行政院國家科學委員會專題研究計畫 成果報告

非線性隨機控制理論研究及其在非傳統控制領域的應用--子計畫二：非線性隨機模糊適應控制與其在無線網路之應用(III)

研究成果報告(精簡版)

計 畫 類 別 ： 整合型
計 畫 編 號 ： NSC 97-2221-E-216-033-
執 行 期 間 ： 97 年 08 月 01 日至 98 年 07 月 31 日
執 行 單 位 ： 中華大學電機工程學系

計 畫 主 持 人 ： 李柏坤

計畫參與人員 ： 碩士班研究生-兼任助理人員：蘇聖翔

報 告 附 件 ： 出席國際會議研究心得報告及發表論文

處 理 方 式 ： 本計畫可公開查詢

中 華 民 國 98 年 09 月 28 日

『非線性隨機控制理論研究及其在非傳統控制領域的應用』子計畫二: 非線性隨機模糊適應控制與其在無線網路之應用(III) Fuzzy Adaptive Control of Nonlinear Stochastic System and Its Applications to Wireless Network (III)

計畫編號:NSC-97-2221-E-216-033
執行期限:97年8月1日至98年7月31日
主持人: 李柏坤 教授 中華大學電機系

一、中文摘要

在此研究中, 我們已完成隨機 T-S 模糊 ARMAX 模式之適應權重最小變異量控制。對於一個隨機 T-S 模糊 ARMAX 模式, 首先我們推導其最佳向前一步的預估模式; 基於此預估模式, 我們使用隨機梯度方法來估測其中之參數。接著於採用之直接適應控制結構下, 我們推導其適應控制律, 使得針對一個參考模式的輸出追蹤誤差的變異量與控制輛間之權衡能夠最佳化。我們推導此適應控制系統的穩定度, 同時藉由模擬研究已驗證所推導之理論。

關鍵詞: 模糊適應控制、隨機T-S 模糊 ARMAX 模式

Abstract

In this study, we attack the weighted adaptive minimum variance control for stochastic T-S fuzzy ARMAX models. From the fuzzy ARMAX model, a fuzzy one-step ahead prediction model is first developed. A stochastic gradient algorithm is then proposed to identify the parameters of the related one-step-ahead predictor. Under the direct adaptive control scheme, the weighted minimum variance control is applied to find the control law to make adaptive control system stable in the sense of mean square stability. Stability of the adaptive stochastic fuzzy control system is rigorously derived. Simulation study is also made to verify the developed results.

Keywords: Fuzzy adaptive control, Stochastic T-S fuzzy ARMAX model

二、緣由與目的

Recently, based on the Takagi-Sugeno model, fuzzy modeling for nonlinear dynamic systems and identification problem are discussed in [1]-[3]. Meanwhile, fuzzy control scheme has been employed for tracking control of nonlinear systems based on the adaptive feedback linearization techniques [4]-[11]. In the previously mentioned literature, the external disturbances or noises are considered to be deterministic for the convenience of control design. However, in many practical applications [12][13], external noises are inevitable and are more adequately described by random processes. In this situation, the systems to be controlled are always modeled by stochastic systems. A nonlinear stochastic system can be approximated by a fuzzy stochastic system [14]-[19]. A stochastic adaptive control scheme for the state-space T-S fuzzy model based on the LQG control theory is proposed in [20]. Non-adaptive LQG fuzzy controllers are also considered in [14] and [15]. On the other hand, the NARMAX (nonlinear ARMAX) model has been presented for modeling nonlinear processes. The NARMAX model can be reduced to a quasi-ARMAX system by linearization or approximation. Fuzzy system identification and nonlinear model predictive control based on the quasi-ARMAX model are discussed in [16][17][18]. Besides the quasi-ARMAX model, the fuzzy ARMAX model has been used to forecast the short-term load of a power system in [19]. However, these results proposed by the mentioned literature are given without vigorous proofs. Solid proof of the stability and tracking performance of the fuzzy ARX model for deterministic systems can be referred to [21]. In the literature [29], the global stochastic stability and tracking performance of the adaptive minimum variance control for stochastic T-S fuzzy ARMAX models are attacked.

In this study, we shall further attack the weighted adaptive minimum variance control for stochastic T-S fuzzy ARMAX models. From the fuzzy ARMAX model, a fuzzy one-step ahead prediction model is first developed. A stochastic gradient algorithm is then proposed to identify the parameters of the related one-step-ahead predictor. Under the direct adaptive control scheme, the weighted minimum variance control is applied to find the control law to make adaptive control system stable in the sense of mean square stability. Stability of the adaptive stochastic fuzzy control system is rigorously derived. Simulation study is also made to verify the developed results.

Notations and Definitions

Let $\|x\|$ be the Euclidean norm of a vector x. Let $A(q^{-1})$ be a polynomial with $A(q^{-1}) = \sum_{i=0}^{n} a_i q^{-i}$. The companion matrix Ξ_A associated with the polynomial $A(q^{-1})$ is defined as

$$\Xi_A = \begin{bmatrix} 0_{(n-1)\times 1} & & I_{n-1} & \\ -a_n & -a_{n-1} & \cdots & -a_1 \end{bmatrix}$$

三、研究方法及成果

A. System modeling and problem formulation

A nonlinear stochastic system can be divided into several local linear stochastic systems according to their operation regions. Local linear systems using ARX, ARMAX, and NARMAX models to approximate nonlinear stochastic systems can be referred to [21][18][19][15]. A fuzzy stochastic model can be employed to interpolate local linear stochastic systems for a nonlinear stochastic system via the smoothing of fuzzy basis functions. This fuzzy stochastic model is described by fuzzy if-then rules and will be used here to deal with the stochastic tracking problem of nonlinear stochastic systems. The i-th rule of this fuzzy stochastic model for nonlinear discrete-time stochastic systems is proposed as the following fuzzy ARMAX form:

Plant Rule i:
If $z_1(k)$ is F_{i1} and $\cdots\cdots$ and $z_{g_0}(k)$ is F_{ig_0},
then $A_i(q^{-1})y(k+1) = B_i(q^{-1})u(k) + C_i(q^{-1})w(k+1)$ (1)

for i = 1, 2,..., L, where F_{ij} is the fuzzy set, $z_1(k)$, $z_2(k)$, $\cdots, z_{g_0}(k)$ are the premise variables, and L is the number of if-then rules. Polynomials $A_i(q^{-1}), B_i(q^{-1})$, and $C_i(q^{-1})$ are defined, respectively, as follows

$$\begin{array}{ll} A_i(q^{-1}) = a_{i0} + a_{i1}q^{-1} + \ldots\ldots + a_{in}q^{-n}, & a_{i0} = 1 \\ B_i(q^{-1}) = b_{i0} + b_{i1}q^{-1} + \ldots\ldots + b_{im}q^{-m}, & \\ C_i(q^{-1}) = c_{i0} + c_{i1}q^{-1} + c_{i2}q^{-2} + \ldots\ldots + c_{il}q^{-l}, & c_{i0} = 1 \end{array} \tag{2}$$

for i = 1, 2,..., L where q^{-1} denotes the delay operator, i.e., $q^{-1}y(k) = y(k-1)$. Without loss of generality, $C_i(q^{-1})$ can be taken to have roots inside the unit circle [12][13]. $y(k)$ is the output measurement, $u(k)$ the control input, and the noise process $w(k)$ will be taken to satisfy the following assumptions [12][13]:

$$E[w(k+1)|\, \digamma_k] = 0,\ a.\ s. \tag{3}$$

$$E[w^2(k+1)|\, \digamma_k] = \sigma_w^2,\ a.\ s. \tag{4}$$

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} w^2(k) \le K_w < \infty,\ a.\ s. \tag{5}$$

where E denotes the expectation, $\digamma_k$ denotes the sub-σ algebra generated from the data set $\{y(s)\}_{s\le k}$. Note that $\digamma_k$ is increasing, i.e., $\digamma_k \subset \digamma_{k+1}$. We also assume that $u(k)$ is $\digamma_k-$measurable. For the premise variables $z_i(k), 1 \le i \le g_0$, we assume that they are $\digamma_k-$measurable, i.e., $z_i(k)$ depends on the data set $\{y(s), u(s)\}_{s\le k}$. Using the smoothing property of the conditional mean [22], conditions (3) and (4) imply that $w(k)$ is also a white process with zero mean and variance σ_w^2. Note that condition (5) implies

$$\frac{1}{N}\sum_{k=1}^{N} w^2(k) \le K_w,\ a.s., \text{ for } N \ge N_w \tag{6}$$

where N_w is a sufficiently large integer.

Given the input/output sequences $\{u(k)\}$ and $\{y(k)\}$, the stochastic fuzzy system (1) is equivalent to

$$\begin{aligned} y(k+1) = \sum_{i=1}^{L} h_i(z(k)) \{(1 - A_i(q^{-1}))y(k+1) \\ +B_i(q^{-1})u(k) + C_i(q^{-1})w(k+1)\} \end{aligned} \tag{7}$$

where $z(k) = [z_1(k)\ z_2(k)\ ...\ z_{g_0}(k)]$ and, for $1 \le i \le L$,

$$\mu_i(z(k)) = \prod_{j=1}^{g_0} F_{ij}(z_j(k)) \tag{8}$$

$$h_i(z(k)) = \frac{\mu_i(z(k))}{\sum_{i=1}^{L} \mu_i(z(k))} \tag{9}$$

Where the function $F_{ij}(z_j(k))$ is the grade of membership of $z_j(k)$ in F_{ij}. For (8) and (9), we assume that

$$h_i(z(k)) \ge 0\ ,\ \sum_{i=1}^{L} h_i(z(k)) = 1 \tag{10}$$

The physical meaning of (7) is that the L local linear stochastic subsystems are interpolated by the fuzzy basis functions $h_i(z(k))$, for $i = 1, 2, \ldots, L$.

In the sequel, we shall first attack the identification problem for estimating the parameters related to the fuzzy ARMAX model (1). To attack this problem, we shall need some mathematical tools concerning the stochastic stability of the T-S stochastic fuzzy system. After obtaining the estimates of the parameters, the design objective for the stochastic fuzzy system in (7) is to determine the adaptive control input $u(k)$ so as to bring the output $y(k+1)$ to optimally track a desired bounded output command $y^*(k+1)$, which is specified beforehand. Based on the identified parameters, the control objective is to choose the input $u(k)$, as a function of $\{y(s), u(s-1)\}_{s\le k}$, to minimize

$$J_1(k+1) = E\{[y(k+1) - y^*(k+1)]^2 + \lambda u^2(k)|\digamma_k\} \tag{11}$$

B. Stability of Stochastic T-S Fuzzy Systems

In order to deal with the identification problem of the T-S stochastic fuzzy system, the stability issue of the stochastic fuzzy system must be addressed first. The results will be used in the sequel for analysis of the optimal predictor and the identification algorithm for the stochastic fuzzy ARMAX model. Since the fuzzy ARMAX model, such as in (7), can be transformed into a state-space stochastic fuzzy model and stability is easier to discuss from the state-space perspective, we consider a forced T-S fuzzy system in the state-space form as follows

$$x(k+1) = \sum_{i=1}^{L} h_i(z(k))A_i x(k) + v(k+1) \tag{12}$$

where $\{v(k)\}$ is the stochastic forced term. It is assumed that $\{v(k)\}$ is $\digamma_k-$measurable. Recall that $z(k)$ is $\digamma_k-$measurable, i.e., $E\{h_i(z(k))|\digamma_k\} = h_i(z(k))$ a.s. Then $x(k)$ is also

$\mathcal{F}_{k-1}-$measurable. Also we make the assumption that $v(k)$ is a process with uniformly bounded average power

$$\sup_k E\left\{\|v(k)\|^2\right\} = \overline{\sigma}_v^2 < \infty \tag{13}$$

Before deriving the stability result for the forced stochastic fuzzy model (12), we need to first consider the following unforced stochastic system

$$x(k+1) = \mathbf{A}(k)x(k) \tag{14}$$

where $\mathbf{A}(k)$ is $\mathcal{F}_k-$measurable and the sequence $\left\{\|\mathbf{A}(k)\|^2\right\}$ is uniformly bounded. A sufficient conditions concerning the stability of the unforced system in (14) is given below.

Theorem 1: If there exists a sequence of symmetric positive definite matrices $\{\mathbf{P}(k)\}$ with $0 < \lambda_P^{\min} I \leq \mathbf{P}(k) \leq \lambda_P^{\max} I < \infty$ and $\mathbf{P}(k)$ being $\mathcal{F}_k-$measurable such that the matrix inequality

$$\lambda\mathbf{P}(k) - \mathbf{A}^T(k)E\left\{\mathbf{P}(k+1)|\mathcal{F}_k\right\}\mathbf{A}(k) > 0, \ \forall k \tag{15}$$

holds for some λ with $0 < \lambda < 1$, then the stochastic fuzzy system (14) is mean square exponentially stable with

$$E\left\{\|x(k)\|^2\right\} \leq \frac{\lambda_P^{\max}}{\lambda_P^{\min}}\lambda^{k-k_0}E\left\{\|x(k_0)\|^2\right\}, \ \forall k \geq k_0 \tag{16}$$

where k_0 is an arbitrary initial time, $x(k_0)$ is an arbitrary initial condition. Furthermore, it is also exponentially stable in the sense that

$$\|x(k)\| \leq c_1(\sqrt{\lambda})^{k-k_0}\|x(k_0)\|, \quad k \geq K_1, \ a.s. \tag{17}$$

for some positive almost surely bounded random variable $c_1 > 0$ and a sufficiently large integer K_1.

Proof: The proof is given in Appendix 4.1. ■

Note that the following unforced system

$$x(k+1) = \sum_{i=1}^{L} h_i(z(k))A_i x(k), \tag{18}$$

which is related to (12), is a special case of (14) by identifying

$$\mathbf{A}(k) = \sum_{i=1}^{L} h_i(z(k))A_i$$

Hence we have the following corollary.

Corollary 1: If there exist symmetric positive definite matrices $P_i, 1 \leq i \leq L$, such that the linear matrix inequalities

$$\Xi_{i,j} \triangleq \begin{bmatrix} \lambda P_i & A_i^T P_j \\ P_j A_i & P_j \end{bmatrix} > 0, \ 1 \leq i, \ j \leq L \tag{19}$$

hold where λ is a positive real number with $0 < \lambda < 1$, then condition (15) is satisfied. Therefore, for the system in (18), the stability properties in (16) and (17) hold.

Proof:
Since P_j is positive definite, by Schur complement, condition (19) is equivalent to

$$\lambda P_i - A_i^T P_j A_i > 0 \tag{20}$$

Define $\mathbf{P}(k) = \sum_{i=1}^{L} h_i(z(k))P_i$. It follows $\mathbf{P}(k)$ is $\mathcal{F}_k-$measurable and $\lambda_P^{\min} I \leq \mathbf{P}(k) \leq \lambda_P^{\max} I$ where $\lambda_P^{\min} = \min_{1\leq i\leq L}(\lambda_{\min}(P_i))$ and $\lambda_P^{\max} = \max_{1\leq i\leq L}(\lambda_{\max}(P_i))$. Due to the properties of $h_i(\cdot)$ in (10), we also have

$$E\left\{h_j(z(k+1))|\mathcal{F}_k\right\} \geq 0 \ , \ \sum_{j=1}^{L} E\left\{h_j(z(k+1))|\mathcal{F}_k\right\} = 1$$

Now applying the operation $\sum_{i=1}^{L}\sum_{j=1}^{L} h_i(z(k))E\left\{h_j(z(k+1))|\mathcal{F}_k\right\}$ to both sides of (20), we have

$$\lambda\mathbf{P}(k) - \mathbf{A}^T(k)E\left\{\mathbf{P}(k+1)|\mathcal{F}_k\right\}\mathbf{A}(k) > 0, \forall k.$$

By using Theorem 1, the proof is completed. □

In the following, we shall see that, provided the matrix inequalities in (15) hold, the stochastic system in (14) behaves like a linear time-varying system. With the system $\mathbf{A}(k)$, the system response of the system in (14) can be described by

$$x(k+1) = \Phi(k+1, k_0)x(k_0), \ k \geq k_0 \geq 0$$

where $\Phi(k+1, k_0)$ can be regarded as the transition matrix [24] and is defined as

$$\Phi(k+1, k_0) \triangleq \mathbf{A}(k)\mathbf{A}(k-1)\cdots\mathbf{A}(k_0) \tag{21}$$

and $\Phi(k,k) \triangleq I$. The following corollary directly follows from the definition of the norm $\|\cdot\|_{ms}$ and inequality (16).

Corollary 2: If there exists a sequence of symmetric positive definite matrices $\{\mathbf{P}(k)\}$ with $0 < \lambda_P^{\min} I \leq \mathbf{P}(k) \leq \lambda_P^{\max} I < \infty$ and $\mathbf{P}(k)$ being $\mathcal{F}_k-$measurable such that the matrix inequality (15) hold for some λ with $0 < \lambda < 1$, then the upper bounds of the induced norm of $\Phi(k, k_0)$ in the mean square and almost sure senses are given by

$$\|\Phi(k, k_0)\|_{ms} \leq \sqrt{\frac{\lambda_P^{\max}}{\lambda_P^{\min}}}(\sqrt{\lambda})^{k-k_0}, \ \forall k \geq k_0 \tag{22}$$

$$\|\Phi(k, k_0)\| \leq c_2(\sqrt{\lambda})^{k-k_0}, \quad k \geq K_1, \ a.s. \tag{23}$$

for some positive almost surly bounded random variable c_2 and a sufficiently large integer K_1.

Now consider the following stochastic system

$$\begin{aligned} x(k+1) &= [\mathbf{A}(k)x(k) + \mathbf{B}(k)u_s(k)] \\ y_s(k) &= [\mathbf{C}(k)x(k) + \mathbf{D}(k)u_s(k)] \end{aligned} \tag{24}$$

where the sequences $\left\{\|\mathbf{A}(k)\|^2\right\}$, $\left\{\|\mathbf{B}(k)\|^2\right\}$, $\left\{\|\mathbf{C}(k)\|^2\right\}$, and $\left\{\|\mathbf{D}(k)\|^2\right\}$ are uniformly bounded.

Theorem 2: For the stochastic system in (24), there exists a sequence of symmetric positive definite matrices $\{\mathbf{P}(k)\}$ with $0 < \lambda_P^{\min} I \leq \mathbf{P}(k) \leq \lambda_P^{\max} I < \infty$ and $\mathbf{P}(k)$ being $\mathcal{F}_k-$measurable such that the matrix inequality (15) hold for some λ with $0 < \lambda < 1$, then we have

$$\frac{1}{N}\sum_{k=1}^{N}\|y_s(k)\|^2 \leq \frac{K_2}{N}\sum_{k=1}^{N}\|u_s(k)\|^2 + \frac{K_3}{N}, \quad a.\ s., \tag{25}$$

for $N \geq K_1$ where K_1 is a sufficiently large number, $0 < K_2 < \infty$, and $0 \leq K_3 < \infty$.

Proof: The proof is given in Appendix 4.2. ■

Note that the following general T-S fuzzy state-space system

$$\begin{aligned} x(k+1) &= \sum_{i=1}^{L} h_i(z(k)) \left[A_i x(k) + B_i u_s(k)\right] \\ y_s(k) &= \sum_{i=1}^{L} h_i(z(k)) \left[C_i x(k) + D_i u_s(k)\right] \end{aligned} \tag{26}$$

is a special case of the systems in (24) by identifying

$$\mathbf{A}(k) = \sum_{i=1}^{L} h_i(z(k))A_i, \quad \mathbf{B}(k) = \sum_{i=1}^{L} h_i(z(k))B_i$$
$$\mathbf{C}(k) = \sum_{i=1}^{L} h_i(z(k))C_i, \quad \mathbf{D}(k) = \sum_{i=1}^{L} h_i(z(k))D_i$$

Therefore, by combining the results in Corollary 1, Corollary 2, and Theorem 2, the following corollary can be easily obtained.

Corollary 3: If there exist symmetric positive definite matrices $P_i, 1 \le i \le L$, such that the matrix inequality (19) holds for some λ with $0 < \lambda < 1$, then, for the stochastic fuzzy system (26), the inequality (25) holds.

Stability analysis of the state-space stochastic fuzzy system (12) is very useful for the optimal tracking design of the fuzzy ARMAX model in (7). For system identification based on the prediction error method [26] for the fuzzy ARMAX model, the optimal fuzzy prediction must be first established as in the next section.

C. Optimal predictor of stochastic fuzzy systems

In this section, the prediction problem of the fuzzy ARMAX model in (7) will be addressed. This will result in a fuzzy predictor model which will be suitable for parameter estimation and optimal tracking design of fuzzy ARMAX systems. The optimal fuzzy predictor for the fuzzy ARMAX model has been studied in [27]. The results in that reference are briefly summarized in the following.

Assumption 1: Let $\Xi_{C,i}$ be the companion matrix associated with the polynomial $C_i(q^{-1})$. Assume that there exist symmetric positive matrices $P_{C,i},\ 1 \le i \le L$, such that the set of matrix inequalities

$$\begin{bmatrix} \lambda_C P_{C,i} & \Xi_{C,i}^T P_{C,j} \\ P_{C,j}\Xi_{C,i} & P_{C,j} \end{bmatrix} > 0,\ 1 \le i,\ j \le L \tag{27}$$

is solvable for some λ_C with $0 < \lambda_C < 1$.

A fuzzy polynomial $\sum_{i=1}^{L} h_i(z(k))C_i(q^{-1})$ with $C_i(q^{-1})$ being manic and $h_i(\cdot)$ satisfying (10) is *stable* if the LMI condition (27) holds.

Let $y^0(k+1|k)$ denote the conditional mean of $y(k+1)$ given the data set $\{u(s),\ y(s)\}_{s\le k}$, i.e., $y^0(k+1|\ k) \triangleq E\left\{y(k+1)|\ F_k\right\}$. Define the polynomial $\alpha_i(q^{-1}),\ 1 \le i \le L$, as

$$C_i(q^{-1}) - A_i(q^{-1}) = q^{-1}\alpha_i(q^{-1}) \tag{28}$$

where

$$\alpha_i(q^{-1}) = \alpha_{i0} + \alpha_{i1}q^{-1} + \cdots + q^{-(\overline{n}-1)},\ \overline{n} = \max(n,\ l)$$

Under **Assumption 1** on the fuzzy ARMAX model (7), the optimal one-step ahead predictor of $y(k+1)$ given the data set $\{u(s),\ y(s)\}_{s\le k}$ is $y^0(k+1|\ k)$ which satisfies the following equation

$$\begin{aligned} y^0(k+1|k) = \sum_{i=1}^{L} h_i(z(k))\{&\left[1 - C_i(q^{-1})\right] y^0(k+1|k) \\ &+ \alpha_i(q^{-1})y(k) + B_i(q^{-1})u(k)\} \end{aligned} \tag{29}$$

with the prediction error

$$y(k+1) - y^0(k+1|k) = w(k+1) \tag{30}$$

Equation (29) defines a unique *one-step ahead fuzzy prediction model* corresponding to the fuzzy ARMAX model (7). From human-operation point of view, the fuzzy prediction model is more feasible than the fuzzy ARMAX model since we can use the current and past measurement data $\{u(s),\ y(s)\}_{s\le k}$ to predict the future response $y(k+1)$ of the stochastic fuzzy system; while using the fuzzy ARMAX model (7), the statistical properties of the noise process $w(k)$ should be specified in advance.

D. Stochastic Gradient Algorithm

Following from the fuzzy prediction model represented by (29), the stochastic gradient algorithm in [13] will be used to identify the parameters. First, rearrange the prediction model (29) as follows

$$y^0(k+1|k) = \sum_{i=1}^{L} h_i(z(k))\chi_0^T(k)\theta_{i0} = \phi_0^T(k)\theta_0 \tag{31}$$

where, for $1 \le i \le L$,

$$\begin{aligned} \chi_0(k) &= \left[-y^0(k|k-1) \cdots - y^0(k-l+1|k-l)\right. \\ &\qquad \left. y(k) \cdots y(k-\overline{n}+1)\ u(k) \cdots u(k-m)\right]^T \\ \theta_{i0} &= \left[c_{i1} \cdots c_{il}\ \alpha_{i0} \cdots \alpha_{i(\overline{n}-1)}\ b_{i0} \cdots b_{im}\right]^T \\ \phi_0(k) &= \left[h_1(z(k))\chi_0^T(k) \quad h_2(z(k))\chi_0^T(k)\right. \\ &\qquad \left. \cdots\cdots h_L(z(k))\chi_0^T(k))\right]^T \\ \theta_0 &= \left[\theta_{10}^T \quad \theta_{20}^T \quad \cdots\cdots \theta_{L0}^T\right]^T \end{aligned}$$

Note that (31) represents a pseudo linear regression form for the fuzzy ARMAX prediction model (29) because the component $y^0(k-i+1|k-i)$ in $\chi_0(k)$ depends on the true parameter vector θ_0. According to the pseudo linear regression form (31), the proposed stochastic gradient algorithm to identify the true parameter vector θ_0 is given by, for $k \ge 1$,

$$\begin{aligned} \widehat{\theta}(k) = \widehat{\theta}(k-1) &+ \frac{\phi(k-1)}{r(k-2) + \phi^T(k-1)\phi(k-1)} \\ &\times \left[y(k) - \phi^T(k-1)\widehat{\theta}(k-1)\right] \end{aligned} \tag{32}$$

where the regression vector $\phi(k)$ and $r\,(k-1)$ are defined as

$$\begin{aligned} \phi(k) &= \left[h_1(z(k))\chi^T(k) \quad h_2(z(k))\chi^T(k)\right. \\ &\qquad \left. \cdots\cdots \quad h_L(z(k))\chi^T(k)\right]^T \end{aligned} \tag{33}$$

$$\begin{aligned} \chi(k) &= [-\overline{y}(k) \cdots - \overline{y}(k-l+1) \\ &\qquad y(k) \cdots y(k-\overline{n}+1)\ u(k) \cdots u(k-m)]^T \end{aligned} \tag{34}$$

$$\overline{y}(k) = \phi^T(k-1)\widehat{\theta}(k) \tag{35}$$

$$r(k-1) = r(k-2) + \phi^T(k-1)\phi(k-1) \tag{36}$$

For the initial conditions, $\widehat{\theta}(0)$ can be arbitrarily chosen and $r(-1)$ must be a positive scalar. By its definition, the variable $\overline{y}(k)$ can be regarded as a posterior estimate of $y(k)$.

Before proceeding to analyze the stochastic gradient algorithm, some useful definitions are made as follows

$$\widehat{y}(k) = \phi^T(k-1)\widehat{\theta}(k-1) \tag{37}$$

$$e(k) = y(k) - \widehat{y}(k) \tag{38}$$

$$\eta(k) = y(k) - \overline{y}(k) \tag{39}$$

$$\varsigma(k) = \eta(k) - w(k) \tag{40}$$

$$\widetilde{\theta}(k) = \widehat{\theta}(k) - \theta_0 \tag{41}$$

$$\beta(k) = -\phi^T(k-1)\widetilde{\theta}(k) \tag{42}$$

The variables $\widehat{y}(k)$ and $\overline{y}(k)$ are the a prior and the a posteriori predictions of $y(k)$, respectively. Accordingly, $e(k)$ and $\eta(k)$ are termed as the a prior and the a posteriori prediction errors, respectively. Using (30), the quantity $\varsigma(k)$ can be rewritten as $\varsigma(k) = y^0(k|k-1) - \overline{y}(k)$ and thus it accounts for the deviation between the optimal prediction $y^0(k|k-1)$ and the a posteriori prediction $\overline{y}(k)$. In the extreme case, if $\varsigma(k) = 0$, then $\phi(k) = \phi_0(k)$ and the pseudo linear regression form (31) becomes $\overline{y}(k+1) = \phi^T(k)\theta_0$ which is a linear regression form.

Lemma 1: For the stochastic gradient algorithm in (32), we have

$$(i)\ \lim_{N\to\infty}\sum_{k=1}^{N}\frac{\phi^T(k-1)\phi(k-1)}{r(k-1)r(k-2)} < \infty \tag{43}$$

$$(ii)\ \eta(k) = \frac{r(k-2)}{r(k-1)}e(k) \tag{44}$$

$$\begin{aligned}(iii)\ &E\left\{\beta(k)w(k)|\ F_{k-1}\right\}\\ &= -\frac{\phi^T(k-1)\phi(k-1)}{r(k-1)}\sigma_w^2,\ a.\ s.\end{aligned} \tag{45}$$

$$(iv)\ \sum_{i=1}^{L} h_i(z(k-1))C_i(q^{-1})\varsigma(k) = \beta(k) \tag{46}$$

Proof: The proof is given in Appendix 4.3. ■

In addition to the results in Lemma 1, we shall need the following assumptions in order to obtain the properties of the parameter estimate $\widehat{\theta}(k)$.

Assumption 2 : For each i, $1 \le i \le L$, system $C_i(q^{-1})$ is input strictly passive (ISP) [13].

In (46), the signals $\varsigma(k)$ and $\beta(k)$ are related by the fuzzy polynomial $\sum_{i=1}^{L} h_i(z(k-1))C_i(q^{-1})$. As shall be shown in the next lemma, **Assumption 2** implies a passivity condition for that fuzzy polynomial.

Lemma 2: Consider the fuzzy system in (46). With **Assumption 2** that $C_i(q^{-1})$ is input strictly passive (ISP), we have

$$\sum_{j=1}^{k}\beta(j)\varsigma(j) - \epsilon\varsigma^2(j) \ge 0,\ for\ k \ge 1 \tag{47}$$

for some $\epsilon > 0$.

Proof: The proof is given in Appendix 4.4. ■

Theorem 3: Under **Assumption 2**, for the stochastic gradient algorithm in (32)-(36), we have the parameter difference convergence

$$\lim_{N\to\infty}\sum_{k=1}^{N}\left\|\widehat{\theta}(k) - \widehat{\theta}(k-1)\right\|^2 < \infty,\ a.\ s. \tag{48}$$

and the normalized prediction error convergence

$$\lim_{N\to\infty}\sum_{k=1}^{N}\frac{[e(k) - w(k)]^2}{r(k-1)} < \infty,\ a.\ s. \tag{49}$$

Proof: The proof is given in Appendix 4.5. ■

With the property in (49), it is possible to attain further results of the stochastic gradient algorithm by imposing an additional key condition. The following *stochastic key technical lemma* is quoted from [13].

Lemma 3: With the property in (49), if there exist positive constants K_{a1}, K_{a2}, and $\overline{N}$ such that, for $N \ge \overline{N}$, $a.\ s.$

$$\frac{1}{N}r(N-1) \le K_{a1} + \frac{K_{a2}}{N}\sum_{k=1}^{N}[e(k) - w(k)]^2, \tag{50}$$

then $a.\ s.$

$$(i)\ \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}[e(k) - w(k)]^2 = 0, \tag{51}$$

$$(ii)\ \limsup_{k\to\infty}\frac{r(N-1)}{N} < \infty, \tag{52}$$

$$\begin{aligned}(iii)\ &\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}E\left\{[y(k) - \widehat{y}(k)]^2 \mid F_{k-1}\right\}\\ &= \sigma_w^2.\end{aligned} \tag{53}$$

Proof: The proof is given in Appendix 4.6. ■

E. Adaptive Weighted Minimum Variance Control

With the stochastic gradient algorithm for identifying parameters in the stochastic fuzzy predictor model, we are ready to propose an adaptive fuzzy controller. The objective of the adaptive control system is to design $u(k)$ to minimize the mean-square error between the output $y(k)$ and the desired output command $y^*(k)$ at any time instance and at the same time to keep the control law at suitable level, i.e., the cost function in (11) is minimized. We shall construct a direct adaptive control. Therefore, we shall first discuss the structure of the weighted minimum variance controller by assuming that the system parameters are given. For the fuzzy stochastic system (7) having the optimal one-step ahead prediction form in (29), the weighted minimum variance tracking control minimizing the cost function $J_1(k+1)$ in (11) is given by [27] as quoted below. First, let

$$\mathbf{b_0}(k) = \sum_{i=1}^{L} h_i(z(k))b_{i0}$$

Theorem 4: ([27]) For the fuzzy stochastic system (7) having the optimal one-step-ahead prediction form in (29), the weighted minimum variance control law minimizing the cost function in (11) is given by

$$\begin{aligned} &u(k) \\ &= \frac{\mathbf{b_0}(k)}{\mathbf{b_0^2}(k)+\lambda}\left\{\sum_{i=1}^{L} h_i(z(k))[C_i(q^{-1})-1]y^0(k+1|k)\right. \\ &+ y^*(k+1) - \sum_{i=1}^{L} h_i(z(k))\alpha_i(q^{-1})y(k) \\ &\left.-\sum_{i=1}^{L} h_i(z(k))[B_i(q^{-1})-b_{i0}]u(k)\right\} \end{aligned} \tag{54}$$

which is equivalent to

$$\mathbf{b_0}(k)\left[y^0(k+1|k)-y^*(k+1)\right]+\lambda u(k)=0 \tag{55}$$

The dynamics of the closed-loop system is governed by

$$u(k)=\frac{\mathbf{b_0}(k)}{\lambda}[w(k+1)-y(k+1)+y^*(k+1)] \tag{56}$$

and

$$\begin{aligned} &\sum_{i=1}^{L} h_i(z(k))\left[A_i(q^{-1})+B_i(q^{-1})\frac{\mathbf{b_0}(k)}{\lambda}\right]y(k+1) \\ &= \sum_{i=1}^{L} h_i(z(k))B_i(q^{-1})\frac{\mathbf{b_0}(k)}{\lambda}y^*(k+1) \\ &+ \sum_{i=1}^{L} h_i(z(k))\left[B_i(q^{-1})\frac{\mathbf{b_0}(k)}{\lambda}+C_i(q^{-1})\right]w(k+1) \end{aligned} \tag{57}$$

Now suppose that the estimated parameters, $\widehat{\alpha}_{ij}(k), \hat{b}_{ij}(k)$, and $\hat{c}_{ij}(k)$ are obtained by using the stochastic gradient algorithm at time k. Accordingly, define the following polynomials

$$\begin{aligned} &\widehat{\alpha}_i(k,q^{-1}) = \widehat{\alpha}_{i0}(k)+\widehat{\alpha}_{i1}(k)q^{-1}+......+\widehat{\alpha}_{i(\overline{n}-1)}(k)q^{-(\overline{n}-1)}, \\ &\hat{B}_i(k,q^{-1}) = \hat{b}_{i0}(K)+\hat{b}_{i1}(k)q^{-1}+......+\hat{b}_{im}(k)q^{-m}, \\ &\hat{C}_i(k,q^{-1}) = 1+\hat{c}_{i1}(k)q^{-1}+\hat{c}_{i2}(k)q^{-2}+......+\hat{c}_{il}(k)q^{-l}, \end{aligned}$$

with $c_{i0}=1$. Let $\widehat{\mathbf{b}}_\mathbf{0}(k)=\sum_{i=1}^{L} h_i(z(k))\hat{b}_{i0}(k)$. Based on the above estimated polynomials, the weighted adaptive minimum variance control law under the certainty equivalent principle is given by

$$\begin{aligned} u(k) = &\frac{\hat{\mathbf{b}}_\mathbf{0}(k)}{\hat{\mathbf{b}}_\mathbf{0}^2(k)+\lambda}\left\{\sum_{i=1}^{L} h_i(z(k))[\widehat{C}_i(q^{-1})-1]\bar{y}(k+1)\right. \\ &+ y^*(k+1) - \sum_{i=1}^{L} h_i(z(k))\hat{\alpha}_i(q^{-1})y(k) \\ &\left.-\sum_{i=1}^{L} h_i(z(k))[\widehat{B}_i(q^{-1})-\hat{b}_{i0}]u(k)\right\} \end{aligned} \tag{58}$$

in which the control law is derived from the following equation

$$\frac{\hat{\mathbf{b}}_\mathbf{0}(k)}{\lambda}\left[\phi^T(k)\widehat{\theta}(k)-y^*(k+1)\right]+u(k)=0 \tag{59}$$

F. Analysis of Stability

In this section, closed-loop stability of the proposed adaptive stochastic fuzzy control system will be discussed. First, we rewrite the equation concerning the adaptive control law in (59) as

$$u(k)=\frac{\hat{\mathbf{b}}_\mathbf{0}(k)}{\lambda}\left[e(k+1)+y^*(k+1)-y(k+1)\right] \tag{60}$$

Now applying the operator $\sum_{i=1}^{L} h_i(z(k))B_i(q^{-1})$ to both sides of the above equation and using the system equation (7), we can obtain

$$\begin{aligned} &\sum_{i=1}^{L} h_i(z(k))\left[A_i(q^{-1})+B_i(q^{-1})\frac{\hat{\mathbf{b}}_\mathbf{0}(k)}{\lambda}\right]y(k+1) \\ &= \sum_{i=1}^{L} h_i(z(k))B_i(q^{-1})\frac{\hat{\mathbf{b}}_\mathbf{0}(k)}{\lambda} \\ &\times[y^*(k+1)+e(k+1)-w(k+1)] \\ &+ \sum_{i=1}^{L} h_i(z(k))\left[C_i(q^{-1})+B_i(q^{-1})\frac{\hat{\mathbf{b}}_\mathbf{0}(k)}{\lambda}\right]w(k+1) \end{aligned} \tag{61}$$

To analyze the closed-loop response of $y(k)$ from the above equation, the estimated term $\hat{\mathbf{b}}_\mathbf{0}(k)$ imposes a difficult issue. Therefore, as done in [28], we assume the following assumption.

Assumption 3: (i) Assume that $b_{0,i}$ is known for $1\le i\le L$ in the fuzzy model and these exists positive number $b_{0,\min}$ and $b_{0,\max}$ such that $0<b_{0,\min}\le b_{0,i}\le b_{0,\max}$ and $0<b_{0,\min}\le |\mathbf{b_0}(k)|\le b_{0,\max}$.

Based on **Assumption 3**, it is possible to find a constant γ_0 such that

$$0<\frac{\lambda}{\lambda+\mathbf{b}_0^2(k)}\le\gamma_0\le 1 \tag{62}$$

Also, the closed-loop system equation in (61) becomes

$$\begin{aligned} &\sum_{i=1}^{L} h_i(z(k))\left[A_i(q^{-1})+B_i(q^{-1})\frac{\mathbf{b_0}(k)}{\lambda}\right]y(k+1) \\ &= \sum_{i=1}^{L} h_i(z(k))B_i(q^{-1})\frac{\mathbf{b_0}(k)}{\lambda} \\ &\times[y^*(k+1)+e(k+1)-w(k+1)] \\ &+ \sum_{i=1}^{L} h_i(z(k))\left[C_i(q^{-1})+B_i(q^{-1})\frac{\mathbf{b_0}(k)}{\lambda}\right]w(k+1) \end{aligned} \tag{63}$$

Without loss of generality, assume that $n=\max(n,m)$. Then, $A_i(q^{-1})$ and $B_i(q^{-1})$ will be regarded as polynomials of degree n where the extra coefficients b_{ij} for $n\ge j>m$ are identified as zeros. Let $\Xi_{\overline{A}_i}$ be the companion matrix associated with the polynomial $\overline{A}_i(q^{-1})\triangleq -\gamma_0 A_i(q^{-1})$ with degree n. Corresponding to the polynomial $B_i(q^{-1})$ with degree n, define a matrix $M_{\overline{B}_i}$ as

$$M_{\overline{B}_i}=\begin{bmatrix} & 0_{(\overline{n}-1)\times\overline{n}} & & \\ -\gamma_0 b_{in} & -\gamma_0 b_{i(n-1)} & \cdots & -\gamma_0 b_{i1} \end{bmatrix}$$

Next we assume that the nonadaptive weighted minimum variance controller defined in 54 is a stabilizing controller for the stochastic T-S fuzzy model defined in (7).

Assumption 4: There exist symmetric positive definite matrices $\overline{P}_i, 1 \le i \le L$, of the form

$$\overline{P}_i = \begin{bmatrix} \overline{P}_i^{11} & 0_{(n-1)\times 1} \\ 0_{1\times(n-1)} & \overline{P}_i^{22} \end{bmatrix} \tag{64}$$

such that the matrix inequalities, for $1 \le i,\ j \le L$,

$$\begin{bmatrix} \lambda_0 \overline{P}_i - \varepsilon \frac{b_{0,\max}^2}{\lambda^2} I_n & \Xi_{\overline{A}_i}^T \overline{P}_j & 0_{n\times n} \\ \overline{P}_j \Xi_{\overline{A}_i} & \overline{P}_j & \overline{P}_j M_{\overline{B}_i} \\ 0_{n\times n} & M_{\overline{B}_i}^T \overline{P}_j & \varepsilon I_n \end{bmatrix} > 0 \tag{65}$$

hold for some $\varepsilon > 0$ and some λ_0 with $0 < \lambda_0 < 1$.

Now, define the following time-varying matrices

$$\begin{aligned} \mathbf{\Xi}_A(k) &= \sum_{i=1}^{L} h_i(z(k)) \Xi_{\overline{A}_i}, \ \ \mathbf{M}_B(k) = \sum_{i=1}^{L} h_i(z(k)) M_{\overline{B}_i} \\ \overline{\mathbf{P}}(k) &= \sum_{i=1}^{L} h_i(z(k)) \overline{P}_i,\ \overline{\mathbf{P}}_+(k) = \sum_{i=1}^{L} h_i(z(k+1)) \overline{P}_i \end{aligned}$$

and let $\Gamma_2(k)$ be a diagonal matrix with

$$\Gamma_2(k) = diag\{\frac{\mathbf{b}_0(k-\overline{n})}{\lambda}, \frac{\mathbf{b}_0(k-\overline{n}+1)}{\lambda}, \cdots, \frac{\mathbf{b}_0(k-1)}{\lambda}\} \tag{66}$$

Note that by the properties of the membership functions in (10) and the definitions of $\mathbf{b}_0(k)$ as well as $b_{0,\max}$, it follows that

$$\|\Gamma_2(k)\| \le \frac{b_{0,\max}}{\lambda} \tag{67}$$

Moreover, from Lemma 7 in [27], we can conclude that

$$\begin{aligned} \left[\mathbf{\Xi}_A(k) + \mathbf{M}_B(k)\Gamma_2(k)\right]^T \overline{\mathbf{P}}_+(k) \left[\mathbf{\Xi}_A(k) + \mathbf{M}_B(k)\Gamma_2(k)\right] \\ < \lambda_0 \overline{\mathbf{P}}(k) \end{aligned} \tag{68}$$

Based on **Assumptions 1-4,** we have the following results which will be used to prove stability of the adaptive control system.

Lemma 4: Under **Assumption 1-Assumption 4**, there exist finite positive constants K_8 to K_{13} such that $a.s.$

$$(i)\ \frac{1}{N}\sum_{k=1}^{N} y^2(k) \le \frac{K_8}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]^2 + K_9, \tag{69}$$

$$(ii)\quad \frac{1}{N}\sum_{k=1}^{N} \overline{y}^2(k) \le \frac{K_{10}}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]^2 + K_{11}, \tag{70}$$

$$(iii)\quad \frac{r(N-1)}{N} \le \frac{K_{a2}}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]^2 + K_{a1}, \tag{71}$$

for $N \ge \overline{N}$ where $\overline{N}$ is a sufficiently large number.

Proof: The proof is given in Appendix (4.7). ∎

Finally, with the above lemma, we have the following tracking performance and global convergence result.

Theorem 5: For the stochastic fuzzy system in (7) with **Assumption 1-Assumption 43**, the weighted adaptive minimum variance control system is stable in the sense that

$$(i)\quad \limsup_{k\to\infty} \frac{1}{N}\sum_{k=1}^{N} y^2(k) < \infty, \quad a.\ s. \tag{72}$$

$$(ii)\quad \limsup_{k\to\infty} \frac{1}{N}\sum_{k=1}^{N} u^2(k) < \infty, \quad a.\ s. \tag{73}$$

Proof:

Since, with (71), the stochastic key technical lemma (Lemma 3) holds, we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]^2 = 0,\ a.\ s.$$

Moreover, we have

$$\limsup_{k\to\infty} \frac{r(N-1)}{N} < \infty,\ a.\ s.$$

which implies (72) and (73). □

G. Simulation Study

In this section, a simulation example is given to verify the proposed adaptive weighted minimum variance control algorithm.

Example 1: Adaptive control for T-S fuzzy systems

Consider the following stochastic fuzzy system:

$$\begin{gathered} If\ z(k)\ is\ F_i,\ then \\ A_i(q^{-1})y(k+1) = B_i(q^{-1})u(k) + C_i(q^{-1})w(k+1) \end{gathered}$$

for $i = 1, 2, \cdots 5$ where

$$\begin{aligned} A_1(q^{-1}) &= 1 - 0.27q^{-1} + 0.011q^{-2}, \\ B_1(q^{-1}) &= 1 - 0.2q^{-1}, \\ C_1(q^{-1}) &= 1 - 0.135q^{-1} \\ A_2(q^{-1}) &= 1 - 0.33q^{-1} + 0.023q^{-2}, \\ B_2(q^{-1}) &= 1 - 0.3q^{-1}, \\ C_2(q^{-1}) &= 1 - 0.165q^{-1} \\ A_3(q^{-1}) &= 1 - 0.362q^{-1} + 0.0288q^{-2}, \\ B_3(q^{-1}) &= 1 - 0.4q^{-1}, \\ C_3(q^{-1}) &= 1 - 0.18q^{-1} \\ A_4(q^{-1}) &= 1 - 0.39q^{-1} + 0.035q^{-2}, \\ B_4(q^{-1}) &= 1 - 0.5q^{-1}, \\ C_4(q^{-1}) &= 1 - 0.195q^{-1} \\ A_5(q^{-1}) &= 1 - 0.44q^{-1} + 0.0468q^{-2}, \\ B_5(q^{-1}) &= 1 - 0.6q^{-1}, \\ C_5(q^{-1}) &= 1 - 0.22q^{-1}, \end{aligned}$$

and $w(k)$ is a zero-mean Gaussian white noise with $\sigma_w = 0.01$. The membership function for the fuzzy logic set F_i is given in Fig 1 and the premise variable $z(k)$ is chosen as $z(k) = y(k)$. We choose $y^*(k+1) = \sin(\frac{2\pi}{100}) + 3\sin\left(\frac{6\pi}{100}\right)$ as the reference signal. to estimation parameter and find the $u(k)$ to yield $y(k+1) = y^*(k+1)$. The weighting constant λ is chosen as $\lambda = 0.01$.The simulation results are shown in Fig 2 and Fig 3. Fig 2 shows the output $y(k)$ and the prediction $\overline{y}(k)$

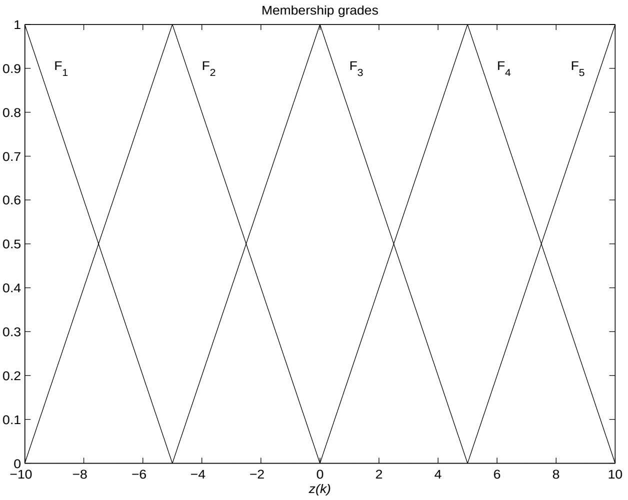

圖 1. Membership functioms of Example 1

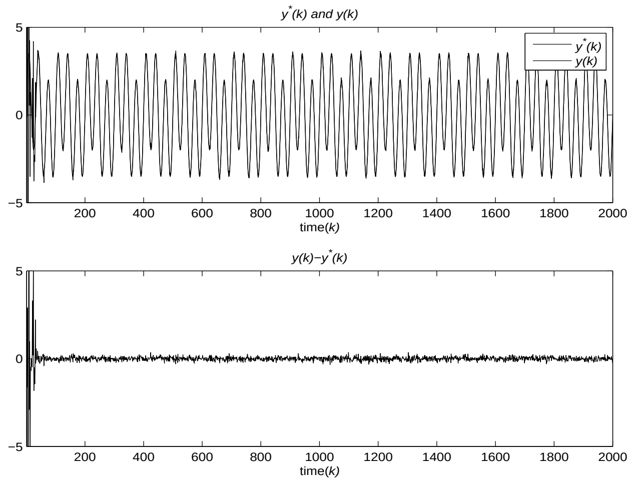

圖 3. The reference signal $y^*(t)$ and the output $y(t)$ of Example 1 is shown in the upper trace. The tracking error is shown in the lower trace.

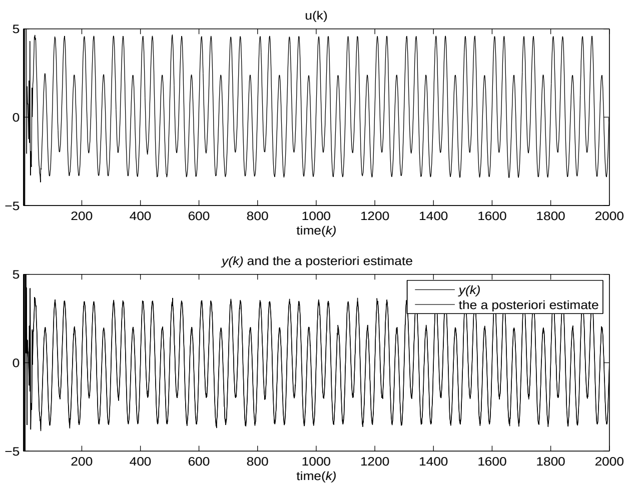

圖 2. Output $y(t)$ and its prediction $\bar{y}(t)$ of Example 1.

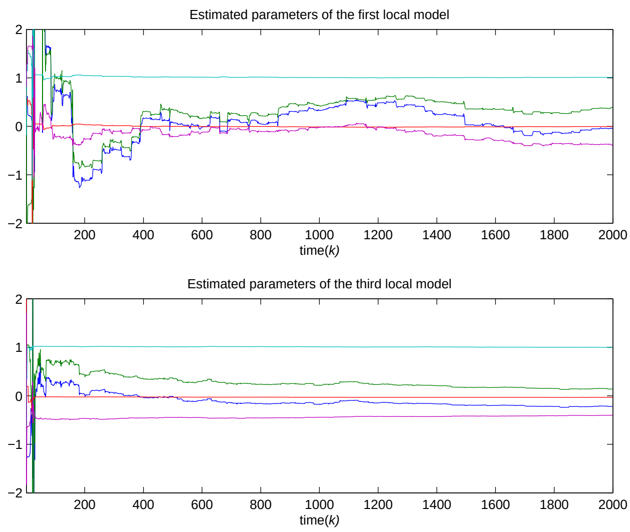

圖 4. The easimated parameters of the first and the third local models.

together with the estimation error. Fig 3 shows the output $y(k)$ and the desired output command $y^*(k)$. The estimated paramers are shown in Fig. 4.

四、結論與討論

Adaptive weighted minimum variance control for stochastic T-S fuzzy ARMAX model is addressed in this study. From the fuzzy ARMAX model, a fuzzy one-step ahead prediction model is first developed. A stochastic gradient algorithm is then proposed to identify the parameters of the related one-step-ahead predictor. Under the direct adaptive control scheme, weighted minimum variance control is applied to find the control law to make the output track a desired reference signal. Stability of the adaptive stochastic fuzzy control system is rigorously derived. Simulation study is also made to verify the developed results.

4.. APPENDIX

A. Proof of Theorem 1

Proof: First define a Lyapunov function as

$$V(x(k)) = x^T(k)\mathbf{P}(k)x(k) \tag{74}$$

which is uniformly positive definite and

$$\lambda_P^{\min} \|x(k)\|^2 \leq V(x(k)) \leq \lambda_P^{\max} \|x(k)\|^2 \tag{75}$$

With the definition of $V(x(k))$, it follows that

$$V(x(k+1)) = x^T(k)\left[\mathbf{A}^T(k)\mathbf{P}(k+1)\mathbf{A}(k)\right]x(k) \tag{76}$$

Note that the terms $x(k)$, $\mathbf{P}(k)$, and $\mathbf{A}(k)$ are all $\mathcal{F}_k-$measurable. Now applying the conditional mean operator $E\{\cdot \mid \mathcal{F}_k\}$ to the both sides of (76) and using (15), we have, almost surely,

$$\begin{array}{rl} & E\left\{V(x(k+1)) \mid \digamma_k\right\} \\ = & x^T(k)\left[\mathbf{A}^T(k)E\left\{\mathbf{P}(k+1)|\digamma_k\right\}\mathbf{A}(k)\right]x(k) \\ \le & \lambda x^T(k)\mathbf{P}(k)x(k) \\ = & \lambda V(x(k)) \end{array} \tag{77}$$

Note that as $E\left\{\|\mathbf{A}(k)\|^2\right\}$ and $E\left\{\|\mathbf{P}(k)\|^2\right\}$are uniformly bounded, $E\left\{V(x(k+1)) \mid \digamma_k\right\}$ and $E\left\{\mathbf{P}(k+1)|\digamma_k\right\}$ are well defined. Apply the conditional expectation operator $E\left\{\cdot \mid \digamma_{k-1}\right\}$ again to the both sides of (77) and recall that the sequence of the $\sigma-$algebra $\digamma_k$ is increasing. With the smoothing properties [13] of conditional mean and inequality (77), it follows that almost surely

$$E\left\{V(x(k+1)) \mid \digamma_{k-1}\right\} \le \lambda^2 V(x(k-1))$$

Continuing this procedure by sequentially applying $E\{\cdot \mid \digamma_{k-2}\}$, $E\{\cdot \mid \digamma_{k-3}\}$, $\cdots$, $E\{\cdot \mid \digamma_{k_0}\}$, one can obtain almost surely

$$E\left\{V(x(k+1)) \mid \digamma_{k_0}\right\} \le \lambda^{k+1-k_0} V(x(k_0)) \tag{78}$$

Now taking expectation of the last inequality to yield

$$E\left\{V(x(k))\right\} \le \lambda^{k-k_0} E\left\{V(x(k_0))\right\}$$

Finally, using the fact of (75), inequality (16) is obtained.

Now we turn to prove the almost sure exponential stability (17). Clearly, it is trivial if $x(k_0) = 0$. Now assume that the initial condition $x(k_0)$ is nonzero. By Chebyshev's inequality [22], for any $\epsilon_k > 0$, we have

$$\begin{array}{rl} & \text{Prob}\left\{\frac{\|x(k)\|}{\|x(k_0)\|} > \epsilon_k\right\} \\ \le & E\left\{\frac{\|x(k)\|^2}{\|x(k_0)\|^2}\right\}/\epsilon_k^2 \\ = & E\left\{\frac{1}{\|x(k_0)\|^2}E\left\{\|x(k)\|^2 \mid \digamma_{k_0}\right\}\right\}/\epsilon_k^2 \end{array} \tag{79}$$

where Prob$\{\mathcal{A}\}$ is the probability measure of the event $\mathcal{A}$. With (75) and (78), one can get

$$E\left\{\|x(k)\|^2 \mid \digamma_{k_0}\right\} \le \frac{\lambda_P^{\max}}{\lambda_P^{\min}}\lambda^{k-k_0}\|x(k_0)\|^2, \ a.s.$$

With the last inequality, (79) can be reduced to

$$\text{Prob}\left\{\frac{\|x(k)\|}{\|x(k_0)\|} > \epsilon_k\right\} \le \frac{1}{\epsilon_k^2}\frac{\lambda_P^{\max}}{\lambda_P^{\min}}\lambda^{k-k_0} \tag{80}$$

Now choose the sequence ϵ_k as $\epsilon_k = \epsilon_0\lambda_1^{(k-k_0)/2}$ for any $\epsilon_0 > 0$ and $\lambda_1 > \lambda$. Then inequality (80) implies that

$$\begin{aligned} & \sum_{k=k_0}^{\infty}\text{Prob}\left\{\|x(k)\| > \epsilon_0\lambda_1^{(k-k_0)/2}\|x(k_0)\|\right\} \\ & \le \frac{1}{\epsilon_0^2}\frac{\lambda_P^{\max}}{\lambda_P^{\min}}\sum_{k=k_0}^{\infty}(\frac{\lambda}{\lambda_1})^{k-k_0} \end{aligned}$$

As $\frac{\lambda}{\lambda_1} < 1$, it follows that

$$\sum_{k=k_0}^{\infty}\text{Prob}\left\{\|x(k)\| > \epsilon_0\lambda_1^{(k-k_0)/2}\|x(k_0)\|\right\} < \infty$$

and consequentially, by the Borel-Cantelli Lemma [22], we obtain that

$$\text{Prob}\left\{\cup_{k\ge K_1}\left\{\|x(k)\| > \epsilon_0\lambda_1^{(k-k_0)/2}\|x(k_0)\|\right\}\right\} = 0$$

for some sufficiently large K_1, any $\epsilon_0 > 0$, and any $\lambda_1 > \lambda$. This means that for any sample path with bounded initial state $x(k_0)$, we have

$$\|x(k)\| \le c_1(\sqrt{\lambda})^{k-k_0}\|x(k_0)\|, \quad k \ge K_1, \ a.s.$$

for any initial condition $x(k_0)$, some positive bounded random variable c_1, and a sufficiently large integer K_1. This completes the proof. ■

B. Proof of Theorem 2

Proof: Suppose that $\|\mathbf{A}(k)\| \le A_L$, $\|\mathbf{B}(k)\| \le B_L$, $\|\mathbf{C}(k)\| \le C_L$, and $\|\mathbf{D}(k)\| \le D_L$ for all k. Using the definition of the transition matrix defined in (21), the response of the output $y_s(k)$ of the fuzzy system in (26) can be represented by

$$\begin{aligned} y_s(k) &= \mathbf{C}(k)\Phi(k,0)x(0) + \mathbf{D}(k)u_s(k) \\ &\quad + \mathbf{C}(k)\sum_{j=0}^{k-1}\Phi(k,j+1)\mathbf{B}(j)u_s(j) \end{aligned}$$

Applying the results in Corollary 2, for $k \ge K_1$, we have

$$\begin{aligned} \|y_s(k)\| &\le C_L c_2\sqrt{\lambda}^k\|x(0)\| + D_L\|u_s(k)\| \\ &\quad + C_L B_L\sum_{j=0}^{k-1}\|\Phi(k,j+1)\|\|u_s(j)\|, \ a.\ s. \end{aligned}$$

By the Cauchy-Schwartz inequality, the last inequality leads to

$$\begin{aligned} \|y_s(k)\|^2 &\le 3\left\{C_L^2c_2^2\lambda^k\|x(0)\|^2 + D_L^2\|u_s(k)\|^2\right. \\ &\quad \left. + C_L^2B_L^2\left[\sum_{j=0}^{k-1}\|\Phi(k,j+1)\|\|u_s(j)\|\right]^2\right. \\ &\le c_3\lambda^k + 3D_L^2\|u_s(k)\|^2 \\ &\quad + 3C_L^2B_L^2\sum_{j=0}^{k-1}\|\Phi(k,j+1)\| \\ &\quad \times\sum_{j=0}^{k-1}\|\Phi(k,j+1)\|\|u_s(j)\|^2, \ a.\ s. \end{aligned} \tag{81}$$

where c_3 is defined as $c_3 = 3C_L^2c_2^2\|x(0)\|^2$. Considering the change of index $i = k-j$, the first summation term in the last inequality can be rearranged as

$$\begin{aligned} \sum_{j=0}^{k-1}\|\Phi(k,j+1)\| &= \sum_{i=1}^{k}\|\Phi(k,k-i+1)\| \\ &= \sum_{i=1}^{K_1}\|\Phi(k,k-i+1)\| \\ &\quad + \sum_{i=K_1+1}^{k}\|\Phi(k,k-i+1)\| \end{aligned} \tag{82}$$

With the transition matrix defined in (21), it follows that $\|\Phi(k,k-i+1)\| \le A_L^{i-1}$ for $i \le K_1$. On the other hand, for $i > K_1$, inequality (23) ensures that $\|\Phi(k,k-i+1)\| \le c_2\sqrt{\lambda}^{i-1}$, $a.s.$ and thus

$$\lim_{k\to\infty}\sum_{j=0}^{k-1}\|\Phi(k,j+1)\| \le \sum_{i=1}^{K_1}A_L^{i-1} + c_2\sum_{i=K_1+1}^{\infty}\sqrt{\lambda}^{i-1} = c_4 < \infty,\ a.\ s. \tag{83}$$

where

$$c_4 = \frac{1-A_L^{K_1}}{1-A_L} + c_2\frac{\sqrt{\lambda}^{K_1}}{1-\sqrt{\lambda}}$$

Taking the summation operation $\frac{1}{N}\sum_{k=1}^{N}$ on both sides of (81) and using (83), one can get

$$\begin{aligned}&\frac{1}{N}\sum_{k=1}^{N}\|y_s(k)\|^2\\ &\le \frac{1}{N}\frac{c_3}{1-\lambda} + \frac{3D_L^2}{N}\sum_{k=1}^{N}\|u_s(k)\|^2\\ &+\Bigg(\frac{3C_L^2B_L^2c_4}{N}\\ &\times\sum_{k=1}^{N}\sum_{j=0}^{k-1}\|\Phi(k,j+1)\|\,\|u_s(j)\|^2\Bigg),\quad a.\ s.\end{aligned} \tag{84}$$

in which the double summation term can be rearranged as follows

$$\sum_{k=1}^{N}\sum_{j=0}^{k-1}\|\Phi(k,j+1)\|\,\|u_s(j)\|^2 = \sum_{j=0}^{N-1}\sum_{k=j+1}^{N}\|\Phi(k,j+1)\|\,\|u_s(j)\|^2$$

With the same argument made from (82) to (83), it is easy to see that

$$\sum_{k=j+1}^{N}\|\Phi(k,j+1)\| \le \sum_{k=j+1}^{\infty}\|\Phi(k,j+1)\| \le c_4 < \infty,\quad a.\ s. \tag{85}$$

Therefore, following from (84) and (85), inequality (25) can be attained with

$$\begin{aligned}K_3 &= \tfrac{c_3}{1-\lambda} + 3C_L^2B_L^2c_4^2\|u_s(0)\|^2\\ &= 3C_L^2\left(\tfrac{c_2^2}{1-\lambda}\|x(0)\|^2 + B_L^2c_4^2\|u_s(0)\|^2\right)\\ K_2 &= \max\left(3D_L^2,\ 3C_L^2B_L^2c_4^2\right)\end{aligned}$$

■

C. Proof of Lemma 1

Proof: (i) Using (36), we can get

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{\phi^T(k-1)\,\phi(k-1)}{r(k-1)\,r(k-2)} &= \sum_{k=1}^{\infty}\frac{r(k-1)-r(k-2)}{r(k-1)\,r(k-2)}\\ &= \sum_{k=1}^{\infty}\frac{1}{r(k-2)} - \frac{1}{r(k-1)}\\ &\le \frac{1}{r(-1)} < \infty \quad for\ r(-1) = r_o\end{aligned}$$

Therefore, condition (43) is valid.

(ii) From equation (32) and using (37), we can obtain

$$\widehat{\theta}(k) = \widehat{\theta}(k-1) + \frac{\phi(k-1)}{r(k-2)+\phi^T(k-1)\phi(k-1)}\left[y(k)-\widehat{y}(k)\right]$$

Then multiply $\phi^T(k-1)$ to both sides of the above equation and use (38) to get

$$\phi^T(k-1)\,\widehat{\theta}(k) = \phi^T(k-1)\,\widehat{\theta}(k-1) + \frac{\phi^T(k-1)\,\phi(k-1)}{r(k-1)}e(k)$$

Subtracting $y(k)$ from both sides of the above equation, we have

$$y(k) - \overline{y}(k) = y(k) - \widehat{y}(k) - \frac{\phi^T(k-1)\,\phi(k-1)}{r(k-1)}e(k)$$

Using (36) and (38), we can get the following equation

$$\begin{aligned}\eta(k) &= \left(1 - \frac{\phi^T(k-1)\,\phi(k-1)}{r(k-1)}\right)e(k)\\ &= \frac{r(k-2)}{r(k-1)}e(k)\end{aligned}$$

This completes the proof.

(iii) From (32), (36), and (37), we can get

$$\widehat{\theta}(k) = \widehat{\theta}(k-1) + \frac{\phi(k-1)}{r(k-1)}e(k) \tag{86}$$

Subtracting θ_0 from (86) and multiplying by $w(k)\,\phi^T(k-1)$ give

$$\begin{aligned}w(k)\,\phi^T(k-1)\,\widetilde{\theta}(k) &= w(k)\,\phi^T(k-1)\,\widetilde{\theta}(k-1)\\ &+\Bigg(\frac{\phi^T(k-1)\,\phi(k-1)}{r(k-1)}\\ &\times\left[e(k)-w(k)+w(k)\right]w(k)\Bigg)\end{aligned} \tag{87}$$

Now taking the conditional mean $E\{\cdot \mid \digamma_{k-1}\}$ on (87), we have

$$\begin{aligned}&E\left\{-\beta(k)\,w(k) \mid \digamma_{k-1}\right\}\\ &= E\left\{w(k)\,\phi^T(k-1)\,\widetilde{\theta}(k-1) \mid \digamma_{k-1}\right\}\\ &+ E\Bigg\{\frac{\phi^T(k-1)\,\phi(k-1)}{r(k-1)}\\ &\times\left[e(k)-w(k)+w(k)\right]w(k) \mid \digamma_{k-1}\Bigg\}\end{aligned} \tag{88}$$

Particularly, by (38), the term $e(k) - w(k)$ can be represented as

$$\begin{aligned}e(k) - w(k) &= y(k) - \widehat{y}(k) - w(k)\\ &= \sum_{i=1}^{L}h_i\left(z(k-1)\right)\left\{(1-A_i(q^{-1}))y(k)\right.\\ &+B_i(q^{-1})u(k-1)\\ &\left.+C_i(q^{-1})w(k)\right\} - \widehat{y}(k) - w(k)\\ &= \sum_{i=1}^{L}h_i\left(z(k-1)\right)\left\{(1-A_i(q^{-1}))y(k)\right.\\ &+B_i(q^{-1})u(k-1)\\ &\left.+\left[C_i(q^{-1})-1\right]w(k)\right\} - \widehat{y}(k)\end{aligned} \tag{89}$$

From (89), the term $e(k) - w(k)$ is $F_{k-1}-$measurable. Therefore, by properties of the noise $w(k)$ in (3) and (4), it follows that (45) is concluded.

(iv) Rewrite (7) to get

$$\begin{aligned}&\sum_{i=1}^{L} h_i(z(k-1))A_i(q^{-1})y(k)\\&=\sum_{i=1}^{L} h_i(z(k-1))\left[B_i(q^{-1})u(k-1)+C_i(q^{-1})w(k)\right]\end{aligned} \tag{90}$$

Substituting (28) into (90), we have

$$\begin{aligned}&\sum_{i=1}^{L} h_i(z(k-1))[C_i(q^{-1})-q^{-1}\alpha_i(q^{-1})]y(k)\\&=\sum_{i=1}^{L} h_i(z(k-1))\left[B_i(q^{-1})u(k-1)\right.\\&\left.+C_i(q^{-1})w(k)\right]\end{aligned}$$

which leads to

$$\begin{aligned}&\sum_{i=1}^{L} h_i(z(k-1))C_i(q^{-1})[y(k)-w(k)]\\&=\sum_{i=1}^{L} h_i(z(k-1))\left[q^{-1}\alpha_i(q^{-1})y(k)\right.\\&\left.+B_i(q^{-1})u(k-1)\right]\end{aligned}$$

From (39) and (40), we subtract $\sum_{i=1}^{L} h_i(z(k-1))C_i(q^{-1})\overline{y}(k)$ from both sides of the above equation to get

$$\begin{aligned}&\sum_{i=1}^{L} h_i(z(k-1))C_i(q^{-1})[y(k)-\overline{y}(k)-w(k)]\\&=\sum_{i=1}^{L} h_i(z(k-1))\left[q^{-1}\alpha_i(q^{-1})y(k)\right.\\&\left.-C_i(q^{-1})\overline{y}(k)+B_i(q^{-1})u(k-1)\right]\end{aligned}$$

and thus

$$\begin{aligned}&\sum_{i=1}^{L} h_i(z(k-1))C_i(q^{-1})\varsigma(k)\\&=\sum_{i=1}^{L} h_i(z(k-1))\left[-(C_i(q^{-1})-1)\overline{y}(k)\right.\\&\left.+q^{-1}\alpha_i(q^{-1})y(k)+B_i(q^{-1})u(k-1)-\overline{y}(k)\right]\end{aligned}$$

Using (34), we can get the following equation

$$\begin{aligned}&\sum_{i=1}^{L} h_i(z(k-1))C_i(q^{-1})\varsigma(k)\\&=\sum_{i=1}^{L} h_i(z(k-1))\chi^T(k-1)\theta_{i0}-\overline{y}(k)\\&=\phi^T(k-1)\theta_0-\phi^T(k-1)\widehat{\theta}(k)=-\phi^T(k-1)\widetilde{\theta}(k)\\&=\beta(k)\end{aligned}$$

This completes the proof. ∎

D. Proof of Lemma 2

Proof: First define $\beta_i(k) = C_i(q^{-1})\varsigma(k)$ for $1 \le i \le L$. With the fuzzy system (46), we have

$$\beta(k)=\sum_{i=1}^{L} h_i(z(k-1))\beta_i(k) \tag{91}$$

As $C_i(q^{-1})$ is ISP [13], for any i, there is a positive number ϵ_i such that

$$\sum_{j=1}^{k}\varsigma(j)\beta_i(j)-\epsilon_i\varsigma^2(j)\ge 0 \tag{92}$$

Taking the operation $\sum_{i=1}^{L} h_i(z(k-1))$ on both side of (92) gives

$$\sum_{j=1}^{k}\left\{\sum_{i=1}^{L} h_i(z(k-1))\left[\varsigma(j)\beta_i(j)-\epsilon_i\varsigma^2(j)\right]\right\}\ge 0 \tag{93}$$

Using equation (91) and letting $\epsilon = \min\limits_{1\le i\le L}\epsilon_i$, we can see that inequality (93) implies the desired property in inequality (47). ∎

E. Proof of Theorem 3

Proof: With (32), (44), and the definitions of $e(k)$ and $\widetilde{\theta}(k)$, we have

$$\widetilde{\theta}(k)-\frac{\phi(k-1)}{r(k-2)}\eta(k)=\widetilde{\theta}(k-1) \tag{94}$$

which leads to

$$\begin{aligned}&\widetilde{\theta}^T(k)\widetilde{\theta}(k)+\frac{2\beta(k)}{r(k-2)}\eta(k)+\frac{\phi^T(k-1)\phi(k-1)}{r^2(k-2)}\eta^2(k)\\&=\widetilde{\theta}^T(k-1)\widetilde{\theta}(k-1)\end{aligned} \tag{95}$$

Now define a quadratic function $V(k) = \widetilde{\theta}^T(k)\widetilde{\theta}(k)$. Thus equation (95) can be rewritten as

$$\begin{aligned}V(k)&=V(k-1)-\frac{2\beta(k)}{r(k-2)}\eta(k)\\&-\frac{\phi^T(k-1)\phi(k-1)}{r^2(k-2)}\eta^2(k)\\&=V(k-1)-\frac{2\beta(k)\varsigma(k)}{r(k-2)}\\&-\frac{2\beta(k)w(k)}{r(k-2)}-\frac{\phi^T(k-1)\phi(k-1)}{r^2(k-2)}\eta^2(k)\\&=V(k-1)-\frac{g_1(k)}{r(k-2)}-\frac{g_2(k)}{r(k-2)}-\\&\frac{2\beta(k)w(k)}{r(k-2)}-\frac{\phi^T(k-1)\phi(k-1)}{r^2(k-2)}\eta^2(k)\end{aligned} \tag{96}$$

where $g_1(k) = 2(\beta(k)\varsigma(k) - \frac{1}{2}\varsigma^2(k))$ and $g_2(k) = \varsigma^2(k)$. In order to make the last equation into a recursive form, we make the following definitions

$$\begin{aligned}S_1(k)&=\textstyle\sum_{j=1}^{k} g_1(j)\ge 0\\S_2(k)&=\textstyle\sum_{j=1}^{k}\frac{g_2(j)}{r(j-2)}\ge 0\\S_3(k)&=\textstyle\sum_{j=1}^{k}\frac{\phi^T(j-1)\phi(j-1)}{r^2(j-2)}\eta^2(j)\ge 0\end{aligned} \tag{97}$$

Note that the fact that $S_1(k) \geq 0$ for all k follows from Lemma 2 under **Assumption 2**. With the above notations, the second term on the right hand side of (96) can be represented as

$$\frac{g_1(k)}{r(k-2)} = \frac{S_1(k) - S_1(k-1)}{r(k-2)} \tag{98}$$

and the third one can be represented by

$$\frac{g_2(k)}{r(k-2)} = S_2(k) - S_2(k-1) \tag{99}$$

and the final one can be represented by

$$\frac{\phi^T(k-1)\phi(k-1)}{r^2(k-2)}\eta^2(k) = S_3(k) - S_3(k-1) \tag{100}$$

By using (97)-(100), inequality (96) implies

$$\begin{aligned} X(k) &= (X(k-1) + S_1(k-1) \\ &\quad \times [\frac{1}{r(k-2)} - \frac{1}{r(k-3)}] - \frac{2\beta(k)w(k)}{r(k-2)}\Big) \end{aligned} \tag{101}$$

$$X(k) \leq X(k-1) - \frac{2\beta(k)w(k)}{r(k-2)} \tag{102}$$

where the nonnegative process $X(k)$ is defined as

$$X(k) = V(k) + \frac{S_1(k)}{r(k-2)} + S_2(k) + S_3(k) \tag{103}$$

Now taking the conditional mean $E\{\cdot \mid \mathcal{F}_{k-1}\}$ on (102) and using (45), we have

$$E\{X(k) \mid \mathcal{F}_{k-1}\} \leq X(k-1) + 2\frac{\phi^T(k-1)\phi(k-1)}{r(k-1)r(k-2)}\sigma_w^2, \ a.\ s. \tag{104}$$

With inequality (43), we can invoke the martingale convergence theorem [13] to obtain that

$$\lim_{k\to\infty} X(k) = X < \infty, \ a.\ s. \tag{105}$$

Almost sure convergence of the nonnegative process $X(k)$ to a bounded nonnegative random variable X implies

$$\lim_{N\to\infty}\sum_{k=1}^{N}\frac{\varsigma^2(k)}{r(k-2)} < \infty, \ a.\ s. \tag{106}$$

$$\lim_{N\to\infty}\sum_{k=1}^{N}\frac{\phi^T(k-1)\phi(k-1)}{r^2(k-2)}\eta^2(k) < \infty, \ a.\ s. \tag{107}$$

Now following from (94), we have

$$\begin{aligned} \left\|\widehat{\theta}(k) - \widehat{\theta}(k-1)\right\|^2 &= \left\|\frac{\phi(k-1)}{r(k-2)}\eta(k)\right\|^2 \\ &= \frac{\phi^T(k-1)\phi(k-1)}{r^2(k-2)}\eta^2(k) < \infty \end{aligned}$$

Therefore, by using (107), inequality (48) can be ensured.

After multiplying $\phi^T(k-1)$ to and subtracting $y(k) - w(k)$ from (94), we can get

$$\varsigma(k) + \frac{\phi^T(k-1)\phi(k-1)}{r(k-2)}\eta(k) = e(k) - w(k)$$

By Cauchy-Schwartz inequality, it follows

$$[e(k) - w(k)]^2 \leq 2\varsigma^2(k) + 2\frac{[\phi^T(k-1)\phi(k-1)]^2}{r^2(k-2)}\eta^2(k)$$

and hence

$$\begin{aligned} &\sum_{k=1}^{N}\frac{(e(k)-w(k))^2}{r(k-1)} \\ &\leq 2\sum_{k=1}^{N}\frac{\varsigma^2(k)}{r(k-1)} + 2\sum_{k=1}^{N}\frac{[\phi^T(k-1)\phi(k-1)]^2}{r(k-1)r^2(k-2)}\eta^2(k) \\ &\leq 2\sum_{k=1}^{N}\frac{\varsigma^2(k)}{r(k-2)} + 2\sum_{k=1}^{N}\frac{\phi^T(k-1)\phi(k-1)}{r^2(k-2)}\eta^2(k) \end{aligned}$$

Therefore, using (106) and (107), the last inequality leads to (49). The completes the proof. ■

F. Proof of Lemma 3

Proof: (i) Assume $r(k-1) < k_q < \infty$, then (49) implies

$$\lim_{N\to\infty}\sum_{k=1}^{N}[e(k) - w(k)]^2 < \infty \quad a.s. \tag{108}$$

With (108), we can apply Kronecker's Lemma in appendix D of [13] to get

$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}k\,[e(k) - w(k)]^2 = 0 \quad a.s.$$

which implies (51).

On the other hand, if $r(k-1)$ is unbounded. By applying Kronecker's Lemma again, we obtain

$$\begin{aligned} &\lim_{N\to\infty}\frac{1}{r(N-1)}\sum_{k=1}^{N}r(k-1)\frac{[e(k)-w(k)]^2}{r(k-1)} \\ &= \lim_{N\to\infty}\frac{N}{r(N-1)} \times \frac{1}{N}\sum_{k=1}^{N}[e(k) - w(k)]^2 = 0 \quad a.s. \end{aligned} \tag{109}$$

Substituting of (50) into (109) gives

$$\lim_{N\to\infty}\frac{\frac{1}{N}\sum_{k=1}^{N}[e(k)-w(k)]^2}{K_{a1} + \frac{K_{a2}}{N}\sum_{k=1}^{N}[e(k)-w(k)]^2} = 0 \quad a.s. \tag{110}$$

From (110), we can easily get (51).

(ii) Substituting (51) into (50), we can get (52).

(iii) For the left hand side of (53), we have

$$\begin{aligned} &E\left\{[y(k) - \widehat{y}(k)]^2 \mid \mathcal{F}_{k-1}\right\} \\ &= E\left\{[y(k) - \widehat{y}(k) - w(k) + w(k)]^2 \mid \mathcal{F}_{k-1}\right\} \\ &= E\Big\{[y(k) - \widehat{y}(k) - w(k)]^2 \\ &\quad + 2[y(k) - \widehat{y}(k) - w(k)]w(k) \\ &\quad + w^2(k) \mid \mathcal{F}_{k-1}\Big\} \ a.s. \end{aligned} \tag{111}$$

Since $y(k) - w(k)$ and $\widehat{y}(k)$ are $\mathcal{F}_{k-1}-$measurable and $E\{w(k) \mid \mathcal{F}_{k-1}\} = 0$, we can obtain

$$\begin{aligned} &E\left\{[y(k) - \widehat{y}(k)]^2 \mid \mathcal{F}_{k-1}\right\} \\ &= [e(k) - w(k)]^2 + E\left\{w^2(k) \mid \mathcal{F}_{k-1}\right\} \ a.s. \end{aligned} \tag{112}$$

From (51) and (4),we can easily derive (53).

(iv) Applying (51), (3), (4), (**??**), and Lemma D.5.2 in Appendix D of [13], we have

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\left[y(k)-\widehat{y}(k)\right]^2\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\left[y\left(k\right)-\widehat{y}\left(k\right)-w(k)+w(k)\right]^2\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]^2\\
&+2\frac{1}{N}\sum_{k=1}^{N}\left[y(k)-\widehat{y}(k)-w(k)\right]w(k)+\frac{1}{N}\sum_{k=1}^{N}w^2(k)\\
&=\sigma_w^2+\lim_{N\to\infty}\frac{2}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]w(k) \qquad (113)
\end{aligned}$$

In (113), define a process $\varpi(k)=\left[e(k)-w(k)\right]w(k)$. Then its is easy to see that $E\left\{\varpi(k)\mid \digamma_{k-1}\right\}=0$ and from (49), (4), and (52), we have

$$\sum_{k=1}^{N}\frac{1}{k^2}E\left\{\varpi^2(k)\mid \digamma_{k-1}\right\}<\infty$$

Using Lemma D.5.1 in [13], we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]w(k)=0 \qquad (114)$$

Substituting (114) into (113), we get (53). ■

G. Proof of Lemma 4

Before presenting the proof of Lemma 4, we shall need a lemma which is quoted from [27].

Lemma 5: Let P be a $m\times m$ symmetric positive definite matrix which is partitioned as

$$P=\begin{bmatrix} P_{11} & P_{12}\\ P_{12}^T & P_{22}\end{bmatrix}$$

where P_{11} and P_{22} are $(m-1)\times(m-1)$ and 1×1 matrices, respectively. Also let Γ be a matrix defined by

$$\Gamma=\begin{bmatrix} I_{m-1} & 0_{(m-1)\times 1}\\ 0_{1\times(m-1)} & r\end{bmatrix},\ 0<|r|\le 1$$

Then $\Gamma^TP\Gamma-P$ is negative semi-definite if and only if $P_{12}=0_{(m-1)\times 1}$.

Proof of Lemma 4:

(i) The equation for the signal $y(k+1)$ in (63) can be rewritten as

$$\begin{aligned}
&\frac{\lambda+\mathbf{b}_0^2(k)}{\lambda}y(k+1)\\
&=-\sum_{i=1}^{L}h_i(z(k))\left[\sum_{j=1}^{n}a_{ij}y(k+1-j)\right]+v_3(k+1)\\
&-\sum_{i=1}^{L}h_i(z(k))\left[\sum_{j=1}^{n}b_{ij}\frac{\mathbf{b}_0(k-j)}{\lambda}y(k+1-j)\right] \qquad (115)
\end{aligned}$$

where the signal $v_3(k+1)$ is defined as

$$\begin{aligned}
&v_3(k+1)\\
&=\sum_{i=1}^{L}h_i(z(k))B_i(q^{-1})\frac{\mathbf{b}_0(k)}{\lambda}\\
&\times\left[y^*(k+1)+e(k+1)-w(k+1)\right]\\
&+\sum_{i=1}^{L}h_i(z(k))\left[B_i(q^{-1})\frac{\mathbf{b}_0(k)}{\lambda}+C_i(q^{-1})\right]w(k+1) \qquad (116)
\end{aligned}$$

Now define a process $\varkappa(k)$ as

$$\varkappa(k)=\frac{\lambda}{\lambda+\mathbf{b}_0^2(k)}\frac{1}{\gamma_0}$$

With the definition of γ_0 in (62), it follows

$$0<\varkappa(k)\le 1 \qquad (117)$$

To analyze the property of the signal $y(k)$, the dynamic equation (115) will be transformed into a state-space form. By constructing a state vector $x_y(k)$ as

$$x_y(k)=\begin{bmatrix} y(k+1-n) & y(k-n+2) & \cdots & y(k)\end{bmatrix}^T,$$

the system defined by equation (115) can be expressed by

$$x_y(k+1)=\widetilde{\mathbf{A}}(k)x_y(k)+v_y(k+1) \qquad (118)$$

where

$$\begin{aligned}
\widetilde{\mathbf{A}}(k) &= \Gamma_1(k)\overline{\mathbf{A}}(k)\\
\overline{\mathbf{A}}(k) &= \mathbf{\Xi}_A(k)+\mathbf{M}_B(k)\Gamma_2(k)\\
v_y(k+1) &= \begin{bmatrix} 0_{1\times(\overline{n}-1)} & \frac{\lambda}{\lambda+\mathbf{b}_0^2(k)}v_3(k+1)\end{bmatrix}^T\\
\Gamma_1(k) &= \begin{bmatrix} I_{\overline{n}-1} & 0_{(\overline{n}-1)\times 1}\\ 0_{1\times(\overline{n}-1)} & \varkappa(k)\end{bmatrix}
\end{aligned}$$

Since, by assumption, there exist symmetric positive definite matrices $\overline{P}_i, 1\le i\le L$, of the form in (64) such that the matrix inequalities (65) hold for some $\varepsilon>0$ and some λ_0 with $0<\lambda_0<1$, and that

$$\overline{\mathbf{A}}^T(k)\overline{\mathbf{P}}_+(k)\overline{\mathbf{A}}(k)<\lambda_0\overline{\mathbf{P}}(k) \qquad (119)$$

for all $k\ge 0$. Since the matrix $\overline{P}_i$ is chosen with the special form shown in (64), we have

$$\begin{aligned}
&\widetilde{\mathbf{A}}^T(k)\overline{\mathbf{P}}_+(k)\widetilde{\mathbf{A}}(k)\\
=\ &\overline{\mathbf{A}}^T(k)\left[\sum_{i=1}^{L}h_i(z(k+1))\Gamma_1^T(k)\overline{P}_i\Gamma_1(k)\right]\overline{\mathbf{A}}(k)\\
\le\ &\overline{\mathbf{A}}^T(k)\left[\sum_{i=1}^{L}h_i(z(k+1))\overline{P}_i\right]\overline{\mathbf{A}}(k)\\
=\ &\overline{\mathbf{A}}^T(k)\overline{\mathbf{P}}_+(k)\overline{\mathbf{A}}(k)\\
<\ &\lambda_0\overline{\mathbf{P}}(k)
\end{aligned}$$

by using Lemma 5 and (119). Therefore, by applying Theorem 2, we have

$$\begin{aligned}
\frac{1}{N}\sum_{k=1}^{N}\|y(k)\|^2&\le\frac{1}{N}\sum_{k=1}^{N}\|x_y(k)\|^2\\
&\le\frac{K_2}{N}\sum_{k=1}^{N}\|v_3(k)\|^2+\frac{K_3}{N},\quad a.\ s.,\ \text{for } N\ge K_1
\end{aligned}$$

where K_1 is a sufficiently large number. Next, by the definition of $v_3(k)$,the boundedness of $y^*(k)$, and the mean-square boundedness of $w(k)$ given in (5), it follows that there are constants K_8 and K_9 such that

$$\frac{1}{N}\sum_{k=1}^{N}y^2(k) \leq \frac{K_8}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]^2 + K_9,\ a.s.,\ \text{for } N \geq K_1$$

(ii) Using (43) and (44), equation (39) gives

$$\begin{aligned}\overline{y}(k) &= e(k) + y^*(k) - \frac{r(k-2)}{r(k-1)}e(k) \\ &= \frac{\phi^T(k-1)\phi(k-1)}{r(k-1)}e(k) + y^*(k) \\ &= \frac{\phi^T(k-1)\phi(k-1)}{r(k-1)}\left[e(k)-w(k)\right] \\ &+ \frac{\phi^T(k-1)\phi(k-1)}{r(k-1)}w(k) + y^*(k) \\ &< \left[e(k)-w(k)\right] + w(k) + y^*(k) \end{aligned} \tag{120}$$

By Cauchy-Schwartz inequality, it follows

$$\overline{y}^2(k) \leq 3\left[e(k)-w(k)\right]^2 + 3w^2(k) + 3{y^*}^2(k)$$

Similar to the proof in part (i), inequality (70) can be concluded.

(iii) First, note that from the equation concerning the adaptive control law $u(k)$ defined in (60), we have

$$\begin{aligned}u(k) &= \frac{\mathbf{b_0}(k)}{\lambda}\left[e(k+1) + y^*(k+1) - y(k+1)\right] \\ &= \frac{\mathbf{b_0}(k)}{\lambda}\left[e(k+1) - w(k+1) + w(k+1)\right. \\ &\left. + y^*(k+1) - y(k+1)\right]\end{aligned}$$

which implies that

$$\begin{aligned}\frac{1}{N}\sum_{k=1}^{N}u^2(k) &\leq \frac{K_4}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]^2 + \frac{K_5}{N}\sum_{k=1}^{N}y^2(k) + K_6,\ a.s. \\ &\leq \frac{K_7}{N}\sum_{k=1}^{N}\left[e(k)-w(k)\right]^2 + K_{7a}\end{aligned} \tag{121}$$

In (36) with $k=N$, we have

$$\begin{aligned}r(N-1) &= r(0) + \sum_{k=1}^{N-1}\phi^T(N-1)\phi(N-1) \\ &= r(0) + \sum_{k=1}^{N-1}\sum_{i=1}^{L}h_i^2(z(k))\chi^T(k)\chi(k) \\ &\leq r(0) + \sum_{k=1}^{N-1}\sum_{i=1}^{L}h_i(z(k))\chi^T(k)\chi(k) \\ &= r(0) + \sum_{k=1}^{N-1}\chi^T(k)\chi(k)\end{aligned} \tag{122}$$

By the definition of $\chi(k)$ in (34), it follows from (121), (69), and (70) that, for $N \geq \overline{N}$,

$$\frac{1}{N}r(N-1) \leq \frac{K_{a2}}{N}\sum_{k=1}^{N-1}\left[e(k)-w(k)\right]^2 + K_{a1}$$

for some positive numbers K_{a2} and K_{a1}. This completes the proof. □

參考文獻

[1] M. Sugeno and K. Tanaka, "Successive identification of Fuzzy model," *Fuzzy Set and Systems,* Vol. 28, pp. 156-33, 1988.

[2] T. Takagi and M. Sugeno,"Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst. , Man, Cybern.*, Vol. 15, pp. 116-462, Jan/Feb, 1985.

[3] M. Sugeno and T. Yasukawa," A fuzzy logic-based approach to qualitative modeling," *IEEE Trans. Fuzzy Systems*, Vol. 1, No. 1, pp.7-31, 1993.

[4] L. X. Wang, *Adaptive Fuzzy Systems and Control: Design and Stability Analysis*, Englewood Cliffs, NJ, Prentice-Hall, 1994.

[5] L. X. Wang, "Stable adaptive fuzzy control of nonlinear systems, "*IEEE Trans. Fuzzy System*, Vol. pp. 146-155, 1993.

[6] L. X. Wang, "Modeling and Control of hierachical systems with fuzzy systems, " *Automatica*, Vol.33, pp. 1041-1053, 1997.

[7] T. K. Yin and C. S. George Lee, " Fuzzy model-reference adaptive control," *IEEE Trans. System, Man, Cybern.*, Vol. 25, pp1606-1615, 1995.

[8] K. M. Passino and S. Yurkovich, *Fuzzy Control*, Addison-Wesley, Monlo Park, California, 1998.

[9] B. Kosko, *Neural Network and Fuzzy System: A dynamical systems approach to machine intelligence*, Prentice-Hall, Englewood Cliffs, N.J. 1992.

[10] I. S. R. Jang, C. T. Sun and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1997

[11] B. S. Chen, C. H. Lee, and Y. C. Chang, "H^∞ tracking design of uncertain nonlinear SISO systems: adaptive fuzzy approach, " *IEEE Trans, Fuzzy Syst.*, Vol.4, pp. 32-43, 1996.

[12] D. Williamson, *Digital Control and Implementation*, Prentice Hall, Englewood Cliffs, N.J. 1991

[13] G. C. Goodwin and K.S. Sin, *Adaptive filtering prediction and control*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1984.

[14] K. Watanabe, "Stochastic fuzzy control. I. Theoretical derivation," *Proceedings of 1995 IEEE International Conference on Fuzzy Systems*, vol.2, pp. 547 -554, 1995.

[15] K. Watanabe, K. Izumi and Fuha Han, "Stochastic fuzzy servo control using multiple linear dynamic models," *Proceedings of the Second International Conference on Knowledge-Based Intelligent Electronic Systems*, vol.3, pp. 474 -482, 1998.

[16] E. G. Laukonen, K. M. Passino, V. Krihnaswami, G.-C. Luh, and G. Rizzoni, "Fault detection and isolation for an experimental internal combustion engine via fuzzy identification," *IEEE Trans. Control System Technology*, vol. 3, no. 3, pp. 347-355, 1995.

[17] J. Hu, K. Kumamaru, and K. Inoue, "A hybrid quasi-ARMAX modeling scheme for identification and control of nonlinear systems, " *Proceedings of the 35th IEEE Conference on Decision and Control*, vol. 2, pp. 1413 -1418, 1996.

[18] J. B. Waller, J. Hu, and K. Kirasawa, "Nonlinear model predictive control utilizing a neuro-fuzzy predictor," *2000 IEEE International Conference on Systems, Man, and Cybernetics*, vol. 5, pp. 3459 -3464, 2000.

[19] H.-T. Yang and C.-M. Huang, "A new short-term load forecasting approach using self-organizing fuzzy ARMAX models," *IEEE Transactions on Power Systems*, vol. 13, no. 1, pp. 217 -225, 1998.

[20] T. T. Ho, "Stochastic fuzzy direct adaptive control ," *Proceedings of the Third IEEE Conference on Fuzzy Systems*, vol.2, pp. 750 -755, 1994.

[21] T. A. Johansen, "Fuzzy model based control: Stability, robustness, and performance issues," *IEEE Trans. on Fuzzy Systems*, vol. 2, no. 3, pp. 221-234, 1994.

[22] A. N. Shiryayev, *Probability*, in Graduate Texts in Mathematics Series , vol. 95, Springer-Verlag, New York, 1984.

[24] C. T. Chen, *Linear System Theory and Design*, Oxford University Press, 3rd Edition, 1999.

[25] P. J. Brockwell and R. A. Davis, *Time Series : Theory and Methods*, in Springer series in statistics, Springer-Verlag, New York, 1991.

[26] L. Ljung, *System Identification: Theory for the User*, Prentice Hall PTR, Upper Saddle River, New Jersey, 1987.

[27] Bor-Sen Chen, Bore-Kuen Lee, and Ling-Bin Guo, "Optimal tracking design for stochastic fuzzy systems," *IEEE Transactions on Fuzzy Systems*, vol. 11, no. 6, pp. 796 - 813, Dec. 2003.

[28] R. Li, "A weighted adaptive one-step-ahead minimum variance controller based on the ELS algorithm," *International Journal of Adaptive Control and Signal Processing*, vol. 11, pp. 461-474, 1997.

[29] Chung-Hung Chiu, *Adaptive Minimum Variance Control of Stochastic Fuzzy Models*, Master Thesis, Institute of Electrical Engineering, Chung Hua University, 2007.

行政院國家科學委員會補助國內專家學者出席國際學術會議報告

98 年 07 月 21 日

附件三

報告人姓名	李柏坤	服務機構及職稱	中華大學電機系 教授
時間 會議 地點	98 年 7 月 12 日至 98 年 7 月 15 日 中國河北保定	本會核定補助文號	國科會 NSC 97-2221-E-216-033 計畫中核定出席國際會議經費五萬元
會議名稱	(中文)2009 年機器學習與人工頭腦學國際研討會 (ICMLC 2009) (英文)2009 International Conference on Machine Learning and Cybernetics		
發表論文題目	(中文)具有狀態相關雜訊的離散時間隨機 T-S 模糊 ARMAX 模式之 H∞輸出回授控制 (英文)H∞ Output Feedback Control of Discrete-Time Stochastic T-S Fuzzy Models with State-Dependent Noise		

一、參加會議經過:

此次2009年機器學習與人工頭腦學國際研討會(2009 International Conference on Machine Learning and Cybernetics ，ICMLC 2009)，由河北大學、華南理工大學、IEEE SMC (System, Man, and Cybernetics) 協會等單位聯合主辦，於 98 年 7 月 12 日到 98 年 7 月 15 日，在中國河北保定市之電谷錦江國際酒店舉行。台灣的學者參與此研討會非常踴躍。

此次研討會所有論文都列入 IEEE Explorer 之資料庫，都屬於 EI Index。研討會之網路首頁為 http://www.icmlc.com/，整個研討會包含三個 Plenary Talk：

[1] Brain-Machine Interfaces, Speaker: Prof. Jose M. Carmena，University of California, Berkeley, USA

[2] Adversarial Pattern Classification, Speaker: Prof. Fabio Roli, University of Cagliari, Italy

[3] On Leveraging Unlabeled Data and Classifier Combination, Speaker: Prof. Zhi-Hua Zhou, Nanjing University, China

另外有兩個 Tutorials：

[1] An Introduction to Graphical Models and Bayesian Nonparametrics, Speaker: Prof. Michael Jordan

[2] Publish or perish (P2): how to successfully publish your research, Speaker: Prof. Witold Pedrycz

此次研討會之主題包含：

1. Adaptive systems
2. Business intelligence
3. Biometrics
4. Bioinformatics
5. Data and web mining
6. Intelligent agent
7. Financial engineering
8. Inductive learning
9. Geoinformatics
10. Pattern recognition
11. Logistics
12. Intelligent control
13. Media computing
14. Neural net and support vector machine
15. Hybrid and nonlinear system
16. Fuzzy set theory, fuzzy control and system
17. Knowledge management
18. Information retrieval

19. Intelligent and knowledge based system
20. Rough and fuzzy rough set
21. Networking and information security
22. Evolutionary computation
23. Ensemble method
24. Information fusion
25. Visual information processing
26. Computational life science

二、與會心得

(1) 從此次研討會所安排之主題來看，比較偏向人工智能於資訊工程之研究，各國有關人工智慧理論都有顯著的研究成果，幾個比較新的主題如 Media computing、Bioinformatics、Computational life science、Business intelligence，非常值得國內學界注意其發展。

(2) 除了認識許多中國之學者外，也認識了很多來自全世界各地的菁英學者，對於將來推動國際學術交流，有相當大的幫助。

(3) 大陸在人工智能領域之研究成果亦有長足之進步，在 IEEE SMC Society 之影響力也已超過台灣相關學界，國內應該即起直追。

三、考察參觀活動(無是項活動者省略)

7 月 14 日上午，應河北大學數學與電腦學院邀請，由元智大學電機工程系林志民教授帶隊的臺灣學術代表團到河北大學參觀訪問。

代表團首先參觀了數學與電腦學院，學院負責人向臺灣代表團介紹了學院的基本情況，臺灣代表團林志民教授、范國清教授、李柏坤教授、陳邱雄教授、劉旭光博士分別介紹了各自學校及所在學院的基本情況。

然後，代表團成員拜會了河北大學副校長哈明虎教授。哈明虎教授向代表團成員簡要介紹了河北大學的歷史和發展現狀，並陪同參觀了坤輿全圖等圖書館藏。

最後，代表團到河北大學新校區參觀了機器學習與計算智慧省級重點實驗室，與實驗室師生進行了學術交流，就共同感興趣的學術問題進行了有益的探討。

四、建議

(1) 台灣應該多爭取舉辦國際研討會，使得全世界各地的菁英學者，能夠共聚一堂。

(2) 大陸學術界的國際化，已經逐步生根，同時以此為基礎邁向國際之競爭，台灣學術界的國際化還有待大家的努力。

五、攜回資料名稱及內容

(1) 完整論文光碟片。

(2) 論文摘要紙本以及部份之論文集。

六、其他

無。