Identification of Enhanced Fuzzy Models with Multidimensional Fuzzy Membership Functions

Özer Ciftcioglu, I.S. Sariyildiz
Delft University of Technology, Faculty of Architecture
Building Technology, 2628 CR Delft, The Netherlands
o.ciftcioglu@bk.tudelft.nl; Tel:31-15 2784485; Fax:31-15 2784127

Abstract – Identification of enhanced data-driven fuzzy models with multidimensional fuzzy membership functions is considered. The enhancement of fuzzy modeling is important since it can be used for modeling nonlinear, uncertain, and complex systems. For this aim a number of proposals appear in the literature. In these works some of the general concerns are local models, global models or the combination of both. The common approach in these works is the use of decomposed fuzzy membership functions from the multidimensional membership functions counterpart. However, because of such decomposition, inevitably a decomposition error creeps in the model development process thereby degrading the model performance. To avoid this error, it is desirable to work with multidimensional membership functions directly, to determine the final fuzzy model outcome. The motivation of this research is due to fuzzy modeling of architectural design data for the development of intelligent design processes where the data are multidimensional and highly nonlinear. From this starting point, this research deals with multidimensional membership functions to form a fuzzy model where the membership functions are modeled by radial basis functions (RBF) network. Comparisons are made with the results presented in the literature and the enhanced fuzzy modeling of this approach is demonstrated.

I. INTRODUCTION

The selection of a set of important fuzzy rules from a given rule base is an important decision-making process for effective fuzzy-rule-based modeling. In a fuzzy model a balance between reducing the fitting error and increasing the model complexity is essential for a satisfactory model that is both accurate and transparent. Since one of the important motivations using fuzzy models is to gain insight into the local behavior of the model, transparency should have a predominant consideration. The article focuses on transparency, compact and accurate fuzzy rule-based model from observation data. Generally, in data driven fuzzy modeling approaches, the Takagi-Sugeno (TS)-type fuzzy model is used [1]. The typical identification of the TS model is accomplished in two steps. In the first step, the fuzzy rule antecedents are determined. In the second step the least squares parameter estimation is applied to determine the consequents. Once the premise has been determined, the consequent parameters of the rules can be obtained as a least-square estimate. One can either carry out the estimate by solving independent weighted least-square problems, one for each rule, or solve for the rules altogether simultaneously. The first method is called local approach, which gives more transparent local models, while the second method is called a global approach, which gives a minimal model error estimate [2]. To maintain model accuracy and interpretability simultaneously, the right number of membership functions with right shapes should be located at the right locations. For single-input-single output (SISO) systems, transparency and model accuracy is relatively easy to maintain [3]. However, in multivariable case (MISO), the situation changes rapidly, due to the curse of dimensionality of the number of fuzzy membership functions, which means increased number of model parameters to be estimated. To maintain the transparency fuzzy modeling one can make use of evolutionary algorithms due to multi-objective optimization properties of the method [4].

After the general observations above having been made, another issue in data-driven modeling is the size of the data samples. Namely, data-driven fuzzy modeling is more interesting compared to neural networks due to its interpretability of the fuzzy membership functions. Otherwise, if there is enough data able to train a neural network, then neural network is a competitive alternative to fuzzy modeling in case the transparency is not a marginal issue. In such cases, universal approximation property of fuzzy modeling is considered apart from the interpretability properties of the model. However, in fuzzy modeling, the multivariable membership functions provide rich aggregated information of data samples. Using this information directly one obtains relatively more accurate results with relatively less number of data samples without any sacrifice in transparency. At this very point, fuzzy modeling gain due attention relative to neural network models. In this research, for the desired enhancement of fuzzy model, the multidimensional membership functions are modeled by means of a neural network of the radial basis functions (RBF) type. Hence, the outcome of a multivariable fuzzy model is based on this compact fuzzy model. For the transparency, one of the conventional decomposition schemes can be used. That is, for the sake of minimized model prediction error by the present approach, transparency is not diminished as trade-off, as this is explained later in the article. The motivation of this research is due to fuzzy modeling of architectural design data for the development of intelligent design processes where the data are multidimensional and highly nonlinear. Even in a simple case, if the right parameters are not given, how
transparency of fuzzy modeling degrades is pointed out before [5]. Therefore the conventional fuzzy modeling approach in a multidimensional case may easily be not accurate enough and model enhancement is a most desirable issue for such advanced modeling tasks. This article aim for this. The organization of the paper is as follows. Section two gives brief descriptions of TS fuzzy modeling, RBF network, OLS training algorithm for the RBF network and modeling of the multivariable fuzzy membership functions by the RBF network. Section three presents the comparative results obtained using different multivariable data including one from the literature. This is followed by the conclusions in section four.

II. FUZZY MODELING

A. Takagi-Sugeno Fuzzy Modeling

Takagi-Sugeno type fuzzy modeling consists of set of fuzzy rules a local input-output relation in a linear form as

\[ R_i : x_i = A_{i1} \text{ and } \ldots \text{and } x_n = A_{in} \]

where \( R_i \) is the ith rule, \( x = [x_1, x_2, \ldots, x_n]^T \)

Then \( \hat{y}_i = a_i x + b_i \), \( i = 1,2,\ldots, K \)

where \( \hat{y}_i \) is the ith rule, \( x=[x_1,x_2,\ldots,x_n]^T \in \mathbb{R} \) is the vector of input variables; \( A_{i1}, A_{i2},\ldots,A_{in} \) are fuzzy sets and \( y_i \) is the rule output; \( K \) is the number of rules. The output of the model is calculated through the weighted average of the rule consequents of the form

\[ \hat{y} = \frac{\sum_{i=1}^{K} \beta_i(x) \hat{y}_i}{\sum_{i=1}^{K} \beta_i(x)} \]

In (2), \( \beta_i(x) \) is the degree of activation of the ith rule

\[ \beta_i(x) = \prod_{j=1}^{n} \mu_{A_{ij}}(x_j), \quad i = 1,2,\ldots,K \]

where \( \mu_{A_{ij}}(x_j) \) is the membership function of the fuzzy set \( A_{ij} \) at the input (antecedent) of \( R_i \).

To form the fuzzy system model from the data set with \( N \) data samples, given by

\[ X=[x_1, x_2, \ldots, x_N]^T, \quad Y=[y_1, y_2, \ldots, y_N]^T \]

where each data sample has a dimension of \( n \) (\( N \gg n \)). First the structure is determined and afterwards the parameters of the structure are identified. The number of rules characterizes the structure of a fuzzy system. The number of rules is determined by clustering methods. Fuzzy clustering in the Cartesian product-space \( \mathbb{R}^n \times \mathbb{R}^n \) is applied to partition the training data. The partitions correspond to the characteristic regions where the system's behaviour is approximated by local linear models in the multidimensional space.

Given the training data \( \Gamma \) and the number of clusters \( K \), a suitable clustering algorithm [5] is applied. One of such clustering algorithms is known as Gustafson-Kessel (GK) [7]. As result of the clustering process a fuzzy partition matrix \( U \) is obtained. The fuzzy sets in the antecedent of the rules are identified by means of the partition matrix \( U \) which has dimensions \([N \times K]\); \( N \) is the size of the data set. The \( ik \)-th element of \( \mu_{ik} \in [0,1] \) is the membership degree of the i-th data item in cluster \( k \); that is, the ith row of \( U \) contains the point wise description of a multidimensional fuzzy set. One-dimensional fuzzy sets \( A_{ij} \) are obtained from the multidimensional fuzzy sets by projections onto the space of the input variables \( x_i \). This is expressed by the point-wise projection operator of the form

\[ \mu_{A_{ij}}(x_i) = \text{proj}_j(\mu_{A_{ij}}) \]

The point-wise defined fuzzy sets \( A_{ij} \) are then approximated by appropriate parametric functions. The consequent parameters for each rule are obtained by means of linear least square estimation. For this, consider the matrices

\[ X=[x_1, \ldots, x_N]^T, \quad X=[X,1] \text{ (extended matrix \([N \times (n+1)]\) ) ; } A_i \text{ (diagonal matrix dimension of \([N \times N]\) ) and} \]

\[ X_E=[(A_1x_1); (A_2x_2); \ldots; (A_Kx_K)] \text{ \([N \times K(n+1)]\) ) (6) \]

where the diagonal matrix \( A_i \) consists of normalized membership degree as its \( k \)-th diagonal element

\[ \hat{\beta}_i(x_i) = \frac{\beta_i(x_i)}{\sum_{j=1}^{K} \beta_j(x_i)} \]

The parameter vector \( \Theta \) dimension of \([K \times (n+1)]\) is given by

\[ \Theta=[\Theta_1^T, \Theta_2^T, \ldots, \Theta_K^T]^T \]

where \( \Theta_i=\left[ a_i^T, b_i \right] \quad (1 \leq i \leq K) \).

Now, if we denote the input and output data sets as \( X_E \) and \( Y \) respectively then, the fuzzy system can be represented as a regression model of the matrix form

\[ Y=X_E \Theta + e \]

B. Radial Basis Function Network

Considering the function approximation by the radial basis functions \( \phi(x,c) \), an approximation to a function \( f(x) \) by radial basis function network is carried out by

\[ y = f(x) = \sum_{i=1}^{M} w_i \phi(||x - c_j||) + e \]

where \( w_i \) are weights; \( M \) is the number of basis functions; \( x \) is the sample vector; \( c_j \) is the RBF center vector; \( e \) is the model error; \( \phi(x,c) \) is the basis function generally based on the Euclidean distance metric defined by

\[ \|x - c_j\|^2 = (x - c_j)^T(x - c_j) \]

so that

\[ \phi = \exp\{-\|x - c_j\|^2 / \sigma_j^2\} \]
where $\sigma_j$ is the j-th width parameter that determines the effective support of the j-th basis function. The symmetric matrix $\Phi$ formed by the elements
\[
\phi_j = \exp\{-\|x_i - c_j\|^2 / \sigma_j^2\}
\]
is referred to as RBF matrix which is used to train the network. For a set of input and output pairs, the approximation model in matrix form can be expressed as
\[
y = \Phi w + e
\]
where $y$ is the desired output vector which is $[N \times q]$ in the multi-output case where q is the number of outputs; $\Phi$ is the regression matrix that consists of regressor vectors and plays the role of $X_E$ in (9), after normalisation similar to that given in (7); $w$ is the parameter vector or output matrix of dimensions $M \times q$ in the multi-output case; $e$ is the error vector.

C. The Orthogonal Least Squares Method

The orthogonal least squares (OLS) method [8,9] makes orthogonal decomposition of $\Phi$. This is generally accomplished by Gram-Schmidt’s orthogonalisation procedure, as described for a single output RBF network:
\[
\Phi = RA
\]
where A is an upper-diagonal matrix and R is an $N \times M$ matrix with orthogonal columns $r_i$ ($i=1,...,M$) such that
\[
R^T R = H, \quad H = \begin{bmatrix} h_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & h_M \end{bmatrix}
\]
$H$ is $M \times M$ diagonal matrix. With the definitions
\[
\hat{R} = \Phi A^{-1}
\]
and
\[
w_R = A w
\]
The matrix equation (13) in this case takes the form
\[
y = \hat{R} w_R + e
\]
which has the solution for the estimation of the transformed parameter vector $w_R$ as
\[
\hat{w}_R = H^{-1} R^T y \quad \text{or,} \quad \hat{w}_R = r_i^T y / r_i^T r_i, \quad i=1,...,M
\]
By means of orthogonalisation process the traditional Gram-Schmidt computes one column of $A$ at a time and orthogonalizes $R$ at the same time. In terms of energy transmitted from input to the output, we can write the following algebraic balance equation for zero-mean output vector:
\[
y = \sum_{j=1}^{M} w_j r_j^T r_i + e^T e
\]
Normalizing by $D^T D$ we obtain the relative energy contribution from each basis function, i.e. from each regressor to the output as
\[
z_i = \frac{w_i^T r_i^T r_i}{y^T y} \quad (i=1,...,M)
\]
which is defined as error reduction ratio.

D. Multivariable Fuzzy Membership Functions Modeled by RBF Network

As result of clustering process, the multivariable fuzzy membership functions (MF) are obtained by point wise projections in the multidimensional space, as described in subsection 2.A, above. The clustering method can be supplemented with a cluster merging algorithm [10], so that appropriate number of clusters with associated multidimensional membership functions are obtained. Note that, the redundancy in the number of membership functions may cause serious detrimental effects on the model being developed. As result of clustering process, next to the fuzzy partition matrix, also the cluster means matrix, cluster covariance matrices, the eigenvalues, and the eigenvectors of the covariance matrices are obtained. From these also the TS model parameters are obtained. Namely, considering the case of two time-discrete input signals $x_1(k)$ and $x_2(k)$ and one time-discrete output signal $y(k)$ representing the system output, TS fuzzy models are characterized by rule statements of the general form
\[
\text{IF } x_1(k) \text{ is } A^{(i)} \text{ AND } x_2(k) \text{ is } B^{(i)} \quad \text{THEN} \\
\hat{y}(k) = a^{(i)} x_1(k) + b^{(i)} x_2(k) + c^{(i)}
\]
where $i, j, r = 1,...,n_r$

Above, $A^{(i)}$ and $B^{(i)}$ represents fuzzy sets of the input variables $x_1$ and $x_2$, and $a^{(i)}, b^{(i)}$ and $c^{(i)}$ are the coefficients of the output models. $n_r$ is the number of clusters.

As a novel approach in this research, first, TS fuzzy model by projections is established for transparency by means of global approach. Second RBF network is established that models the multivariable membership functions. The training of the RBF network is done by OLS algorithm where the fuzzy model inputs as data samples are provided at the RBF network input and the corresponding membership degrees of these samples are provided at the RBF network output. For example, for three model variables $x=[x_1 x_2 x_3]$ and two clusters, the RBF network has three inputs and two outputs.

Based on the membership degree information the local polynomial information, i.e., $a^{(i)}, b^{(i)}$ and $c^{(i)}$ obtained from the result of GK clustering. Namely, the local linear model coefficients can be obtained by using the cluster centers and the eigenvectors of the covariance matrices information thereby identifying the fuzzy model. Alternatively, the model parameters can be obtained using multidimensional fuzzy membership information and the method of global least squares.

III. SIMULATION RESULTS AND COMPARISON
Investigations are carried out by means of simulations taken from the literature so that comparison with the results obtained from different fuzzy modeling approaches is made possible. The simulation data set subject to fuzzy modeling belongs to a MISO system with two inputs and one output [11]. The system is given by

\[ y(t) = F(x_1(t), x_2(t)) \]

with the a priori unknown nonlinear mapping

\[ F(x_1, x_2) = x_1^2 x_2 \]

to be modeled. After providing the system with the harmonic input signals

\[
x_1(t) = 1 + \cos(2\pi t) \\
\]
\[
x_2(t) = \sin\left(\frac{\pi}{10} t\right)
\]

for the dimensionless time range \(0 \leq t \leq 10\), the \(N=105\) time discretely measured data records with sampling time interval \(0.095\). The input and the output values of the system constitute the basis of the fuzzy modeling procedure. Application of the GK fuzzy clustering algorithm equipped with the cluster merging procedure yielded \(c=2\) identified clusters. The system-specific surface formed by the data of the input product space and the output space, as well as the positions of the identified cluster centers, are illustrated in figure 1.

The fuzzy MFs obtained by means of pointwise projection and with smooth gaussian-shaped functions are shown in Fig.2. The multivariable fuzzy modeling outcome based on the these MFs is shown in Fig.3 where uppermost plot is the model output together with the given data set of 105 samples. Two plots almost coincide. The middle plot represent the model output with the twice as much reduced sampling interval thereby amounting to 210 data samples. The model performance is reasonably fine while the situation is nearly the same as presented for 105 data samples above it.

In a further step, the fuzzy model with the normal rule base is provided with input signals that differ from those used for the procedure of identification of the fuzzy model. The new signals are applied, differing from the former input signals \(x_1(t)\) and \(x_2(t)\) in frequency and magnitude, respectively

\[
x_1(t) = 1 + \cos(\pi t) \\
x_2(t) = \frac{3}{4}\sin\left(\frac{\pi}{10} t\right)
\]

The simulation results in figure 4 (lowermost) illustrate that the fuzzy model with the conventional, normal rule base
with two clusters still shows reasonably fine results although the input signals differ from those used to identify the fuzzy model. This can be accepted already as desired degree of universality of the fuzzy model without any further elaboration, for instance, increasing the number of clusters, ad hoc shaping MF etc. The multidimensional fuzzy membership function approach for fuzzy modeling, as described above as a novel feature of this research gave also similar results, i.e. the same quality of modeling results as presented in figure 3. The study above can be considered as conventional fuzzy modeling for comparison with the results reported in literature [11] based on the same data set different modeling approach. However, in the above example the system-specific surface of defined by the data is only slightly nonlinear. In contrast with this, a relatively more nonlinear surface is considered for the same modeling processes using multidimensional MF as a novel approach being presented in this article. These two different system-specific surfaces are shown together in figure 4, where lower plot corresponds to the former case.

Fig.4. System-specific surfaces considered in this work.

The resulting model outcomes are presented in figure 5 as the counterpart of figure 1. As seen from figure 5, the number of data points is 100 and they form a grid and the number of clusters is six. The conventional fuzzy model outcomes with the projected fuzzy sets as to each model variable are shown in figure 6. The fuzzy model is tested with a grid of input data samples 14×14, that is altogether 196 data samples which do not include the data samples used for fuzzy modeling. The identified surface is considerably distorted with respect to the true surface subject to identification; namely the upper plot in figure 4. The model outcome for each data sample together with the true model output are represented in the lower plot in the same figure. The gross differences resulting in gross distortions in 3D representation are clear to observe. With less number of clusters, the situation becomes worse.

Fig.5: System specific surface and the location of the identified surface

Fig.6: Fuzzy model outcomes of model testing process with projected fuzzy sets as to each model input variable.

This basic investigation reveals that, the conventional multivariable fuzzy modeling in the case of highly nonlinear model surface, needs more attention than those used for simple cases, i.e., almost linear system-specific surfaces. This can be implemented by fuzzy modeling using the multidimensional fuzzy sets directly. In this case each fuzzy set form a multidimensional function and such a function is difficult to identify in an analytical form. Therefore, in general this situation poses a problem to use multidimensional fuzzy sets in fuzzy modeling. Decomposing the multidimensional fuzzy MF with respect to the model input variables conventionally circumvents this. At the same time, such decomposition is justified due to the interpretability issues of fuzzy modeling; that is, fuzzy
modeling is generally praised with the interpretability of the model with respect to model variables. This property is termed as transparency of the model. However, as noted from figure 7, the transparency is badly traded off by the model accuracy. The basic solution to the multivariable function representation can be obtained by means of an RBF network, since such a network has desirable interpolating properties perfectly matching the present case as described elsewhere [12]. The multidimensional MFs can altogether be modeled by an RBF network especially using OLS algorithm for training. By using the same number of hidden layer nodes as the data samples, the perfect representation of the data set following the training is guaranteed. Such a model is dense in the sense that it provides a continuous functional representation in a multidimensional space. This is a desirable property for interpolation exactly fitting the present goals. The results obtained by using multidimensional MF directly are presented in Fig.7 as a counterpart of figure 6. Namely, the model is tested by means of the same data set which does not include the data samples used for fuzzy model forming.

A marked improvement is observed in figure 7 relative to that shown in figure 6. The improvement still becomes better while employing more cluster centers, thereby introducing more fuzzy sets in the model.

IV. CONCLUSIONS

Fuzzy modeling with multidimensional fuzzy sets is presented. Although conventional fuzzy modeling with decomposed fuzzy sets gives satisfactory performance for simple cases, for highly nonlinear surfaces in a multidimensional space, the same modeling process degrades and the approach needs to be enhanced. Such desired enhancement is accomplished in this work by means of integrating an RBF network to the modeling process. The enhancement achieved is demonstrated. The basic explanation of this modeling degradation mentioned above is due to irrecoverable information loss due to decomposition process of the multidimensional fuzzy sets. That is, due to the projection made onto each variable. As result of this, this information loss is reflected on the model outcome as systematic modeling error. This error is termed as reconstruction error. By working multidimensional fuzzy sets, such systematic error is eliminated and thereby the model performance enhancement. The price paid for this gain is some transparency loss of the model. To regain the transparency, still the decomposition of the multidimensional sets can be carried out and this information can be used for transparency although this is an isolated action playing no role in the modeling approach presented above.

V. REFERENCES