MODELOS DIFUSOS DE TAKAGI Y SUGENO

Estos modelos se caracterizan por relaciones basadas en reglas difusas, donde las premisas de cada regla representan subespacios difusos y las consecuencias son una relación lineal de entrada-salida (Takagi y Sugeno, 1995).

Las variables de entrada en las premisas de cada regla son relacionadas por operadores "y" y la variable de salida es función de las variables de estado, en general, una función lineal.

Por lo tanto, las reglas del modelo tienen la siguiente forma:

$$R_i : \text{Si } X1 \text{ es } Al_i \text{ y} \ldots \text{y } Xk \text{ es } Ak_i$$
$$\text{entonces } Y_i = f_i(X1, \ldots, Xk)$$

donde

X1, ..., Xk son las variables de entrada o premisas de las reglas,
$A1_i$, ..., Ak_i son los conjuntos difusos asociados a las variables de entrada,
$p_0^1, \ldots, p_k^i$ son los parámetros de la regla i, e
Y_i es la salida de la regla i.
f_i es una función que generalmente es lineal, es decir:

$$Y_i = p_0^i + p_1^i X1 + \ldots + p_k^i Xk$$

Por lo tanto, la salida del modelo, Y, se obtiene ponderando la salida de cada regla por su respectivo grado de cumplimiento W_i, es decir:

$$Y = \sum_{i=1}^{M} (W_i \, Y_i) / (\sum_{i=1}^{M} W_i)$$

donde M es el número de reglas del modelo y W_i corresponde al grado de activación de la regla i, que se define como:

$$w_i = \text{oper}\left(\mu_{A1_i}, \ldots, \mu_{Am_i}, \ldots, \mu_{Ak_i}\right)$$

donde "oper" es el operador mínimo o el producto, y μ_{Am_i} es el grado de pertenencia de la variable de entrada Xm al conjunto difuso Am_i para $m = 1, \ldots, k$.

En particular, los modelos dinámicos de tiempo discreto de Takagi y Sugeno están dados por:

$$\begin{aligned} R_i: \text{Si } & y(t-1) \text{ es } Al_i \text{ y } \ldots \text{ y } y(t-ny) \text{ es } Any_i \text{ y} \\ & u(t-1) \text{ es } B1_i \text{ y } \ldots \text{ y } u(t-nu) \text{ es } Bnu_i \\ \text{entonces } & y_i(t) = a_1^i y(t-1) + \ldots + a_{ny}^i y(t-ny) + \\ & b_1^i u(t-1) + \ldots + b_{nu}^i u(t-nu) + c^i \end{aligned}$$

donde a_j^i, b_j^i y c^i son parámetros de los modelos de las consecuencias.

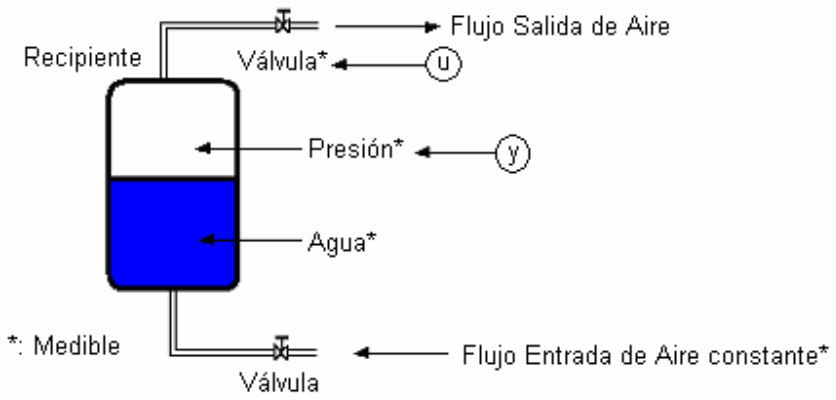
También los modelos dinámicos de Takagi y Sugeno se pueden representar en variables de estado, es decir:

$$R_i: \text{ Si } x_1(t) \text{ es } Al_i \text{ y } \ldots \text{ y } x_n(t) \text{ es } An_i$$
$$\text{entonces } x_i(t+1) = A^i x(t) + B^i u(t) + C^i$$

donde $x = [x_1, \ldots, x_n]$ es un vector de variables de estado del proceso, y A^i, B^i y C^i son las matrices de los modelos lineales en variables de estado de las consecuencias.

Ejemplo Modelo difuso de Takagi y Sugeno para un fermentador de alimentación.

La presión en el estanque de fermentación puede ser controlada a través del cambio de flujo de aire de salida manteniendo constante el flujo de aire de entrada.



A continuación se presentan la base de reglas del modelo de Takagi y Sugeno que caracterizan al fermentador.

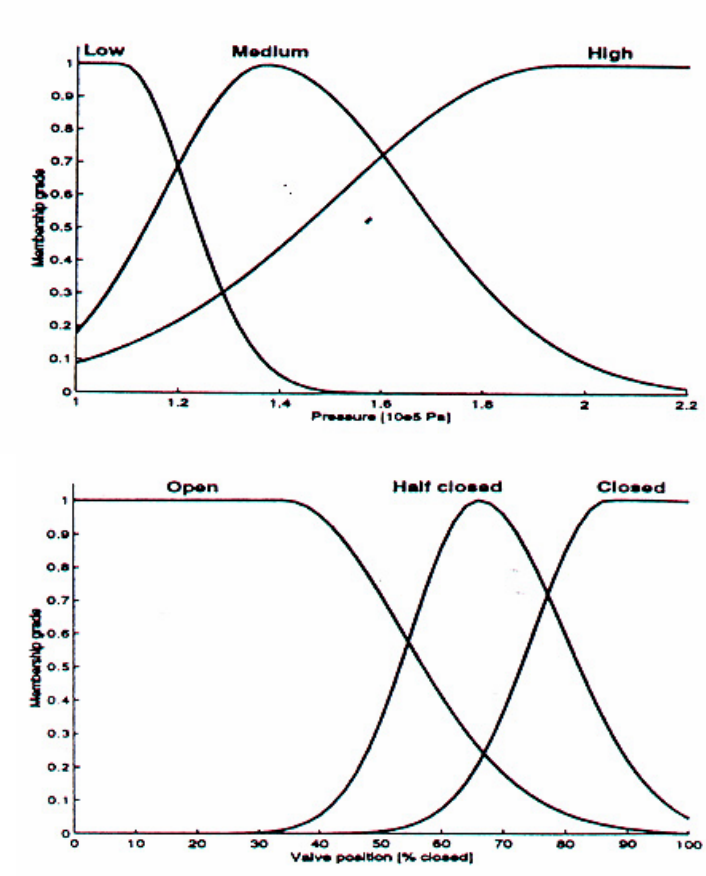
R1:

*Si la presión y(k) es **LOW** y la válvula u(k) está **OPEN***
Entonces y(k+1)=0.67y(k)+0.0007u(k)+0.35

R2:
*Si la presión y(k) es **MEDIUM** y la válvula u(k) está **HALF-CLOSED***
Entonces y(k+1)=0.80y(k)+0.0028u(k)+0.07

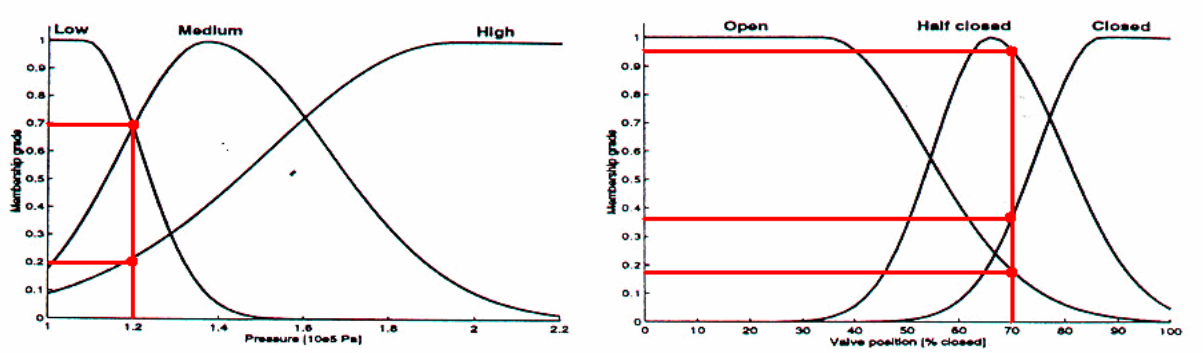
R3:
*Si la presión y(k) es **HIGH** y la válvula u(k) está **CLOSED***
Entonces y(k+1)=0.90y(k)+0.0071u(k)-0.39

Funciones de pertenencia del modelo de Takagi y Sugeno que caracterizan al fermentador:



Por ejemplo, calcule la estimación del modelo si y(k)=1.2 u(k)=70.

EL761 D. Sáez, C. Cortés, A. Nuñez (2009).



$$\mu_{LOW}(1.2)=0.7,\ \mu_{MEDIUM}(1.2)=0.7,\ \mu_{HIGH}(1.2)=0.2.$$

$$\mu_{OPEN}(70)=0.18,\ \mu_{HALF-CLOSED}(70)=0.95,\ \mu_{CLOSED}(70)=0.35.$$

Luego se calcula el grado de activación de cada regla. Suponiendo producto como el operador AND.

$$W_1=\mu_{LOW}(1.2)\cdot\mu_{OPEN}(70)=0.7\cdot 0.18=0.126$$

$$W_2=\mu_{MEDIUM}(1.2)\cdot\mu_{HALF-CLOSED}(70)=0.7\cdot 0.95=0.665$$

$$W_3=\mu_{HIGH}(1.2)\,\mu_{CLOSED}(70)=0.2\cdot 0.35=0.07$$

Ahora se calcula la salida de cada regla.

$$Y^1=0.67\cdot 1.2+0.0007\cdot 70+0.35=1.2030$$

$$Y^2=0.80\cdot 1.2+0.0028\cdot 70+0.07=1.2260$$

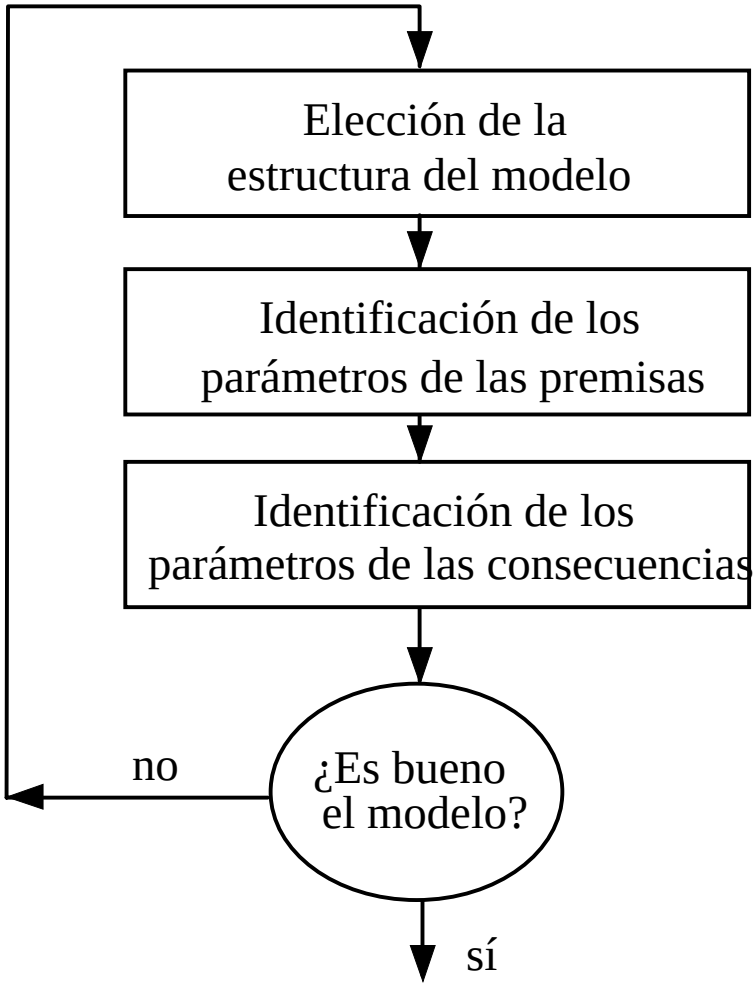
$$Y^3=0.90\cdot 1.2+0.0071\cdot 70-0.39=1.1870$$

Finalmente la salida del modelo está dada por:

$$y(k+1)=\frac{W_1\cdot Y^1+W_2\cdot Y^2+W_3\cdot Y^3}{W_1+W_2+W_3}=1.2195.$$

IDENTIFICACIÓN DE MODELOS TAKAGI-SUGENO

En la siguiente figura se presenta un diagrama del método de identificación:



MÉTODO DE IDENTIFICACIÓN

1) Elección de la estructura del modelo

- Método heurístico
- Análisis de sensibilidades

2) Identificación de parámetros de las premisas

- Conocimiento del proceso
- “Clustering” difuso

3) Identificación de parámetros de las consecuencias

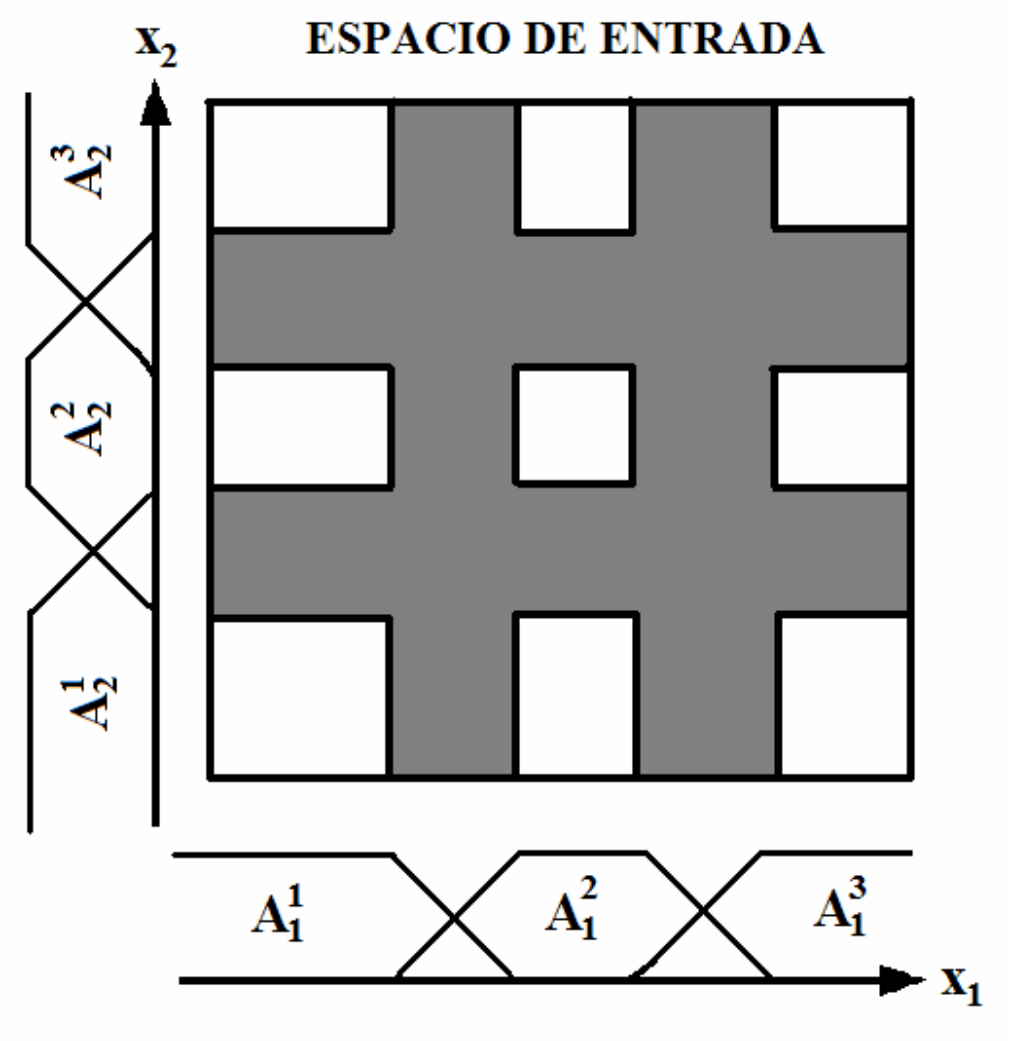
- Mínimos cuadrados.

IDENTIFICACIÓN DE LOS PARÁMETROS DE LAS PREMISAS: "CLUSTERING" DIFUSO

En esta estructura los conjuntos difusos $A1_i$, ..., Ak_i y sus respectivas funciones de pertenencia pueden ser determinadas basándose en un conocimiento previo del proceso o por métodos más complejos como "clustering" difuso.

El número óptimo de reglas y conjuntos difusos del modelo se determina haciendo una partición del universo de la variable de salida (y(k+1)) y luego proyectándolo al espacio de entrada (y(k),u(k), etc). (Sugeno, 1993).

Como se presenta en la siguiente figura, generalmente para la modelación difusa, se pone primero atención en las premisas de las reglas y se encuentra la partición de las variables de entrada. Esta partición se realiza independientemente para cada una de las variables, encontrándose las reglas difusas al realizar las diversas combinaciones posibles entre los conjuntos difusos obtenidos.



En *"Clustering"* difuso, el número óptimo de reglas y conjuntos difusos del modelo se determina haciendo una partición del universo de la variable de salida y luego proyectándolo al espacio de entrada (Sugeno, 1993).

Para obtener los conjuntos difusos de la salida o los clusters asociados a la salida, el criterio utilizado es considerar una partición del espacio de salida tal que cada conjunto o cluster sea lo más distinto posible a los otros pero que sus elementos sean lo más parecido posible entre sí. Esto es que la varianza entre clusters debe ser máxima pero mínima al interior de éstos:

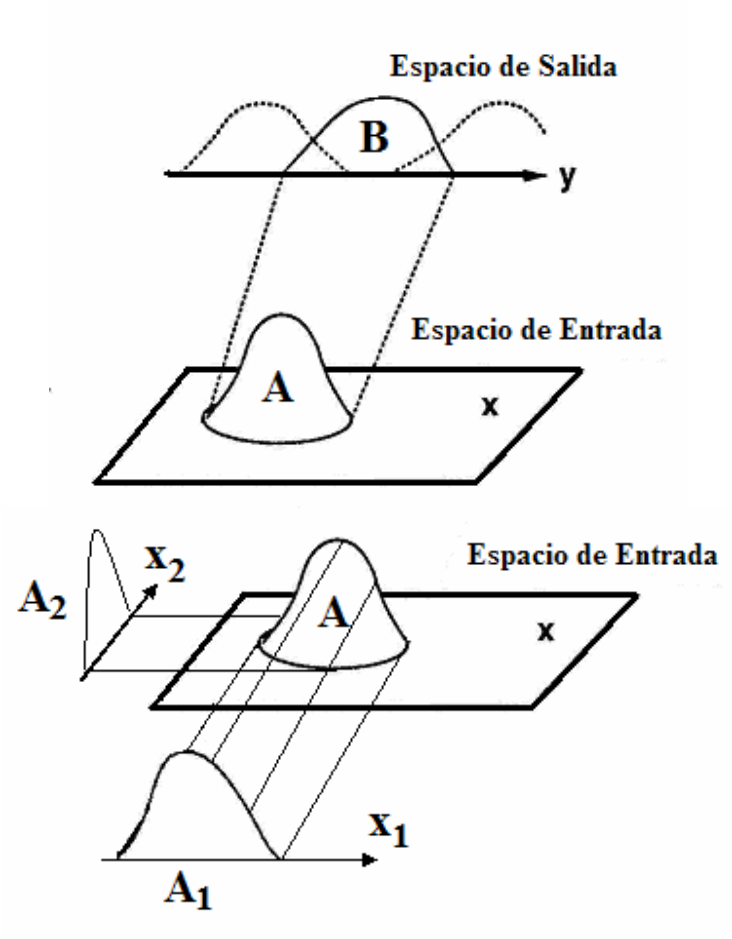
$$S(c) = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_{ik})^m \left(\left\| x_k - v_i \right\|^2 - \left\| v_i - \bar{x} \right\|^2 \right)$$

donde n número de datos para el clustering,
c número de clusters,
x_k k-ésimo vector de datos,
$\bar{x}$ promedio de los datos,
v_i centro del cluster i-ésimo,
μ_{ik} grado del k-ésimo dato perteneciente al i-ésimo cluster,
m peso ajustable.

Por lo tanto, se obtiene que todas las salidas tienen asociado un grado de pertenencia al conjunto difuso Bi.

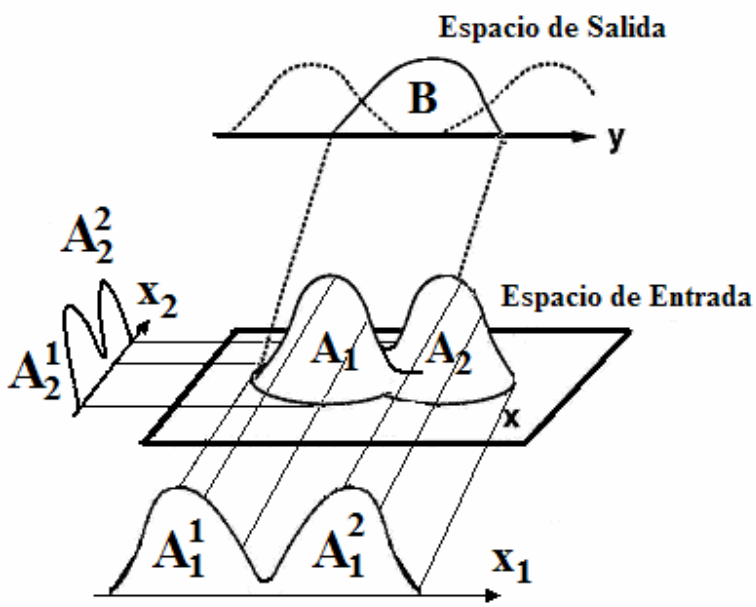
Como se muestra en la siguiente figura , el conjunto B del espacio de salida ha sido proyectado en los conjuntos difusos A_j del espacio de entrada, posibilitando de esta manera establecer la siguiente regla difusa:

$R: \text{Si } x_1 \text{ es } A_1 \text{ y } x_2 \text{ es } A_2 \text{ entonces y es B}$



Sin embargo, se debe considerar que el conjunto A sea no convexo como se muestra en la siguiente figura. En este caso el conjunto difuso del espacio de salida se proyecta en más de un cluster en el espacio de entradas. Por lo tanto, se obtiene más de una regla para el mismo cluster del espacio de salida, es decir:

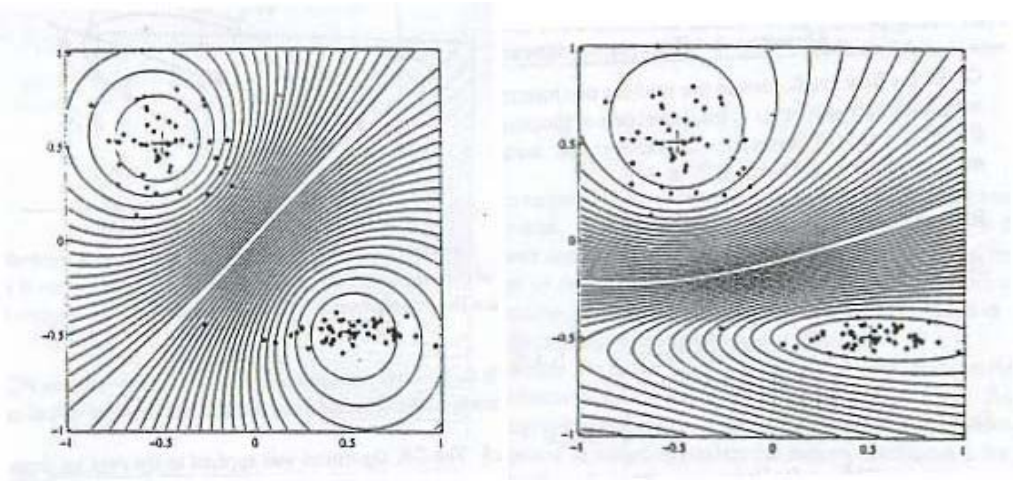
$$R^1: \text{Si } x_1 \text{ es } A_1^1 \text{ y } x_2 \text{ es } A_2^1 \text{ entonces y es B}$$
$$R^2: \text{Si } x_1 \text{ es } A_1^2 \text{ y } x_2 \text{ es } A_2^2 \text{ entonces y es B}$$



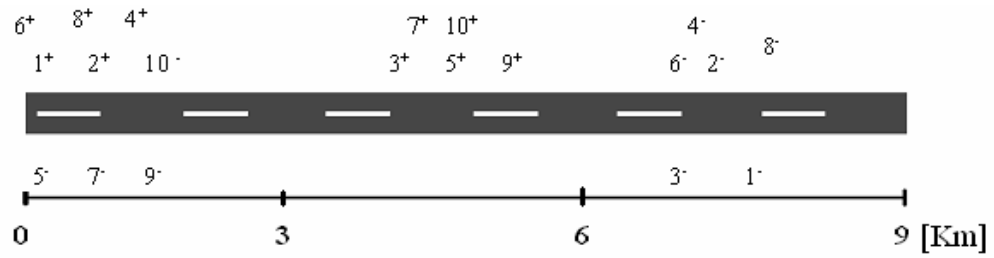
De esta manera, el clustering difuso entrega el número óptimo de reglas difusas y conjuntos difusos.

Uno de los métodos mas usados en Clustering es el Fuzzy C-Means (FCM) en el cual los clusters generados son circulares. (MATLAB: FCM)

En la identificación de modelos Takagi-Sugeno el método de clustering usualmente ocupado es Gustafson-Kessel (GKclus), en el cual los clusters generados dependen de la geometría de los datos (pueden ser elipsoides).



Ejemplo: En el horario punta-mañana ocurren requerimientos de traslado desde un punto a otro en una calle de 9[km].

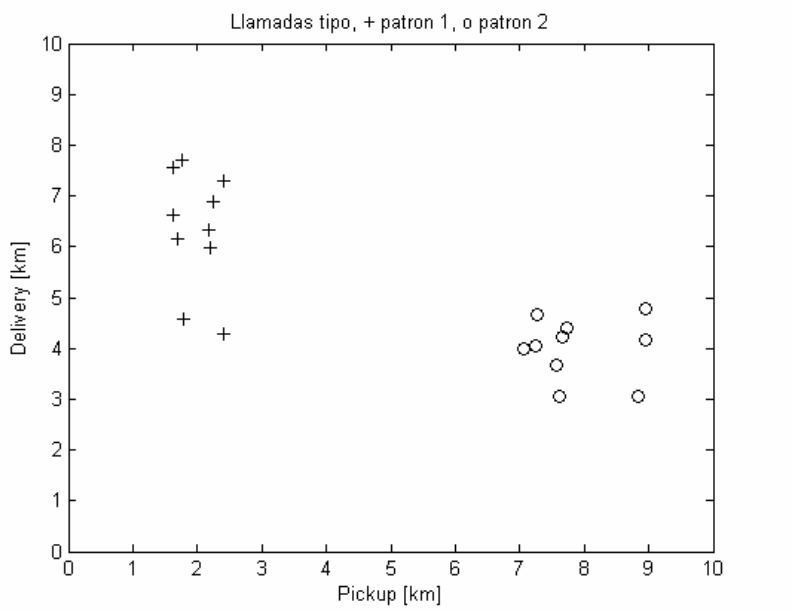


El registro de llamadas está en la siguiente tabla:

Pasajero	Pickup	Delivery	Pasajero	Pickup	Delivery
1	7.6	3.7	**11**	7.7	4.4
2	2.4	4.3	**12**	1.8	7.7
3	7.1	4	**13**	7.6	3.1
4	1.7	6.2	**14**	2.2	6.9
5	7.7	4.2	**15**	7.2	4.1
6	1.8	4.6	**16**	1.6	7.6
7	9	4.8	**17**	8.8	3.1
8	8.9	4.2	**18**	2.4	7.3
9	2.2	6.3	**19**	7.3	4.7
10	1.6	6.6	**20**	2.2	6

Identifique los dos patrones de viaje con FCM (centros de los clusters). A continuación indique el grado de pertenencia al patrón 1 y al patrón 2 del requerimiento de viaje que va desde el punto 2[km] al 8[km].

Sol:



Ocupando Matlab se genera un matriz “data” cuya fila i tiene la coordenada de pickup y delivery del pasajero i.

A continuación con fcm:

```
Nclus=2; % Número de clusters.
[center,U,obj_fcn] = fcm(data, Nclus);
```

De “center” se obtienen los centros de los clusters, lo que nos entrega los patrones de viaje:
Patrón 1: pickup 7.866[km], delivery 4.0319[km].
Patrón 2: pickup 1.9866[km], delivery 6.4108[km].

De “U” se obtienen los grados de pertenencia a cada cluster i (fila i) de cada dato j (columna j).

Se proyecta en la coordenada de pickup y en la coordenada de delivery. Graficamente:

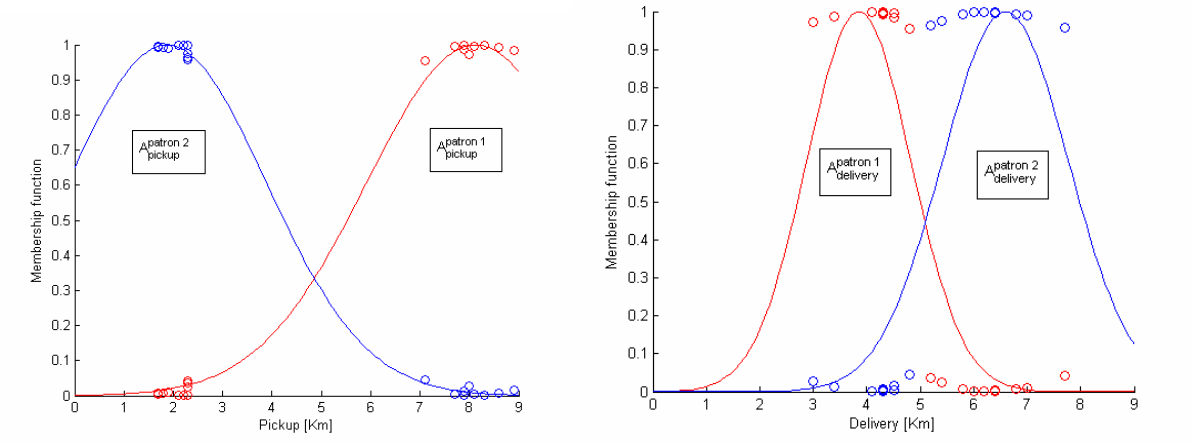
```
plot(data(:,1),U(1,:),'k+')
hold on
plot(data(:,1),U(2,:),'ko')


plot(data(:,2),U(1,:),'k+')
hold on
plot(data(:,2),U(2,:),'ko')
```

Ahora se escoge el tipo de función de pertenencia que ocuparemos. Pueden ser gaussianas, triangulares, etc. Seleccionaremos gausianas de la forma:

$$\mu_1(x) = e^{-\frac{(x-c)^2}{2\sigma^2}} = gaussmf\left(x;[\sigma,c]\right).$$

Para determinar los parámetros que mejor se ajusten a los datos se puede ocupar la función “nlinfit” (non-linear-fit). Es importante darle un buen valor inicial para asegurar convergencia.

```
modelFun =  @(p,x)gaussmf(x,[p(1) p(2)]);
startingVals = [1 5]; %Valor inicial
coefEsts11 = nlinfit(data(:,1), U(1,:)', modelFun, startingVals)
coefEsts12 = nlinfit(data(:,1), U(2,:)', modelFun, startingVals)
coefEsts21 = nlinfit(data(:,2), U(1,:)', modelFun, startingVals)
coefEsts22 = nlinfit(data(:,2), U(2,:)', modelFun, startingVals)
```



Entonces resulta:

(x,y) es Patron 1 si: x es $A_{pickup}^{patron1}$ and y es $A_{delivery}^{patron1}$

(x,y) es Patron 2 si: x es $A_{pickup}^{patron2}$ and y es $A_{delivery}^{patron2}$

x=2, y =8, es patron 1:

```
modelFun(coefEsts11,x)*modelFun(coefEsts21,y)= 2.418e-6
```

x=2, y =8, es patron 2:

```
modelFun(coefEsts12,x)*modelFun(coefEsts22,y)=0.4885.
```

Se verifica que el requerimiento pickup en 2 y delivery en 8 es mas similar al patron tipo 2 que el tipo 1.

IDENTIFICACIÓN DE LOS PARÁMETROS DE LAS CONSECUENCIAS: "MÍNIMOS CUADRADOS ORDINARIOS"

Una vez obtenidas las premisas, solo falta determinar los parámetros $p_0^i, \ldots, p_k^i$ de las consecuencias. Dado que el modelo Takagi-Sugeno es no-lineal pero lineal en los parámetros, estos se pueden obtener por el método de mínimos cuadrados, es decir, se minimiza el índice de error dado por:

$$e^2 = \sum_{p=1}^{N} (y_p - \hat{y}_p)^2$$

donde y_p es la salida real del proceso,
$\hat{y}_p$ es la salida del modelo difuso, considerando las mismas entradas del proceso, y
N es el número de muestras.

Ejemplo. Suponga que se tienen N datos entrada salida del proceso:

$$y(k) = \theta_1 y(k-1) + \theta_2 u(k-1) + \theta_o + w(k)$$

Donde w(k) es ruido-blanco. Para identificar los parámetros se escriben las ecuaciones disponibles:

$$Y = X\theta$$

$$Y = \begin{bmatrix} y(2) \\ y(3) \\ \vdots \\ y(N) \end{bmatrix}, \quad X = \begin{bmatrix} y(1) & u(1) & 1 \\ y(2) & u(2) & 1 \\ \vdots & \vdots & \vdots \\ y(N-1) & u(N-1) & 1 \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_0 \end{bmatrix}.$$

Luego, según mínimos cuadrados:

$$\theta^* = \left(X^T X \right)^{-1} X^T Y \ .$$

Ahora suponga que se sabe que el proceso tiene la forma:

$$y(k) = \theta_1 \, y^3(k-1) + \theta_2 \cos\left(u(k-1)\right) + \theta_o \ln\left(y(k-1)^2 + 1\right) + w(k)$$

Este sigue siendo lineal en los parámetros. Para identificarlo se hace el mismo procedimiento anterior solo que ahora:

$$X = \begin{bmatrix} y^3(1) & \cos\left(u(1)\right) & \ln\left(y^2(1)+1\right) \\ y^3(2) & \cos\left(u(2)\right) & \ln\left(y^2(2)+1\right) \\ \vdots & \vdots & \vdots \\ y^3(N-1) & \cos\left(u(N-1)\right) & \ln\left(y^2(N-1)+1\right) \end{bmatrix},$$

Volvamos al caso del modelo Takagi-Sugeno.

Entonces dado un conjunto de N muestras de entrada/salida, el algoritmo calcula, para la muestra j, los grados de cumplimiento de cada regla según la definición dada anteriormente, lo que corresponde a:

$$W_{ij} = \text{oper}\ (\mu_{A1_i}(X_{1j}); \mu_{A2_i}(X_{2j}); \ldots; \mu_{Ak_i}(X_{kj}))$$

con: $1 \le i \le M$, $1 \le j \le N$.
M es el número de reglas, N es el número de muestras y k es el número de variables.

La salida del modelo es:

$$y_j = \frac{\sum_{i=1}^{M} (p_o^i + p_1^i X_{1j} + \ldots + p_k^i X_{kj}) W_{ij}}{\sum_{i=1}^{M} W_{ij}}$$

Se define

$$B_{ij} = \frac{W_{ij}}{\sum_{h=1}^{M} W_{hj}}$$

entonces, la predicción de la salida j está dada por la siguiente expresión:

$$y_j = \sum_{i=1}^{M} (p_o^i B_{ij} + p_1^i B_{ij} X_{1j} + \ldots + p_k^i B_{ij} X_{kj})$$

Si se construye la matriz x' de dimensiones N,M*(k+1)

$$x' = \begin{bmatrix} B_{11} \ldots B_{M1} & X_{11}B_{11} \ldots X_{11}B_{M1} & \ldots & X_{k1}B_{11} \ldots X_{k1}B_{M1} \\ \vdots & \vdots & \vdots & \vdots \\ B_{1N} \ldots B_{MN} & X_{1N}B_{1N} \ldots X_{1N}B_{MN} & \ldots & X_{kN}B_{1N} \ldots X_{kN}B_{MN} \end{bmatrix}$$

el vector de salida $y' = [y_1,\ldots,y_N]^T$ se expresa de la forma:

$$y' = x'P$$

donde $P = \left[p_0^1, \ldots, p_0^M, p_1^1, \ldots, p_1^M, \ldots, p_k^1, \ldots, p_k^M \right]^T$ es el vector de parámetros.

La identificación se reduce, entonces, a resolver la ecuación anterior encontrando el vector P mediante un algoritmo de mínimos cuadrados.

$$y' = x'P \qquad /x'^T$$
$$x'^T y' = x'^T x' P$$

$$P = \left(x'^T x'\right)^{-1} x'^T y'$$

Finalmente, con los parámetros óptimos de las consecuencias ya determinados se puede alterar la estructura del modelo o las funciones de pertenencia obtenidas, de manera de reducir el índice de error.

Este método de identificación puede no ajustar apropiadamente datos que no son suficientemente ruidosos (singularidad en matriz de covariancia).
Una técnica usada en la literatura recomienda identificar cada regla por separado. Para eso se multiplica el vector de salida por el grado de pertenencia a la regla j.

$$W_{ij} y_j = p_o^i W_{ij} + p_1^i X_{1j} W_{ij} + \ldots + p_k^i X_{kj} W_{ij}$$

Y se aplica mínimos cuadrados. Los datos que poseen un grado de activación menor que un número pequeño $W_{ij} \leq \delta$ se eliminan de la identificación de dicha regla.

OPTIMIZACIÓN ESTRUCTURAL: "ANALISIS DE SENSIBILIDAD"

En este caso, la estructura del modelo difuso se define como la selección de las variables de entrada significativas.

Método heurístico (Sugeno, 1993)

En general, para el modelo difuso definido por:

$$\begin{aligned}
&\text{if } y(k-1) \text{ is } A_1^r \text{ and} \cdots \text{and } y(k-na) \text{ is } A_{na}^r \text{ and} \\
&u(k-nk-1) \text{ is } A_{na+1}^r \text{ and} \cdots \text{and } u(k-nb-nk) \text{ is } A_{na+nb}^r \\
&\text{then } y_r(k) = g_0^r + g_1^r y(k-1) + \ldots + g_{na}^r y(k-na) \\
&\quad + g_{na+1}^r u(k-nk-1) + \ldots + g_{na+nb}^r u(k-nb-nk)
\end{aligned}$$

Se tienen (na+nb) variables candidatas de entrada, de este modo el total de posibles modelos considerando todas las entradas es: $2^{na+nb} - 1$.

Similarmente al método de regresión por pasos, el método heurístico consiste en seleccionar algunas variables de entrada dentro de todas las variables de entrada candidatas, incrementando el número de entradas de una en una, de acuerdo a un cierto criterio. Por ejemplo, el siguiente criterio puede ser usado:

$$RC = \frac{\left[\sum_{i=1}^{k_A}\left(y^A(i) - y^{AB}(i)\right)^2 + \sum_{i=1}^{k_B}\left(y^B(i) - y^{BA}(i)\right)^2\right]}{2}$$

donde k_A y k_B son el número de datos de dos grupos del conjunto de ajuste, $y^A(i)$ y $y^B(i)$ son las salidas de los grupos A y B, $y^{AB}(i)$ es la salida estimada para el grupo A con el modelo identificado usando el

conjunto de datos B, e $y^{BA}(i)$ es la salida estimada para el grupo B con el modelo identificado usando el conjunto de datos A.
Pasos del algoritmo:

- Identificar (na+nb) modelos difusos con una variable de entrada cada uno.
- Calcular el criterio RC para cada modelo y seleccionar el modelo con menor RC.
- Fijar la variable seleccionada en el paso anterior y agregar otra variable de entrada de las restantes candidatas.
- Continuar hasta que RC aumente.

Análisis de Sensibilidades

Se considera el siguiente modelo no lineal:

$$y(k) = f(\mathbf{x}(k))$$

con

$$\mathbf{x}(k) = \begin{bmatrix} y(k-1) \\ \vdots \\ y(k-na) \\ u(k-nk-1) \\ \vdots \\ u(k-nk-nb) \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_{na} \\ x_{na+1} \\ \vdots \\ x_{na+nb} \end{bmatrix}$$

La función *f* no lineal con modelos difusos de Takagi&Sugeno está dada por:

$$\text{if } y(k-1) \text{ is } A_1^r \text{ and} \cdots \text{and } y(k-na) \text{ is } A_{na}^r \text{ and}$$
$$u(k-nk-1) \text{ is } A_{na+1}^r \text{ and} \cdots \text{and } u(k-nb-nk) \text{ is } A_{na+nb}^r$$
$$\text{then } y_r(k) = g_0^r + g_1^r y(k-1) + ... + g_{na}^r y(k-na)$$
$$+ g_{na+1}^r u(k-nk-1) + ... + g_{na+nb}^r u(k-nb-nk)$$

En general, la sensibilidad de las variables de entrada de un modelo no lineal está dada por:

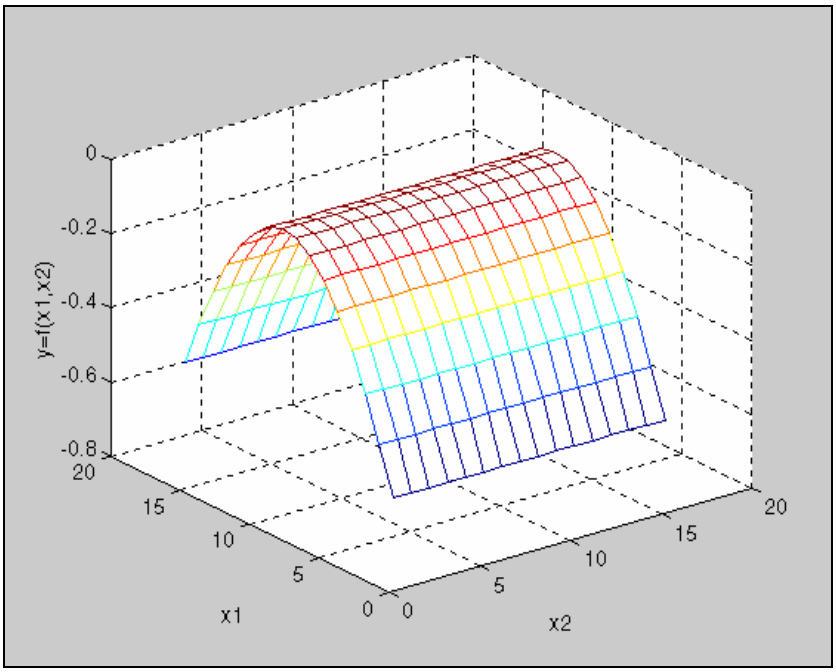
$$\xi_i(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x_i}$$

con

$$\mathbf{x}(k) = \begin{bmatrix} y(k-1) \\ \vdots \\ y(k-na) \\ u(k-nk-1) \\ \vdots \\ u(k-nk-nb) \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_{na} \\ x_{na+1} \\ \vdots \\ x_{na+nb} \end{bmatrix}$$

La sensibilidad de cada variable de entrada representa la relevancia de esta entrada en la salida del modelo.

Por ejemplo, en la figura se muestra que la variable de entrada x_2 es irrelevante para la salida del modelo, es decir, $\frac{\partial f(x_1, x_2)}{\partial x_2} = 0.$



Función no lineal $y = f(x_1,x_2)$

En este caso, la función no lineal f para los modelos difusos de Takagi (1995):

$$\text{Si } y(k-1) \text{ es } A_1^r \text{ y} \cdots \text{y } y(k-na) \text{ es } A_{na}^r \text{ y}$$

$$u(k-nk-1) \text{ es } A_{na+1}^r \text{ y} \cdots \text{y } u(k-nb-nk) \text{ es } A_{na+nb}^r \text{ entonces}$$

$$y_r(k) = g_0^r + g_1^r y(k-1) + \ldots + g_{na}^r y(k-na) + g_{na+1}^r u(k-nk-1) + \ldots + g_{na+nb}^r u(k-nb-nk)$$

La salida del modelo difuso es:

$$y(k) = \frac{\sum_{r=1}^{N_r} w_r y_r(k)}{\sum_{r=1}^{Nr} w_r}$$

con el grado de activación de las reglas dado por:

$$w_r = \mu_1^r \cdots \mu_i^r \cdots \mu_{na+nb}^r$$

Se considera la siguiente función de pertenencia para el conjunto difuso A_i^r, pues es diferenciable:

$$\mu_i^r = \exp\left(-0.5\left(a_i^r\left(x_i - b_i^r\right)\right)^2\right)$$

Para estos modelos difusos definidos, la sensibilidad de una variable de entrada está dada por:

$$\xi_i(\mathbf{x}) = \frac{\sum_{r=1}^{N_r}\left(\frac{\partial w_r}{\partial x_i} y_r + \frac{\partial y_r}{\partial x_i} w_r\right)\sum_{r=1}^{N_r} w_r - \sum_{r=1}^{N_r}\left(\frac{\partial w_r}{\partial x_i}\right)\sum_{r=1}^{N_r}\left(w_r y_r\right)}{\left(\sum_{r=1}^{N_r} w_r\right)^2}$$

donde

$$\frac{\partial w_r}{\partial x_i} = \frac{\partial \mu_i^r}{\partial x_i} \times \mu_1^r \times \cdots \times \mu_{i-1}^r \times \mu_{i+1}^r \times \cdots \times \mu_{na+nb}^r$$

$$\frac{\partial \mu_i^r}{\partial x_i} = \mu_i^r \times c_i^r$$

$$c_i^r = -\left(a_i^r \times \left(x_i - b_i^r\right)\right) \times a_i^r$$

$$\frac{\partial y_r}{\partial x_i} = g_i^r$$

Entonces, la sensibilidad ξ_i con respecto a la variable de entrada x_i de un modelo difuso es:

$$\xi_i(\mathbf{x}) = \frac{\sum_{r=1}^{Nr}\left(w_r c_i^r y_r + g_i^r w_r\right)\sum_{r=1}^{Nr} w_r - \sum_{r=1}^{Nr}\left(w_r c_i^r\right)\sum_{r=1}^{Nr}\left(w_r y_r\right)}{\left(\sum_{r=1}^{Nr} w_r\right)^2}$$

Las sensibilidades ξ_i depende de las variables de entrada **x**, siendo evaluadas usando el conjunto de entrenamiento. Para comparar las sensibilidades de las variables de entrada, el siguiente índice de sensibilidades es propuesto:

$$I_i = \mu_i^2 + \sigma_i^2$$

donde μ_i y σ_i es la media y la desviación estándar de las sensibilidades, respectivamente. Entonces, las variables de entrada con menores índices I_i serán eliminadas.

Por último, el modelo difuso es obtenido usando sólo las variables de entrada con mayores sensibilidades.

Ejemplo: Sistema dinámico no lineal Chen (1989)

$$\begin{aligned} y(k) &= \left(0.8 - 0.5\exp\left(-y^2(k-1)\right)\right)y(k-1) \\ &- \left(0.3 + 0.9\exp\left(-y^2(k-1)\right)\right)y(k-2) + u(k-1) \\ &+ 0.2u(k-2) + 0.1u(k-1)u(k-2) + \varepsilon(k) \end{aligned}$$

con u(k) una serie independiente e idénticamente distribuida ($\mu = 0, \sigma = 1$), y $\varepsilon(k)$ es un ruido blanco ($\mu = 0, \sigma = 0.2$).

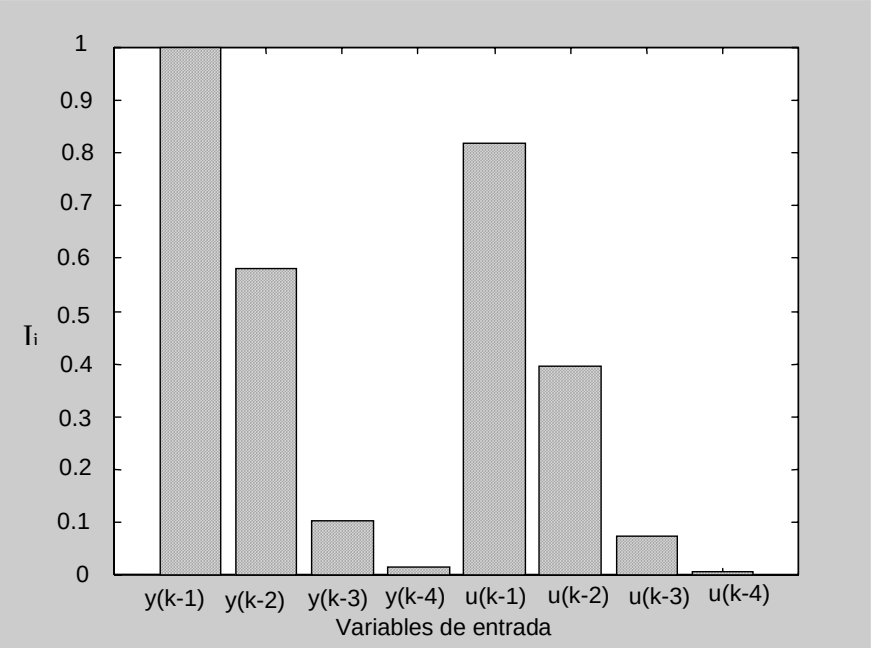
EL761 D. Sáez, C. Cortés, A. Nuñez (2009).

250 datos para el conjunto de entrenamiento 250 datos para el conjunto de validación.

Análisis de sensibilidades

La estructura del modelo inicial difuso considera las 8 variables de entrada $y(k-1)$, ..., $y(k-4)$, $u(k-1)$, ..., $u(k-4)$.

La siguiente figura presenta los índices de las sensibilidades del modelo inicial para las 8 variables de entrada. En el gráfico, se observa que las variables $y(k-3)$, $y(k-4)$, $u(k-3)$, $u(k-4)$ presentan menores índices de las sensibilidades por lo cual no se serán incluidos en el modelo difuso. De esta manera, la estructura obtenida considera las variables de entrada $y(k-1)$, $y(k-2)$, $u(k-1)$ y $u(k-2)$.



Índices de sensibilidades

Modelo	Variables de entrada	e^2
1	$y(k-1) - y(k-2) - y(k-3) - y(k-4)$ $u(k-1) - u(k-2) - u(k-3) - u(k-4)$	0.5802
2	$y(k-1) - y(k-2) - u(k-1) - u(k-2)$	**0.4696**

Valores de los índices de error, utilizando el análisis de sensibilidades

Análisis Comparativo

El método de sensibilidades para identificación de estructura de modelos difusos permite estudiar el universo completo de modelos posibles, dentro de una complejidad de estructura definida por el orden del modelo inicial.

Además, el método propuesto permite obtener la mejor estructura ajustando menos modelos (dos) que con el método heurístico, en el cual es necesario construir varios modelos.