A New ANFIS Model based on Multi-Input Hamacher T-norm and Subtract Clustering

Feng-Yi Zhang* and Zhi-Gao Liao

Department of Management, Guangxi University of Science and Technology, Liuzhou, Guangxi, 545000, China

Abstract: This paper proposed a novel adaptive neuro-fuzzy inference system (ANFIS), which combines subtract clustering, employs adaptive Hamacher T-norm and improves the prediction ability of ANFIS. The expression of multi-input Hamacher T-norm and its relative feather has been originally given, which supports the operation of the proposed system. Empirical study has testified that the proposed model overweighs early work in the aspect of benchmark Box-Jenkins dataset and may provide a practical way to measure the importance of each rule.

Keywords: ANFIS, Hamacher T-norm, Subtract Clustering, T-norm.

1. INTRODUCTION

Takagi, Sugeno and Kang established a method called the Takagi-Sugeno-Kang (TSK) method [1-3]. This neural-network-based fuzzy reasoning scheme is capable of learning the membership function of the "IF" part and determining the amount of control in the "THEN" part of the inference rules. Moreover, it is nicely suited to mathematical analysis and usually works well with optimization and adaptive techniques. Subsequently, many improved algorithms and extensions were developed for the TSK model. In particular, the adaptive neuro-fuzzy inference system (ANFIS) is an important approach to implement the TSK fuzzy system, which has been put forward by Jang in 1993 [4]. However among all the works which have been conducted earlier, only Iliadis *et al*. replaced the algebraic product T-norm with other fuzzy T-norms to handle intersection operation [5]. But they did not explain the reason why they selected the one. Besides, in the case of input dimension increasing, the number of rules increase with the input dimension exponentially, which inevitably leads the conventional ANFIS structure dimension to disaster. In order to improve the online access speed of ANFIS T-S rules for complex system, various clustering algorithms have been used to construct a new multidimensional structure of ANFIS, which combines mechanism of T-S fuzzy inference and clustering algorithm from the perspective of knowledge discovery. In this paper, Hamacher T-norm has been selected to tackle the intersection operation for two reasons:

1. Algebraic product T-norm is used widely in ANFIS, and when λ equals to 1. The Hamacher T-norm is actually an algebraic product T-norm, which means that it is not conflicting with the regular ANFIS.
2. Hamacher product T-norm, a clustering of fuzzy product T-norms, differs depending on λ. So, to select the most suitable fuzzy T-norm by changing the numberical parameter λ is advisable. Subtract clustering, which could obtain the amount and value of clustering center, was used to determine the If part of each rule, for its wide application in ANFIS center determination.

The rest of the paper is organized as follows. Section 2 provides some necessary background information, and the proposed system and its essential interference are discussed in Section 3. Section 4 presents the simulation results for benchmark Box-Jenkins dataset. Finally, the summary of this paper is given in Section 5.

2. BACKGROUND

In this section, the basic theory of ANFIS model and normalization method which has been used in this experiment is introduced.

2.1. Adaptive Network Based Fuzzy Inference System (ANFIS)

Both artificial neural network and fuzzy logic are used in ANFIS architecture. ANFIS consists of if-then rules and couples of input-output. For ANFIS training, learning algorithms of neural network are also used. To simplify the explanations, the fuzzy inference system under consideration is assumed to have two inputs (x and y) and one output (f). For a regular ANFIS model, a typical rule set with basic fuzzy if-then rules can be expressed as if x is A_1 and y is B_1, then

$$f_1 = p_1 x + q_1 y + r_1 \tag{1}$$

where p is linear output parameter. The ANFIS architecture with two inputs and one output are shown in Fig. (**1**).

*Address correspondence to this author at the Department of Management, Guangxi University of Science and Technology, Liuzhou, Guangxi, 545000, China; E-mail: 798118862@qq.com

1874-155X/14 2014 Bentham Open

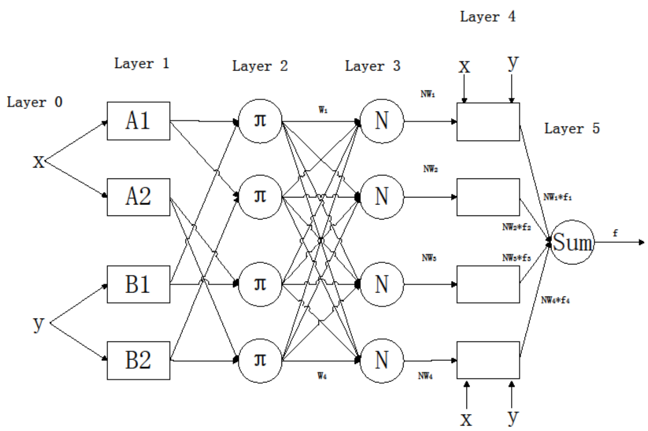

Fig. (1). The structure of regular ANFIS.

This architecture is formed by five layers and nine if-then rules:

Layer-1: Every node i in this layer is a square node with a node function.

$$O_{1,i} = \mu_{A_i}(x), O_{1,3+j} = \mu_{B_j}(y) \quad i,j = 1,2 \tag{2}$$

where x and y are inputs to node i, and A_i and B_i are linguistic labels for inputs. In other words, $O_{1,i}$ is the membership function of A_i and B_i. Usually μ_{Ai} (x) and μ_{Bi} (y) are chosen to be bell-shaped with maximum equaling to 1 and minimum equaling to 0, such as

$$\mu_{A_i}(x) = exp(-((x - a_i)/(c_i))^2) \tag{3}$$

where a_i, c_i is the parameter set. These parameters in this layer are referred to as premise parameters.

Layer-2: Every node in this layer multiplies the incoming signals and sends the product out. For instance,

$$O_{2,2(i-1)+j} = \mu_{A_i}(x) * \mu_{B_j}(y), i,j = 1,2 \tag{4}$$

Each node output represents the firing strength of a rule.

Layer-3: Every node in this layer calculates the ratio of the rule's firing strength to the sum of all ruless firing strengths:

$$O_{3,i} = \tilde{w}_i = w_i / (w_1 + w_2 + \cdots + w_9), i = 1,2,\cdots,4 \tag{5}$$

Layer-4: Every node i in this layer is a square node with a node function

$$O_{4,i} = \tilde{w}_i * f_i = \tilde{w}_i(p_{i,1}x_1 + p_{i,2}x_2), i = 1,2,\cdots,4 \tag{6}$$

where w_i is the output of layer 3 and $p_{i,1}$, $p_{i,2}$, $p_{i,3}$ is the parameter set. Parameters in this layer are referred to as consequent parameters.

Layer-5: The single node in this layer computes the overall output as the summation of all incoming signals:

$$O_{5,i} = \sum \tilde{w}_i * f_i = \frac{\sum w_i * f_i}{\sum w_i} \tag{7}$$

2.2. Hamacher T-Norm

Hamacher T-norm as a kind of T-norm with parameter satisfies the boundary conditions, commutativity, associativity and monotonicity. The parameter of Hamacher T-norm is also monotonous, and its expression is given below:

$$T_\lambda(x,y) = \frac{xy}{\lambda + (1-\lambda)(x + y - xy)} \tag{8}$$

where $\lambda > 0$. Especially, when $\lambda = 1$, Hamacher T-norm equals to algebraic product T-norm.

It is easy to recognise that algebraic product T-norm is a special Hamacher T-norm which has a constant parameter λ. However, employing a constant parameter λ is not always appropriate. For any rule, there must be a corresponding parameter λ suited for it. It is wise to use back-propagation algorithm to determine the corresponding λ.

3. PROPOSED SYSTEM

The output of layer-2 $O_{2,2(i-1)+j}$ refers to the result of intersection operation between $\mu_{A_i}(x)$ and $\mu_{B_j}(y)$, which means the membership degree that x_1 belongs to A_i and x_2 belongs to B_j. It is common to use algebraic product T-norm "$*$" to deal with the membership degree in intersection operation, but as is well-known that algebraic product T-norm is not proper in any situation. A study shows that algebraic product T-norm is a special Hamacher T-norm whose parameter is constant to 1 (8). So modifying the parameter to suit to the data pairs is a meaningful way to overcome the dilemma. It is not easy to determine the value of λ that should be served in Hamacher T-norm to handle intersection operation. Iliadis *et al.* tried to use other constant λ to obtain better performance but not all always resulted in good situation [6]. It is a good solution to make ANFIS to adaptively select its own λ for each rule. If ANFIS could select λ for each rule respectively, according to the training data pairs, it is more likely to fit to the performance curve and close to the inherent law. Back-propagation algorithm could be adopted in the process of determining the parameter of each rule, but this method needs to obtain $\frac{\partial T_\lambda(x,y)}{\partial x}$ and $\frac{\partial T_\lambda(x,y)}{\partial \lambda}$ which is the gradient of $T_\lambda(x,y)$.

3.1. Multi-Input Hamacher T-Norm

Calculating $\frac{\partial T_\lambda(x,y)}{\partial x}$ and $\frac{\partial T_\lambda(x,y)}{\partial \lambda}$ is easy; ANFIS may have more than two inputs and how to calculate their gradients is a real problem. More attention should be paid on how to calculate their gradients with more than 2 inputs. Now the definition of multi-input Hamacher T-norm is given below.

$T_\lambda(A_n)$ is multi-input Hamacher T-norm on A_n which has n elements, where $A_n = \{a_1, a_2, \cdots, a_n\}$ and $\forall i \in N^+, 2 \le n, 0 \le a_i \le 1$. $T_\lambda(A_n) = T_\lambda(T_\lambda(A_{n-1}), a_n)$. Especially, when n=2, $T_\lambda(A_2) = T_\lambda(a_1, a_2)$.

The definition given above is recursive definition. In other words, the meaning of upper layer is corresponding to the lower one and the lowest is clarified. To express it clearly, a useful tool $\chi^j(A_n)$ has been used. The definition and features are given below: $\chi^j(A_n) = \sum_{c_1,\cdots,c_j \in \{1,\cdots,n\}, c_1 \ne \cdots \ne c_j} a_{c_1} a_{c_2} a_{c_3} \cdots a_{c_j}$, where $n \in N^+$, n ≥ 2 and $j \in N$, $j \le n$. Especially, $\chi^0(A_n) = 1$. For example, $\chi^2(A_4) = a_1\ a_2 + a_1\ a_3 + a_1\ a_4 + a_2\ a_3 + a_2\ a_4 + a_3\ a_4$.

Corollary 1 When $j \ne n$,

$$\chi^j(A_n) = \chi^j(A_{n-1}) + a_n \chi^{j-1}(A_{n-1})$$

The proof is given in Appendix A.

Corollary 2 When j = n,

$$\chi^n(A_n) = a_n \chi^{n-1}(A_{n-1})$$

The proof is given in Appendix A.

Corollary 3 $\frac{\partial \chi^j(A_n)}{\partial a_k} = \chi^{j-1}(A_n \setminus a_k)$, where j,k $\in$ N and j,k$\ne$ n. $A_n \setminus a_k = \{a_1, a_2, \cdots, a_{k-1}, a_{k+1}, \cdots, a_n\}$

The proof is given in Appendix A.

One evident feature of multi-input Hamacher T-norm is the monotonicity with respect to λ. After concise proof, the feature and the expression of $T_\lambda(A_n)$ are confirmed below:

Proposition 1 $T_\lambda(A_n)$ is decreasing with respect to λ. Especially, when $\forall i \in [1, n+1]$ $a_i \ne 1$ and $a_i \ne 0$, $T_\lambda(A_n)$ are strictly decreasing with respect to λ.

The proof is given in Appendix B.

Proposition 2 $T_\lambda(A_n) =$

$$\frac{\chi^n(A_n)}{\lambda^{n-1} + \sum_{j=1}^{n-1} \lambda^{n-j-1}(1-\lambda)^j \chi^j(A_n) - \sum_{i=1}^{n-1}(1-\lambda)^i \chi^n(A_n)}$$

The proof is given in Appendix B

Proposition 3 $\frac{\partial T_\lambda(A_n)}{\partial \lambda} = \frac{-\chi^n(A_n) R_n}{Q_n^2}$, where

$$Q_n = \lambda^{n-1} + \sum_{i=1}^{n-1} \lambda^{n-i-1}(1-\lambda)^i \chi^i(A_n) - \sum_{i=1}^{n-1}(1-\lambda)^i \chi^n(A_n)$$

and

$$R_n = \begin{matrix} (n-1)\lambda^{n-2} + \sum_{i=1}^{n-1} \lambda^{n-i-2}(1-\lambda)^{i-1}[(n-1)-(n-1)\lambda - i]\chi^i(A_n) + \\ \sum_{i=1}^{n-1} i(1-\lambda)^{i-1} \chi^n(A_n) \end{matrix}$$

The proof is given in Appendix B.

3.2. Hamacher T-Norm and Subtract Clustering Based ANFIS

The proposed model Fig. (**2**) differs from regular ANFIS in two points:

1. It makes Hamacher parameter variable and adaptive by adopting back-propagation algorithm, and needs to calculate the gradient with respect to each parameter and input.
2. It is combined with subtract clustering and employs it to determine the amount and value of each rule.

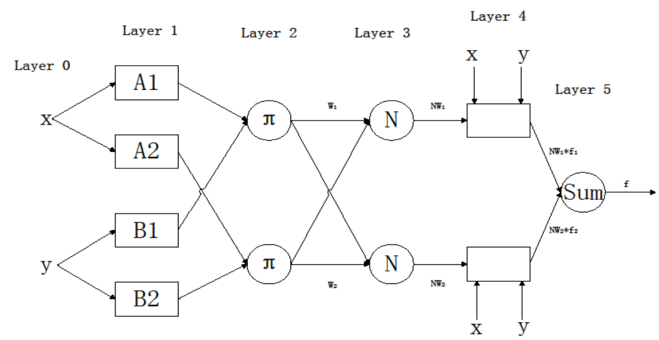

Fig. (2). The structure of ANFIS combined with subtract clustering.

As provided above, the gradient of Hamacher T-norm's parameter and inputs have been achieved by Proposition 3 and Proposition 4. Different from the regular ANFIS, the output of layer-2 for a proposed model which has 3 inputs is given below:

$$O_{2,3^2(i-1)+3(j-1)+k} = T_\lambda(A_3) \tag{9}$$

where $A_3 = \{\mu_{A_i}(x), \mu_{B_j}(y), \mu_{C_k}(z)\}$.

Each rule has same position in regular ANFIS, because their λ have been uniformly set to 1, which hammers the system to find the most significant rule adaptively. However, proposed model's each rule with different λ in the end can lay the foundation for measuring its importance. Both the IF part and the THEN part correlate to the λ and the principle of updating λ is to minimize the error, which guarantees that updated λ is harmonious to the system. w_i is the weight of i th rule and decreases with respect to λ according to Proposition 1 and Equation (9). It means that the less λ leads to larger w_i, larger $\tilde{w}_i$, so the i th rule plays a more important role in the proposed model.

In addition, this model involves the field that the others have never touched upon. This field is attached to the improvement in fuzzy reasoning, and it could be combined with the improvement both in fuzzy reasoning and in other process, because it provides a new methodology for handling intersection operation.

With the variable and adaptive parameter, the prediction ability of proposed model may be improved; the parameter is modified according to the gradient and so as to fit to the inherent law. Empirical study is given in the next section, which proves that the proposed one overweighs the regular ANFIS.

4. EXPERIMENT

Famous Box-Jenkins dataset is the benchmark dataset to validate the performance of fitting method. The Box-Jenkins dataset represents the CO_2 concentration as output, y(t), in terms of input gas flow rate, u(t), forming a combustion process of a methane air mixture. A number of works have been carried out earlier on fitting Box-Jenkins dataset. Among them, 7 input-type has been widely used:(A) u(k-4), y(k-1); (B) u(k-3), y(k-1); (C) u(k-3), u(k-4), y(k-1); (D) u(k), u(k-1), y(k-1), y(k-2); (E) u(k-1), u(k-2), y(k-1), y(k-2); (F) u(k), u(k-1), u(k-2), y(k-1), y(k-2), y(k-3); (G) u(k-1), u(k-2), u(k-3), y(k-1), y(k-2), y(k-3).

The whole experiment was undertaken in the environment of Matlab 7.8.0. The results of proposed model and early works are listed in Table **1**, where it can be found that the proposed model has an outstanding performance.

Table 1. The results of each model.

Model	Input-Type	MSE
PMDE [7]	F	0.1247
PMGA [7]	F	0.3508
TS-GMDH1 [8]	G	0.1299
TS-GMDH2 [8]	G	0.2197
TS-GMDH3 [8]	G	0.3310
Mejias [4]	B	0.3129
Yue [9]	A	0.1480
Yue [9]	B	0.1240
Yue [9]	C	0.1030
Yue [9]	F	0.0460
Yue [9]	G	0.0420
Proposed Model	F	0.0325
Proposed Model	G	0.0301

CONCLUSION

Hamacher T-norm is one of the most influential T-norms. In this paper, the feasibility of applying ANFIS implemented with Hamacher T-norm is investigated. While employing benchmark Box-Jenkins dataset, the proposed methods have a more competitive performance in prediction accuracy compared to early work.

There are two main advantages of the proposed model: on one hand, it is the extention of ANFIS in fuzzy reasoning, which makes it possible to improve when implemented with other improvement in fuzzification, defuzzification, even training method and other optimal algorithms such as GA and PSO; on the other hand, it provides a very vital parameter λ to infer the importance of each rule, but the normal form of inferring and measuring has not been proposed. The study of all above expectations is in progress.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

The authors would like to thank the support by innovation Project of Guangxi Graduate Education (YCSZ2014203).

APPENDIX A. PROOF OF COROLLARY

Proof of Corollary 1 When $0 \le j < n$, $1 \le i \le n$

$$\chi^j(A_n) = \sum_{c_1,\cdots,c_j \in \{1,\cdots,i-1,i+1,\cdots,n\}, c_1 \ne \cdots \ne c_j} (a_{c_1} a_{c_2} a_{c_3} \cdots a_{c_j})$$

$$= \sum_{c_1,\cdots,c_j \in \{1,\cdots,i-1,i+1,\cdots,n\}, c_1 \ne \cdots \ne c_j} (a_{c_1} a_{c_2} a_{c_3} \cdots a_{c_j})$$

$$+ a_i \sum_{c_1,\cdots,c_j-1 \in \{1,\cdots,i-1,i+1,\cdots,n\}, c_1 \ne \cdots \ne c_{j-1}} (a_{c_1} a_{c_2} a_{c_3} \cdots a_{c_{j-1}})$$

$$= \chi^j(A_n \setminus a_i) + a_i \chi^{j-1}(A_n \setminus a_i)$$

which completes the proof.

Proof of Corollary 2 When $j \ne n$,

$$\chi^n(A_n) = a_1 a_2 a_3 \cdots a_n$$

$$= a_n (a_1 a_2 a_3 \cdots a_n)$$

$$= a_n \chi^n(A_n)$$

which completes the proof.

Proof of Corollary 3 When $0 \le j < n$, $1 \le i \le n$

$$\frac{\partial \chi^j(A_n)}{\partial a_i} = \frac{\partial \chi^j(A_n \setminus a_i)}{\partial a_i} + \frac{\partial (a_i \chi^{j-1}(A_n \setminus a_i))}{\partial a_i}$$

$$= \chi^{j-1}(A_n \setminus a_i)$$

which completes the proof.

APPENDIX B. PROOF OF PROPOSITION

Proof of Proposition 1 $\forall$ $\lambda_1, \lambda_2 \in [0, +\infty]$ and $\lambda_1 < \lambda_2$, when n = 1, $T_{\lambda_1}(A_2) = T_{\lambda_1}(a_1, a_2) \ge T_{\lambda_2}(a_1, a_2) \ge T_{\lambda_2}(A_2)$. Especially, when $a_1, a_2 \ne 1$ and $a_1, a_2 \ne 0$, $T_\lambda(A_n)$ is strictly decreasing with respect to λ. The proposition is confirmed.

Assuming when n=t-1, the proposition is right too, then when n=t, $T_{\lambda_1}(A_{t+1}) = T_{\lambda_1}(T_{\lambda_1}(A_t), a_{t+1}) \ge T_{\lambda_1}(T_{\lambda_2}(A_t), a_{t+1})$

$\geq T_{\lambda_2}(T_{\lambda_2}(A_t),a_{t+1}) = T_{\lambda_2}(A_{t+1})$. Especially, when $\forall i\in[0,n+1]$ $a_i\neq 1$ and $a_i\neq 0$, $T_\lambda(A_n)$ strictly decreased with respect to λ.

which completes the proof.

Proof of Proposition 2 The proof is given below:

Let $Q_n=\lambda^{n-1}+\sum_{j=1}^{n-1}\lambda^{n-j-1}(1-\lambda)^j\chi^j(A_n)-\sum_{i=1}^{n-1}(1-\lambda)^i\chi^n(A_n)$,

so $T_\lambda(A_n)=\dfrac{\chi^n(A_n)}{Q_n}$.

when $n=2$,

$$T_\lambda(A_2)=\frac{a_1a_2}{\lambda+(1-\lambda)(a_1+a_2-a_1a_2)}$$

$$=\frac{a_1a_2}{\lambda+(1-\lambda)(a_1+a_2)+(\lambda-1)a_1a_2}$$

$$=\frac{\chi^2(A_2)}{\lambda^{2-1}+\sum_{j=1}^{2-1}\lambda^{1-j}(1-\lambda)^j\chi^j(A_2)-\sum_{i=1}^{2-1}(\lambda-1)^i\chi^2(A_2)}$$

$$=\frac{\chi^2(A_2)}{Q_2}$$

The proposition is right.

Assuming when n = t, the proposition is right too. So, $T_\lambda(A_t)=\dfrac{\chi^t(A_t)}{Q_t}$.

$$T_\lambda(A_{t+1})=\frac{\dfrac{\chi^t(A_t)}{Q_t}a_{t+1}}{\lambda+(1-\lambda)(\dfrac{\chi^t(A_t)}{Q_t}+a_{t+1})-(1-\lambda)a_{t+1}\dfrac{\chi^t(A_t)}{Q_t}}$$

$$=\frac{\chi^{t+1}(A_{t+1})}{\lambda Q_t+(1-\lambda)\chi^t(A_t)-(1-\lambda)\chi^{t+1}(A_{t+1})+(1-\lambda)a_{t+1}Q_t}$$

$$\lambda Q_t=\lambda\sum_{j=0}^{t-1}\lambda^{t-j-1}(1-\lambda)^j\chi^j(A_t)+\lambda\frac{(1-\lambda)^t-(1-\lambda)}{\lambda}\chi^t(A_t)$$

$$=\sum_{j=0}^{t-1}\lambda^{t-j}(1-\lambda)^j\chi^j(A_t)+[(1-\lambda)^t-(1-\lambda)]\chi^t(A_t)$$

$(1-\lambda)a_{t+1}Q_t$

$$=\sum_{j=0}^{t-1}\lambda^{t-j-1}(1-\lambda)^{j+1}a_{t+1}\chi^{j+1-1}(A_t)+\frac{[(1-\lambda)^{t+1}-(1-\lambda)^2]}{\lambda}a_{t+1}\chi^t(A_t)$$

$$=\sum_{j=1}^{t}\lambda^{t-j}(1-\lambda)^j a_{t+1}\chi^{j-1}(A_t)+\frac{[(1-\lambda)^{t+1}-(1-\lambda)^2]}{\lambda}\chi^{t+1}(A_{t+1})$$

$\lambda Q_t+(1-\lambda)\chi^t(A_t)-(1-\lambda)\chi^{t+1}(A_{t+1})+(1-\lambda)a_{t+1}Q_t$

$$=\lambda^t+\sum_{j=1}^{t-1}\lambda^{t-j}(1-\lambda)^j\chi^j(A_t)+\sum_{j=1}^{t}\lambda^{t-j}(1-\lambda)^j a_{t+1}\chi^{j-1}(A_t)$$

$$+[(1-\lambda)^t-(1-\lambda)]\chi^t(A_t)+\frac{[(1-\lambda)^{t+1}-(1-\lambda)^2]}{\lambda}a_{t+1}\chi^{t+1}(A_{t+1})$$

$$+(1-\lambda)\chi^t(A_t)-(1-\lambda)\chi^{t+1}(A_{t+1}))$$

$$=\lambda^t+\sum_{j=1}^{t-1}\lambda^{t-j}(1-\lambda)^j(\chi^j(A_t)+a_{n+1}\chi^{j-1}(A_t)$$

$$+(1-\lambda)^t f(a_{t+1})\chi_{t-1,f(\cdot)}^{t-1}(A_t)+(1-\lambda)^t\chi^t(A_t)-(1-\lambda)\chi^{t+1}(A_{t+1})$$

$$+\frac{[(1-\lambda)^{t+1}-(1-\lambda)^2]}{\lambda}\chi^{t+1}(A_{t+1})$$

$$=\lambda^t+\sum_{j=1}^{t-1}\lambda^{t-j}(1-\lambda)^j\chi^j(A_{t+1})$$

$$+(1-\lambda)^t(\chi^t(A_{t+1})-\chi^t(A_t)+(1-\lambda)^t\chi^t(A_t)$$

$$+\frac{(1-\lambda)^{t+1}-(1-\lambda)^2-\lambda+\lambda^2}{\lambda}\chi^{t+1}(A_{t+1})$$

$$=\lambda^t+\sum_{j=1}^{t-1}\lambda^{t-j}(1-\lambda)^j\chi^j(A_{t+1})+(1-\lambda)^t\chi^t(A_{t+1})+$$

$$\frac{(1-\lambda)^{t+1}-(1-\lambda)}{\lambda}\chi^{t+1}(A_{t+1})$$

$$=\lambda^t+\sum_{j=1}^{t}\lambda^{t-j}(1-\lambda)^j\chi^j(A_{t+1})+\frac{(1-\lambda)^{t+1}-(1-\lambda)}{\lambda}\chi^{t+1}(A_{t+1})$$

$$=\lambda^t+\sum_{j=1}^{t}\lambda^{t-j}(1-\lambda)^j\chi^j(A_{t+1})-\sum_{i=1}^{t}(1-\lambda)^i\chi^{t+1}(A_{t+1})$$

$$=Q_{t+1}$$

So, $T_\lambda(A_{t+1})=\dfrac{\chi^{t+1}}{Q_{t+1}}$.

In conclusion, when $n\in N^+$,

$$T_\lambda(A_n)=\frac{\chi^n(A_n)}{\lambda^{n-1}+\sum_{j=1}^{n-1}\lambda^{n-j-1}(1-\lambda)^j\chi^j(A_n)-\sum_{i=1}^{n-1}(1-\lambda)^i\chi^n(A_n)}$$

which completes the proof.

Proof of Proposition 3

$$\frac{\partial Q_n}{\partial\lambda}=(n-1)\lambda^{n-2}+\sum_{i=1}^{n-1}(n-i-1)\lambda^{n-i-2}(1-\lambda)^i\chi^i(A_n)-\lambda^{n-i-1}(1-\lambda)^{i-1}\chi^iA_n$$

$$+\sum_{i=1}^{n-1}i(1-\lambda)^{i-1}\chi^n(A_n)$$

$$=(n-1)\lambda^{n-2}+\sum_{i=1}^{n-1}\lambda^{n-i-2}(1-\lambda)^{i-1}((n-1)-(n-1)\lambda-i)\chi^i(A_n)+\sum_{i=1}^{n-1}i(1-\lambda)^{i-1}\chi^n(A_n)$$

$$=R_n$$

$$\frac{\partial T_n(A_n)}{\partial \lambda} = \frac{\partial \dfrac{\chi^n(A_n)}{Q_n}}{\partial \lambda} = \frac{-\chi^n(A_n)R_n}{Q_n^2},$$

which completes the proof.

REFERENCES

[1] T. Takagi, and I. Hayashi, "NN-driven fuzzy reasoning," *Int. J. Approx. Reason*., vol. 5, no. 3, pp. 191-212, 1991.

[2] M. Sugeno, and G. T. Kang, "Structure identification of fuzzy model," *Fuzzy Sets Syst.*, vol.28, no.1, pp. 15-33, 1988.

[3] M. Sugeno, and K. Tanaka, "Successive identification of a fuzzy model and its applications to prediction of a complex system," *Fuzzy Sets Syst.*, vol. 42, no.3, pp. 315-334, 1991.

[4] A. Mejas, S. Romero, and F.J. Moreno, "A Systematic Methodology to Obtain a Fuzzy Model Using an Adaptive Neuro Fuzzy Inference System. Application for Generating a Model for Gas-Furnace Problem," *Distributed Computing and Artificial Intelligence*, 452-460. DOI: 10.1007/978-3-540-85863-8_53, 2008.

[5] L.S. Iliadis, S. Spartalis and S. Tachos, "Application of fuzzy T-norms towards a new Artificial Neural Networks evaluation framework: A case from wood industry," *Information Sciences*, vol.178, 3828-3839, 2008.

[6] J.S. R, ANFIS: "Adaptive-network- based fuzzy inference systems," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no.3, pp. 665-685, 1993.

[7] S. Ozer. and H. Zorlu, "Identification of bilinear systems using differential evolu-tion algorithm," *Sadhana-academy Pro-ceedings in Engineering Sciences*, vol. 36, no. 3, pp. 281-292, 2011. DOI: 10.1007/s12046-011-0022-8.

[8] B. Zhu, C. Z. He, P. Liatsisb, and X.Y. Li, "A GMDH-based fuzzy modeling approach for constructing TS-model," *Fuzzy Sets and Systems*, vol. 189, pp. 19-29, 2012. DOI: 10.1016/j.fss.2011.08.004

[9] J.H. Yue, J.Z. Liu, X.J. Liu, and W. Tan, "Identification of Nonlinear System Based on ANFIS with Subtractive Clustering," In:*Proceedings of the 6th World Congress on Intelligent Control and Automation*, vol. 1, pp. 1852-1856, 2006. DOI: 10.1109/WCI CA.2006.1712675

Received: December 8, 2014 Revised: December 15, 2014 Accepted: December 16, 2014