

Time-varying non-linear system identification using dynamical adaptive fuzzy systems

M. Cerrada¹, J. Aguilar², E. Colina³ y A. Titli⁴

¹Departamento de Sistemas de Control, CEMISID, e-mail: cerradam@ing.ula.ve

²Departamento de Computación, CEMISID, e-mail: aguilar@ing.ula.ve

³Postgrado de Control y Automatización, e-mail: ecolina@ing.ula.ve

Universidad de Los Andes, Facultad de Ingeniería, Avenida Tulio Febres Cordero, Mérida 5101, Venezuela. Telf: 58-274-2402983 Fax: 58-274-2402846.

⁴Laboratoire d'Analysis et d'Architecture de Systèmes (LAAS-CNRS), Groupe DISCO, e-mail: titli@laas.fr 7, avenue du Colonel Roche-31077, Toulouse, Cedex, France.

Phone: 33-5-61336200 Fax: 33-5-61553577

Abstract

In processes control, it is required to have a model that describes the process behaviour. Identification techniques provide an approximate representation of the real system behaviour, by using a linguistic or mathematical expression, or an algorithm. In this work, a dynamical adaptive fuzzy model is used in order to propose the input-output identification model of a time-varying non-linear system. This class of fuzzy models permit a generic adaptive approach of the process behaviour through the adaptation of the membership functions to the universes of discourse of the model variables.

Key words: System identification, adaptive fuzzy systems, dynamical fuzzy systems, fuzzy logic.

Identificación de un sistema no lineal variante en el tiempo usando sistemas difusos adaptativos dinámicos

Resumen

Para realizar las tareas de control de procesos se requiere de un modelo que describa el comportamiento aproximado del proceso en particular bajo estudio. Las técnicas de identificación proporcionan una representación del comportamiento de un sistema real, basándose en una expresión lingüística, matemática o un algoritmo. En este trabajo se propone el uso de modelos difusos adaptativos dinámicos para la obtención de un modelo de identificación entrada-salida de un sistema no lineal variante en el tiempo. La ventaja de usar esta clase de modelos difusos es la de obtener un aproximación genérica del comportamiento del proceso a identificar a partir de la adaptación constante de las funciones de membresía de sus variables sobre todo el universo de discurso.

Palabras clave: Identificación de sistemas, sistemas difusos adaptativos, sistemas difusos dinámicos, lógica difusa.

1. Introducción

La identificación de sistemas comprende un conjunto de modelos y métodos que permiten aproximar el comportamiento real de un sistema desconocido a través de un modelo matemático

válido capaz de describir las propiedades esenciales del sistema. Dichos modelos pueden ser usados posteriormente en tareas de control, manejo de fallas, o cualquier otra actividad en la que se requiera conocer la dinámica del sistema. En la literatura se encuentra una gran variedad de

métodos de identificación, los cuales van desde los basados en la teoría clásica [1], hasta las nuevas técnicas basadas en la computación inteligente [2-6].

El uso de la Lógica Difusa se ha extendido en esta área como herramienta de modelado, utilizando el conocimiento experto para derivar una serie de reglas difusas SI-ENTONCES que permiten obtener un modelo lingüístico aproximado del proceso con la capacidad de manipular la incertidumbre [7]. Sin embargo, si el conocimiento experto es pobre, el modelo difuso obtenido no tendrá un desempeño adecuado. Esta limitación ha dado lugar a la proposición de modelos difusos adaptativos paramétricos que permiten incorporar algoritmos de aprendizaje para la adaptación de sus parámetros [8]. En [9] se propone un modelo difuso adaptativo que es dinámico. Esta propiedad incorpora en las funciones de membresía la dinámica temporal de las variables del modelo difuso, lo que les permite adaptarse a los cambios en los dominios de discurso de sus variables. Esta característica permite mejorar algunas limitaciones encontradas en los sistemas difusos adaptativos clásicos [10].

El uso de un modelo difuso con esta propiedad en el área de identificación de sistemas es interesante, puesto que permite incorporar en el modelo los cambios que puedan ocurrir en los dominios de las variables del sistema, debido a perturbaciones internas o externas o a dinámicas variantes en el tiempo, logrando un modelo adecuado para aplicaciones posteriores en control, manejo de fallas, etc.

En este trabajo presentamos una aplicación de los modelos difusos adaptativos dinámicos en la identificación entrada-salida de un sistema no lineal variante en el tiempo, permitiendo proponer un modelo capaz de aproximar el comportamiento del sistema aun bajo cambios bruscos en las señales de entrada. El sistema seleccionado para demostrar las capacidades de estos modelos difusos adaptativos dinámicos es el de un proceso de reacción que se lleva a cabo en un tanque reactor continuamente agitado [11].

En la segunda sección, se presentan los aspectos teóricos de base sobre el modelo difuso adaptativo dinámico propuesto y su aplicación en la identificación de sistemas. En la tercera parte, se presentan los experimentos realizados

para obtener el modelo de identificación del sistema bajo estudio y finalmente, en la cuarta parte, se presentan las conclusiones.

2. El Modelo Difuso Adaptativo Dinámico

A continuación se dan los aspectos teóricos de base sobre los modelos difusos adaptativos dinámicos (MDAD).

2.1. Preliminares

Definición 1. Un modelo lógico difuso de múltiples entradas y una salida (MIMO, por sus siglas en inglés), es un modelo lingüístico descrito por una base de M reglas difusas de la forma:

R^l : Si x_1 es F_1^l y ... y x_n es F_n^l entonces y es G^l (1)

donde $\underline{X} = (x_1 \ x_2 \dots \ x_n)^T$ es un vector de variables lingüísticas de entrada x_i definidas sobre un universo de discurso U_i ; y es la variable lingüística de salida definida sobre un universo de discurso V ; F_i^l y G^l son conjuntos difusos sobre U_i y V , respectivamente, $i=1, \dots, n$; $l=1, \dots, M$, definidos cada uno por sus funciones de membresía.

Definición 2. Una función de membresía es **dinámica** si su estructura o parámetros cambian dinámicamente en el tiempo.

Definición 3. Un modelo difuso es **dinámico** si sus funciones de membresía son **dinámicas**.

Definición 4. Sea $x(t_j)$ el valor de la variable difusa x en el instante de tiempo t_j . La función de membresía gaussiana dinámica $\mu_F(x, t_j)$ que determina el grado de pertenencia de $x(t_j)$ al conjunto difuso F sobre un universo de discurso U , se define como:

$$\mu_F(x, t_j) = \exp \left[- \frac{(x(t_j) - \alpha(\underline{v}, t_j))^2}{\beta(\underline{w}, t_j)} \right] \quad (2)$$

donde $\alpha(\underline{v}, t_j)$ y $\beta(\underline{w}, t_j)$ son funciones dependientes del tiempo, \underline{v} y \underline{w} son vectores de parámetros ajustables, $\underline{v} \in \mathfrak{R}^q$, $\underline{w} \in \mathfrak{R}^r$. En la Figura 1 se ilustra la idea sobre las funciones de membresía gaussianas dinámicas.

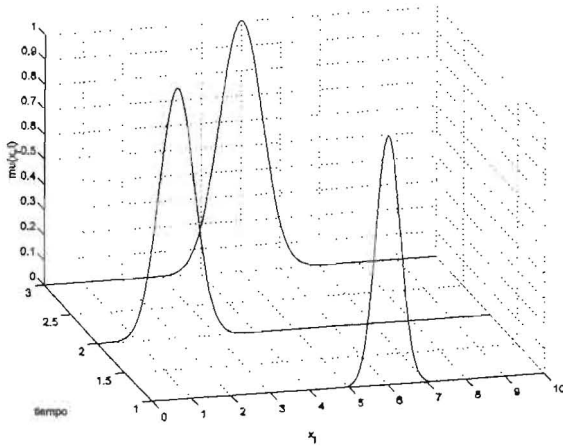


Figura 1. Funciones de membresía dinámicas para una variable x_i dada.

Definición 5. Sea $y(t_j)$ el valores de la variable difusa y en el instante de tiempo t_j . El centro de la función de membresía dinámica que determina el grado de pertenencia de $y(t_j)$ al conjunto difuso G sobre un universo de discurso V , es definida como $\gamma(\underline{u}, t_j)$, la cual es una función dependiente del tiempo, \underline{u} es un vector de parámetros ajustables, $\underline{u} \in \mathbb{R}^p$.

Usando los definiciones anteriores, el MDAD obtenido a partir de la base de reglas (1), usando un mecanismos de inferencia que considera valores difusos de entrada descritas por "singleton" difusos, funciones de membresía gaussianas para los conjuntos difusos asociados a las variables difusas de entrada y el método de difusificación de centro promedio ("center-average method", en inglés), es descrito por Cerrada [9]:

$$y(X, t_j) = \frac{\sum_{l=1}^M \gamma^l(\underline{u}^l, t_j) \left[\prod_{i=1}^n \exp \left[-\frac{(x_i(t_j) - \alpha_i^l(\underline{v}_i^l, t_j))^2}{\beta_i^l(\underline{w}_i^l, t_j)} \right] \right]}{\sum_{l=1}^M \left[\prod_{i=1}^n \exp \left[-\frac{(x_i(t_j) - \alpha_i^l(\underline{v}_i^l, t_j))^2}{\beta_i^l(\underline{w}_i^l, t_j)} \right] \right]} \quad (3)$$

donde X es un vector de valores de las variables de entrada x_i , y es la variable de salida; $\alpha_i^l(\underline{v}_i^l, t)$ y $\beta_i^l(\underline{w}_i^l, t)$ son funciones que definen la media y la varianza, respectivamente, de la función de

membresía gaussiana del conjunto difuso F_i^l ; $\gamma^l(\underline{u}^l, t)$ es la función que define el centro del conjunto difuso de salida G^l ; \underline{u}^l es un vector de parámetros u_p^l de la función γ^l ; \underline{v}_i^l es un vector de parámetros v_{iq}^l de la función α_i^l ; \underline{w}_i^l es un vector de parámetros w_{ir}^l de la función β_i^l , $p=1, \dots, P$; $q=1, \dots, Q$; $r=1, \dots, R$. u_p^l, v_{iq}^l y w_{ir}^l son parámetros ajustables.

Definición 6. Sean $x_i(t_j)$ los valores de las variables de entrada x_i al MDAD en el instante de tiempo t_j para generar la salida $y(t_j)$. La estructura genérica de las funciones $\alpha_i^l(\underline{v}_i^l, t_j)$, $\beta_i^l(\underline{w}_i^l, t_j)$ y $\gamma^l(\underline{u}^l, t_j)$ del MDAD (3), vienen definidas por las ecuaciones siguientes:

$$\alpha(\underline{v}_i^l, t_j) = v_{i1}^l \bar{x}_i(t_j) \quad (4)$$

$$\beta_i^l(\underline{w}_i^l, t_j) = w_{i1}^l (\sigma_i^2(t_j) + w_{i2}^l) \quad (5)$$

$$\gamma^l(\underline{u}, t_j) = u_1^l \bar{y}(t_j) \quad (6)$$

donde:

$$\bar{x}_i(t_j) = \frac{\sum_{k=1}^j x_i(t_k)}{j} \quad (7)$$

$$\sigma_i^l(t_j) = \frac{\sum_{k=1}^j (x_i(t_k) - \bar{x}_i(t_k))^2}{j} \quad (8)$$

$$\bar{y}(t_j) = \frac{\sum_{k=j-\delta}^{j-1} y(t_k)}{\delta}, \quad \delta \in \mathbb{N} \quad (9)$$

o, alternativamente:

$$\bar{y}(t_j) = \frac{\sum_{k=1}^{j-1} y(t_k)}{j-1} \quad (10)$$

En [9] se estudia con más detalles los aspectos relacionados con los MDAD.

2.2. Aplicación del MDAD en la identificación de sistemas

El MDAD definido en la sección anterior puede usarse para proponer modelos de identificación entrada-salida [1]. De esta manera, el ob-

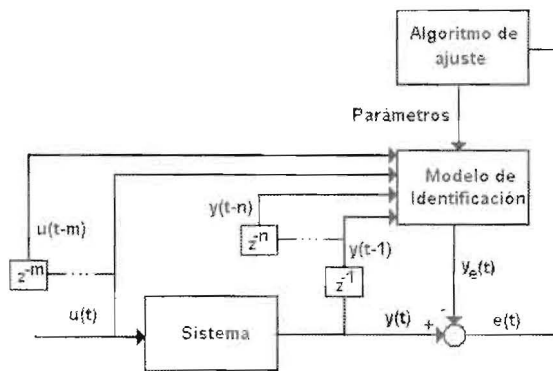


Figura 2. Esquema de identificación entrada-salida.

jetivo es ajustar los parámetros del MDAD con el fin de minimizar el error entre la salida real del sistema y la salida estimada por el modelo (usualmente llamado "error de identificación" o "error de predicción"). La Figura 2 muestra el esquema clásico que se sigue en tareas de identificación de sistemas basadas en ese enfoque.

De esta manera, usando este esquema de identificación, el vector de entrada \underline{X} al MDAD descrito por (3) está formado por los valores de $u(t)$ al sistema real en t y en m instantes anteriores y los valores de su salida $y(t)$ en $t-1$ y en n instantes anteriores. Esto es:

$$\underline{X} = [u(t_j) \ u(t_{j-1}) \ u(t_{j-m}) \ y(t_{j-1}) \ y(t_{j-2}) \ y(t_{j-n})] \quad (11)$$

En este trabajo, se usa un algoritmo de aprendizaje supervisado fuera de línea basado en el descenso del gradiente (DG), como método de ajuste de los parámetros u_p^l , v_{iq}^l y w_r^l del MDAD.

Algoritmo de ajuste de parámetros basado en el error de predicción y DG. Sea el promedio de error cuadrático medio ecm entre la salida real $y(t_j)$ y la salida $y_e(t_j)$ estimada por el modelo difuso, sobre una colección de patrones históricos $\{(X(t_i), y(t_i)), i=1, \dots, N\}$, del sistema a modelar:

$$ecm = \frac{1}{T} \sum_{j=1}^T \frac{1}{2} (y_e(t_j) - y(t_j))^2 \quad (12)$$

Usando el método del descenso del gradiente, las leyes de actualización de los parámetros viene dada por las ecuaciones (13), (14), (15):

$$u_p^l(K+1) = u_p^l(K) - \rho_1 \left. \frac{\partial ecm}{\partial u_p^l} \right|_K \quad (13)$$

$$v_{iq}^l(K+1) = v_{iq}^l(K) - \rho_2 \left. \frac{\partial ecm}{\partial v_{iq}^l} \right|_K \quad (14)$$

$$w_r^l(K+1) = w_r^l(K) - \rho_3 \left. \frac{\partial ecm}{\partial w_r^l} \right|_K \quad (15)$$

donde p, q, r son los identificadores de los parámetros; i es el identificador de la variable de entrada; l es el identificador de la regla; K es el número de iteración para la actualización de los parámetros y ρ_j es la tasa de aprendizaje, $j=1,2,3$.

Desarrollando las expresiones anteriores, se tiene que:

$$\left. \frac{\partial ecm}{\partial u_p^l} \right|_K = \frac{1}{T} \sum_{j=1}^T \left[\frac{y_e(t_j) - y(t_j)}{d(t_j)} h^l(t_j) \frac{\partial \gamma^l(\underline{u}^l, t_j)}{\partial u_p^l} \right]_K \quad (16)$$

$$\left. \frac{\partial ecm}{\partial v_{iq}^l} \right|_K = \frac{1}{T} \sum_{j=1}^T \left[2 \frac{y_e(t_j) - y(t_j)}{d(t_j)} (\gamma^l(\underline{u}^l, t_j) - y_e(t_j)) \cdot \right.$$

$$\left. h^l(t_j) \frac{(x_i - \alpha_i^l(v_i^l, t_j))}{\beta_i^l(w_i^l, t_j)} \frac{\partial \alpha_i^l(v_i^l, t_j)}{\partial v_{iq}^l} \right]_K \quad (17)$$

$$\left. \frac{\partial ecm}{\partial w_r^l} \right|_K = \frac{1}{T} \sum_{j=1}^T \left[\frac{y_e(t_j) - y(t_j)}{d(t_j)} (\gamma^l(\underline{u}^l, t_j) - y_e(t_j)) \cdot \right.$$

$$\left. h^l(t_j) \frac{(x_i - \alpha_i^l(v_i^l, t_j))^2}{(\beta_i^l(w_i^l, t_j))^2} \frac{\partial \beta_i^l(w_i^l, t_j)}{\partial w_r^l} \right]_K \quad (18)$$

donde:

$$d(t_j) = \sum_{i=1}^M h^l(t_j)$$

$$h(t_j) = \prod_{i=1}^n \exp \left[- \frac{(x_i - \alpha_i^l(v_i^l, t))^2}{\beta_i^l(w_i^l, t)} \right]$$

Sustituyendo (16), (17) y (18) respectivamente en (13), (14) y (15), se obtienen las leyes de adaptación de los parámetros. Para el caso particular de las funciones propuestas en (4), (5) y (6), se tiene que:

$$\frac{\partial \alpha_i^l(w_i^l, t_j)}{\partial w_{i1}^l} = \bar{x}_i(t_j) \quad (19)$$

$$\frac{\partial \beta_i^l(w_i^l, t_j)}{\partial w_{i1}^l} = s_i^2(t_j) + w_{i2}^l \quad (20)$$

$$\frac{\partial \beta_i^l(w_i^l, t_j)}{\partial w_{i2}^l} = w_{i1}^l \quad (21)$$

$$\frac{\partial \gamma^l(u^l, t_j)}{\partial u_1^l} = \bar{y}(t_j) \quad (22)$$

3. Identificación de un Proceso de Reacción No Lineal Variante en el Tiempo

Considérese un tanque reactor de volumen constante V_r , es cual es alimentado con una tasa volumétrica F [11]. Los reactantes dentro del reactor son continuamente agitados con el fin de mantener una mezcla homogénea, produciéndose una reacción química sucesiva isotérmica $A \rightleftharpoons B \rightarrow C$. Los productos son tomados en el fondo del tanque, según se ilustra en la Figura 3.

En presencia de un sitio de reacción separada sobre el catalizador, la cinética de esos sitios es usualmente diferente y variante en el tiempo. La dinámica de este sistema es descrita por las siguientes ecuaciones:

$$\begin{aligned} \frac{\partial z_1}{\partial t} &= 1 - D_{a1}z_1 + D_{a2}z_2^2 \\ \frac{\partial z_2}{\partial t} &= z_2 - D_{a1}z_1 - D_{a2}z_2^2 - D_{a3}d_2(t)z_2^2 + u \\ \frac{\partial z_3}{\partial t} &= -z_3 - D_{a3}d_2(t)z_2^2 \\ y &= z_3 \end{aligned} \quad (23)$$

donde:

$D_{a1} = k_1 d_1 V_r / F = 3$; $D_{a2} = k_2 d_1 V_r / F = 0.5$; $D_{a3} = k_3 V_r / F = 1$; F es la tasa de alimentación volumétrica; d_1 es el primer sitio de reacción sobre

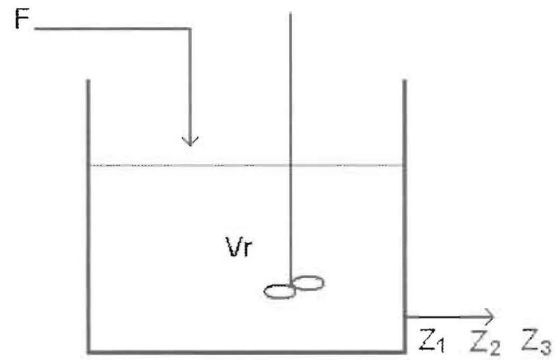


Figura 3. Tanque reactor continuamente agitado.

el catalizador; $d_2(t)$ es el segundo sitio de reacción sobre el catalizador ($d_2(t) = 1 + 0.1 \sin t$); k_1 es la tasa de reacción constante de primer orden; k_2 y k_3 son tasas de reacción constante de segundo orden; V_r es el volumen del reactor; z_1 es la conversión del reactante A; z_2 es la conversión del reactante medio B; z_3 es la conversión del producto C; u es la señal de control (entrada al sistema) y y es la variable de salida.

Por otro lado, la tasa de reacción k_3 es desconocida, lo que se considera como una incertidumbre paramétrica en este sistema.

Las características dinámicas de este sistema (sistema con incertidumbre y variante en el tiempo), lo hace particularmente interesante para probar las capacidades de MDAD como identificador difuso.

3.1. Fase experimental

A continuación, se presentan un conjunto de experimentos que han sido realizados con el fin de proponer un MDAD adecuado para el proceso bajo estudio. En cada experimento se ponen a prueba los elementos fundamentales que definen a un MDAD: la estructura de las funciones de membresía correspondientes a las variables difusas de entrada y la estructura de las funciones que definen a los centros de los conjuntos difusos de salida.

Los resultados que se presentan en esta sección han sido simulados usando la herramienta de computación MatLab[®], usando una tasa de muestreo de 0.1 seg. En la realización de