A new approach to fuzzy estimation of Takagi–Sugeno model and its applications to optimal control for nonlinear systems

Basil M. Al-Hadithi, Agustín Jiménez , Fernando Matía

A B S T R A C T

An efficient approach is presented to improve the local and global approximation and modelling capability of Takagi–Sugeno (T–S) fuzzy model. The main aim is obtaining high function approximation accuracy. The main problem is that T–S identification method cannot be applied when the membership functions are overlapped by pairs. This restricts the use of the T–S method because this type of membership function has been widely used during the last two decades in the stability, controller design and are popular in industrial control applications. The approach developed here can be considered as a generalized version of T–S method with optimized performance in approximating nonlinear functions. A simple approach with few computational effort, based on the well known parameters' weighting method is suggested for tuning T–S parameters to improve the choice of the performance index and minimize it. A global fuzzy controller (FC) based Linear Quadratic Regulator (LQR) is proposed in order to show the effectiveness of the estimation method developed here in control applications. Illustrative examples of an inverted pendulum and Van der Pol system are chosen to evaluate the robustness and remarkable performance of the proposed method and the high accuracy obtained in approximating nonlinear and unstable systems locally and globally in comparison with the original T–S model. Simulation results indicate the potential, simplicity and generality of the algorithm.

1. Introduction

Nonlinear control systems based on the T–S fuzzy model [35–37] have attracted lots of attention during the last twenty years (e.g., see [5,8,13–16,18,20,28,29,31,39,44]. It provides a powerful solution for development of function approximation, systematic techniques to stability analysis and controller design of fuzzy control systems in view of fruitful conventional control theory and techniques. They also allow relatively easy application of powerful learning techniques for their identification from data [11]. This model is formed using a set of fuzzy rules to represent a nonlinear system as a set of local affine models which are connected by fuzzy membership functions [6]. This fuzzy modelling method presents an alternative technique to represent complex nonlinear systems [12,39,46,48] and reduces the number of rules in modelling higher order nonlinear systems [13,37]. T–S fuzzy models are proved to be universal function approximators as they are able to approximate any smooth nonlinear functions to any degree of accuracy in any convex compact region [12,22,28,39,46,48]. This result provides a theoretical foundation for applying T–S fuzzy models to represent complex nonlinear systems approximately.

Great attention has been paid to the identification of T–S fuzzy models and several results have been obtained [7,22,30,40,47]. They are based upon two kinds of approaches, one is to linearize the original nonlinear system in various operating points when the model of the system is known, and the other is based on the input–output data collected from the original nonlinear system when its model is unknown. The authors in [7] use a fuzzy clustering method to identify T–S fuzzy models, including identification of the number of fuzzy rules and parameters of fuzzy membership functions, and identification of parameters of local linear models by using a least squares method [33,43]. The goal is to minimize the error between T–S fuzzy models and the corresponding original nonlinear systems. In [24], Klawonn et al. explained how fuzzy clustering techniques could be applied to design a fuzzy controller from the training data. In [17], Hong and Lee have analyzed that the disadvantages of most fuzzy systems are that the membership functions and fuzzy rules should be predefined to map numerical data into linguistic terms and to make fuzzy reasoning work. They suggested a method based on the fuzzy clustering technique and the decision tables to derive membership functions and fuzzy rules from numerical data. However, Hong and Lee's algorithm presented in [17] needs to predefine the membership functions of the input

linguistic variables and it simplifies fuzzy rules by a series of merge operations. As the number of variables becomes larger, the decision table will grow tremendously and the process of the rule simplification based on the decision tables becomes more complicated. The authors in [22] suggest a method to identify T–S fuzzy models. Their method aims at improving the local and global approximation of T–S model. However, this complicates the approximation in order to obtain both targets. It has been shown that constrained and regularized identification methods may improve interpretability of constituent local models as local linearizations, and locally weighted least squares method may explicitly address the trade-off between the local and global accuracy of T–S fuzzy models.

In [33] a new method of interval fuzzy model identification was developed. The method combines a fuzzy identification methodology with some ideas from linear programming theory. The idea is then extended to modelling the optimal lower and upper bound functions that define the band which contains all the measurement values. This results in lower and upper fuzzy models or a fuzzy model with a set of lower and upper parameters. This approach can also be used to compress information in the case of large amount of data and in the case of robust system identification. The method can be efficiently used in the case of the approximation of the nonlinear functions family. The paper focuses on the development of an interval L-norm function approximation methodology problem using the LP technique and the TS fuzzy logic approach. This results in lower and upper fuzzy models or a fuzzy model with lower and upper parameters.

In [29] a constructive method to synthesize a MISO TS fuzzy logic system imposing the requested derivative constraints on the function representing its behavior is presented. The values of that function and its partial derivatives on the grid points of the input space permit to define a suitable interpolator of the function itself.

In [23], a new approach to fuzzy modelling using the relevance vector learning mechanism (RVM) based on a kernel-based Bayesian estimation is introduced. The main concern is to find the best structure of the TS fuzzy model for modelling nonlinear dynamic systems with measurement error. The number of rules and the parameter values of membership functions can be found as optimizing the marginal likelihood of the RVM in the proposed FIS. Because the RVM is not necessary to satisfy Mercer's condition, selection of kernel function is beyond the limit of the positive definite continuous symmetric function of SVM. The relaxed condition of kernel function can satisfy various types of membership functions in fuzzy model. The RVM which was compared with support vector learning mechanism in examples had the small model capacity and described good generalization. Simulated results showed the effectiveness of the proposed FIS for modelling of nonlinear dynamic systems with noise.

In [45] a singular value-based method for reducing a given fuzzy rule set was discussed. The method conducts singular value decomposition of the rule consequents and generates certain linear combinations of the original membership functions to form new ones for the reduced set. The work characterizes membership functions by the conditions of sum normalization (SN), nonnegativeness (NN), and normality (NO). Algorithms to preserve the SN and NN conditions in the new membership functions were presented. Preservation of the NO condition relates to a high-dimensional convex hull problem and is not always feasible in which case a closed-to-NO solution may be sought. The proposed method is applicable regardless of the adopted inference paradigms. With product-sum-gravity inference and singleton support fuzzy rule base, output errors between the full and reduced fuzzy set are bounded by the sum of the discarded singular values. The present work discusses three specific applications of fuzzy reduction: fuzzy rule base with singleton support, fuzzy rule base with nonsingleton support (which includes the case of missing rules), and the T–S

model. Numerical examples are presented to illustrate the reduction process.

Jang [19] proposed a fuzzy neural network model – the Adaptive Neural Fuzzy Inference System or semantically equivalently, Adaptive Network-based Fuzzy Inference System (ANFIS). It is one of the most popular algorithms to integrate the best features of fuzzy systems and neural networks. He reported that the ANFIS architecture can be employed to model nonlinear functions, identify and estimate nonlinear components on-line in a control system, and predict a chaotic time series. It is a hybrid neuro-fuzzy technique that brings learning capabilities of neural networks to fuzzy inference systems. The learning algorithm tunes the membership functions of a Sugeno-type Fuzzy Inference System using the training input/output data using. Its learning method is based on gradient descent and least square estimation [1]. The limitation of ANFIS is that it cannot be applied to estimate fuzzy systems when the membership functions are overlapped by pairs.

In [37], the authors develop an interesting method to identify nonlinear systems using input–output data. They divide the identification process in three steps; premise variables, membership functions and consequent parameters. With respect to membership functions, they apply nonlinear programming technique using the complex method for the minimization of the performance index.

In spite of such important works, prior to 1989, most FLS's were still designed with preselected structures and the adjustment of membership functions (MF's) was carried out by trial and error. Since 1989, a number of techniques for structure and/or parameter identification from I/O data have been suggested in the literature. In 1989, Sugeno and Tanaka presented a new method for a successive identification method for constructing a T–S fuzzy model [33]. In this method, structure determination was implemented by a combination of least-squares estimation (LSE), the complex method, and an unbiasedness criterion; while parameter adjustment was achieved using fuzzy adjustment rules and a weighted recursive least-squares algorithm.

In 1991, Wang and Mendel developed a method for generating fuzzy rules by learning from examples [43] and proved that a fuzzy inference system is a universal approximator by the Stone–Weierstrass theorem [41].

In 1995, Wang proposed a new state-space analytical approach to fuzzy identification of nonlinear dynamical systems [42]. In 1996, Langari and Wang proposed achieving structure identification of a T–S fuzzy model using a combination of fuzzy c-means clustering technique and a fuzzy discretization technique [26].

In [32], Nozaki et al. presented a heuristic method for generating T–S fuzzy rules from numerical data, and then converted the consequent parts of T–S fuzzy rules into linguistic representation.

In [25], a study has outlined a new min–max approach to the fuzzy clustering, estimation, and identification with uncertain data. The proposed approach minimizes the worst-case effect of data uncertainties and modelling errors on estimation performance without making any statistical assumption and requiring a priori knowledge of uncertainties. Simulation studies have been provided to show the better performance of the proposed method in comparison to the standard techniques. The developed fuzzy estimation theory was applied to a real world application of physical fitness classification and modelling.

In [9] a fuzzy linear control design method is introduced for nonlinear systems with optimal H^∞ robustness performance. First, the Takagi and Sugeno fuzzy linear model is used to approximate a nonlinear system. Then, based on the fuzzy linear model, a fuzzy controller is developed to stabilize the nonlinear system, and at the same time the effect of external disturbance on control performance is reduced to a minimum level. Thus, based on the fuzzy linear model, H^∞ performance design can be achieved in nonlinear control systems. In the suggested fuzzy linear control

method, the fuzzy linear model provides rough control to approximate the nonlinear control system, while the H^∞ scheme provides an accurate control to obtain the optimal robustness performance. Linear matrix inequality (LMI) techniques are employed to solve this robust fuzzy control problem. In the case that state variables are unavailable, a fuzzy observer-based H^∞ control is also proposed to achieve a robust optimization design for nonlinear systems.

A new fuzzy system containing a dynamic rule base is proposed in [10]. The characteristic of the proposed system is in the dynamic nature of its rule base which has a fixed number of rules and allows the fuzzy sets to dynamically change or move with the inputs. The number of the rules in the proposed system can be small, and chosen by the designer. The focus of article is mainly on the approximation capability of this fuzzy system. The proposed system is capable of approximating any continuous function on an arbitrarily large compact domain. Moreover, it can even approximate any uniformly continuous function on infinite domains. This paper addresses existence conditions, and as well provides constructive sufficient conditions so that the new fuzzy system can approximate any continuous function with bounded partial derivatives.

As we will demonstrate in this article, the T–S model cannot be applied when the membership functions are overlapped by pairs. This limits the usage of the model because as it was shown in the last two decades that the major part of the results obtained in the field of stability and controller synthesis are based on this type of membership functions. Moreover, the method presented here is characterized by the high accuracy obtained in approximating nonlinear systems locally and globally in comparison with the original T–S model. The rest of the paper is organized as follows. In Section 2, the estimation of fuzzy T–S model's parameters is described. Section 3 presents the restrictions of T–S identification method. Section 4 demonstrates the proposed approach to improve and generalize the T–S model. In Section 5, identification of vectorial functions is presented. The design of a global optimal controller based on the proposed estimation method is detailed in Section 6. Section 7 entails two nonlinear examples of an inverted pendulum and Van der Pol system to demonstrate the validity of the proposed approach. The results show that the proposed approach is less conservative than those based on (standard) T–S model and illustrate the utility of the proposed approach in comparison with T–S model. The conclusions of the effectiveness and validity of the proposed approach are explained in Section 8.

2. Estimation of fuzzy T–S model's parameters

An interesting method of identification is presented in [37]. The idea is based on estimating the nonlinear system parameters minimizing a quadratic performance index. The method is based on the identification of functions of the following form:

$$f : n\Re \to \Re$$

$$y = f(x_1, x_2, \ldots, x_n)$$

Each IF–THEN rule $R^{i_1 \cdots i_n}$, for an nth order system can be written as follows:

$$S^{(i_1 \cdots i_n)} : \text{If } x_1 \text{ is } M_1^{i_1} \text{ and} \ldots x_n \text{ is } M_n^{i_n} \text{ then}$$
$$\hat{y} = p_0^{(i_1 \cdots i_n)} + p_1^{(i_1 \cdots i_n)} x_1 + p_2^{(i_1 \cdots i_n)} x_2 + \cdots + p_n^{(i_1 \cdots i_n)} x_n \tag{1}$$

where the fuzzy estimation of the output is:

$$\hat{y} = \frac{\sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} w^{(i_1 \cdots i_n)}(\mathbf{x}) \left[p_0^{(i_1 \cdots i_n)} + p_1^{(i_1 \cdots i_n)} x_1 + p_2^{(i_1 \cdots i_n)} x_2 + \cdots + p_n^{(i_1 \cdots i_n)} x_n \right]}{\sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} w^{(i_1 \cdots i_n)}(\mathbf{x})} \tag{2}$$
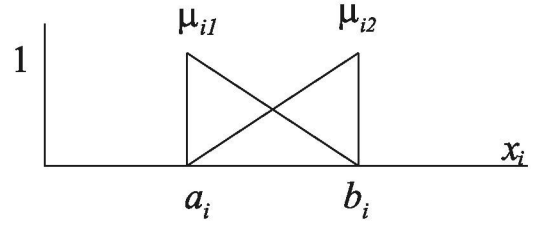

Fig. 1. Membership functions of the fuzzy system.

where

$$w^{(i_1 \cdots i_n)}(\mathbf{x}) = \mu_{1i_1}(x_1)\mu_{2i_2}(x_2) \ldots \mu_{ni_n}(x_n)$$

being $\mu_{ji_j}(x_j)$ the membership function that corresponds to the fuzzy set $M_j^{i_j}$.

Let m be a set of input/output system samples $\{x_{1k}, x_{2k}, \ldots, x_{nk}, y_k\}$. The parameters of the fuzzy system can be calculated as a result of minimizing a quadratic performance index:

$$J = \sum_{k=1}^{m} (y_k - \hat{y}_k)^2 = \left\| Y - XP \right\|^2 \tag{3}$$

where

$$
\begin{aligned}
Y &= \begin{bmatrix} y_1 & y_2 & \cdots & y_m \end{bmatrix}^t \\
P &= \begin{bmatrix} p_0^{(1 \cdots 1)} p_1^{(1 \cdots 1)} p_2^{(1 \cdots 1)} & \cdots & p_n^{(1 \cdots 1)} \cdots & p_0^{(r_1 \cdots r_n)} \cdots p_n^{(r_1 \cdots r_n)} \end{bmatrix}^t \\
X &= \begin{bmatrix} \beta_1^{(1 \cdots 1)} \beta_1^{(1 \cdots 1)} x_{11} \ldots \beta_1^{(1 \cdots 1)} x_{n1} & \cdots & \beta_1^{(r_1 \cdots r_n)} \ldots \beta_1^{(r_1 \cdots r_n)} x_{n1} \\ \beta_m^{1 \cdots 1} \beta_m^{(1 \cdots 1)} x_{1m} \ldots \beta_m^{(1 \cdots 1)} x_{nm} & \cdots & \beta_m^{(r_1 \cdots r_n)} \ldots \beta_m^{(r_1 \cdots r_n)} x_{nm} \end{bmatrix}
\end{aligned} \tag{4}
$$

and

$$\beta_k^{(i_1 \cdots i_n)} = \frac{w^{(i_1 \cdots i_n)}(\mathbf{x}_k)}{\sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} w^{(i_1 \cdots i_n)}(\mathbf{x}_k)} \tag{5}$$

If X is a matrix of full rank, the solution is obtained as follows:

$$
\begin{aligned}
J &= \left\| Y - XP \right\|^2 = (Y - XP)^t (Y - XP) \\
\nabla J &= X^t(Y - XP) = X^t Y - X^t XP = 0 \\
P &= (X^t X)^{-1} X^t Y
\end{aligned} \tag{6}
$$

3. Restrictions of T–S identification method

The method proposed in [37] arises serious problems as it cannot be applied in the most common case where the membership functions are those shown in Fig. 1

The membership functions $\mu_{i1}(x_i) = (b_i - x_i)/(b_i - a_i)$ and $\mu_{i2}(x_i) = (x_i - a_i)/(b_i - a_i)$ are defined in an interval $[a_i, b_i]$ which should verify:

$$
\begin{aligned}
\mu_{i1}(a_i) &= 1 & \mu_{i1}(b_i) &= 0 \\
\mu_{i2}(a_i) &= 0 & \mu_{i2}(b_i) &= 1 \\
\mu_{i1}(x_i) + \mu_{i2}(x_i) &= 1
\end{aligned}
$$

For this case which is widely used, it can be easily demonstrated [21] that the matrix X is not of full rank and therefore $X^t X$ is not invertible, which makes the mentioned method of T–S invalid. This result can be easily proven as follows:

Supposing that it exists:

$$f : \Re \to \Re$$

$$y = f(x)$$

in which each row of the matrix X is of the form:

$$X_k = \begin{bmatrix} \mu_1(x_k) & \mu_1(x_k)x_k & \mu_2(x_k) & \mu_2(x_k)x_k \end{bmatrix}$$
$$= \begin{bmatrix} \dfrac{b-x_k}{b-a} & \dfrac{b-x_k}{b-a}x_k & \dfrac{x_k-a}{b-a} & \dfrac{x_k-a}{b-a}x_k \end{bmatrix} \quad (7)$$

verifying that:

$$\begin{bmatrix} \dfrac{b-x_k}{b-a} & \dfrac{b-x_k}{b-a}x_k & \dfrac{x_k-a}{b-a} & \dfrac{x_k-a}{b-a}x_k \end{bmatrix} \begin{bmatrix} -a \\ 1 \\ -b \\ 1 \end{bmatrix} = 0$$

The rank of X in this case is 3. In other words, the columns of X are linearly dependent which in turn makes impossible the use of the above mentioned identification method proposed in [37]. Analyzing another example of two variables:

$$f : 2\Re \to \Re$$

$$y = f(x_1, x_2)$$

Each row of the matrix X is of the form:

$$X_k = \begin{bmatrix} \mu_{11}\mu_{21} & \mu_{11}\mu_{21}x_{1k} & \mu_{11}\mu_{21}x_{2k} & \mu_{11}\mu_{22} & \mu_{11}\mu_{22}x_{1k} & \mu_{11}\mu_{22}x_{2k} \\ \mu_{12}\mu_{21} & \mu_{12}\mu_{21}x_{1k} & \mu_{12}\mu_{21}x_{2k} & \mu_{12}\mu_{22} & \mu_{12}\mu_{22}x_{1k} & \mu_{12}\mu_{22}x_{2k} \end{bmatrix} \quad (8)$$

It can be noticed that the columns 1, 3, 4 and 6 have the same form as in the previous example multiplied by a constant μ_{11} and therefore they are linearly dependent as well. The same thing happens with the columns 6, 9, 10 and 12, etc. In fact, the rank of the matrix in this case is 8.

The solution proposed in [37] avoids the occurrence of this situation. In order to identify a function in the interval $[a_i, b_i]$ using T–S method, certain intermediate points are chosen of the form:

$$a_i^* \in [a_i, b_i] \text{ and } b_i^* \in [a_i, b_i]$$

and they use membership functions which verify:

$$\mu_{i1}(x) = \begin{cases} \dfrac{x_i - b_i^*}{a_i - b_i^*} & a_i \le x \le b_i^* \\ 0 & b_i^* \le x \le b_i \end{cases} \quad (9)$$

$$\mu_{i2}(x) = \begin{cases} 0 & a_i \le x \le a_i^* \\ \dfrac{x_i - a_i^*}{b_i - a_i^*} & a_i^* \le x \le b_i \end{cases} \quad (10)$$

and thus:

$$\mu_{i1}(a_i) = 1 \quad \mu_{i1}(b_i^*) = 0$$
$$\mu_{i2}(a_i^*) = 0 \quad \mu_{i2}(b_i) = 1$$

which impedes that the domains of these functions being overlapped and therefore it can be observed that, except for some isolated points,

$$\mu_{i1}(x_i) + \mu_{i2}(x_i) \ne 1$$

and thus, in general, the matrix X will be of full rank and the method is applicable. This solution can be clearly seen in [37] where the authors find the optimum membership functions minimizing the performance index and reducing the problem to a nonlinear programming one. For this reason, they use the well-known complex method for the minimization. This can obviously be observed in the illustrative examples selected by the authors in [37] where all the identified memberships are non overlapping ones.

4. Proposed approach

The restriction of T–S identification method for the case presented in the previous section does not mean the non-existence of solutions rather than an incentive for their search. As it has been seen, the problem comes from the fact that the solution should fulfil:

$$\nabla J = X^t Y - X^t X P = 0 \quad (11)$$

But as it was shown above, the columns of the matrix X are linearly dependent and consequently $X^t X$ is not an invertible matrix, therefore it is impossible to calculate P through:

$$P = (X^t X)^{-1} X^t Y \quad (12)$$

Nevertheless, as the rows of X^t are linearly dependent, the independent term in Eq. (11) $X^t Y$ will have the same dependence among its rows and thereupon the rank of the system matrix will be the same as the rank of the extended matrix with the independent term.

$$\text{rank}(X^t X) = \text{rank}(X^t X | X^t Y)$$

And so the system has solution. In other words, the system is a compatible indeterminate one, that is, if P is a solution of (11) and K belongs to the kernel of $X^t X$

$$K \in Ker(X^t X)$$

then $P^* = P + K$ will also be a solution. Therefore, the problem is not the lack of a solution rather the existence of infinite solutions and the key idea is the ability to find one of them. Several proposals can be made to select a solution. In our case, the aim is to find solutions with lower norm.

4.1. Parameters' weighting method

An effective approach with few computational efforts, based on the well known parameters' weighting method, is proposed. The main target is to improve the choice of the performance index and minimize it. It is characterized by extending the objective function by including a weighting γ of the norm of P vector.

$$J = \sum_{k=1}^{m} (y_k - \hat{y}_k)^2 + \gamma^2 \sum_j p_j^2$$
$$= \left\| Y - XP \right\|^2 + \gamma^2 \left\| P \right\|^2 \quad (13)$$

This can be rewritten as follows:

$$J = \left\| Y - XP \right\|^2 + \gamma^2 \left\| P \right\|^2$$
$$= \left\| \begin{bmatrix} Y \\ 0 \end{bmatrix} - \begin{bmatrix} X \\ \gamma I \end{bmatrix} P \right\|^2 \quad (14)$$
$$= \left\| Y_a - X_a P \right\|^2$$

Now the extended matrix X_a is of full rank, and the vector P can be computed as:

$$P = (X_a^t X_a)^{-1} X_a^t Y_a \quad (15)$$

Obviously the solution that minimizes this index is not optimum. However, for a small value of γ, it will be close to the optimum one and it will be unique as well.

The weighting of γ does not need to have a unique value, rather than some values can be weighted more than others to choose the most suitable one.

$$J = \sum_{k=1}^{m}(y_k - \hat{y}_k)^2 + \sum_{j}\gamma_i^2 p_j^2$$
$$= \left\|Y - XP\right\|^2 + \left\|\Gamma P\right\|^2 \qquad (16)$$
$$= \left\|\begin{bmatrix} Y \\ 0 \end{bmatrix} - \begin{bmatrix} X \\ \Gamma \end{bmatrix} P\right\|^2$$

where Γ is a diagonal matrix formed by the values γ^2 which should necessarily have a value other than zero to guarantee that the extended matrix is invertible.

4.2. Example

In the following example, we will compare the proposed method with the one proposed by T–S in [37]. A non-linear function will be proposed and the fuzzy models will be obtained assigning an interval $[a_i, b_i]$ for each variable x_i. In this interval, two fuzzy sets are defined whose membership functions are:

$$\mu_{i1}(x_i) = \frac{b_i - x_i}{b_i - a_i}, \qquad \mu_{i2}(x_i) = \frac{x_i - a_i}{b_i - a_i}$$

However, as demonstrated above, these membership functions cannot be used directly in the method of T–S, as the resulting matrix X would not be full rank.

In order to compare these methods, we use a factor $0 \le \alpha \le 1$, which determines two points within the interval. This means:

$$a_i^* = b_i - \alpha(b_i - a_i)$$

$$b_i^* = a_i + \alpha(b_i - a_i)$$

And so, we define two fuzzy sets whose membership functions are:

$$\mu_{i1}(x) = \begin{cases} \dfrac{x_i - b_i^*}{a_i - b_i^*} & a_i \le x \le b_i^* \\ 0 & b_i^* \le x \le b_i \end{cases} \qquad (17)$$

$$\mu_{i2}(x) = \begin{cases} 0 & a_i \le x \le a_i^* \\ \dfrac{x_i - a_i^*}{b_i - a_i^*} & a_i^* \le x \le b_i \end{cases} \qquad (18)$$

which are those used with the direct method of T–S. As a measure of error for comparing these methods, the maximum of the absolute values of the errors is used, which is the same method applied by T–S in [37].

Example 4.1. Consider the following simple nonlinear system:

$$\dot{x} = x^2$$

It is aimed to estimate this system:

$$y = x^2 \quad x \in [0, 1]$$

Let us suppose that we define in this interval two fuzzy sets with their corresponding membership functions as follows:

$$\mu_1(x) = 1 - x$$

$$\mu_2(x) = x$$

The objective is to calculate the corresponding fuzzy model in an optimum form:

$$\begin{aligned} S^1 &: \text{ If } x \text{ is } M_1^1 \text{ then } y = p_0^1 + p_1^1 x \\ S^2 &: \text{ If } x \text{ is } M_1^2 \text{ then } y = p_0^2 + p_1^2 x \end{aligned} \qquad (19)$$

In order to identify the nonlinear function, we take 6 points uniformly distributed $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$ in the interval $[0, 1]$. In the original T–S method [37], The X matrix becomes:

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.8 & 0.16 & 0.2 & 0.04 \\ 0.6 & 0.24 & 0.4 & 0.16 \\ 0.4 & 0.24 & 0.6 & 0.36 \\ 0.2 & 0.16 & 0.8 & 0.64 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Its rank is 3 (the condition number is $8.1154e + 015$). The $X^T X$ product has the following value:

$$X^T X = \begin{bmatrix} 2.2 & 0.4 & 0.8 & 0.4 \\ 0.4 & 0.1664 & 0.4 & 0.2336 \\ 0.8 & 0.4 & 2.2 & 1.8 \\ 0.4 & 0.2336 & 1.8 & 1.5664 \end{bmatrix}$$

whose determinant is $5.5707e - 017$.

Applying the proposed weighting of parameters approach for various values of γ to examine the trade-off between the accuracy of the identification of the model and the condition number, the following results are obtained:

For $\gamma = 0.1$, the obtained fuzzy rules are:

$$\begin{aligned} S^1 &: \text{ If } x \text{ is } M_1^1 \text{ then } y = -0.004 - 0.3167x \\ S^2 &: \text{ If } x \text{ is } M_1^2 \text{ then } y = 0.3392 + 0.6559x \end{aligned} \qquad (20)$$

The error between the original system and the identified fuzzy model is with an error of 0.0050.

For $\gamma = 0.01$, the obtained fuzzy rules are:

$$\begin{aligned} S^1 &: \text{ If } x \text{ is } M_1^1 \text{ then } y = -0.3332x \\ S^2 &: \text{ If } x \text{ is } M_1^2 \text{ then } y = 0.3334 + 0.66666x \end{aligned} \qquad (21)$$

The error between the original system and the identified fuzzy model is with an error of $5.1410e - 005$.

For $\gamma = 0.001$, the obtained fuzzy rules are:

$$\begin{aligned} S^1 &: \text{ If } x \text{ is } M_1^1 \text{ then } y = -0.3333x \\ S^2 &: \text{ If } x \text{ is } M_1^2 \text{ then } y = 0.3333 + 0.6667x \end{aligned} \qquad (22)$$

The error between the original system and the identified fuzzy model is with an error of $5.1409e - 007$.

As it can be seen that although it is not an optimal solution, but it is obviously very close and presents a good estimation. In addition to that, it can be clearly noticed that as γ is reduced, the accuracy of the identified model is tangibly increased though the condition number of the matrix X_a is about $3.6608e + 003$ which is excessively high. In other words, X_a is ill-conditioned and this leads to an unreliable results.

In the method of T–S, let us suppose that $\alpha = 0.7$. The fuzzy model is:

$$\begin{aligned} S^1 &: \text{ If } x \text{ is } M_1^1 \text{ then } y = -0.0389 + 0.5238x \\ S^2 &: \text{ If } x \text{ is } M_1^2 \text{ then } y = -0.5151 + 01.4762x \end{aligned} \qquad (23)$$

The identification error is 0.0201, which is higher than the one obtained in the proposed two methods. Increasing $\alpha = 0.8$, the error is reduced to 0.0160, i.e., the identification is improved. With $\alpha = 0.9$, the error becomes 0.0081, where the identification is again improved. The same occurs for $\alpha = 0.95$ where the error is reduced to 0.0044. In other words, as the factor α is approaching unity the identification is improved, but it cannot reach the optimum which precisely occurs when $\alpha = 1$. Since at this value, the matrix X is not of full rank and therefore the matrix $X_t X$ is not invertible. Even when α is approaching unity, the condition number of the matrix X starts increasing which indicates that $X_t X$ is approaching the singularity and therefore its inverse is no longer numerically reliable. For instance, when $\alpha = 0.999$, the condition number of X is $5.2523e + 004$ which shows clearly a non reliable result.

4.3. Parameters tuning using the parameters' weighting method

The parameters' weighting method can also be used for parameters tuning of T–S model from local parameters obtained through the identification of a system in an operating region or from any physical input/output data.

We suppose that in this case we have a first estimation

$$P_0 = [p_0^0 \ p_1^0 \ p_2^0 \ldots p_n^0]^T$$

of the T–S model parameters. In order to obtain such an estimation, the classical least square method can be used around the equilibrium point. The objective is to obtain a global approximation of the system.

$$\hat{y} = p_0^0 + p_1^0 x_1 + p_2^0 x_2 + \cdots + p_n^0 x_n$$

Let us analyze a set of input/output system samples $\{x_{1k}, x_{2k}, \ldots, x_{nk}, y_k\}$. The parameters of the global approximation can be calculated by minimizing the following quadratic performance index:

$$J = \sum_{k=1}^{m}(y_k - \hat{y}_k)^2 = \left\| Y - X_g P_0 \right\|^2 \tag{24}$$

where

$$
\begin{aligned}
Y &= [y_1 \ y_2 \ y_3 \ \ldots \ y_m]^T \\
P_0 &= [p_0^0 \ p_1^0 \ p_2^0 \ \ldots \ p_n^0]^T \\
X_g &= \begin{bmatrix} 1 & x_{11} & x_{21} & \ldots & x_{n1} \\ 1 & x_{12} & x_{22} & \ldots & x_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1m} & x_{2m} & \ldots & x_{nm} \end{bmatrix}
\end{aligned} \tag{25}
$$

In this case, if we select a sufficient number of points distributed in the region where it is required to obtain the approximation, then the matrix X_g will be of a full rank and therefore, the solution becomes unique, which can be calculated as follows:

$$P_0 = (X_g^T X_g)^{-1} X_g^T Y \tag{26}$$

This first approximation can be utilized as reference parameters for all the subsystems. Then, the parameters' vector of the fuzzy model can be obtained minimizing:

$$
\begin{aligned}
J &= \sum_{k=1}^{m}(y_k - \hat{y}_k)^2 + \gamma^2 \sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} \sum_{j=0}^{n}(p_j^0 - p_j^{(i_1 \ldots i_n)})^2 \\
&= \left\| Y - Xp \right\|^2 + \gamma^2 \| p_0 - p \|^2 \\
&= \left\| \begin{bmatrix} Y \\ \gamma p_0 \end{bmatrix} - \begin{bmatrix} X \\ \gamma I \end{bmatrix} p \right\|^2 \\
&= \left\| Y_a - X_a p \right\|^2
\end{aligned} \tag{27}
$$

where

$$p_0 = \underbrace{[P_0 \ P_0 \ldots P_0]}_{r_1, r_2 \ldots r_n}{}^T$$

In this case, the factor γ represents the degree of confidence of the parameters initially estimated. In a similar way to the previous case, different weight factors of $\gamma_j^{(i_1 \ldots i_n)}$ can be used to each one of the parameters $p_j^{(i_1 \ldots i_n)}$ depending on the reliability of the initial parameter p_j^0 in the specific rule.

$$
\begin{aligned}
J &= \sum_{k=1}^{m}(y_k - \hat{y}_k)^2 + \sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} \sum_{j=0}^{n}\gamma_j^{(i_1 \ldots i_n)2}(p_j^0 - p_j^{(i_1 \ldots i_n)})^2 \\
&= \left\| Y - Xp \right\|^2 + \left\| \Gamma(p_0 - p) \right\|^2 \\
&= \left\| \begin{bmatrix} Y \\ \Gamma p_0 \end{bmatrix} - \begin{bmatrix} X \\ \Gamma \end{bmatrix} p \right\|^2 \\
&= \left\| Y_a - X_a p \right\|^2
\end{aligned} \tag{28}
$$

where Γ is a diagonal matrix with the weight factor $\gamma_j^{(i_1 \ldots i_n)}$. It is not necessary to apply this process for all the parameters. If the values of some of them are known, they can be fixed beforehand or we can assign them a weighting factor $\gamma_j^{(i_1 \ldots i_n)}$ comparatively high.

5. Identification of vectorial functions

In the case of vectorial functions,

$$f : \Re^n \to \Re^m$$

$$y = f(x_1, x_2, \ldots, x_n) \quad y \in \Re^m$$

This function is equivalent to m scalar functions

$$y_j = f_j(x_1, x_2, \ldots, x_n) \quad y \in \Re, \quad j = 1 \ldots m$$

Each one of these functions can be modeled as follows:

$$S_j^{(i_1 \ldots i_n)} : \text{If } x_1 \text{ is } M_{j1}^{i_1} \text{ and } x_2 \text{ is } M_{j2}^{i_2} \text{ and } \ldots x_n \text{ is } M_{jn}^{i_n} \text{ then}$$
$$\hat{y}_j = p_{j0}^{(i_1 \ldots i_n)} + p_{j1}^{(i_1 \ldots i_n)}x_1 + p_{j2}^{(i_1 \ldots i_n)}x_2 + \cdots + p_{jn}^{(i_1 \ldots i_n)}x_n \tag{29}$$

Applying the previous method for each one of these models, we get in basic case:

$$
\begin{aligned}
Y_j &= \begin{bmatrix} y_{j1} & y_{j2} & \ldots & y_{jN} \end{bmatrix}^t \\
P_j &= \begin{bmatrix} p_{j0}^{(1 \ldots 1)} & p_{j1}^{(1 \ldots 1)} \ldots p_{jn}^{(1 \ldots 1)} \ldots p_{j0}^{(r_1 \ldots r_n)} & p_{j1}^{(r_1 \ldots r_n)} \ldots p_{jn}^{(r_1 \ldots r_n)} \end{bmatrix}^t \\
X &= \begin{bmatrix} \beta_{j1}^{(1 \ldots 1)} & \beta_{j1}^{(1 \ldots 1)}x_{11} \ldots \beta_{j1}^{(1 \ldots 1)}x_{n1} \ldots \beta_{j1}^{(r_1 \ldots r_n)} & \beta_{j1}^{(r_1 \ldots r_n)}x_{11} \ldots \beta_{j1}^{(r_1 \ldots r_n)}x_{n1} \\ \beta_{jN}^{(1 \ldots 1)} & \beta_{jN}^{(1 \ldots 1)}x_{1N} \ldots \beta_{jN}^{(1 \ldots 1)}x_{nN} \ldots \beta_{jN}^{(r_1 \ldots r_n)} & \beta_{jN}^{(r_1 \ldots r_n)}x_{1N} \ldots \beta_{jN}^{(r_1 \ldots r_n)}x_{nN} \end{bmatrix}
\end{aligned} \tag{30}
$$

and the formulation for parameters tuning becomes

$$J_j = \left\| \begin{bmatrix} Y_j \\ \Gamma_j p_{j0} \end{bmatrix} - \begin{bmatrix} X_j \\ \Gamma_j \end{bmatrix} p_j \right\|^2 \tag{31}$$

It is quite usual in the modelling of T–S model, that the fuzzy sets of the premise part of all the subsystems are the same, fulfilling in this case that:

$$X_1 = X_2 = \cdots = X_m = X$$

In addition to that, if the weight factors $\gamma_j^{(i_1 \cdots i_n)}$ of the matrix Γ are the same for all the subsystems, they can be all grouped as follows:

$$
J = \left\| \begin{bmatrix} Y_1 & Y_2 & & Y_m \\ \Gamma p_{10} & \Gamma p_{20} & & \Gamma p_{m0} \end{bmatrix} - \begin{bmatrix} X \\ \Gamma \end{bmatrix} \begin{bmatrix} p_1 & p_2 & & p_m \end{bmatrix} \right\|^2
$$
$$
= \left\| Y_a - X_a p \right\|^2 \tag{32}
$$

and the solution is still:

$$p = (X_a^t X_a)^{-1} X_a^t Y_a \tag{33}$$

6. Design of an optimal controller

In order to show the effectiveness of the proposed estimation methods, a design of an optimal controller is carried out for a dynamic system whose discrete model is following:

$$x(k+1) = f(x(k), u(k)) \quad x \in n\Re \quad u \in m\Re$$

Applying the proposed estimation method mentioned above to a set of points, the T–S model can be adjusted as follows:

$$
S^{(i_1 \cdots i_n)} \quad : \quad \text{If } x_1(k) \text{ is } M_1^{i_1} \text{ and } x_2(k) \text{ is } M_2^{i_2} \text{ and} \dots \text{ and } x_n(k) \text{ is } M_n^{i_n} \\
\text{then } x(k+1) = a_0^{(i_1 \cdots i_n)} + A^{(i_1 \cdots i_n)} x(k) + B^{(i_1 \cdots i_n)} u(k) \tag{34}
$$

$$a_0^{(i_1 \cdots i_n)} \in n\Re \quad A^{(i_1 \cdots i_n)} \in n \times n\Re \quad B^{(i_1 \cdots i_n)} \in n \times m\Re$$

where $M_1^{i_1}$ ($i_1 = 1, 2, \dots, r_1$) are fuzzy sets for $x_1(k)$, $M_2^{i_2}$ ($i_2 = 1, 2, \dots, r_2$) are fuzzy sets for $x_2(k)$ and $M_n^{i_n}$ ($i_n = 1, 2, \dots, r_n$) are fuzzy sets for $x_n(k)$. The fuzzy system is described as:

$$
x(k+1) = \frac{\sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} w^{(i_1 \cdots i_n)}(x(k)) \left[a_0^{(i_1 \cdots i_n)} + A^{(i_1 \cdots i_n)} x(k) \right]}{\sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} w^{(i_1 \cdots i_n)}(x(k))} \\
+ \frac{\sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} w^{(i_1 \cdots i_n)}(x(k)) \left[B^{(i_1 \cdots i_n)} u(k) \right]}{\sum_{i_1=1}^{r_1} \cdots \sum_{i_n=1}^{r_n} w^{(i_1 \cdots i_n)}(x(k))} \tag{35}
$$

which is equivalent to the system:

$$x(k+1) = a_0(x(k)) + A(x(k))x(k) + B(x(k))u(k) \tag{36}$$

where the matrices are determined by the fuzzy system as follows:

$$
S^{(i_1, \cdots, i_n)} \quad : \quad \text{If } x_1(k) \text{ is } M_1^{i_1} \text{ and } x_2(k) \text{ is } M_2^{i_2} \text{ and} \dots \text{ and } x_n(k) \text{ is } M_n^{i_n} \\
\text{then } a_0 = a_0^{(i_1 \cdots i_n)} \text{ and } A = A^{(i_1 \cdots i_n)} \text{ and } B = B^{(i_1 \cdots i_n)} \tag{37}
$$

In order to calculate the rest of control coefficients, any control design methodology through state feedback control can be applied. However, these methodologies can lead to systems where the control action exceeds the allowed environment limits. Selecting closed loop poles with great negative real parts make the dynamic response of the system to be quick while the control effort to be greater than permissible levels. If selection of the closed loop poles saturate of control signals, dynamic behavior of the system will not be as same as the desired behavior even it may become unstable. Therefore optimal selection of closed loop poles will lead to a trade-off between speed of dynamic response and control effort. Together with the proposed estimation method, the well known Linear Quadratic Regulator (LQR) method might be an appropriate choice. With this method we might not guarantee of stability, which needs to be analyzed a posteriori, gaining in return a balance between static and dynamic behavior of the system with admissible control actions.

The objective is to find the control action $u(k)$ to transfer the system from any initial state $x(k_0)$ to some final state $x(k_1) = 0$ in a time interval, minimizing a quadratic performance index of the form:

$$J = \sum_{k=k_0}^{k_1-1} x^t(k+1)Qx(k+1) + u^t(k)Ru(k) \tag{38}$$

Q is a non-negative definite matrix which is called weighting matrix. On the other hand, R is a positive definite weighting matrix. Optimal selection of these matrices determines the dynamic speed of the controller as well as amplitudes of state variables and control signals. For example if R is selected small while Q is selected large, more stability will be attained with large control efforts.

The solution is well known [2–4,27,49] when the system is a linear one. However, it is not easy enough for the nonlinear system obtained before (see Eq. (36). In order to solve this problem, it is proposed to minimize the cost of each step instead of the global cost. The solution will be a suboptimal one but with a great advantage due to its easy calculation. It is suggested in each step to apply the control action that minimizes:

$$J_k = x^t(k+1)Qx(k+1) + u^t(k)Ru(k) \tag{39}$$

Substituting (36) in (39), we obtain:

$$
J_k = [a_0(x(k)) + A(x(k))x(k) + B(x(k))u(k)]^t \\
\times Q[a_0(x(k)) + A(x(k))x(k) + B(x(k))u(k)] + u^t(k)Ru(k) \tag{40}
$$

Applying the gradient and equating to zero to minimize the performance index:

$$
\nabla_{u(k)} J_k = 2[a_0(x(k)) + A(x(k))x(k) + B(x(k))u(k)]^t \\
\times QB(x(k)) + 2u^t(k)R = 0
$$

Thus,

$$
Ru(k) + B^t(x(k))Q[a_0(x(k)) + A(x(k))x(k) + B(x(k))u(k)] = B^t(x(k)) \\
\times Q[a_0(x(k)) + A(x(k))x(k)] + [R + B^t(x(k))QB(x(k))]u(k) \tag{41}
$$

The optimal control law is then computed as follows:

$$u(k) = -[R + B^t(x(k))QB(x(k))]^{-1}B^t(x(k))Q[a_0(x(k)) + A(x(k))x(k)]$$

7. Illustrative example

In this section the proposed estimation method and its application to design a FC-LQR is illustrated by two examples of an inverted pendulum and Van der Pol oscillator respectively.

Example 7.1. Inverted pendulum:

Consider the problem of stabilizing and balancing of swing up of an inverted pendulum (see Fig. 2). The control of this system is a widely used performance measure of a controller, since this system is unstable and highly nonlinear. The objective is to maintain the inverted pendulum upright with θ despite small disturbances due to wind or system noises. The inverted pendulum can be represented as follows:

$$\ddot{\theta} = \frac{g\sin\theta - \cos\theta((u + ml\dot{\theta}^2 \sin\theta)/(M+m))}{l(4/3 - (m\cos^2\theta/M + m)} \tag{42}$$

where θ denotes the angular position (in radians) deviated from the equilibrium position (vertical axis) of the pendulum and $\dot{\theta}$ is the angular velocity, g (gravityacceleration) = 9.8 m/s^2, M (mass) of the cart = 1 kg, m (mass) of the pole = 0.1 kg, l is the distance from the center of the mass (m) of the pole to the cart = 0.5 m.
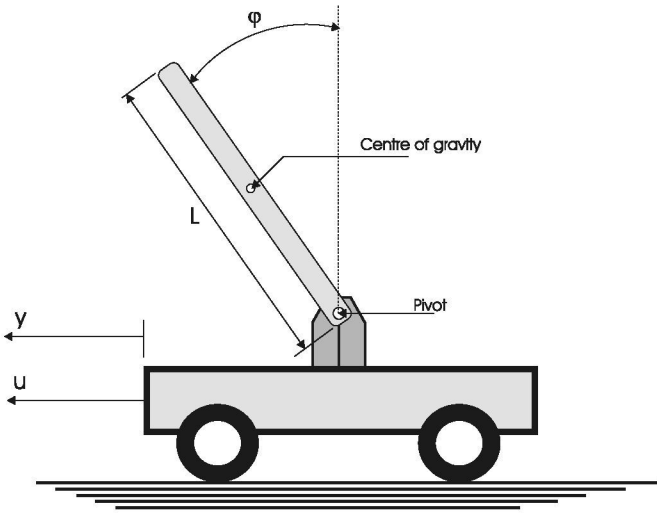
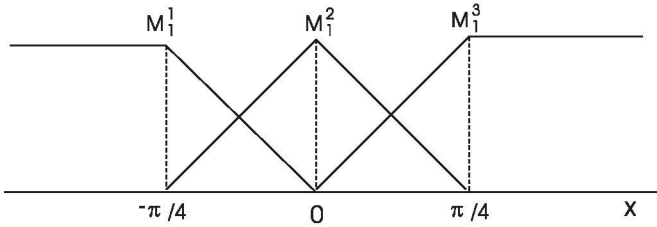**Fig. 2.** Inverted pendulum system.

Fig. 3. Membership functions for the angle x of the inverted pendulum.

Assuming that $x_1 = \theta$ and $x_2 = \dot{\theta}$, then (42) can be rewritten in state space form as follows:

$$x_1 = \theta$$
$$\dot{x}_1 = x_2$$
$$x_2 = \dot{\theta} \qquad (43)$$
$$\dot{x}_2 = \frac{g\,\sin(x_1) - \cos(x_1)((u + mlx_2^2\,\sin(x_1))/(M+m))}{l(4/3 - (m\cos^2(x_1)/(M+m)))}$$

The aim is to move the pendulum to its instable equilibrium position, i.e., $x_1 = x_2 = 0$. In order to control the system, a discrete controller is proposed with a sampling time of 0.1 s and a sample-hold device is used for the control action. It is suggested to model the equivalent discrete system as:

$$x(k+1) = a_0(x(k)) + A(x(k))x(k) + B(x(k))u(k) \qquad (44)$$

$$x \in 2\Re \quad u \in \Re \quad a_0 \in 2\Re \quad A \in 2 \times 2\Re \quad B \in 2\Re$$

Using the proposed methodology, supposing the universe of discourse of the angle is $[-\pi/4, \pi/4]\,\text{rad}$. and the one of the angular velocities is $[-5, 5]\,\text{rad/seg}$. Both membership functions for the angle x and its derivative x′ are shown in Figs. 3 and 4 respectively.
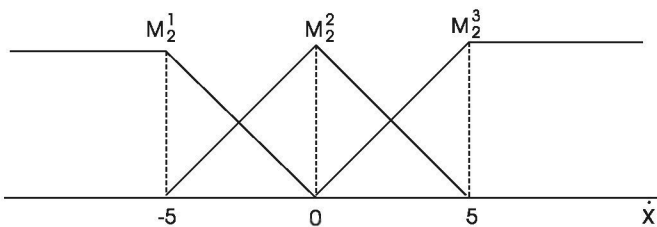
Fig. 4. Membership functions for the angular velocity of the inverted pendulum.

If we apply the T–S method directly to this example, then the condition number of the matrix X is $3.4148e+015$, which shows clearly a non reliable result.

By using the identification with the classical least square method in an interval close to the equilibrium point

$$x_1 \in \left[\frac{-\pi}{40}, \frac{\pi}{40}\right] \quad x_2 \in [-0.5, 0.5]$$

The resultant linear model of the system in this interval is:

$$x(k+1) = \begin{bmatrix} 1.0790 & 0.1026 \\ 1.6154 & 1.0788 \end{bmatrix} x(k) + \begin{bmatrix} -0.0073 \\ -0.1499 \end{bmatrix} u(k)$$

This approximation gives us an idea of the quantitative relation among various variables. By applying the tuning of parameters method, taking as a reference the parameters obtained through least square method with a weighting factor $\gamma = 0.001 * np$, where np is the number of samples used in the identification, we obtain:

S^{11} : If $(x_1$ is $M_1^1)$ and $(x_2$ is $M_2^1)$ then
$$x(k+1) = \begin{bmatrix} 0.0037 \\ 0.0865 \end{bmatrix} + \begin{bmatrix} 1.0696 & 0.1010 \\ 1.3784 & 1.0388 \end{bmatrix} x(k) + \begin{bmatrix} -0.0041 \\ -0.0707 \end{bmatrix} u(k)$$

S^{12} : If $(x_1$ is $M_1^1)$ and $(x_2$ is $M_2^2)$ then
$$x(k+1) = \begin{bmatrix} 0.0029 \\ 0.0610 \end{bmatrix} + \begin{bmatrix} 1.0724 & 0.1022 \\ 1.4757 & 1.0665 \end{bmatrix} x(k) + \begin{bmatrix} -0.0053 \\ -0.1074 \end{bmatrix} u(k)$$

S^{13} : If $(x_1$ is $M_1^1)$ and $(x_2$ is $M_2^3)$ then
$$x(k+1) = \begin{bmatrix} 0.0034 \\ 0.0795 \end{bmatrix} + \begin{bmatrix} 1.0786 & 0.1035 \\ 1.6547 & 1.0958 \end{bmatrix} x(k) + \begin{bmatrix} -0.0062 \\ -0.1361 \end{bmatrix} u(k)$$

S^{21} : If $(x_1$ is $M_1^2)$ and $(x_2$ is $M_2^1)$ then
$$x(k+1) = \begin{bmatrix} 0.0001 \\ 0.0021 \end{bmatrix} + \begin{bmatrix} 1.0831 & 0.1023 \\ 1.6833 & 1.0687 \end{bmatrix} x(k) + \begin{bmatrix} -0.0075 \\ -0.1499 \end{bmatrix} u(k)$$

S^{22} : If $(x_1$ is $M_1^2)$ and $(x_2$ is $M_2^2)$ then
$$x(k+1) = \begin{bmatrix} -0.0000 \\ -0.0000 \end{bmatrix} + \begin{bmatrix} 1.0875 & 0.1028 \\ 1.7990 & 1.0861 \end{bmatrix} x(k) + \begin{bmatrix} -0.0078 \\ -0.1599 \end{bmatrix} u(k)$$

S^{23} : If $(x_1$ is $M_1^2)$ and $(x_2$ is $M_2^3)$ then
$$x(k+1) = \begin{bmatrix} -0.0000 \\ -0.0006 \end{bmatrix} + \begin{bmatrix} 1.0831 & 0.1022 \\ 1.6833 & 1.0683 \end{bmatrix} x(k) + \begin{bmatrix} -0.0075 \\ -0.1494 \end{bmatrix} u(k)$$

S^{31} : If $(x_1$ is $M_1^3)$ and $(x_2$ is $M_2^1)$ then
$$x(k+1) = \begin{bmatrix} -0.0035 \\ -0.0802 \end{bmatrix} + \begin{bmatrix} 1.0786 & 0.1034 \\ 1.6536 & 1.0954 \end{bmatrix} x(k) + \begin{bmatrix} -0.0062 \\ -0.1362 \end{bmatrix} u(k)$$

S^{32} : If $(x_1$ is $M_1^3)$ and $(x_2$ is $M_2^2)$ then
$$x(k+1) = \begin{bmatrix} -0.0029 \\ -0.0612 \end{bmatrix} + \begin{bmatrix} 1.0724 & 0.1022 \\ 1.4756 & 1.0663 \end{bmatrix} x(k) + \begin{bmatrix} -0.0053 \\ -0.1066 \end{bmatrix} u(k)$$

S^{33} : If $(x_1$ is $M_1^3)$ and $(x_2$ is $M_2^3)$ then
$$x(k+1) = \begin{bmatrix} -0.0037 \\ -0.0873 \end{bmatrix} + \begin{bmatrix} 1.0696 & 0.1010 \\ 1.3774 & 1.0392 \end{bmatrix} x(k) + \begin{bmatrix} -0.0042 \\ -0.0714 \end{bmatrix} u(k)$$

which has a mean square error of $7.9608e-006$. It has been taken into account that the system has an instable equilibrium point in $[0\,0]$, therefore, it should fulfill that

$$a_0^{(2,2)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

To achieve this target, we can either fix the value of this parameter or assign it a very large weighting factor as explained in Section (4.3). In this case, a weighting factor of 100γ has been assigned to this parameter.

For the control system, it is required to minimize the following performance index:

$$J_k = x^t(k+1) \begin{bmatrix} 100 & 0 \\ 0 & 1 \end{bmatrix} x(k+1) + u^t(k)[0.01]u(k)$$

Fig. 5 shows the controlled system behavior subjected initial conditions $[15°0]^t$ and $[-10°0]^t$. Fig. 6 shows the controlled system
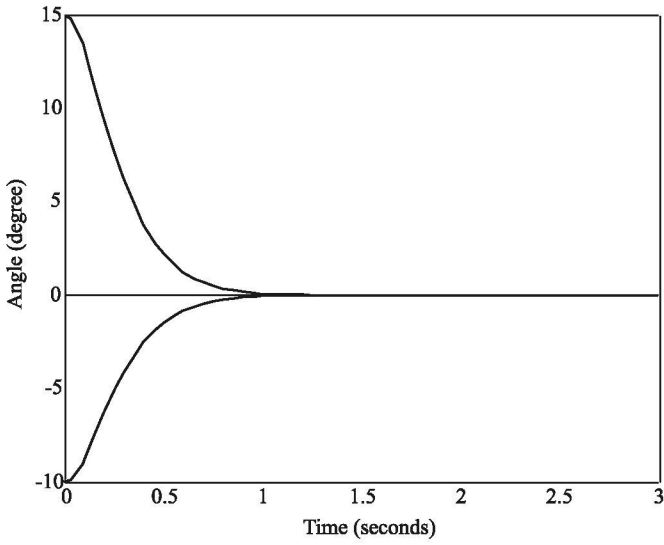
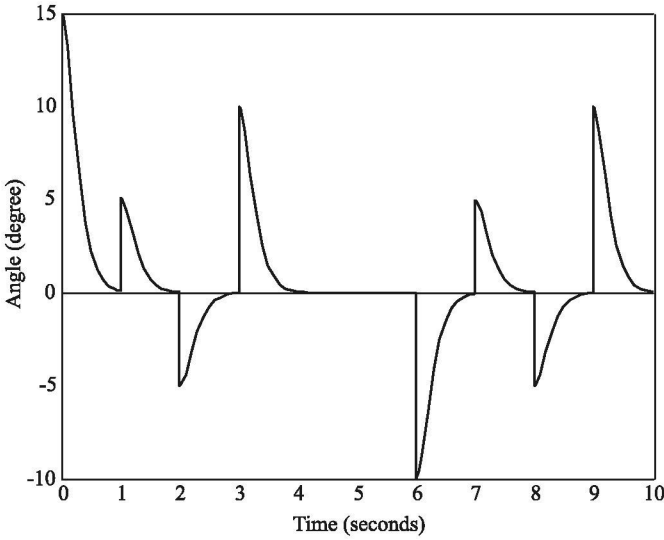**Fig. 5.** Transient response of the inverted pendulum controlled by the fuzzy LQR.

Fig. 6. Transient response of the inverted pendulum controlled by the fuzzy LQR subjected to disturbances.
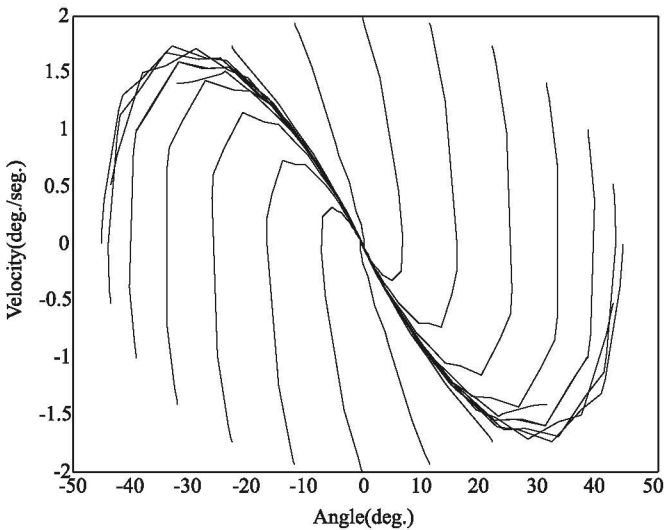

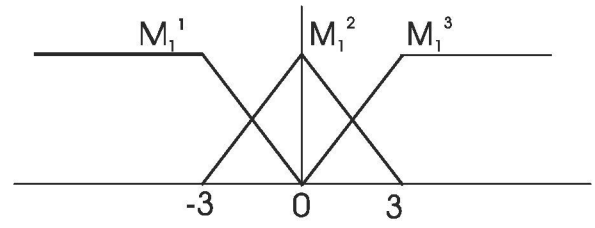**Fig. 7.** Trajectories of the system for several initial conditions.

Fig. 8. Membership functions for x.

behavior subjected to disturbances. Fig. 7 shows several trajectories of the system for several initial conditions.

Example 7.2. Van der Pol Oscillator:

The second order nonlinear differential equation:

$$\ddot{y} + (y^2 - 1)\dot{y} + y = u$$

$$x_1 = y$$

$$x_2 = \dot{y}$$

$$\dot{x} = F(x, u)$$

is the famous Van der Pol equation [34]. It is well known that for $u = 0$, the equation has a unique singular point (0, 0) which is an unstable focus. Moreover, when $u = 0$, the system displays a sustained oscillation independent of initial conditions known as limit cycle of amplitude $A = 2$ and natural frequency $\omega = 1$.

The objective is to apply the methodology used in the previous example to guide the system to its equilibrium point. Using the proposed methodology with a universe of discourse of x is $[-3, 3]$ and the one of its derivatives is $[-5, 5]$. Three membership functions overlapped by pairs are assigned for x and its derivative as in the previous example. Both membership functions for the x and its derivative $\dot{x}$ are shown in Figs. 8 and 9 respectively.

If we apply the T–S method directly to this example, then the condition number of the matrix X is $1.8716e + 016$, which shows clearly a non reliable result

By using the identification with the classical least square method in an interval close to the equilibrium point, the resultant linear model of the system in this interval is:

$$x(k+1) = \begin{bmatrix} 0.9949 & 0.1048 \\ -1.1055 & 1.0966 \end{bmatrix} x(k) + \begin{bmatrix} 0.0051 \\ 0.1046 \end{bmatrix} u(k)$$
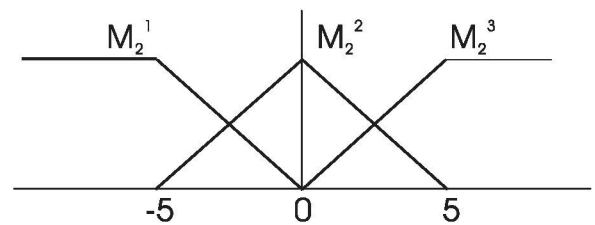

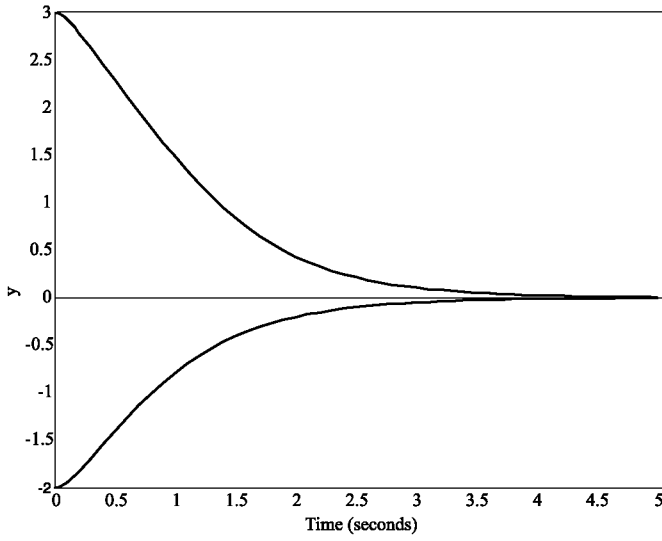**Fig. 9.** Membership functions for $\dot{x}$.

Fig. 10. Transient response of the Van der Pol system controlled by the fuzzy LQR.

Fig. 11. Transient response of the Van der Pol system controlled by the fuzzy LQR subjected to various disturbances.

By applying the tuning of parameters method, taking as a reference the parameters obtained through least square method with a weighting factor $\gamma = 0.3$, we obtain:

S^{11} : If $(x_1$ is $M_1^1)$ and $(x_2$ is $M_2^1)$ then

$$x(k+1) = \begin{bmatrix} 0.0028 \\ 0.0809 \end{bmatrix} + \begin{bmatrix} 0.9509 & 0.0933 \\ -0.7242 & 0.7771 \end{bmatrix} x(k) + \begin{bmatrix} 0.0036 \\ 0.0582 \end{bmatrix} u(k)$$

S^{12} : If $(x_1$ is $M_1^1)$ and $(x_2$ is $M_2^2)$ then

$$x(k+1) = \begin{bmatrix} -0.0004 \\ -0.0123 \end{bmatrix} + \begin{bmatrix} 0.9957 & 0.0965 \\ -0.0847 & 0.8581 \end{bmatrix} x(k) + \begin{bmatrix} 0.0041 \\ 0.0744 \end{bmatrix} u(k)$$

S^{13} : If $(x_1$ is $M_1^1)$ and $(x_2$ is $M_2^3)$ then

$$x(k+1) = \begin{bmatrix} -0.0021 \\ -0.0602 \end{bmatrix} + \begin{bmatrix} 1.0442 & 0.0998 \\ 0.6911 & 0.9489 \end{bmatrix} x(k) + \begin{bmatrix} 0.0043 \\ 0.0834 \end{bmatrix} u(k)$$

S^{21} : If $(x_1$ is $M_1^2)$ and $(x_2$ is $M_2^1)$ then

$$x(k+1) = \begin{bmatrix} -0.0003 \\ -0.0078 \end{bmatrix} + \begin{bmatrix} 0.9904 & 0.1056 \\ -0.2668 & 1.1140 \end{bmatrix} x(k) + \begin{bmatrix} 0.0054 \\ 0.1116 \end{bmatrix} u(k)$$

S^{22} : If $(x_1$ is $M_1^2)$ and $(x_2$ is $M_2^2)$ then

$$x(k+1) = \begin{bmatrix} 0.0000 \\ 0.0000 \end{bmatrix} + \begin{bmatrix} 0.9952 & 0.1059 \\ -0.0898 & 1.1291 \end{bmatrix} x(k) + \begin{bmatrix} 0.0053 \\ 0.1093 \end{bmatrix} u(k)$$

S^{23} : If $(x_1$ is $M_1^2)$ and $(x_2$ is $M_2^3)$ then

$$x(k+1) = \begin{bmatrix} 0.0002 \\ 0.0054 \end{bmatrix} + \begin{bmatrix} 0.9903 & 0.1055 \\ -0.2680 & 1.1130 \end{bmatrix} x(k) + \begin{bmatrix} 0.0055 \\ 0.1139 \end{bmatrix} u(k)$$

S^{31} : If $(x_1$ is $M_1^3)$ and $(x_2$ is $M_2^1)$ then

$$x(k+1) = \begin{bmatrix} 0.0030 \\ 0.0828 \end{bmatrix} + \begin{bmatrix} 1.0438 & 0.0997 \\ 0.6801 & 0.9476 \end{bmatrix} x(k) + \begin{bmatrix} 0.0042 \\ 0.0800 \end{bmatrix} u(k)$$

S^{32} : If $(x_1$ is $M_1^3)$ and $(x_2$ is $M_2^2)$ then

$$x(k+1) = \begin{bmatrix} 0.0002 \\ 0.0078 \end{bmatrix} + \begin{bmatrix} 0.9957 & 0.0965 \\ -0.0821 & 0.8594 \end{bmatrix} x(k) + \begin{bmatrix} 0.0041 \\ 0.0746 \end{bmatrix} u(k)$$

S^{33} : If $(x_1$ is $M_1^3)$ and $(x_2$ is $M_2^3)$ then

$$x(k+1) = \begin{bmatrix} -0.0020 \\ -0.0583 \end{bmatrix} + \begin{bmatrix} 0.9505 & 0.0933 \\ -0.7340 & 0.7773 \end{bmatrix} x(k) + \begin{bmatrix} 0.0035 \\ 0.0566 \end{bmatrix} u(k)$$

which has a mean square error of $4.3502e - 004$.

For the control system, it is required to minimize the following performance index:

$$J_k = x^t(k+1) \begin{bmatrix} 20 & 0 \\ 0 & 1 \end{bmatrix} x(k+1) + u^t(k)[0.01]u(k)$$

Fig. 10 shows transient response of the Van der Pol system controlled by the fuzzy LQR subjected to initial conditions $[3\,0]^t$ and $[-2\,0]^t$. Fig. 11 shows the transient response of the system
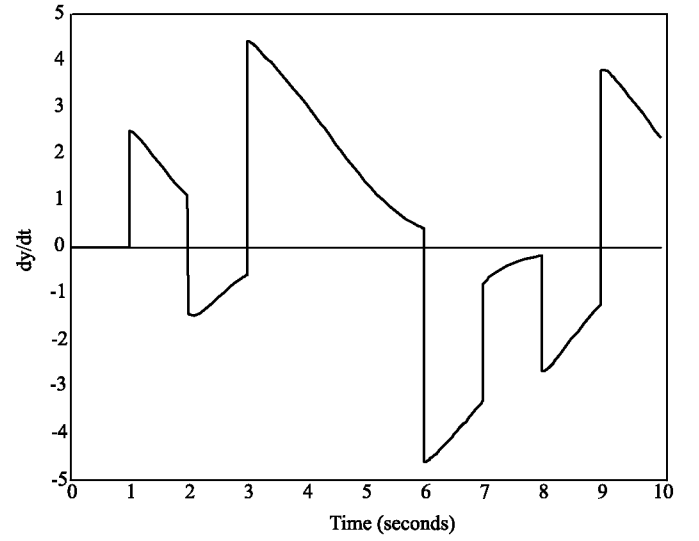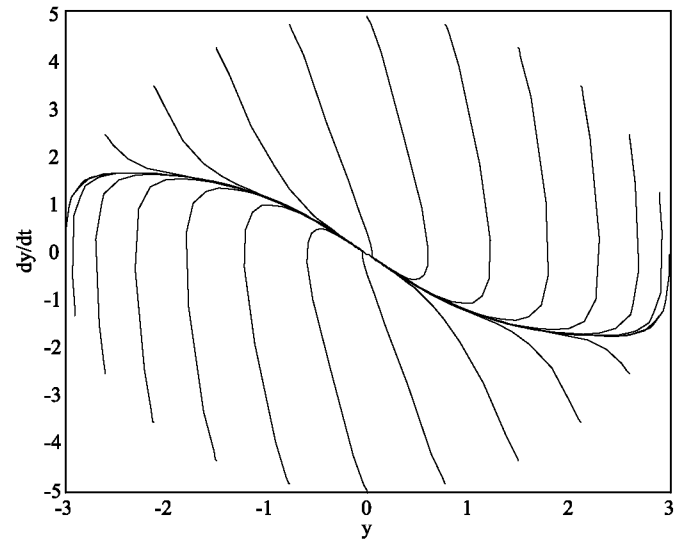

Fig. 12. Trajectories of the Van der Pol system for several initial conditions.

subjected to various disturbances. Fig. 12 shows several trajectories in state space form of the system for several initial conditions.

8. Conclusions

A new efficient approach has been presented to improve the local and global approximation of Takagi–Sugeno (T–S) fuzzy model. The main problem of T–S identification method is that it cannot be applied when the membership functions are overlapped by pairs. This restricts the use of the T–S method because this type of membership function has been widely used during the last two decades in the stability, controller design and are popular in industrial control applications. The approach developed here can be considered as a generalized version of T–S method with optimized performance in approximating nonlinear functions. An easy approach, based on the known parameters' weighting method has been applied for tuning T–S parameters to improve the choice of the performance index and minimize it.

A global fuzzy controller (FC) based Linear Quadratic Regulator (LQR) has been proposed in order to show the effectiveness of the estimation method developed here in control applications.

Illustrative examples of an inverted pendulum and Van del Pol system have been chosen to evaluate the robustness and remarkable performance of the proposed method and the high accuracy obtained in approximating nonlinear and unstable systems locally and globally in comparison with the original T–S model. Simulation results have shown the potential, simplicity and generality of the algorithm. As future objectives, it is aimed to extend the proposed estimation method for multivariable nonlinear systems. The second target is to examine other control methodologies and compare them with the proposed LQR. It will also be an interesting issue to examine the validity of the proposed control method for tracking control problems for several typical reference inputs such as step, ramp, etc.

References

[1] S. Alavandar, M.J. Nigam, Adaptive neuro-fuzzy inference system based control of six DOF robot manipulator, J. Eng. Sci. Technol. Rev. 1 (2008) 106–111.
[2] A.E. Bryson, Applied optimal control: optimization, estimation and control, Hemisphere (2008).
[3] B.D.O. Anderson, J.B. Moore, in: Optimal Control: Linear Quadratic Methods, Dover Publications, 2007.
[4] M. Athans, P. Falb, in: Optimal Control: An Introduction to the Theory and Its Applications, Dover Publications, 2006.
[5] P. Baranyi, SVD-based reduction to MISO TS models, IEEE Trans. Ind. Electron. 51 (February (1)) (2003) 232–242.
[6] S.G. Cao, N.W. Rees, G. Feng, Stability analysis of fuzzy control systems, IEEE Trans. Syst. Man Cybern. B Cybern. 26 (February (1)) (1996) 201–204.
[7] S.G. Cao, N.W. Rees, G. Feng, Analysis and design for a class of complex control systems – Part I: fuzzy modeling and identification, Automatica 33 (1997) 1017–1028.
[8] B. Chen, Xiaoping Liu, Shaocheng Tong, Adaptive fuzzy output tracking control of MIMO nonlinear uncertain systems, IEEE Trans. Fuzzy Syst. 15 (April (2)) (2007) 287–300.
[9] B.-S. Chen, C.-S. Tseng, H.-J. Uang, Robustness design of nonlinear dynamic systems via fuzzy linear control, IEEE Trans. Fuzzy Syst. 7 (October (5)) (1999) 571–585.
[10] W. Chen, M. Saif, A novel fuzzy system with dynamic rule base, IEEE Trans. Fuzzy Syst. 13 (October (5)) (2005) 569–582.
[11] O. Cordon, F. Herrera, L. Magdalena, P. Villar, A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base, Inform. Sci. 136 (2001) 85–107.
[12] C. Fanutzzi, R. Rovatti, On the approximation capabilities of the homogeneous Takagi–Sugeno model, in: in: Proc. 5th IEEE Int. Conf. Fuzzy Systems, New Orleans, LA, 1996, pp. 1067–1072.
[13] F. Gang, A survey on analysis and design of model-based fuzzy control systems, IEEE Trans. Fuzzy Syst. 14 (October (5)) (2006) 676–697.
[14] T.M. Guerra, L. Vermeiren, LMI-based relaxed nonquadratic stabilization conditions for nonlinear systems in the Takagi–Sugeno's form, Automatica 40 (2004) 823–829.
[15] Y. Hou, Jacek M. Zurada, W. Karwowski, W.S. Marras, K. Davis, Identification of key variables using fuzzy average with fuzzy cluster distribution, IEEE Trans. Fuzzy Syst. 15 (August (4)) (2007) 673–685.
[16] T. Hseng, S. Li, S.-H. Tsai, Fuzzy bilinear model and fuzzy controller design for a class of nonlinear systems, IEEE Trans. Fuzzy Syst. 15 (June (3)) (2007) 494–506.
[17] T.P. Hong, C.Y. Lee, Induction of fuzzy rules and membership functions from training examples, Fuzzy Sets Syst. 84 (1) (1996) 33–47.
[18] K. Jae-Hun, C.-H. Hyun, E. Kim, M. Park, New adaptive synchronization of uncertain chaotic systems based on T–S fuzzy model, IEEE Trans. Fuzzy Syst. 15 (June (3)) (2007) 359–369.
[19] J.S.R. Jang, ANFIS: adaptive-network-based fuzzy inference systems, IEEE Trans. Syst. Man Cybern. 23 (03) (1993) 665–685.
[20] X. Jiang, Q.-L. Han, Robust H control for uncertain Takagi–Sugeno fuzzy systems with interval time-varying delay, IEEE Trans. Fuzzy Syst. 15 (April (2)) (2007) 321–331.
[21] A. Jimnez, B.M. Al-Hadithi, F. Mata, An optimal T–S model for the estimation and identification of nonlinear functions, WSEAS Trans. Syst. Control October (3) (10) (2008) 897–906.
[22] T.A. Johansen, R. Shorten, R. Murray-Smith, On the interpretation and identification of dynamic Takagi–Sugeno models, IEEE Trans. Fuzzy Syst. 8 (June (3)) (2000) 297–313.
[23] J. Kim, Y. Suga, S. Won, A new approach to fuzzy modeling of nonlinear dynamic systems with noise: relevance vector learning mechanism, IEEE Trans. Fuzzy Syst. 14 (April (2)) (2006) 222–231.
[24] F. Klawonn, R. Kruse, Constructing a fuzzy controller from data, Fuzzy Sets Syst. 85 (2) (1997) 177–193.
[25] M. Kumar, R. Stoll, N. Stoll, A min–max approach to fuzzy clustering, estimation, and identification, IEEE Trans. Fuzzy Syst. 14 (April (2)) (2006) 248–262.
[26] R. Langari, L. Wang, Complex systems modeling via fuzzy logic, IEEE Trans. Syst. Man Cybern. B Cybern. 26 (February) (1996) 100–106.
[27] F.L. Lewis, in: Optimal and Robust Estimation: With An Introduction to Strochastic Control Theory, CRC Press, 2008.
[28] K.-Y. Lian, C.-H. Su, C.-S. Huang, Performance enhancement for T–S fuzzy control using neural networks, IEEE Trans. Fuzzy Syst. 14 (October (5)) (2006) 619–627.
[29] A. Mencattini, M. Salmeri, A. Salsano, Sufficient conditions to impose derivative constraints on MISO Takagi–Sugeno fuzzy logic systems, IEEE Trans. Fuzzy Syst. 13 (August (4)) (2005).
[30] S. Mollov, R. Babuska, J. Abonyi, H.B. Verbruggen, Effective optimization for fuzzy model predictive control, IEEE Trans. Fuzzy Syst. 12 (October (5)) (2004) 661–675.
[31] O. Nelles, in: Nonlinear System Identification, Springer, Berlin, 2001, ISBN: 3540673695.
[32] K. Nozaki, H. Ishibuchi, H. Tanaka, A simple but powerful heuristic method for generating fuzzy rules from numerical data, Fuzzy Sets Syst. 86 (3) (1997) 251–270.
[33] I. Skrjanc, S. Blazic, O. Agamennoni, Interval fuzzy model identification using l-norm, IEEE Trans. Fuzzy Syst. 13 (October (5)) (2005).
[34] J.-J.E. Slotine, in: Applied Nonlinear Systems, Prentice-Hall, 1991.
[35] M. Sugeno, G. Kang, Structure identification of fuzzy model, Fuzzy Sets Syst. 28 (1988) 15–33.
[36] M. Sugeno, K. Tanaka, Successive identification of a fuzzy model and its applications to prediction of a complex system, Fuzzy Sets Syst. 42 (1991) 315–334.
[37] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Trans. Syst. Man Cybern. SMC-15 (January (1)) (1985) 116–132.
[39] K. Tanaka, T. Hori, H.O. Wang, A multiple Lyapunov function approach to stabilization of fuzzy control systems, IEEE Trans. Fuzzy Syst. 11 (August (4)) (2003) 582–589.
[39] K. Tanaka, H.O. Wang, in: Fuzzy Control Systems Design and Analysis: A Linear Matrix Inequality Approach, Wiley, New York, 2001.
[40] M.C.M. Teixeira, E. Assuno, R.G. Avellar, On relaxed LMI based designs for fuzzy regulators and fuzzy observers, IEEE Trans. Fuzzy Syst. 11 (October (5)) (2003) 613–623.
[41] L. Wang, in: Adaptive Fuzzy Systems and Control: Design and Stability Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1994, pp. 27–28.
[42] L.X. Wang, J.M. Mendel, Design and analysis of fuzzy identifiers of nonlinear dynamic systems, IEEE Trans. Automat. Control 40 (January) (1995) 11–23.
[43] L.X. Wang, J.M. Mendel, Generating fuzzy rules by learning from examples, IEEE Trans. Syst. Man Cybern. 22 (1992) 1414–1427.
[44] P. Xian-Tu, Generating rules for fuzzy logic controllers by functions, Fuzzy Sets Syst. 36 (1990) 83–89.
[45] Y. Yam, P. Baranyi, C.-T. Yangand, Reduction of fuzzy rule base via singular value decomposition, IEEE Trans. Fuzzy Syst. 712 (April (25)) (1999), 120–132675.
[46] H. Ying, General SISO Takagi–Sugeno fuzzy systems with linear rule consequent are universal approximators, IEEE Trans. Fuzzy Syst. 6 (November (4)) (1998) 582–587.
[47] W. Yu, X.O. Li, Fuzzy identification using fuzzy neural networks with stable learning algorithms, IEEE Trans. Fuzzy Syst. 12 (June (3)) (2004) 411–420.
[48] K. Zeng, N.Y. Zhang, W.L. Xu, A comparative study on sufficient conditions for Takagi–Sugeno fuzzy systems as universal approximators, IEEE Trans. Fuzzy Syst. 8 (December (6)) (2000) 773–780.
[49] K. Zhou, in: Robust and Optimal Control, Prentice-Hall, 1996.