N-Dimensional Structure Identification and Fuzzy Reasoning

Kemal Kilic
Department of Mechanical and Industrial Engineering, University of Toronto
kilic@mie.utoronto.ca

I. Burhan Turksen
Department of Mechanical and Industrial Engineering, University of Toronto
turksen@mie.utoronto.ca

Abstract

In this paper a new fuzzy modeling algorithm is proposed and a new inference schema is developed for fuzzy reasoning that suits to the proposed approach

Keywords: Fuzzy System Modeling, k-nearest neighbor, fuzzy inference

1. Introduction

Fuzzy system modeling consists of two stages. The first stage is the structure identification stage where the fuzzy if-then rules are determined, and the second stage is the fuzzy inference where a fuzzy output is inferred from the given fuzzy if-then rules and a fuzzy premise.

In the early days of fuzzy system modeling the *structure* was determined a priori subjectively from other sources such as experts' opinion. Recent research on fuzzy system modeling is based on the objective determination of the *structure* from historical data. One of the well known generic frames proposed for fuzzy system modeling, is the Sugeno-Yasukawa model [4] which was further investigated by Nakanishi et. al. [2]. This approach has been slightly modified by Turksen and Bazoon and successfully applied to pharmacological problems [3]. A typical fuzzy rule base is as follows;

$Rule_1$: IF X_1 isr $A_{1,1}$ AND ... AND X_{NV} isr $A_{1,NV}$ THEN Y isr B_1

...

$Rule_c$: IF X_1 isr $A_{c,1}$ AND ... AND X_{NV} isr $A_{1,NV}$ THEN Y isr B_c

where X_j's are the fuzzy linguistic input variables in universe of discourse U_j, Y is the fuzzy linguistic output variable in universe of discourse V, $A_{i,j}$'s and B_i's are fuzzy sets. NV stands for number of input variables and c denotes the number of rules

In order to determine the fuzzy if then rules from the historical data, Sugeno-Yasukawa [4] proposes first to cluster the outputs (to determine B_i's) and then projecting the output clusters onto each input variable (to form the $A_{i,j}$'s). Therefore, if an individual has an output membership degree of $\mu_{Bi}(y)$ then the input features that leads to such an output value also have a membership degree of $\mu_{Bi}(y)$. After projecting the output clusters onto input space, in order to form the input fuzzy clusters ($A_{i,j}$'s) a convex trapezoid is fitted. The formation of the fuzzy rule base is known the structure identification level.

Next step is fuzzy inference; given a new data with known input obtain the output. Hence the question is given "X_1 isr A_1\` AND $X_2$ isr $A_2$\` ... AND X_{NV} isr A_{NV}\`" what is going to be B\` for Y. Usually, in real life one desires to obtain a crisp output, and the new data vector provided has crisp input values. Hence A_j\` is a crisp set such as $A_j$\`(x_j\`)=1 and 0 otherwise.

The following general schema achieves the fuzzy inference for crisp input cases which is known as first infer then aggregate (FITA) approach,

1. Determine degree of firing for each rule $\tau_i(X`)$
2. Infer by using an implication operator
3. Aggregate the outputs of all the rules
4. Defuzzify the output

Briefly, Sugeno-Yasukawa proposes to determine the degree of the firing of the new data for each rule and obtain the output by taking the weighted average of the center of gravities of the output fuzzy sets where the weights are the degree of firings. These weights are determined for each rule *i* by determining the membership degree of x_j\`'s in fuzzy set $A_{i,j}$ for each input variable j, and "anding" the membership degrees, with algebraic product, i.e., $\tau_i(X')=A_{i,1}(x_1')\times\ldots\times A_{i,NV}(x_{NV}')$ where X' is an n-dimensional vector of $[x_1', x_2',\ldots,x_{NV}']$.

Note that this fuzzy reasoning schema is only applicable to rule bases where the input space variables are treated independently, i.e., the output fuzzy clusters are projected to the input space one by one. However, usually each data vector belongs to an *object*, hence the value of the attributes are dependent on each other. The major problem with the structure identification approaches in the current literature is that the natural links between the input variables are ignored and they are assumed to be independent. Furthermore the existing algorithms assume a single convex input fuzzy set for each fuzzy output set. It has been demonstrated that both assumptions (independence of input variables, and convexity of input fuzzy set) produce invalid rule structures [1]. In this paper a new structure identification is presented where the output fuzzy clusters are projected onto n-dimensional input space.

2. The Proposed Algorithm

The basic steps of the proposed approach is similar to the Sugeno-Yasukawa.

1. Fuzzy clustering of the output
2. Determination of the significance of the input variables
3. Input membership assignment
4. Inference

The first three steps are the structure identification parts and the fourth step is the reasoning part. Let's now explain briefly each step of the proposed algorithm.

2.1. Fuzzy Clustering of the output

The proposed fuzzy clustering algorithm, first determines the cluster centers (v_i) by applying any clustering algorithm in the literature, and sorting the cluster centers in ascending order such as $v_{[i]} < v_{[i+1]}$ and determine the membership degree of output value of the k^{th} data point (y_k) in the i^{th} cluster (B_i) and in $i+1^{th}$ cluster (B_{i+1}) is as follows,

if $v_{[i]} \leq y_k \leq v_{[i+1]}$

$$\mu_{(Bi)}(y_k) = \|y_k - v_{[i+1]}\| / \| v_{[i]} - v_{[i+1]}\|$$
$$\mu_{(Bi+1)}(y_k) = \|y_k - v_{[i]}\| / \| v_{[i]} - v_{[i+1]}\|$$

otherwise, $\mu_{(Bi)}(y_k) = \mu_{(Bi+1)}(y_k)=0$

With this methodology the output space is clustered into triangular fuzzy sets where each output variable is assigned a total membership degree of one in two consecutive fuzzy clusters. Number of clusters are determined such that be number of rules that minimizes the modeling error of the training data set.

2.2. Input membership assignment

Next step is the determination of the significance degrees of the input variables. However since the proposed approach requires to build the rule base for this purpose, we will first discuss the proposed input membership assignment methodology. The input membership degrees are obtained by first projection of the output membership degrees onto NV-dimensional input space. Then they are assigned the input membership degrees for each input data vector X_k. This is achieved by setting $\mu_{Ai}(X_k)= \mu_{Bi}(y_k)$ for the i^{th} rule. Hence a fuzzy rule base as follows is developed,

Rule 1: IF X isr A_1 THEN Y isr B_1

...

Rule c: IF X isr A_c THEN Y isr B_c

Note that the main difference in this rule base is the antecedents are not "multi inputs" but "multi dimensional inputs".

2.3. Determination of the significance of the input variables

In the proposed approach the input variable selection concept is fuzzified and a weight is associated with each input variable. The proposed algorithm is a local heuristic search with the objective of minimizing the modeling error. For this purpose a probabilistic hill climbing algorithm is used.

Initially all the weights of the input variables are set to be same and assigned $1/NV$. In each iteration of the algorithm, weight of each input variable is increased with a small ε and the weights associated with the remaining $NV-1$ input variables are decreased with $\varepsilon/(NV-1)$ and the modeling error is calculated. Modeling error is determined by inferring an output for each individual data in the training data set from a model developed by excluding it and repeating this methodology for each data in the training set. The modeling error is the sum of each prediction error obtained for each data in the training set. This is repeated *NV* times until all of the input variables' weights are increased with ε once and the corresponding modeling error calculated. In the next step among the best candidates, i.e., among the increases that produced smallest modeling error, select one of the input variables randomly and update the significance vector associated with the input variables with respect to the selection. Repeat the above described steps until a termination criteria is satisfied. In this paper, the termination criteria used is setting a maximum number of iterations. Recall that the number of clusters is selected with respect to the model error. Hence, the above algorithm is iterated for each possible cluster size.

2.4. Inference

The first three sections described is the setup stage for system modeling where the fuzzy if-then rule structure for the data is identified. The final phase of fuzzy system modeling is the reasoning where new information is inferred given the fuzzy if-then rules and a premise (test data) where only the input variables are known and the corresponding output is the point of interest.

The first step in the proposed reasoning algorithm is the determination of the degree of matching with each rule, $\tau_i(X')$, and next to infer a crisp output with the proposed schema.

Note that since the fuzzy if-then rules are not "multi inputs" any more, we cannot use the existing methodologies in order to determine the degree of firing of the i^{th} rule for the test data X' ($\tau_i(X')$).

A *k-nearest neighborhood algorithm* is proposed in order to determine the membership degree of the given input in the *NV*-dimensional antecedent input cluster. Basically the degree of matching is estimated by first determining the k^*-*nn* training data vectors to the test data, based on the weighted Euclidean distance measure where the weight of each dimension is the significance degree associated with the corresponding input variable. After the k^*-*nn* training data is determined the degree of firing, a vector $\mu_A(X')$ ($[\tau_1(X')\ldots\tau_c(X')]$) is determined as follows,

$$\mu_A(X') = [(\Sigma_{k=1..k^*} \mu_{A1}(X^{[k]}))/k^*, \ldots, (\Sigma_{k=1..k^*} \mu_{Ac}(X^{[k]}))/k^*)]$$

where $X^{[k]}$ denotes the closest k^{th} training data vector, and each dimension of the vector corresponds to the degree of firing of a different rule. Since the proposed output clustering approach constraints the belongingness to only two consecutive clusters for the training data, we propose to constraint the test data as well and select the two consecutive clusters that have maximum total degree of firing. Finally the degrees of firing of these consecutive two clusters are determined by weighted average of each clusters degree of firing.
Final step is the inference where a crisp output is calculated. The proposed inference schema is similar to the position gradient methodology proposed by Sugeno-Yasukawa [4]. The crisp output y` is obtained as follows,

$$y' = \Sigma_{i=1..c} (\tau_i(X') \times v_{[i]})$$

where $v_{[i]}$ is the cluster center of the i^{th} rule and $\tau_i(X')$ is the degree of match (or firing) of the test data, i.e., estimated membership degree of X' to the *n*-dimensional cluster of A_i, antecedent of the i^{th} rule.

3. Experimental Analysis

The proposed algorithm is applied to the modeling of a daily price of a stock in a stock market. This data is a benchmark data provided by Sugeno-Yasukawa [4]. There is 100 data vectors and 10 input variables. The input variables are,

x_1 : past change of moving average (1) over a middle period
x_2 : present change of moving average (1) over a middle period
x_3 : past separation ratio(1) with respect to moving average over a middle period
x_4 : present separation ratio(1) with respect to moving average over a middle period
x_5 : present change of moving average(2) over a short period
x_6 : past change of price(1), for instance, change on one day before
x_7 : present change of price(1)
x_8 : past separation ratio(2) with respect to moving average over a short period
x_9 : present change of moving average(3) over a long period

x_{10} : present separation ratio(3) with respect to moving average over a short period
y : prediction of stock price

Nakanishi et. al. [2] applied six different reasoning methodologies to the data set. These reasoning algorithms are namely Sugeno (P) which is the position type reasoning, Sugeno (PG) which is the position gradient type reasoning, Sugeno (P&PG) where either P or PG is used based on the maximum membership values of the rules, Mamdani type reasoning, Turksen's point valued approximate analogical reasoning (PVAAR) and Turksen's interval valued approximate analogical reasoning (IVS). The proposed algorithm is compared with these algorithms as well as the Turksen-Bazoon approach [3] which is an improved version of Sugeno-Yasukawa algorithm, based on the RMSE of the prediction. For the comparison the same 50 input data is used as training set and the remaining 50 data is used as the test data with Nakanishi et. al. [2].

In Table 1, the significance degrees are shown as they are determined by the proposed algorithm. In Table 2, we present the comparison of the algorithms in terms of RMSE. First six rows of Table 2 is borrowed from Nakanishi et. al. [2] The actual *vs.* predicted daily stock price for the test data is depicted in Figure 1.

Table 1. The significance degrees determined by the proposed algorithm

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
.21	.03	.10	.19	0	0	.09	.01	.03	.33

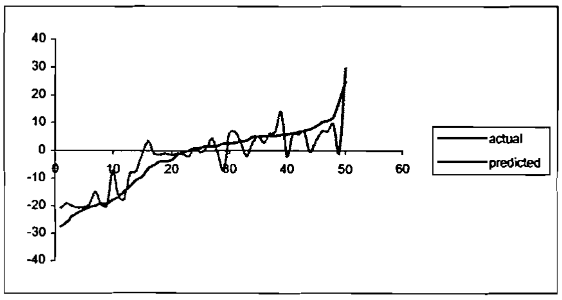

Figure 1. The actual daily stock prices *vs.* the predicted daily stock prices obtained by the proposed algorithm

In Table 2., it is shown that the proposed approach provides the better performance for this data set in terms of the root mean square error of the prediction. Among the remaining modeling approaches Turksen's PVAAR provides the best predictive performance.

Table 2. The comparison of the RMSE of the predictions of the daily stock prices.

Sugeno (P)	8.12
Sugeno (PG)	13.0
Sugeno (P&PG)	9.70
Mamdani	6.40
Turksen (PVAAR)	5.99
Turksen (IVAAR)	6.40
Turksen-Bazoon	6.38
Proposed approach	5.01

4. Conclusion

In this paper, we proposed a new structure identification approach that is based on projection onto *n*-dimensional input space rather than projecting onto each input variable separately. Also a new approach for the significance degrees of the input variables is developed. Rather than specifying the input variables as either significant or insignificant, the new approach fuzzifies this concept and assigns a relative significance to each input dimension. Furthermore the cluster validity problem is answered with the most suitable solution, which is based on the minimization of the modeling error. A new clustering algorithm that solves some problems of FCM is provided. Finally an inference schema is developed that is suitable for the structure identification that is proposed. The proposed algorithm is successfully applied to prediction of daily stock price.

References

[1] K. Kilic, B.A. Sproule, I.B. Turksen, C.A. Naranjo, Fuzzy system modeling in pharmacology: A New algorithm in *Fuzzy Sets and System*, (accepted 2001)

[2] H. Nakanishi, I.B. Turksen, M. Sugeno, A review and comparison of six reasoning methods in *Fuzzy Sets and Systems*, Volume 57, pages 257-294, 1993

[3] B.A. Sproule, M. Bazoon, K.I. Shulman, I.B. Turksen, C.A. Naranjo, Fuzzy logic pharmacokinetic modeling: Application to lithium concentration prediction in *Clinical Pharmacology Therapy*, Volume 62, pages 29-40, 1997

[4] M. Sugeno, T.A. Yasukawa, A Fuzzy Logic Based Approach to Qualitative Modeling, in *IEEE Transactions on Fuzzy Systems*, Volume 31, pages 7-31, 1993