Pergamon

PII: S0005-1098(97)00026-5

Automatica, Vol. 33, No. 7, pp. 1333-1338, 1997

Printed in Great Britain
0005-1098/97 $17.00 + 0.00

Brief Paper

Necessary Conditions for Some Typical Fuzzy Systems as Universal Approximators*

HAO YING† and GUANRONG CHEN‡

Key Words—Fuzzy systems; function approximation; fuzzy control; fuzzy modeling.

Abstract—We investigate necessary conditions for general single-input single-output fuzzy systems and a class of typical multiple-input single-output fuzzy systems as universal approximators for continuous functions defined on compact domains with arbitrarily small uniform approximation error bounds. Considering the case where the only available data about the continuous function to be approximated are a finite set of its extrema, we have established some necessary conditions for the fuzzy systems to be universal approximators of the function. The conditions can be used practically to determine input fuzzy sets, output fuzzy sets and fuzzy rules of the fuzzy systems. Furthermore, these necessary conditions provide a basis for insightful analysis of the strengths as well as the limitations of the fuzzy systems. The main strength is that only a small number of fuzzy rules may be needed to uniformly approximate continuous functions that have a complicated formulation but a relatively small number of extrema. The limitation is that, in order to approximate highly oscillatory continuous functions, the number of fuzzy rules must be large. © 1997 Elsevier Science Ltd.

1. *Introduction*

Fuzzy systems have been developed for many different practical purposes, such as emulating human expert control strategy in fuzzy control and representing the behavior of physical systems in fuzzy modeling (Marks, 1993). From a mathematical point of view, fuzzy systems can be regarded as useful and, more importantly, practical function approximators. Using the Stone–Weierstrass theorem, it has been proved by several researchers that some particular fuzzy systems are universal approximators, in the sense that they can uniformly approximate any continuous functions, defined on compact domains, to any degree of accuracy (Buckley, 1992; Kosko, 1992; Wang, 1992; Wang and Mendel, 1992; Lee and Chae, 1993; Zeng and Singh, 1994). In a constructive way, we have proved that general and typical fuzzy systems are universal approximators, and have established some sufficient conditions for uniform approximation with a pre-specified approximation error bound (Ying, 1994).

From a practical viewpoint, necessary conditions for function approximation are more important, since they can be used to determine restrictions on a selection of fuzzy systems components. To date, there exists no such necessary condition in the literature. The objectives of this research are (i) to study necessary conditions for general SISO fuzzy systems and a class of typical MISO fuzzy systems as universal approximators, and (ii) to explore, in a mathematically explicit and rigorous way, the strengths and limitations of fuzzy systems so that they can be better understood and utilized in applications.

2. *Configuration of general SISO fuzzy systems and a class of typical MISO fuzzy systems*

2.1. *General SISO fuzzy systems.* The interval $[a, b]$, on which a continuous function $y = f(x)$ is defined, is divided into N (>1) subintervals:

$$a = C_0 < C_1 < C_2 < \ldots < C_{N-1} < C_N = b.$$

For SISO fuzzy systems, there are $N + 1$ convex, normal and continuous input fuzzy sets defined on $[a, b]$ to fuzzify the input variable x (Fig. 1). Each input fuzzy set is denoted by A_i ($0 \le i \le N$). Among them, A_i ($1 \le i \le N-1$) has a membership function $\mu_i(x)$ whose value is nonzero only on $[C_{i-1}, C_{i+1}]$. More specifically, $\mu_i(x)$ is zero at $x = C_{i-1}$, increases monotonically on $[C_{i-1}, C_{i-1} + \alpha_{i-1}]$, where $0 < \alpha_{i-1} \le C_i - C_{i-1}$, and reaches one at $x = C_{i-1} + \alpha_{i-1}$. $\mu_i(x)$ is one on $[C_{i-1} + \alpha_{i-1}, C_{i+1} - \beta_{i+1}]$, where $0 < \beta_{i+1} \le C_{i+1} - C_i$, and decreases monotonically and becomes zero at $x = C_{i+1}$. $\mu_i(x)$ is zero when x is outside $[C_{i-1}, C_{i+1}]$. For the end points $C_0 = a$ and $C_N = b$, $\mu_0(x)$ is one on $[C_0, C_1 - \beta_1]$, where $0 < \beta_1 \le C_1 - C_0$, and then decreases monotonically to zero at $x = C_1$. $\mu_N(x)$ is zero at $x = C_{N-1}$ increases monotonically to one at $x = C_{N-1} + \alpha_{N-1}$, where $0 \le \alpha_{N-1} \le C_N - C_{N-1}$, and remains one until $x = C_N$. $\mu_0(x)$ and $\mu_N(x)$ are zero elsewhere. $\mu_i(x)$ intersects with only $\mu_{i-1}(x)$ and $\mu_{i+1}(x)$, and only once.

There are K (>1) fuzzy rules in the form

$$\text{IF } x \text{ is } A_i \text{ THEN } y \text{ is } B_i, \quad i = 1, 2, \ldots, K. \tag{1}$$

B_i is a singleton output fuzzy set for the output variable y, whose membership value is one only at $y = V_i$ (an arbitrary constant) and is zero elsewhere. V_i need not be different in different fuzzy rules. Because of SISO, fuzzy logic AND is not needed in fuzzy rules. After being defuzzified by a typical centroid defuzzifier, the output of the SISO fuzzy system is

$$F(x) = y = \frac{\sum_i \mu_i(x) V_i}{\sum_i \mu_i(x)}. \tag{2}$$

The SISO fuzzy systems defined here are general ones because

(i) the input fuzzy sets are quite arbitrary;

(ii) there is no restriction on the type of fuzzy rules used;

(iii) the output fuzzy sets are the widely used singleton ones without constraint;

(iv) the defuzzifier is a popular centroid one.

2.2. *A class of typical MISO fuzzy systems.* There are r independent input variables represented by an input variable

* Received 16 June 1994; revised 1 November 1995; received in final form 26 October 1996. This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor K. P. White under the direction of Editor A. P. Sage.
Corresponding author Dr Hao Ying. Tel. +1 409 772 8415; Fax +1 409 772 0751; E-mail hying@marlin.utmb.edu.

† Department of Physiology and Biophysics, Biomedical Engineering Center, University of Texas Medical Branch, Galveston, TX 77555-0456, U.S.A.

‡ Department of Electrical Engineering, University of Houston, Houston, TX 77204, U.S.A.

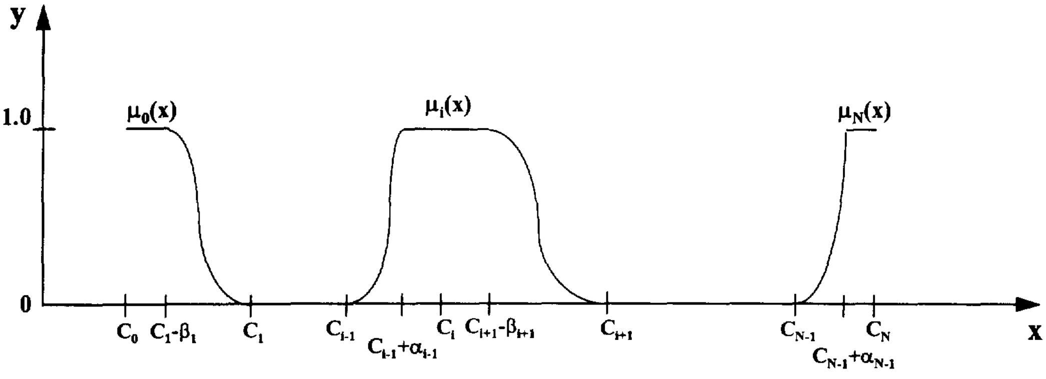

Fig. 1. Illustrative definition of input fuzzy sets for general SISO fuzzy systems.

vector

$$\underset{\sim}{x} = (x_1, x_2, \ldots, x_r),$$

where the domain of x_i is $\Theta_i = [a_i, b_i]$. It follows that the domain of $\underset{\sim}{x}$ is

$$\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_r = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_r, b_r].$$

The interval $[a_i, b_i]$ is divided into N_i subintervals:

$$a_i = C_0^i < C_1^i < C_2^i < \cdots < C_{N_i-1}^i < C_{N_i}^i = b_i, \quad 1 \le i \le r.$$

On each interval Θ_i $(1 \le i \le r)$, $N_i + 1$ $(N_i > 0)$ continuous triangular input fuzzy sets, denoted by A_j^i $(0 \le j \le N_i)$, are defined to fuzzify x_i. The membership function of A_j^i is denoted by $\mu_j^i(x_i)$ and is shown in Fig. 2. It is noted that (i) the two sides of the triangular membership functions are not required to be symmetrical; and (ii) the following holds:

$$\mu_j^i(x_i) + \mu_{j+1}^i(x_i) = 1 \quad \text{for all } i \text{ and } j.$$

We stress that these fuzzy sets are not restrictive. Indeed, they have been widely employed in many practical fuzzy systems (e.g., fuzzy controllers and fuzzy models).

There are K (>1) fuzzy rules in the form

$$\text{IF } x_1 \text{ is } A_{h_1}^1 \text{ AND } x_2 \text{ is } A_{h_2}^2 \text{ AND} \cdots \text{AND } x_r \text{ is } A_{h_r}^r \text{ THEN } y \text{ is } B_{h_1,h_2,\ldots,h_r}, \quad (3)$$

where $B_{h_1,h_2,\ldots,h_r}$ is a singleton output fuzzy set for the output variable y. $B_{h_1,h_2,\ldots,h_r}$ is nonzero only at $y = V_{h_1,h_2,\ldots,h_r}$ (an arbitrary constant). The product fuzzy logic AND is employed to evaluate the ANDs in the fuzzy rules. We use the same centroid defuzzifier given in (2) to obtain the output of the MISO fuzzy systems as follows:

$$F(\underset{\sim}{x}) = y = \frac{\sum (\mu_{h_1}^1 \mu_{h_2}^2 \cdots \mu_{h_r}^r) V_{h_1,h_2,\ldots,h_r}}{\sum (\mu_{h_1}^1 \mu_{h_2}^2 \cdots \mu_{h_r}^r)}. \quad (4)$$

These MISO fuzzy systems are typical and widely used ones in the theory and practice of fuzzy systems (Lee, 1990, 1993; Raju and Zhou, 1993; Qin and Borders, 1994; Li and Gatland, 1995) because arbitrary fuzzy rules, typical triangular input fuzzy sets, arbitrary singleton output fuzzy sets and the centroid defuzzifier are employed.

We should also point out that the SISO and MISO fuzzy systems studied in this paper are not restrictive, not only because they are typical and general fuzzy systems, but also, more importantly, because they are universal approximators (Ying, 1994). For any function approximators, it is always significant to investigate necessary conditions for function approximation, as is well known in the field of mathematical approximation theory.

3. *Statement of the approximation problems for fuzzy systems*

For a MISO fuzzy system (SISO is a special case of MISO), the approximation problem can be stated as follows. Designate $\bar{C}_f$ as the family of r-input one-output continuous functions that have a finite number of extrema defined on the r-dimensional compact domain Θ. Suppose that one is given

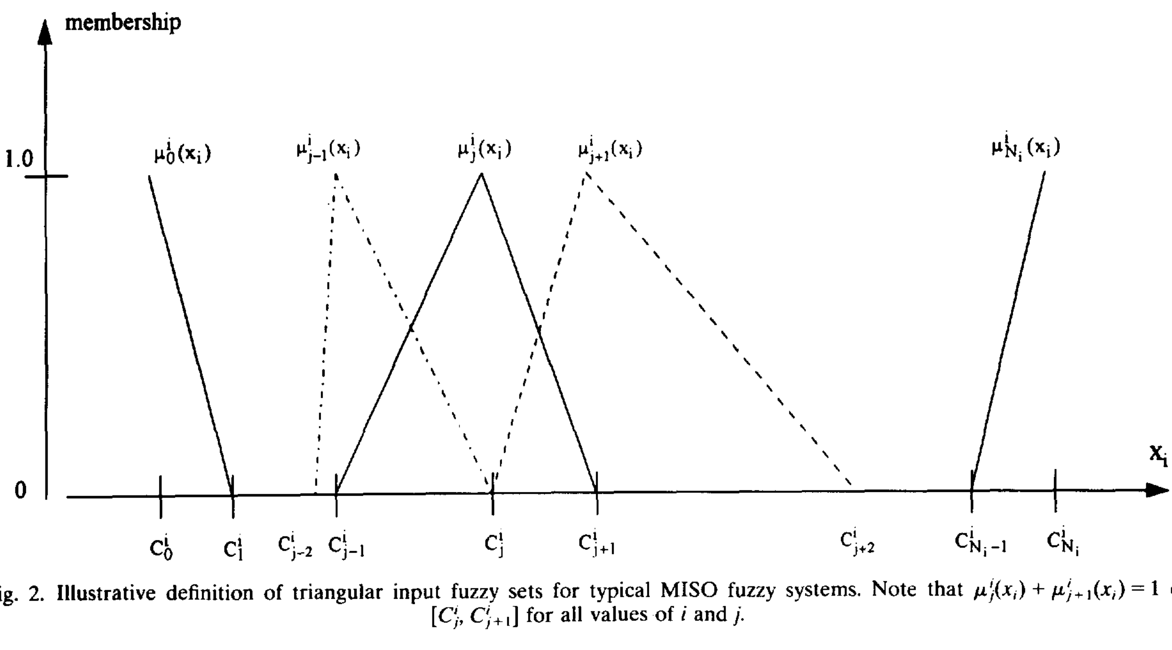

Fig. 2. Illustrative definition of triangular input fuzzy sets for typical MISO fuzzy systems. Note that $\mu_j^i(x_i) + \mu_{j+1}^i(x_i) = 1$ on $[C_j^i, C_{j+1}^i]$ for all values of i and j.

(i) an arbitrarily small approximation error bound $\varepsilon > 0$;

(ii) a set of extrema, $\{m_j\}$, of an arbitrarily selected function $f(\underline{x}) \in \bar{C}_f$, at $\underline{x} = \underline{H}_j = (h_j^1, h_j^2, \ldots, h_j^r) \in (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_r, b_r)$, $j = 1, 2, \ldots, M$;

(iii) the values of $f(\underline{x})$ at $x_i = a_i$ and $x_i = b_i$ for all $i = 1, 2, \ldots, r$ (there are a total of $2r$ such values).

Then, the question is as follows: What are necessary conditions under which there always exists an r-input one-output fuzzy system as defined in Section 2.2, with output $F(\underline{x})$, that satisfies

$$\max_{\underline{x} \in \Theta} |F(\underline{x}) - f(\underline{x})| \leq \varepsilon? \tag{5}$$

As always, one needs to impose certain assumptions on the continuous function to be approximated before establishing necessary conditions for fuzzy systems, or any other types of approximators. Different assumptions require different amount of information be available. If too much information is required, fuzzy-system technology may not be necessary in the first place, since many other well-developed classical functional approximators, such as spline functions, can be used to perform the approximation more efficiently. Bearing this point in mind, we wanted our assumptions to be as less restrictive as possible, and yet still practically sensible. In this paper, we only assume that a set of extrema of $f(x)$ is known, which is probably the minimum amount of information necessary for characterizing major features of a well-behaved continuous function. More importantly, such information can be gained in practice if the continuous function is readily measurable.

Another related question of interest is the following (here we only describe the SISO case, but we study both the SISO and MISO cases in this paper): Given all the extrema of a sequence of continuous functions $f_k(x)$, which can have different numbers of extrema, and a sequence of uniform approximation error bounds ε_k, where $\varepsilon_k > 0$ and $\varepsilon_k \to 0$ as $k \to 0$, what are the necessary conditions under which there always exists a SISO fuzzy system as defined in Section 2.1, with output $F(x)$, that satisfies

$$\max_{x \in [a,b]} |F(x) - f_k(x)| \leq \varepsilon_k \quad \text{for all } k = 1, 2, \ldots. \tag{6}$$

4. *Main results*

4.1. *Necessary conditions for general SISO fuzzy systems as universal approximators.*

Lemma 1. The function $F(x)$ defined in (2) is continuous on $[a, b]$ if and only if the two fuzzy rules 'IF x is A_i THEN y is B_i' and 'IF x is A_{i+1} THEN y is B_{i+1}' are assigned to each subinterval $[C_i, C_{i+1}]$ for all $i = 0, 1, \ldots, N-1$.

Proof. Without losss of generality, we assume that $x \in [C_i, C_{i+1}]$. Only the fuzzy sets A_i and A_{i+1} have nonzero values. Hence only the following two fuzzy rules are activated:

$$\begin{aligned} &\text{IF } x \text{ is } A_i \text{ THEN } y \text{ is } B_i, \\ &\text{IF } x \text{ is } A_{i+1} \text{ THEN } y \text{ is } B_{i+1}. \end{aligned} \tag{7}$$

The output of the fuzzy systems is

$$\begin{aligned} F(x) &= \frac{\mu_i(x)V_i + \mu_{i+1}(x)V_{i+1}}{\mu_i(x) + \mu_{i+1}(x)} \\ &= V_{i+1} + \varphi(x)(V_i - V_{i+1}), \end{aligned} \tag{8}$$

where

$$\varphi(x) = \frac{\mu_i(x)}{\mu_i(x) + \mu_{i+1}(x)} = \frac{1}{1 + \mu_{i+1}(x)/\mu_i(x)}, \tag{9}$$

which satifises $0 \leq \varphi(x) \leq 1$ for all $x \in [C_i, C_{i+1}]$. Because $\mu_i(x)$ and $\mu_{i+1}(x)$ are continuous on $[C_i, C_{i+1}]$, $F(x)$ is continuous on $[C_i, C_{i+1}]$. □

Lemma 2. Suppose that $[a, b]$ is divided into N subintervals $[C_i, C_{i+1}]$, where $i = 0, 1, \ldots, N-1$. The function $F(x)$ defined in (2) is continuous on $[a, b]$ if and only if the two fuzzy rules in the form of (7) are assigned to each of the N subintervals, that is, $K = N + 1$.

Proof. If there exists one subinterval to which no more than one fuzzy rule is assigned then the input x on this subinterval cannot be mapped to $F(x)$. Therefore if $F(x)$ is continuous, in order for the SISO fuzzy systems to map $x \in [a, b]$ to $F(x)$, it is necessary to have two fuzzy rules of (7) to map $x \in [C_i, C_{i+1}]$ to $F(x)$ for all $i = 0, 1, \ldots, N-1$. That means that $N + 1$ distinct fuzzy rules are needed, i.e. $K = N + 1$.

On the other hand, once each subinterval has been assigned a fuzzy rule, $F(x)$ is guaranteed to be continuous by Lemma 1. □

Lemma 3. Assume that $V_i \neq V_{i+1}$. When the conditions set in Lemma 1 are met, the function $F(x)$ defined in (2) is either monotonically increasing or decreasing on $[C_i, C_{i+1}]$ for all $i = 0, 1, \ldots, N-1$.

Proof. As x increases from C_i to C_{i+1}, $\mu_i(x)$ decreases monotonically from 1 to 0 and $\mu_{i+1}(x)$ increases monotonically from 0 to 1, resulting in a monotonic increase of $\mu_{i+1}(x)/\mu_i(x)$ in (9). Consequently the function $\varphi(x)$ decreases monotonically from 1 to 0. According to (8), if $V_i > V_{i+1}$, $F(x)$ decreases monotonically as x increases from C_i to C_{i+1}. Otherwise, $F(x)$ increases monotonically as x increases. □

We remark that $F(x)$ is a continuous convex function on $[C_i, C_{i+1}]$ for all $i = 0, 1, \ldots, N-1$. In Lemma 3, we assume that $V_i \neq V_{i+1}$. This is needed because if $V_i = V_{i+1}$, the two fuzzy rules in (7) will have the same output fuzzy set, namely B_i, for two different input fuzzy sets, resulting in $F(x) \equiv V_i$ instead of a monotonic function on $[C_i, C_{i+1}]$. This situation is avoided in Lemma 3 except when $f(x)$ is a constant on $[C_i, C_{i+1}]$.

Lemma 4. At $x = C_i$, $F(x) = V_i$ for all $i = 0, 1, \ldots, N$.

Proof. When $x = C_i$, where $i = 0, 1, \ldots, N-1$, $\mu_{i+1}(x) = 0$, resulting in $\varphi(x) = 1$. As a result, $F(x) = V_i$. When $x = C_N$, $\mu_{N-1}(x) = 0$, yielding $\varphi(x) = 0$, and consequently $F(x) = V_N$. □

We are now in a position to establish some necessary conditions for general SISO fuzzy systems as universal approximators.

Theorem 1. Given the number of extrema of the continuous function, M, the following necessary conditions must simultaneously be satisfied in order for the general SISO fuzzy systems to achieve the approximation (5):

(a) $[a, b]$ must be divided into at least $M + 1$ subintervals, that is, $N \geq M + 1$;

(b) $M + 1$ of the N subintervals must be such that $C_0 = a$, $C_i = H_i$ $(1 \leq i \leq M)$ and $C_{M+1} = b$;

(c) both fuzzy rules in (7) must be assigned to each subinterval $[C_i, C_{i+1}]$ for all $i = 0, 1, \ldots, M$. That is, the number of fuzzy rules $K = M + 2$. Also, V_i of the output fuzzy set B_i must be so chosen that it satisfies

$$m_i - \varepsilon \leq V_i \leq m_i + \varepsilon \quad \text{for } i = 0, 1, \ldots, M + 1.$$

Proof. We use a contradiction argument to show that $N \geq M + 1$ (condition (a)) is necessary. Suppose that $N < M + 1$. Let the partition of $[a, b]$ be

$$a = D_0 < D_1 < D_2 < \cdots < D_{N-1} < D_N = b.$$

Then there must exist at least one subinterval, say $[D_j, D_{j+1}]$, on which the function $f(x)$ is non-monotonic. That is, $f(x)$ has at least one extremum inside the subinterval but not at the two end points D_j and D_{j+1}. Without losing generality, let us assume that there is one maximum, m_h $(1 \le h \le M)$, at $x = D^* \in (D_j, D_{j+1})$. Suppose, for an arbitrarily small approximation error bound $\varepsilon > 0$, that the following inequality holds:

$$\max_{x \in [D_j, D_{j+1}]} |F(x) - f(x)| \le \varepsilon. \quad (10)$$

This implies that the following three inequalities must hold simultaneously:

$$|F(D_j) - f(D_j)| \le \varepsilon, \quad (11)$$

$$|F(D^*) - m_h| \le \varepsilon, \quad (12)$$

$$|F(D_{j+1}) - f(D_{j+1})| \le \varepsilon. \quad (13)$$

However, that m_h is a maximum means $m_h > f(D_j)$ and $m_h > (D_{j+1})$. Hence $f(x)$ increases monotonically on $[D_j, D^*]$ and decreases monotonically on $[D^*, D_{j+1}]$. According to Lemma 3, $F(x)$ is monotonic on $[D_j, D_{j+1}]$. If $F(x)$ increases monotonically then the inequalities (11) and (12) may hold for any ε, but the inequality (13) cannot hold at the same time if ε is small enough. Similarly, if $F(x)$ is decreasing monotonically, then the inequalities (12) and (13) may hold for an arbitrarily small ε, but the inequality (11) cannot be true simultaneously if ε is small enough. In either case, the inequalities (11)–(13) cannot be true simultaneously. This analysis is also clear geometrically. This contradiction means that $[a, b]$ must be divided into at least $M + 1$ subintervals. That is, $N \ge M + 1$, which is the necessary condition (a).

Furthermore, according to the above analysis, when $N = M + 1$, the subintervals must be divided in such a way that $f(x)$ reaches its M extrema only at $x = H_i$ $(i = 1, 2, \ldots, M)$. That requires $C_i = H_i$ to form $M + 1$ subintervals $[C_i, C_{i+1}]$ for $i = 0, \ldots, M$, which is the necessary condition (b).

Now let us analyze the necessity of condition (c). First of all, according to Lemma 2, it is necessary to assign both fuzzy rules in (7) to each subinterval $[C_i, C_{i+1}]$ for $i = 0, 1, \ldots, M$ so that $F(x)$ is a continuous function. Moreover, to realize the approximation (5), the following inequality must be satisfied:

$$|F(C_i) - f(C_i)| = |F(C_i) - m_i| \le \varepsilon \quad \text{for all } i = 0, 1, \ldots, M + 1. \quad (14)$$

According to Lemma 4, (14) can be rewritten as

$$|V_i - m_i| \le \varepsilon,$$

or

$$m_i - \varepsilon \le V_i \le m_i + \varepsilon \quad \text{for all } i = 0, 1, \ldots, M + 1.$$

□

From Theorem 1, one can see that the selection of V_i of the singleton output fuzzy set B_i depends directly on the given approximation error bound ε. The smaller the value of ε, the narrower the range of the value of V_i. As a limit, $V_i = m_i$ for all i when $\varepsilon = 0$.

Theorem 1 sheds some light on both strength and limitation of general SISO fuzzy systems as universal approximators. We observe that N is related to M in a certain way. Specifically, the number of fuzzy rules needed, N, increases with increasing number of extrema of the continuous function, M. Therefore if M is a small number then N can be a small number. Better yet, N does not relate to the approximation error bound ε. These facts suggest that, even if a given ε is very small, a small number of fuzzy rules may suffice to uniformly approximate those continuous functions that have a complicated formulation but a relatively small number of extrema. This insightful analysis offers a possible explanation of the fact that the majority of practically successful fuzzy controllers and fuzzy models only have to use a small number of fuzzy rules to achieve successful applications.

On the other hand, the limitation of fuzzy systems is also exposed by the fact that the number of fuzzy rules needed increases with increasing number of extrema of the continuous function. A large number of fuzzy rules is necessary for uniform approximation of functions that are simple but have many extrema. For instance, a simple function like $f(x) = \sin(nx)$ has $2|n|$ extrema on $[0, 2\pi]$. If $|n|$ is large then a large number of fuzzy rules is needed by the fuzzy system, according to Theorem 1. This means that fuzzy systems are not ideal function approximators for highly oscillatory functions. Generally, fuzzy systems are not desirable function approximators for applications involving continuous functions that have high-frequency oscillations.

Although, in general, ε and N are not related to each other, as we have pointed out earlier, there are some special situations where ε and n can be related. The use of triangular input fuzzy sets (Fig. 2) is one such special situation. We shall state the result in Corollary 1 below, whose proof requires the following lemma.

Lemma 5. If the triangular input fuzzy sets are used, $F(x)$ defined in (2) either increases linearly, if $V_{i+1} > V_i$, or decreases linearly, if $V_{i+1} < V_i$, on $[C_i, C_{i+1}]$.

Proof. Note that when triangular input fuzzy sets are used, $\mu_i(x) + \mu_{i+1}(x) = 1$ on $[C_i, C_{i+1}]$ for all i. Hence $\varphi(x) = \mu_i(x)$, which is linearly decreasing on $[C_i, C_{i+1}]$. Consequently, (8) becomes

$$F(x) = V_{i+1} + \mu_i(x)(V_i - V_{i+1}), \quad x \in [C_i, C_{i+1}],$$

$F(x)$ increases linearly if $V_{i+1} > V_i$ and decreases linearly if $V_{i+1} < V_i$. □

Corollary 1. When triangular input fuzzy sets are used, a necessary condition for general SISO fuzzy systems to achieve the approximation (6) is that $N \to \infty$.

Proof. According to Lemma 5, when triangular input fuzzy sets are used, $F(x)$ is a linear function in $[C_i, C_{i+1}]$ for all i. Hence if there is at least one subinterval, say $[C_p, C_{p+1}]$, on which $f(x)$ is not a linear function then the subinterval must be further divided into smaller and smaller subintervals to satisfy smaller and smaller approximation error bounds ε_k. Hence, as $\varepsilon_k \to 0$ $(k \to \infty)$, we must have $N \to \infty$. □

We remark that this condition may seem intuitive and trivial when it is considered as a sufficient condition. But we should emphasize that it is also a necessary one. This is significant, since triangular membership functions are commonly used in many practical fuzzy systems reported in the literature. Based on this corollary, we suggest that triangular membership functions may not be used in constructing SISO fuzzy systems as high-precision function approximators such as high-performance fuzzy controllers or accurate fuzzy models.

4.2. *Necessary conditions for typical MISO fuzzy systems as universal approximators*

The following lemma deals with the continuity of the output of the MISO fuzzy systems $F(\underline{x})$ defined in (4) in relation to the assignment of fuzzy rules.

Lemma 6. If the 2^r fuzzy rules in the form of (3) are assigned to each of the $N_1 \times N_2 \times \cdots \times N_r$ different combinations of r subintervals $[C^i_j, C^i_{j+1}]$ (for all i and j) then $F(\underline{x})$ is continuous on the entire compact domain Θ.

Lemma 7. When the conditions in Lemma 6 are met, the denominator of (4) is always equal to one.

Lemma 8. When the conditions in Lemma 6 are met, $F(\underline{x})$ is monotonic on $[C^1_{j_1}, C^1_{j_1+1}] \times [C^2_{j_2}, C^2_{j_2+1}] \times \cdots \times [C^r_{j_r}, C^r_{j_r+1}]$ for all $j_i = 1, 2, \ldots, N_i - 1$, where $i = 1, 2, \ldots, r$.

Proof. For simplicity, we shall only prove the $r = 2$ case. Without loss of generality, we assume that $x_1 \in [C^1_{j_1}, C^1_{j_1+1}]$ and $x_2 \in [C^2_{j_2}, C^2_{j_2+1}]$. Fuzzification of x_1 generates the following:

$$\mu^1_{j_1}(x_1) = \frac{C^1_{j_1+1} - x_1}{C^1_{j_1+1} - C^1_{j_1}}, \quad \mu^1_{j_1+1}(x_1) = 1 - \mu^1_{j_1}(x_1),$$

while fuzzification of x_2 produces

$$\mu^2_{j_2}(x_2) = \frac{C^2_{j_2+1} - x_2}{C^2_{j_2+1} - C^2_{j_2}}, \quad \mu^2_{j_2+1}(x_2) = 1 - \mu^2_{j_2}(x_2).$$

Only four fuzzy rules that relate to these nonzero memberships are executed. After using the product fuzzy logic AND, we get

$$F(\underline{x}) = p_1 x_1 + p_2 x_1 x_2 + p_3 x_2 + p_4, \tag{15}$$

where p_1, p_2, p_3 and p_4 are constants:

$$p_1 = \frac{-V_{j_1,j_2}C^2_{j_2+1} + V_{j_1,j_2+1}C^2_{j_2} + V_{j_1+1,j_2}C^2_{j_2+1} - V_{j_1+1,j_2+1}C^2_{j_2}}{(C^1_{j_1+1} - C^1_{j_1})(C^2_{j_2+1} - C^2_{j_2})},$$

$$p_2 = \frac{V_{j_1,j_2} - V_{j_1,j_2+1} - V_{j_1+1,j_2} + V_{j_1+1,j_2+1}}{(C^1_{j_1+1} - C^1_{j_1})(C^2_{j_2+1} - C^2_{j_2})},$$

$$p_3 = \frac{-V_{j_1,j_2}C^1_{j_1+1} + V_{j_1,j_2+1}C^1_{j_1+1} + V_{j_1+1,j_2}C^1_{j_1} - V_{j_1+1,j_2+1}C^1_{j_1}}{(C^1_{j_1+1} - C^1_{j_1})(C^2_{j_2+1} - C^2_{j_2})},$$

$$p_4 = \frac{\begin{array}{l} V_{j_1,j_2}C^1_{j_1+1}C^2_{j_2+1} - V_{j_1,j_2+1}C^1_{j_1+1}C^2_{j_2} \\ \quad - V_{j_1+1,j_2}C^1_{j_1}C^2_{j_2+1} + V_{j_1+1,j_2+1}C^1_{j_1}C^2_{j_2} \end{array}}{(C^1_{j_1+1} - C^1_{j_1})(C^2_{j_2+1} - C^2_{j_2})}. \tag{16}$$

$F(\underline{x})$ has no extremum on $[C^1_{j_1}, C^1_{j_1+1}] \times [C^2_{j_2}, C^2_{j_2+1}]$, because

$$\Delta = \frac{\partial^2 F}{\partial x_1^2}\frac{\partial^2 F}{\partial x_2^2} - \left(\frac{\partial^2 F}{\partial x_1\, \partial x_2}\right)^2 = -p_2^2 \le 0.$$

For any $p_2 \ne 0$, $\Delta < 0$. If $p_2 = 0$, $F(\underline{x})$ in (15) becomes

$$F(\underline{x}) = p_1 x_1 + p_3 x_2 + p_4,$$

which is a two-dimensional plane. We conclude that $F(\underline{x})$ has no extrema, and hence is monotonic on $[C^1_{j_1}, C^1_{j_1+1}] \times [C^2_{j_2}, C^2_{j_2+1}]$. □

Recall that $F(x)$ for general SISO fuzzy systems using triangular input fuzzy sets is a piecwise-linear continuous function. However, according to (15), $F(\underline{x})$ for MISO fuzzy systems employing triangular input fuzzy sets is generally not a piecewise-linear function. $F(\underline{x})$ becomes piecewise-linear if and only if $p_2 = 0$ in (15). Let $p_2 = 0$ in (16); then we get

$$V_{j_1,j_2} - V_{j_1,j_2+1} - V_{j_1+1,j_2} + V_{j_1+1,j_2+1} = 0. \tag{17}$$

That is, only when the fuzzy rules satisfy (17) does $F(\underline{x})$ become a piecewise-linear function. Equation (17) can be rewritten as

$$V_{j_1+1,j_2+1} - V_{j_1,j_2+1} = V_{j_1+1,j_2} - V_{j_1,j_2} \tag{18}$$

or

$$V_{j_1+1,j_2+1} - V_{j_1+1,j_2} = V_{j_1,j_2+1} - V_{j_1,j_2}, \tag{19}$$

which imply linearity in the output fuzzy sets of the fuzzy rules. The only type of fuzzy rule that satisfies (18) and (19) (or, equivalently, (17)) is the linear fuzzy rule (Ying, 1993)

$$\text{IF } x_1 \text{ is } A^1_{j_1} \text{ AND } x_2 \text{ is } A^2_{j_2} \text{ THEN } y \text{ is } B_{\alpha j_1 + \beta j_2 + \gamma}, \tag{20}$$

where α, β and γ are three arbitrary constants that make $\alpha j_1 + \beta j_2 + \gamma$ an integer, since the indices of the output fuzzy sets must be integers. We have previously proved that a fuzzy system (e.g., a controller) using the product fuzzy logic AND and linear fuzzy rules (20) is a globally linear system (Buckley and Ying, 1990). In Buckley and Ying (1990), however, the linear fuzzy rule (20) was used only as a sufficient condition. Here, we show that it is also a necessary one.

Theorem 2. The output $F(\underline{x})$ of typical MISO fuzzy systems becomes a globally linear function if and only if the linear fuzzy rule (20) is employed.

In the light of Theorem 2, we remark that linear fuzzy rules should not be used for MISO fuzzy systems unless the function to be approximated is a piecewise-linear one—but in this case fuzzy-system approximation may not need to be used in the first place.

The following is the last lemma that we need to establish some necessary conditions for using typical MISO fuzzy systems as universal approximators.

Lemma 9. $F(\underline{x}) = V_{j_1,j_2,\ldots,j_r}$ at $\underline{x} = (C^1_{j_1}, C^2_{j_2}, \ldots, C^r_{j_r})$ for all $j_1 = 1, 2, \ldots, N_i$, where $i = 1, 2, \ldots, r$.

Proof. At $\underline{x} = (C^1_{j_1}, C^2_{j_2}, \ldots, C^r_{j_r})$, $\mu^i_{j_i}(x_i) = 1$ and $\mu^i_{j_i+1}(x_i) = 0$, for all $j_i = 1, 2, \ldots, N_i - 1$. Therefore only one fuzzy rule is activated:

IF x_1 is $A^1_{j_1}$ AND x_2 is $A^2_{j_2}$ AND ··· AND x_r is $A^r_{j_r}$ THEN y is $B_{j_1,j_2,\ldots,j_r}$.

After using the product fuzzy logic AND, $B_{j_1,j_2,\ldots,j_r}$ is assigned a membership value of 1, and consequently $F(\underline{x}) = V_{j_1,j_2,\ldots,j_r}$. It is then straightforward to prove that $F(\underline{x}) = V_{N_1,N_2,\ldots,N_r}$ at $\underline{x} = (C^1_{N_1}, C^2_{N_2}, \ldots, C^r_{N_r})$. □

Recall that when formulating (5), we assumed that M extrema of $f(\underline{x})$, at $\underline{x} = \underline{H}_j = (h^1_j, h^2_j, \ldots, h^r_j)$, where $j = 1, 2, \ldots, M$, are given. Note that, for a fixed $1 \le i \le r$, we have M points inside the interval $[a_i, b_i]$ but not at the two end points: $S_i = (h^i_1, h^i_2, \ldots, h^i_M)$, where the values of some of the points may be the same. We shall keep only the distinct points in S_i and rearrange them in the ascending order to form the following new sets:

$$T_i = (\theta^i_1, \theta^i_2, \ldots, \theta^i_{K_i}), \quad i = 1, 2, \ldots, r, \tag{21}$$

where $\theta^i_1 < \theta^i_2 < \cdots < \theta^i_{K_i}$. Here we suppose that T_i has K_i $(1 \le K_i \le M)$ distinct points. Obviously, T_i divides $[a_i, b_i]$ into $K_i + 1$ subintervals.

Now, using the notation T_i defined in (21) and the lemmas on MISO fuzzy systems, we establish the following necessary conditions for MISO fuzzy systems.

Theorem 3. Given K_i $(i = 1, 2, \ldots, r)$, the following necessary conditions for typical MISO fuzzy systems to achieve the approximation (5) must be simultaneously satisfied:

(a) for all $i = 1, 2, \ldots, r$, the interval $[a_i, b_i]$ must be divided into at least $K_i + 1$ subintervals, that is, $N_i \ge K_i + 1$;

(b) $K_i + 1$ of the N_i subintervals must be formed such that $C^i_0 = a_i$, $C^i_j = \theta^i_j$ $(1 \le j \le K_i)$ and $C^i_{K_i+1} = b_i$;

(c) 2^r fuzzy rules of (3) must be assigned to each of the r-dimensional cubes $[C^1_{j_1}, C^1_{j_1+1}] \times [C^2_{j_2}, C^2_{j_2+1}] \times \cdots \times [C^r_{j_r}, C^r_{j_r+1}]$ for all $j_i = 0, 1, \ldots, K$, that is, the number of fuzzy rules $K = \prod_{i=1}^r (K_i + 2)$; also, $V_{j_1,j_2,\ldots,j_r}$ of the singleton output fuzzy set $B_{j_1,j_2,\ldots,j_r}$ must be so chosen that it satisfies

$$m_{j_1,j_2,\ldots,j_r} - \varepsilon \le V_{j_1,j_2,\ldots,j_r} \le m_{j_1,j_2,\ldots,j_r} + \varepsilon.$$

Proof. The proof is similar to that of Theorem 1, and hence is omitted. □

The limitation of general SISO fuzzy systems as universal approximators, pointed out after the proof of Theorem 1, is entirely extensible to typical MISO fuzzy systems because of the shared nature of Theorems 3 and 1. That is to say, typical MISO fuzzy systems are not ideal approximators for highly oscillatory continuous functions.

5. *Conclusions*

We have established some necessary conditions for general SISO fuzzy systems and a class of typical MISO fuzzy systems as universal approximators. We have demonstrated how input fuzzy sets, output fuzzy sets and fuzzy rules of fuzzy systems should be determined according to a given approximation error bound and the extrema of the continuous function to be approximated so that the uniform approximation *may* be achieved. Because the assumption that we imposed on the continuous function is minimal and realistic, the necessary conditions derived in this paper can be used in many applications.

Our insightful analysis utilizing the necessary conditions established here reveals the strengths as well as the limitations of fuzzy systems. On the one hand, even if a given approximation error bound is very small, a small number of fuzzy rules may suffice to uniformly approximate those continuous functions that have a complicated formulation but a relatively small number of extrema. On the other hand, a large number of fuzzy rules are necessary for uniform approximation of functions that are simple but have many extrema. Compared with other approximation techniques, fuzzy systems are not desirable function approximators for applications (e.g. control and modeling) involving continuous functions that have high-frequency oscillations. This is especially true for SISO and MISO fuzzy systems that use triangular input fuzzy sets.

Whether the necessary conditions obtained in this paper are applicable or extendible to other MISO fuzzy systems is an open question at present, and warrants further research. Our results can serve as a stepping stone toward the answer to this important question, particularly for MISO fuzzy systems.

References

Buckley, J. J. (1992) Universal fuzzy controllers. *Automatica* **28,** 1245–1248.

Buckley, J. J. and Ying, H. (1990) Linear fuzzy controller: it is a linear non-fuzzy controller. *Inf. Sci.* **51,** 183–192.

Kosko, B. (1992) Fuzzy systems as universal approximators. In *Proc. IEEE International Conf. on Fuzzy Systems,* San Diego, CA, pp. 1153–1162.

Lee, C. C. (1990) Fuzzy logic in control systems: fuzzy logic controller—Parts I & II. *IEEE Trans. Syst. Man. Cyber.* **SMC-20,** 404–435.

Lee, J. (1993) On methods for improving performance of PI-type fuzzy logic controllers. *IEEE Trans. Fuzzy Syst.* **FS-1,** 298–301.

Lee, J. and Chae, S. (1993) Analysis on function duplicating capabilities of fuzzy controllers. *Fuzzy Sets Syst.* **56,** 127–143.

Li, H. and Gatland, H. B. (1995) A new methodology for designing a fuzzy logic controller. *IEEE Trans. Syst. Man, Cyber.* **SMC-25,** 505–512.

Marks, R. J. II (ed.) (1993) *Fuzzy Logic Technology and Applications.* IEEE Press, New York.

Qin, S. J. and Borders, G. (1994) A multiregion fuzzy logic controller for nonlinear process control. *IEEE Trans. Fuzzy Syst.* **FS-2,** 74–81.

Raju, G. V. S. and Zhou, J. (1993) Adaptive hierarchical fuzzy control. *IEEE Trans. Syst. Man, Cyber.* **SMC-23,** 973–980.

Wang, L. X. (1992) Fuzzy systems are universal approximators. In *Proc. IEEE International Conf. on Fuzzy Systems,* San Diego, CA, pp. 1163–1170.

Wang, L. X. and Mendel, J. M. (1992) Fuzzy basis functions, universal approximation, and orthogonal least-squares learning. *IEEE Trans. Neural Networks* **NN-3,** 807–814.

Ying, H. (1993) General analytical structure of typical fuzzy controllers and their limiting structure theorems. *Automatica* **29,** 1139–1143.

Ying, H. (1994) Sufficient conditions on general fuzzy systems as function approximators. *Automatica* **30,** 521–525.

Zeng, X. J. and Singh, M. G. (1994) Approximation theory of fuzzy systems—SISO case. *IEEE Trans. Fuzzy Syst.* **FS-2,** 162–176.