceedings of the IEEE TENCON'98*, volume 1, pages 230-233, New Delhi, December 1998.