

ESTUDIO PARA APLICACIÓN DE TÉCNICAS DE IDENTIFICACIÓN DE LAZO CERRADO A ALGORITMOS DE CLUSTERING DIFUSO

Trabajo de Tesis
presentado al
Departamento de Ingeniería Electrónica

por

Diana Marcela González Chaparro

Director: Alain Gauthier, PhD

Asesor: Victor H. Grisales, PhD(c)

Para optar al título de
Magister en Ingeniería Electrónica y de Computadores

Ingeniería Electrónica
Universidad de Los Andes
Agosto 2006

**ESTUDIO PARA APLICACIÓN DE TÉCNICAS
DE IDENTIFICACIÓN DE LAZO CERRADO A
ALGORITMOS DE CLUSTERING DIFUSO**

Aprobado por:

Alain Gauthier, PhD

Victor H. Grisales, PhD(c)

Fernando Jiménez, PhD

Fecha de Aprobación _____

*A Dios por que he aprendido a confiar en tí cada día más y sin tí no
habría llegado donde estoy.*

A mi gordis por que sin ella esto no habría sido posible

A mi familia por toda la confianza que han depositado en mí.

sólo espero no defraudarlos.

A Alexander por tu amor, tu apoyo y tu confianza.

Prefacio

La mayor motivación para este proyecto fue el deseo de mejorar y continuar con el trabajo de investigación realizado en el pregrado donde se analizaron las técnicas de clustering y se diseñó un Algoritmo Aglomerativo de Clustering Robusto con Fines de Modelamiento Difuso que presentó resultados destacados conforme al desempeño de algoritmos existentes, en sistemas no lineales, multivariados, dinámicos y en presencia de ruido.

Con base a estos resultados se pretendía analizar el desempeño de las técnicas de clustering en sistemas con realimentación o en lazo cerrado y plantear una metodología que permitiera la obtención de un modelo difuso tipo TKS de la planta en lazo abierto. Con el objetivo de poder manipular sistemas que por su complejidad en la operación o por comportamientos inestables no permitiera abrir dicho lazo de realimentación.

Reconocimientos

Quiero agradecer al ingeniero Victor Hugo Grisales, quién confió en mi investigación de pregrado y me brindó sus aportes y sugerencias durante este trabajo.

Al ingeniero Rubén Darío Moreno por su paciencia, enseñanzas y apoyo. Se ha convertido para mí en un ejemplo a seguir.

Al ingeniero Jorge Posada por sus ideas y aportes a este trabajo.

A mi familia quienes con dificultad aprendieron a aceptar mi ausencia, que todo el tiempo dedicado a mi trabajo y a mi estudio es para que cada día se sientan más orgullosos de mí.

A mis amigos a quienes hace tiempo no veo pero que me han hecho saber que están presentes con cada llamada o cada mensaje de aliento.

Tabla de Contenido

Dedicatoria	III
Prefacio	IV
Reconocimientos	v
Lista de Tablas	VIII
Lista de Figuras	IX
Resumen	XI
I. INTRODUCCIÓN	1
II. APROXIMACIÓN EN LAZO CERRADO	3
2.1. Teoría y Práctica de la Identificación en Lazo Cerrado	4
2.1.1. Aproximación Directa	4
2.1.2. Aproximación Indirecta	5
2.1.3. Aproximación Entrada-Salida Comunes	6
2.2. Identificación Recursiva de Modelos Lineales en Lazo Cerrado	7
2.3. Algoritmos de Error de Salida en Lazo Cerrado	12
2.3.1. CLOE Closed Loop Output Error	12
2.3.2. F-CLOE Algoritmo Error de Salida en Lazo Cerrado Filtrado	13
2.3.3. X-CLOE Algoritmo del Error de Salida en Lazo Cerrado Ex- tendido	14
III. RESULTADOS EXPERIMENTALES EN SISTEMAS LINEALES	17
3.1. Planta 1	18
3.2. Planta 2	19

3.3. Planta 3	20
3.4. Planta 4	22
IV. MODELAMIENTO E IDENTIFICACIÓN DIFUSOS	24
4.1. Modelos Takagi Sugeno Kang	24
4.2. Técnicas de Clustering	26
4.2.1. Algoritmo Gustafson-Kessel	27
4.3. Algoritmo RoFeR	30
4.3.1. Formulación Robusta y de Aglomeración Competitiva	32
4.4. Derivación de Modelos TSK a Partir de los Parámetros Obtenidos en los Métodos de Clustering	33
4.4.1. Derivación de Consecuentes	33
4.5. Derivación de Antecedentes	35
4.6. Modelos de Entrada - Salida	38
V. MODELAMIENTO DIFUSO DE PLANTAS NO LINEALES EN LAZO CERRADO	40
5.1. Metodología Planteada	41
5.2. Aplicación	44
5.2.1. Análisis de la Realimentación del Sistema Difuso	49
5.2.2. Aplicación de Realimentación Positiva	54
5.2.3. Sensibilidad del Modelo a la Escala	55
VI. Conclusiones	60

Lista de Tablas

1. Resultados obtenidos con los diferentes algoritmos para el sistema lineal 1 19
2. Resultados obtenidos con los diferentes algoritmos para el sistema lineal 2 19
3. Resultados obtenidos con los diferentes algoritmos para el sistema lineal 3 21
4. Resultados obtenidos con los diferentes algoritmos para el sistema lineal 4 23

Lista de Figuras

1.	Aproximaciones de identificación en lazo cerrado	5
2.	Estructuras par identificación en lazo cerrado. a) excitación como referencia, b) excitación a la salida del controlador	8
3.	Controlador RST	17
4.	Señales de excitación (superior) y salida del modelo 1 (inferior)	18
5.	Señales de excitación (superior) y salida del modelo 2 (inferior)	20
6.	Señales de excitación (superior) y salida del modelo 2 (inferior)	21
7.	Determinación del orden del sistema	22
8.	Señales de excitación (superior) y salida del modelo 2 (inferior)	23
9.	La estructura propia de la matriz de covarianza difusa provee información acerca de la forma y orientación del cluster	34
10.	Aproximación mediante una ecuación paramétrica de datos proyectados	36
11.	Error por descomposición. (a) Cluster obtenido (b) Cluster reconstruido después de la proyección	37
12.	Representación clásica de un sistema de lazo cerrado	41
13.	Diagrama que representa la metodología propuesta	43
14.	Cuatro Tanques en Cascada	45
15.	Sistema de Cuatro Tanques Interconectados	46
16.	Sistema de Dos Tanques - Controlador PID	47
17.	Respuesta a una entrada escalón para diseño del controlador PID . . .	47
18.	Identificación del sistema de 4 tanques interconectados, con el algoritmo de Gustafson-Kessel. Toolbox FMID. La señal azul indica la salida real del sistema y la señal discontinua roja la aproximación GK.	48

19.	Ejemplo de las reglas que arroja el modelo difuso obtenido para el sistema de 4 tanques interconectados con el Toolbox de R. Babuska .	48
20.	Salida del Modelo Difuso obtenido para el sistema de dos tanques con controlador	49
21.	Funciones de pertenencia de los antecedentes del modelo obtenido. . .	50
22.	Diagrama de bloques para el sistema realimentado	51
23.	Respuesta en lazo abierto; Entrada (Rojo Señal Continua) y Salida (Azul Señal Discontinua)	51
24.	Respuesta en lazo cerrado; Entrada (Rojo Señal Continua) y Salida (Azul Señal Discontinua)	52
25.	Diagrama de bloques para la aproximación difusa con lazo de realimentación	52
26.	Comparación de la salida del sistema real realimentado (Rojo Continua) y Salida del modelo difuso obtenido en lazo cerrado(Azul Discontinua).	53
27.	Salida del Sistema (Rojo) y Salida del Modelo Difuso (Azul), en lazo abierto.	53
28.	Entrada (Rojo)y Salida (Azul Discontinua) del sistema de lazo directo	55
29.	Diagrama de bloques con realimentación positiva	56
30.	Entrada y Salida del sistema con los dos lazos de realimentación . . .	56
31.	Diagrama de bloques para el sistema Difuso con realimentación positiva	56
32.	Salida del sistema(Azul) y Salida del modelo difuso (Rojo) en lazo directo	57
33.	Salida del sistema(Azul) y Salida del modelo difuso (Rojo) con los dos lazos de realimentación	57
34.	Salida del Modelo (Rojo)y Salida de la Planta(Azul)	58
35.	Salida del Modelo (Rojo)y Salida de la Planta(Azul)	59
36.	Salida del Modelo (Rojo)y Salida de la Planta(Azul)	59

Resumen

El desarrollo de un modelo que represente adecuadamente la realidad del problema es una tarea importante. Si el modelo no es suficientemente exacto los pasos subsecuentes de análisis, predicción, control, entre otros, pueden no ser exitosos. El modelamiento y la identificación son importantes pasos en el diseño de control, supervisión y detección de fallas de sistemas. Pero en la realidad muchos sistemas no son tratables mediante aproximaciones de modelamiento en lazo abierto convencional debido a la dificultad de aislar la planta del controlador o la inestabilidad del sistema una vez se haya eliminado la realimentación.

En este trabajo, se analizaron metodologías para la obtención de modelos lineales en lazo cerrado [15]. Se realizaron pruebas de simulación de varios sistemas lineales para evaluar su desempeño obteniendo buenas aproximaciones a los sistemas reales.

Luego de ese estudio, se trabajó en el desarrollo de una metodología para la identificación de sistemas no lineales en lazo cerrado, a partir de algoritmos de clustering con el fin de obtener modelos difusos afines Takagi Sugeno Kang (TSK) en tiempo discreto, para los cuales dinámicas locales en diferentes regiones del espacio de estado son representadas por submodelos lineales. Los algoritmos de clustering utilizados para este trabajo fueron GK [2] [5] y RoFer [9].

Finalmente se evaluó la metodología planteada, con una aplicación específica

basados en la planta de tanques interconectadas presentada por R. Babuska en [3] y [5] que presenta condiciones de no linealidad.

Capítulo I

INTRODUCCIÓN

La automática consiste en el estudio de sistemas reales de diferentes disciplinas científicas (Electrónica, mecánica, térmica, química, ecología, biología, economía), con miras al análisis, la predicción, la supervisión, el control y/o la optimización de sistemas. La condición necesaria para esto es la obtención de un modelo matemático del sistema real (modelamiento). Un sistema es un objeto en el que las variables de diferentes tipos se integran y producen las señales observables. Cuando el modelo del sistema no es conocido, es necesario proceder a su identificación.

En efecto, el modelamiento se realiza a partir de un comportamiento del sistema, consiste en determinar la estructura que rige el comportamiento de ese sistema y también, en fijar, un valor a priori de sus parámetros (longitudes, masas, inercias). Pero, frecuentemente es imposible obtener un conocimiento a priori completo y preciso de todos los parámetros del modelo.

La identificación de sistemas controlados ayudan a calcular las necesidades en lugar de describir sus propiedades, al menos ayuda en la medida de valores (tabla de valores) o de representaciones gráficas, con el fin de determinar su respuesta a un escalón o a una frecuencia (identificación no paramétrica). Para los sistemas complejos, es necesario utilizar los modelos matemáticos (identificación paramétrica). La utilización de modelos matemáticos es inherente a todos los dominios de técnicas de la ingeniería y de la física. Ayuda a la concepción de sistemas y es un instrumento de simulación y de predicción que es utilizado en todos los dominios.

En automatización, se distinguen dos aproximaciones para la identificación de sistemas: aproximación en lazo abierto y aproximación en lazo cerrado. El problema

fundamental con los datos en lazo cerrado, es la correlación entre el ruido no medible y la entrada. Por esta razón, la mayoría de los métodos utilizados en lazo abierto no funciona cuando son aplicados a los datos de lazo cerrado. La identificación en lazo cerrado debe permitir obtener un modelo que describa lo más fielmente posible el comportamiento del sistema real con un regulador dado. También debe permitir mejorar el desempeño del sistema enlazado a través de la utilización de sistemas identificados en lazo cerrado. El objetivo de la operación es encontrar el mejor modelo del procedimiento que minimice la diferencia entre la salida predicha y la salida real del sistema (error de predicción).

Estas aproximaciones son impuestas en los casos donde el sistema es inestable en lazo abierto, presenta una desviación importante del punto de funcionamiento o cuando no se puede abrir el lazo (seguridad, producción). Estas técnicas buscan obtener un mejor modelo para el control de sistemas, la motivación del poder trabajar con restricciones prácticas, tales como realizar el mantenimiento de un controlador existente sin abrir el lazo o reajustarlo en el sitio, también, con el tiempo, que los parámetros y el modelo de un proceso pueda ser modificado. Adicionalmente, permitir reducir el orden del modelo, con el fin de reconsiderar el control y conservar los desempeños deseados en lazo cerrado.

Se pretende encontrar una extensión de los métodos de estimación de parámetros presentes en los algoritmos de lazo cerrado de control convencional a las técnicas de identificación difusa, en concreto a los algoritmos de clustering que permiten la obtención de modelos difusos del tipo Takagi-Sugeno (TS). Lo anterior, gracias a que las técnicas de clustering difuso se han presentado como una efectiva herramienta para la aproximación de sistemas no-lineales inciertos basándose en datos medidos y el modelo TS ha atraído mayor atención debido a su buen desempeño en varias aplicaciones. Seguido del análisis para el descuento del lazo de realimentación, así como del controlador, para lograr obtener una representación difusa de la dinámica de la planta.

Capítulo II

APROXIMACIÓN EN LAZO CERRADO

Esta es una alternativa que a menudo es imposible evitar en un círculo industrial. Es necesario la implementación del lazo de realimentación en casos donde:

- halla requisitos de la producción, de seguridad o por razones económicas.
- si el proceso es inestable en lazo abierto o presenta un comportamiento integrador.
- La capacidad de un regulador de linealizar el comportamiento del proceso alrededor de un punto de la operación a circuito cerrado constituye un punto fuerte.

La aproximación en lazo cerrado, considerada como difícil ha sido el objeto de interés en los últimos años con el desarrollo de algoritmos y una comprensión mejor del tema. Sus ventajas y características son:

- Motivación por las restricciones prácticas, por ejemplo realizar el mantenimiento de un regulador existente sin abrir el lazo o realizar un reajuste en sitio. Poder modificar con el tiempo, los parámetros y el modelo de un proceso.
- Permite obtener mejores modelos para el control de sistemas (una identificación en lazo cerrado da mejores resultados para un sistema que funciona normalmente en lazo cerrado.),
- Validación del control de un sistema: la comparación entre los desempeños deseados y reales en el dominio temporal y frecuencial, el cálculo de polos y las pruebas estadísticas.

- Es particularmente favorable en la síntesis de un regulador robusto.
- Permite reducir el orden de un modelo, para resintonizar el control (modificación del controlador) preservando los funcionamientos deseados en lazo cerrado.
- La identificación en tiempo real en lazo cerrado y la adaptación del controlador.

2.1. Teoría y Práctica de la Identificación en Lazo Cerrado

El problema fundamental con los datos en lazo cerrado, es la correlación entre el ruido no medible y la entrada. Por esta razón, la mayoría de los métodos usados en lazo abierto no funcionan cuando se aplican a los datos del lazo cerrado. Sin embargo, el inconveniente mayor de los métodos tradicionales, es que son sensibles al ruido, por otra parte, los modelos utilizados carecen de precisión.

La identificación en lazo cerrado debe permitir obtener un modelo del proceso que describa lo más exactamente posible el comportamiento del sistema real con un controlador dado. El objetivo de la operación es encontrar el mejor modelo del proceso que reduzca al mínimo la diferencia entre la salida predicha y la salida verdadera del sistema (error de la predicción). [15].

La figura 1 muestra algunas maneras de aproximar un sistema en lazo cerrado que pueden ser clasificados en tres aproximaciones distintas.

2.1.1. Aproximación Directa

Es cuestión de aplicar el método de predicción del error. Se utiliza la salida del proceso de $y(t)$ y su entrada $u(t)$, de la misma manera que cuando funciona en lazo abierto, siendo ignorada la señal de retroalimentación así como la referencia $r(t)$. Las ventajas de esta aproximación que se pueden considerar como naturales para el estudio en lazo cerrado son:

- Funciona sin la necesidad preocuparse por la complejidad del controlador, ni de conocer la realimentación.
- Ningún algoritmo, ni lógicas específicas son requeridas.

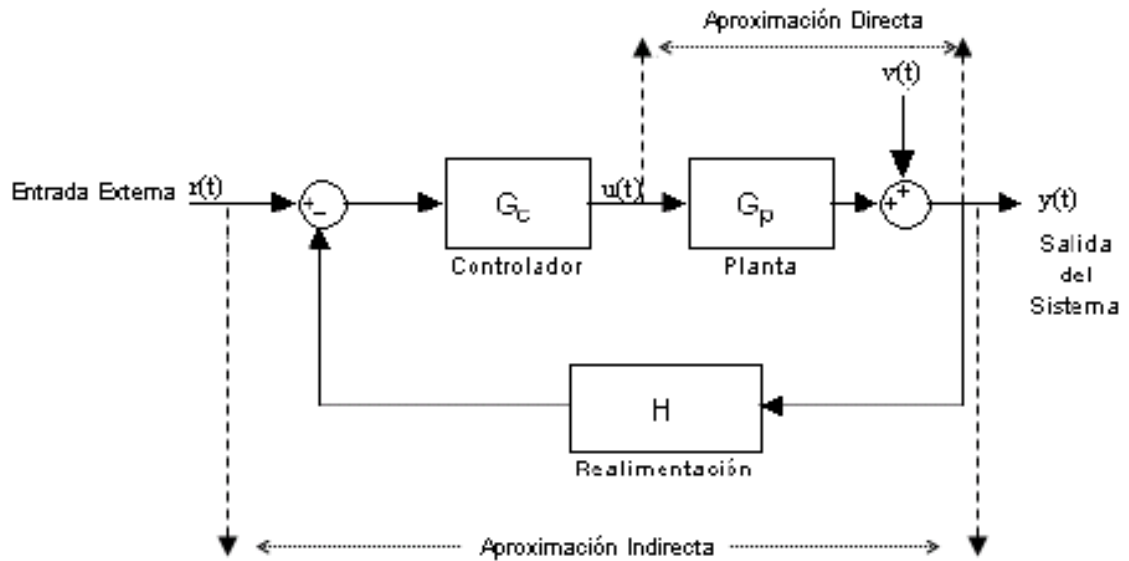


Figura 1: Aproximaciones de identificación en lazo cerrado

- Si la estructura del modelo contiene el modelo real incluyendo las características del ruido, se obtienen muy buenos resultados.
- Los sistemas inestables son soportados sin problemas mientras el lazo cerrado y el predictor sean estables.

La única desventaja es por lo tanto la necesidad de tener un buen modelo del ruido para una buena estimación de los parámetros (ausencia de bias). Se elegirá generalmente utilizar los modelos del tipo ARX o ARMAX con esta aproximación.

2.1.2. Aproximación Indirecta

Es cuestión de identificar el sistema del lazo cerrado a partir de la referencia externa de la entrada $r(t)$ y de la salida $y(t)$ y a partir de estos datos encontrar el sistema del lazo abierto usando el conocimiento del controlador. La ventaja de este método es que es posible aplicar al sistema los métodos usados en lazo abierto. El problema principal de esta aproximación es que cualquier error en el lazo de realimentación será reflejado en el modelo \hat{G}_{ld} . En este acercamiento, el modelo en el lazo cerrado G_{lc} se estima a partir de medidas de $y(t)$ y de $u(t)$. Es entonces

posible calcular \widehat{G}_{ld} , la función de transferencia en lazo abierto:

$$\widehat{G}_{lc} = \frac{\widehat{G}_{ld}}{1 + \widehat{G}_{ld}H_y} \quad (1)$$

donde el subíndice ld indica la ganancia de lazo abierto. Si se tiene un buen conocimiento del lazo de realimentación y se conoce el controlador, es preferible utilizar este método que no requiere un conocimiento perfecto del ruido.

2.1.3. Aproximación Entrada-Salida Comunes

Para este caso se observa $y(t)$ y $u(t)$ como las salidas de un sistema que tiene como entrada $r(t)$ (referencia) y el ruido y encontrar el modelo del sistema y del controlador a partir de esta aproximación conjunta. De hecho, se mira a y y a u como un sistema multivariable, series de tiempo de dimensión $(n_u + n_y)$, considerando un modelo de estructura apropiada para ésta. Con esta aproximación, el proceso en lazo abierto así como el controlador, pueden ser identificados. El conocimiento del controlador no es necesario aquí, que es una ventaja comparada con método indirecto. Si se utiliza el método de predicción de error suponiendo que el regulador es lineal y que e y d son fuentes independientes de ruido, la salida y y la entrada u pueden ser descritas como sigue:

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = G_0(q)r(t) + H_0(q) \begin{bmatrix} e(t) \\ d(t) \end{bmatrix} \quad (2)$$

con H representando el modelo del ruido. Esta aproximación es conocida también como el método de dos etapas, que consiste en:

1. Identificar la función de la sensibilidad S_0 usando las medidas de u y de r .
2. Construir la señal e e identificar el sistema en lazo abierto.

Esta aproximación genera buenas estimaciones del sistema reestableciéndolo en lazo abierto a una condición que la regresión sea lineal.

2.2. Identificación Recursiva de Modelos Lineales en Lazo Cerrado

En la práctica, para muchas plantas puede ser difícil su operación en lazo abierto llevando a cabo un protocolo experimental para identificación (por ejemplo: comportamiento integrador, lazo abierto inestable). En algunas situaciones, un controlador puede ya existir y no existir ninguna razón para abrir el lazo cuando el objetivo de la identificación sea obtener un mejor modelo para cualquier diseño de un nuevo controlador o resintonizar el controlador existente.

En control adaptivo, también se tiene que estimar los parámetros del modelo de la planta en operación de lazo cerrado (note que, en este caso, el controlador será con parámetros variantes en el tiempo dependiendo de los parámetros del modelo estimado). En el contexto de control adaptivo, el estudio de identificación recursiva de modelos en lazo cerrado hace énfasis en el manejo de la importancia de los filtros para ser usados con algoritmos de identificación recursiva en lazo cerrado y por otro lado, el uso de algoritmos específicos dedicados a identificación recursiva en lazo cerrado.

El problema de la identificación de modelos en lazo cerrado es de crucial importancia en el contexto de la combinación iterativa de la identificación en lazo cerrado y el rediseño de un control robusto. Este puede ser como un esquema de control adaptivo en el cual la adquisición de datos y la identificación de parámetros son hechas en cada muestra, pero la actualización del controlador es hecha paso a paso. La figura 2.a ilustra las bases de este procedimiento iterativo (para el caso de un controlador digital RST). La parte superior representa el sistema en lazo cerrado real y la parte inferior represente el sistema de diseño. El objetivo es minimizar el error entre los dos sistemas usando nuevos datos adquiridos en operación en lazo cerrado. Ya que el sistema no es directamente manejable, la idea es primero mejorar la estimación del modelo de la planta y luego rediseñar el controlador basado en el nuevo modelo. Esta secuencia de operaciones es llevado a cabo una o varias veces. Sin embargo, un punto clave es que la estimación del nuevo modelo podría ser hecha para reducir el error entre los dos sistemas. De hecho, el objetivo es obtener un mejor predictor para lazo cerrado por medio de una mejor estimación del modelo.

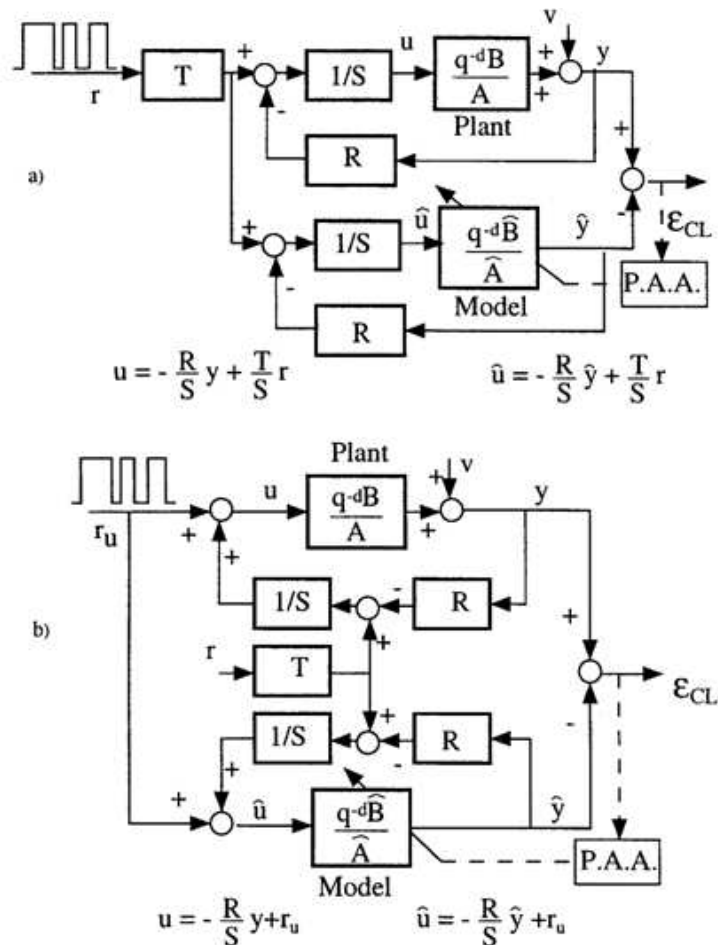


Figura 2: Estructuras par identificación en lazo cerrado. a) excitación como referencia, b) excitación a la salida del controlador

En el contexto de algoritmos recursivos, el problema de identificación de modelo en lazo cerrado puede ser formulado como sigue. Bajo la suposición que el controlador es constante, se identifica un modelo tal que:

1. La estabilidad asintótica global es asegurada para cualquier estimación inicial de parámetros y error inicial entre la salida del sistema real y el del predictor en lazo cerrado (en ausencia de ruido).
2. Obtener un predictor asintóticamente óptimo del sistema real en lazo cerrado.
3. estimaciones no influenciadas (unbiased) de los parámetros de los modelos son asintóticamente obtenidas bajo apropiadas condiciones de bondad (en la presencia

de ruido)

Se asume que la sección de entrada-salida de la planta es identificada como parte del modelo establecido.

El problema de identificación del modelo de una planta en lazo cerrado será visto en dos modos distintos dirigido sin embargo a tipos de algoritmos similares.

La excitación externa puede ser aplicada o en la referencia figura 2.a o agregada a la salida del controlador figura 2. Esto no cambia la estructura del algoritmo, sólo afecta la cantidad externa agregada a la entrada de la planta (o del modelo estimado).

De este esquema básico dos familias de algoritmos recursivos para identificación de modelos en lazo cerrado pueden ser derivadas. La primera aproximación comienza de la reformulación de la ecuación de salida del sistema en lazo cerrado en términos del regresor del predictor (es decir, en términos de $\hat{u}(t)$ y $\hat{y}(t)$ y los valores previos del error predictor en lazo cerrado (ver figura 2). Esta aproximación genera la familia de algoritmos de identificación recursiva de error de salida en lazo cerrado [15].

La segunda aproximación comienza de la reformulación de la ecuación de salida del predictor en términos del vector de medida (es decir, en términos de $u(t)$ y $y(t)$ y los valores previos del error de predicción en lazo cerrado. Esta aproximación conduce a una expresión del error de predicción en lazo cerrado como un error de predicción tipo mínimos cuadrados filtrado para identificación de medidas de entrada-salida. Esto permite usar muchos de los algoritmos de identificación recursiva con un apropiado filtrado de los datos. Esta familia será llamada *algoritmos de identificación recursiva filtrados en lazo cerrado* (FOL). Note que éstos son los tipos de métodos los cuales han sido usados en control adaptivo indirecto para estimación de parámetros de modelos.

El objetivo es estimar los parámetros del modelo definidos por el operador de transferencia:

$$G(q^{-1}) = \frac{q^{-d}B(q^{-1})}{A(q^{-1})} \quad (3)$$

Donde

$$\begin{aligned} B(q^{-1}) &= b_1 q^{-1} + \dots + b_{n_B} q^{-n_B} = q^{-1} B^*(q^{-1}) \\ A(q^{-1}) &= 1 + a_1 q^{-1} + \dots + a_{n_A} q^{-n_A} = 1 + q^{-1} A^*(q^{-1}) \end{aligned} \quad (4)$$

La planta es operada en lazo cerrado con un control digital RST (sin falta de generalidad). La salida de la planta operando en lazo cerrado esta dada por (ver figura 2.b):

$$y(t+1) = -A^* y(t) + B^* u(t-d) + Av(t+1) = \theta^T \varphi(t) + Av(t+1) \quad (5)$$

Donde $u(t)$ es la entrada de la planta, $y(t)$ es la salida planta, $v(t)$ es la salida de perturbación de ruido y:

$$\theta T = [a_1 \dots a_{n_A}, b_1 \dots b_{n_B}] \quad (6)$$

$$\varphi T(t) = [-y(t) \dots, -y(t-n_A+1), u(t-d) \dots, u(t-n_B+1-d)] \quad (7)$$

$$u(t) = -\frac{R}{S} y(t) + r_u \quad (8)$$

Donde r_u es la excitación externa agregada a la salida del controlador (r_u es igual a $r \frac{T}{S}$ si la excitación es aplicada en la referencia).

Para un valor fijo de los parámetros estimados, el predictor del lazo cerrado (el sistema de diseño) puede ser expresado como:

$$\hat{y}(t+1) = -\hat{A}^* \hat{y}(t) + \hat{B}^* \hat{u}(t-d) + Av(t+1) = \hat{\theta}^T \varphi(t) \quad (9)$$

Donde

$$\hat{\theta} T = [\hat{a}_1 \dots \hat{a}_{n_A}, \hat{b}_1 \dots \hat{b}_{n_B}] \quad (10)$$

$$\varphi T(t) = [-\hat{y}(t) \dots, -\hat{y}(t-n_A+1), \hat{u}(t-d) \dots, \hat{u}(t-n_B+1-d)] \quad (11)$$

$$u(t) = -\frac{R}{S} \hat{y}(t) + r_u \quad (12)$$

El error de predicción de lazo cerrado (output) es definida como:

$$\varepsilon_{CL}(t+1) = y(t+1) - \hat{y}(t+1) \quad (13)$$

Note como claramente éste resulta de la figura 2, para valores constantes de los parámetros estimados, el vector regresor del predictor $\phi(t)$ depende solamente de la excitación externa. Por eso bajo la suposición que la excitación externa (r o r_u) y la perturbación estocástica v son independientes, $\phi(t)$ y $v(t)$ no son correlacionados (también como $\phi(t)$ y $\varepsilon_{CL}(t+1)$). El esquema tiene la estructura de una predicción error de salida. Si además un modelo para el ruido $v(t)$ es asumido, entonces un predictor puede ser construido tal que bajo el exacto conocimiento del modelo entrada-salida y del modelo de ruido, el error predicción lazo cerrado tiende asintóticamente a un ruido blanco.

Para todos los métodos el algoritmo de adaptación de parámetros (PAA) tiene la forma general:

$$\hat{\theta}(t+1) = \hat{\theta}(t) + F(t)\Phi(t)v(t+1) \quad (14)$$

$$F^{-1}(t+1) = \lambda_1(t)F^{-1}(t) + \lambda_2(t)\Phi(t)\Phi^T(t) \quad (15)$$

$0 < \lambda_1(t) \leq 1; 0 < \lambda_2(t) \leq 2; F(0) > 0; F^{-1}(t) > \alpha F^{-1}(0); 0 < \alpha < \infty$

$$F(t+1) = \frac{1}{\lambda_1(t)} \left[F(t) - \frac{F(t)\Phi(t)\Phi^T(t)F(t)}{\frac{\lambda_1(t)}{\lambda_2} + \Phi^T(t)F(t)\Phi(t)} \right] \quad (16)$$

$$v(t+1) = \frac{v^o(t+1)}{1 + \Phi^T(t)F(t)\Phi(t)} \quad (17)$$

Donde $v^o(t+1) = f_1(\hat{\Theta}(t), \hat{\Theta}(t-1), \dots, y(t+1), v(t), v(t-1), \dots)$ es el error de adaptación *a priori*, $v(t+1) = f_2(\Theta(t+1), \Theta(t), \dots, y(t+1), v(t), v(t-1), \dots)$ es el error de adaptación *a posteriori* y $\Phi(t)$ es el vector de observación.

Para cada algoritmo de identificación recursiva $\Theta, \Phi, v^o(t+1)$ tendrán una específica expresión. Note que la secuencia $\lambda_1(t)$ y $\lambda_2(t)$ dan diferentes leyes para el tiempo de evolución de la ganancia de adaptación $F(t)$. Por análisis de convergencia en el ambiente estocástico, se asume que un PAA es usado con una ganancia de adaptación disminuyendo (es decir, $\lambda_1(t) \equiv 1$ y $\lambda_2(t) > 0$ o $\lambda_2(t) \equiv \lambda_2 > 0$)

2.3. Algoritmos de Error de Salida en Lazo Cerrado

2.3.1. CLOE Closed Loop Output Error

La observación clave es que la salida del sistema en lazo cerrado dada por (5) puede ser expresada como:

$$y(t+1) = \theta^T \varphi(t) + Av(t+1) = \theta^T \varphi(t) - \left[A^* + \frac{q^{-d} B^* R}{S} \right] \varepsilon_{CL}(t) + Av(t+1) \quad (18)$$

y

$$1 + q^{-1} \left[\frac{A^* + q^{-d} B^* R}{S} \right] = \frac{AS + q^{-d} BR}{S} = \frac{P}{S} \quad (19)$$

Donde $P = AS + q^{-d} BR$ define los polos del sistema real en lazo cerrado.

Restando la ecuación (9) de (18) se obtiene en el caso determinístico ($v(t+1) \equiv 0$):

$$\varepsilon_{CL}(t+1) = [\theta - \hat{\theta}]^T \phi(t) - \left[\frac{A^* + q^{-d} B^* R}{S} \right] \varepsilon_{CL}(t) \quad (20)$$

Tomando en cuenta la relación (19), la ecuación (20) puede ser rescrita como:

$$\varepsilon_{CL}(t+1) = \frac{S}{P} [\theta - \hat{\theta}]^T \phi(t) \quad (21)$$

Nótese que en el caso lineal con parámetros conocidos debido a que $\phi(t)$ y $y(t)$ son no correlacionados, es obtenido un predictor óptimo minimizando $E\varepsilon_{CL}^2(t+1)$ para $\hat{\theta} = \theta$.

Reemplazando ahora el predictor fijo del lazo cerrado dado en la ecuación (14) por un predictor ajustable, es decir:

Salida predicha a priori:

$$\hat{y}^o(t+1) = \hat{y}[t+1|\hat{\theta}(t)] = \theta^T(t)\phi(t) \quad (22)$$

Salida predicha a posteriori:

$$\hat{y}(t+1) = \hat{y}[t+1|\hat{\theta}(t+1)] = \theta^T(t+1)\phi(t) \quad (23)$$

y definiendo el error de predicción a priori como:

$$\varepsilon^o_{CL}(t+1) = y(t+1) - \hat{y}^o(t+1) \quad (24)$$

y el error de predicción a posteriori como:

$$\varepsilon_{CL}(t+1) = y(t+1) - \hat{y}(t+1) \quad (25)$$

Se sugiere un PAA con :

$$\begin{aligned} \Theta &= \hat{\theta}(t) \\ \hat{\Phi} &= \hat{\phi}(t) \\ v^o(t+1) &= \varepsilon^o_{CL}(t+1) \end{aligned} \quad (26)$$

Este es llamado el algoritmo error de salida en lazo cerrado (CLOE) [15]. Usando los resultados del PAA se puede mostrar que en un ambiente determinístico la condición de suficiencia para $\lim_{t \rightarrow \infty} \varepsilon_{CL}(t+1) = 0$ junto con el límite de $\varepsilon_{CL}(t+1)$ para cualquier condición inicial es que:

$$H'(z^{-1}) = \frac{S(z^{-1})}{P(z^{-1})} - \frac{\lambda_2}{2} \quad (27)$$

Es estrictamente real positivo (donde $\lambda_2(t) \leq \lambda_2 < 0$). Ya que $\varepsilon_{CL}(t)$ está limitado, esto permite concluir que $\hat{u}(t)$ y $\hat{y}(t)$ están limitados (bajo la hipótesis que S es estable y que el lazo cerrado es inestable) y por eso también $\lim_{t \rightarrow \infty} \varepsilon^o_{CL}(t+1) = 0$. En un ambiente estocástico usando el método del promedio, la ecuación del error de predicción de lazo cerrado para un valor fijo de parámetros $\hat{\theta}$ estimados, toma la forma:

$$\varepsilon_{CL}(t+1, \hat{\theta}) = \frac{S}{P} [\theta - \hat{\theta}]^T \phi(t, \hat{\theta}) + \frac{AS}{P} v(t+1) \quad (28)$$

Bajo la hipótesis que $v(t+1)$ es independiente con respecto a la excitación externa $r_u, \phi(t, \hat{\theta}) y v(t+1)$ son independientes y el método del promedio puede ser usado ($\lambda_1(t) \equiv 1, \lambda_1(t) > 0$). La misma condición real positiva de la ecuación (27) asegura asintóticamente estimaciones no influenciadas (unbiased) bajo condiciones de bondad [15]

2.3.2. F-CLOE Algoritmo Error de Salida en Lazo Cerrado Filtrado

La ecuación (25) puede también ser rescrita como:

$$\varepsilon_{CL}(t+1) = \frac{S}{P} \frac{P}{S} [\theta - \hat{\theta}] \frac{S}{\hat{P}} \phi(t) = \frac{P}{P} [\theta - \hat{\theta}] \phi_f(t) \quad (29)$$

donde

$$\phi_f(t) = \frac{S}{\hat{P}}\phi(t) \quad (30)$$

$$P = AS + q^{-d}BR \quad (31)$$

en la ecuación 31 \hat{P} es una estimación de los polos reales de lazo cerrado basada en una estimación inicial del modelo de la planta (por ejemplo usando un experimento de lazo abierto). Esta formulación conduce a un *algoritmo de error de salida en lazo cerrado filtrado* (F-CLOE) [15] el cual usa el mismo predictor ajustable como CLOE (ver ecuaciones (22) y (23)) y el PAA con:

$$\begin{aligned} \hat{\Theta}(t) &= \hat{\theta}(t) \\ \Phi(t) &= \phi_f(t) \\ v^o(t+1) &= \varepsilon^o_{CL}(t+1) \end{aligned} \quad (32)$$

Se puede mostrar que el descuido de la no conmutatividad de los operadores variantes con el tiempo (un exacto algoritmo puede ser derivado), bajo la condición suficiente que:

$$H'(z^{-1}) = \frac{\hat{P}(z^{-1})}{P(z^{-1})} - \frac{\lambda_2}{2} \quad (33)$$

es estrictamente real positiva, la estabilidad asintótica en un ambiente determinístico y la no influenciabilidad asintótica en un ambiente estocástico son aseguradas.

Se puede además relajar la condición de la ecuación (33) filtrando $\phi(t)$ a través de un filtro variante en el tiempo $S/\hat{P}(t)$ donde $\hat{P}(t)$ corresponde a la estimación actual del lazo cerrado dado por $\hat{P}(t) = \hat{A}(t)S + q^{-d}\hat{B}(t)R$ donde $\hat{A}(t)$ y $\hat{B}(t)$ son la estimaciones actuales de los polinomios A y B (El algoritmoAF-CLOE)

2.3.3. X-CLOE Algoritmo del Error de Salida en Lazo Cerrado Extendido

Para el caso $v(t+1) = \frac{C}{A}e(t+1)$, donde $e(t+1)$ es un ruido blanco gaussiano con media cero y $C(q^{-1}) = 1 + q^{-1}C^*(q^{-1})$ es un polinomio asintóticamente estable, un modelo de predicción de error de salida extendido puede ser definido:

$$\begin{aligned} \hat{y}(t+1) &= -\hat{A}^*\hat{y}(t) + \hat{B}^*\hat{u}(t-d) + \hat{H}^*\frac{\varepsilon_{CL}(t)}{S} \\ &= \hat{\theta}^T\phi(t) + \hat{H}^*\frac{\varepsilon_{CL}(t)}{S} = \hat{\theta}_e^T\phi_e(t) \end{aligned} \quad (34)$$

La ecuación 5 para la salida de la planta se convierte en este caso:

$$y(t+1) = \hat{\theta}^T \phi(t) + \hat{H}^* \frac{\varepsilon_{CL}(t)}{S} - C^* \varepsilon_{Cl}(t) + C^* e(t+1) \quad (35)$$

donde

$$H^* = h_1 + h_2 q^{-1} + \dots + h_{n_H} q^{-n_H+1} = C^* S - A^* S - q^{-d} B^* R \quad (36)$$

$$H = 1 + q^{-1} H^* = 1 + C^* S - P \quad (37)$$

y restando (34) de (35), se obtienen las siguientes expresiones para el error de predicción en lazo cerrado:

$$\varepsilon_{CL}(t+1) = \frac{1}{c} [\theta_e - \hat{\theta}_e]^T \phi_e(t) + e(t+1) \quad (38)$$

donde

$$\theta_e^T = [\theta^T, h_1, \dots, h_{n_H}] \quad (39)$$

$$\hat{\theta}_e^T = [\hat{\theta}^T, \hat{h}_1, \dots, \hat{h}_{n_H}] \quad (40)$$

$$\phi_e^T(t) = [\phi^T, \varepsilon_{CLf}(t), \dots, \varepsilon_{CLf}(t - n_H + 1)] \quad (41)$$

$$\varepsilon_{CLf}(t) = \frac{1}{S} \varepsilon_{CL}(t) \quad (42)$$

la ecuación (2.3.2) claramente muestra que para $\hat{\theta}_e = \theta_e$ el error de predicción para lazo cerrado tiende asintóticamente hacia $e(t+1)$.

Reemplazando el predictor fijo de (34) con uno ajustable, un algoritmo de identificación recursiva (X-CLOE) puede ser obtenido usando un PAA con:

$$\begin{aligned} \hat{\Theta}(t) &= \hat{\theta}_e(t) \\ \Phi(t) &= \phi_e(t) \end{aligned} \quad (43)$$

$$v^o(t+1) = \varepsilon^o_{CL}(t+1) = y(t+1) - \hat{\theta}_e(t) \phi_e(t)$$

el análisis en el caso determinístico usando el método del promedio muestra que la esta estabilidad asintótica global es asegurada sin alguna condición real positiva (ya que la ecuación del error de predicción a posteriori en lazo cerrado en este caso es $\varepsilon_{CL} = [\theta_e - \hat{\theta}_e(t+1)]^T \phi_e(t)$). En el ambiente estocástico la ecuación del error de predicción en lazo cerrado toma la forma:

$$\varepsilon_{CL}(t+1) = \frac{1}{c} [\theta_e - \hat{\theta}_e]^T \phi_e(t) + e(t+1)$$

estimaciones asintóticamente no influenciadas (un-biased) en un ambiente estocástico pueden ser obtenidas bajo la condición de suficiente:

$$H'(z^{-1}) = \frac{1}{C(z^{-1})} - \frac{\lambda_2}{2} \quad (44)$$

es estrictamente real positiva (donde $\max_t \lambda_2(t) \leq \lambda_2 < 2$) usando cualquier método promedio o la aproximación martingala.

Capítulo III

RESULTADOS EXPERIMENTALES EN SISTEMAS LINEALES

Considere un controlador RST como el mostrado en la figura 1.

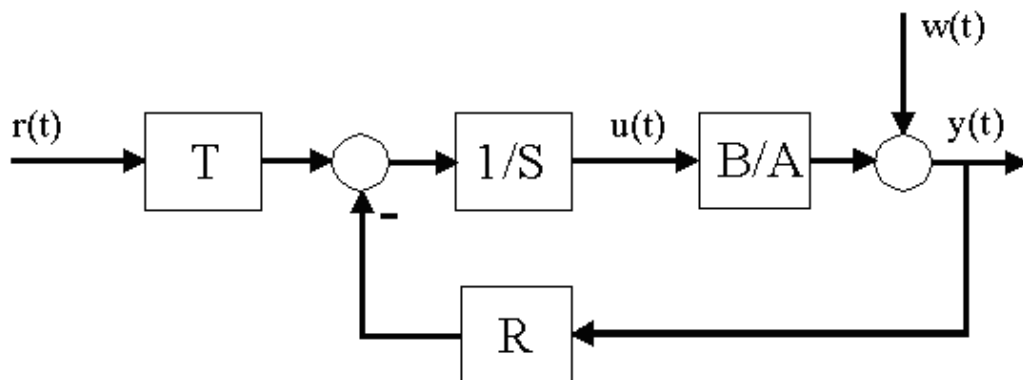


Figura 3: Controlador RST

En esta sección se realizarán algunos experimentos en los cuales se supondrá que la planta B/A es desconocida y se intenta identificar. Para identificar la planta se utilizarán los algoritmos CLOE, FCLOE, AFCLOE y XCLOE, los cuales son explicados con anterioridad.

3.1. Planta 1

Se estudiará un sistema como el mostrado en la figura 3, con las siguientes características:

$$\begin{aligned}
 A(q^{-1}) &= 1 - 1,5q^{-1} + 0,7q^{-2} \\
 B(q^{-1}) &= q^{-1} + 0,5q^{-2} \\
 R(q^{-1}) &= 0,8659 - 1,2763q^{-1} + 0,5204q^{-2} \\
 S(q^{-1}) &= 1 - 0,6283q^{-1} - 0,3717q^{-2} \\
 T(q^{-1}) &= 0,11
 \end{aligned} \tag{45}$$

La excitación $r(t)$ y la salida $y(t)$ son mostradas en la figura 4. La salida está contaminada por ruido blanco con potencia promedio 0.05. Esto se hace para simular los diferentes tipos de perturbaciones que pueden surgir al tomar las mediciones.

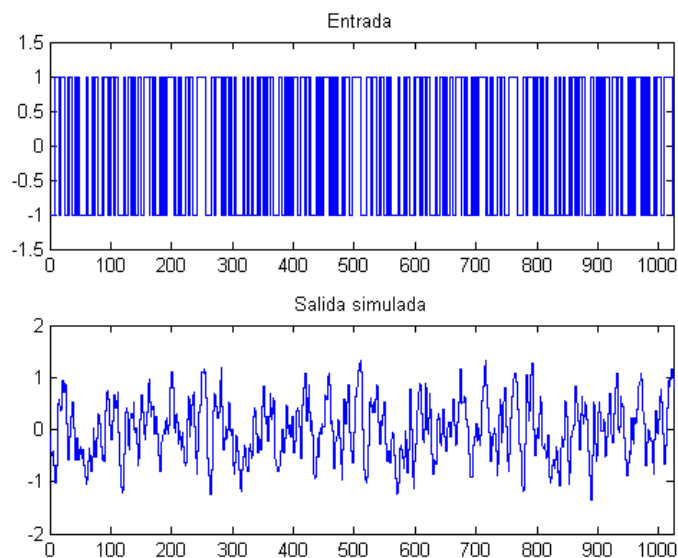


Figura 4: Señales de excitación (superior) y salida del modelo 1 (inferior)

Usando está entrada y salida se obtienen los parámetros mostrados en la tabla 1 para cada uno de los algoritmos estudiados:

Donde el intervalo Vinnicombe es una métrica que es usada en el contexto de control robusto para mostrar la distancia entre dos funciones de transferencia. Es en cierto sentido un criterio que permite evaluar la aproximación hecha por el algoritmo.

Algoritmo	A Estimada	B Estimada	Intervalo de Vinnicombe
CLOE	[1 -1.4808 0.6716]	[0 0.9527 0.4900]	0.0230
FCLOE	[1 -1.4750 0.6744]	[0 0.9532 0.5074]	0.0146
AFCLOE	[1 -1.4844 0.6821]	[0 0.9684 0.4722]	0.0213
XCLOE	[1 -1.5138 0.6990]	[0 0.7447 0.5155]	0.1209

Tabla 1: Resultados obtenidos con los diferentes algoritmos para el sistema lineal 1

Algoritmo	A Estimada	B Estimada	Intervalo de Vinnicombe
CLOE	[1 0.1931 0.1305]	[0 -0.0040]	0.3447
FCLOE	[1 0.2752 0.1529]	[0 -0.0028]	0.3448
AFCLOE	[1 0.0097 0.0667]	[0 -0.0040]	0.3447
XCLOE	[1 -1.8783 0.9613]	[0 0.0889]	0.9923

Tabla 2: Resultados obtenidos con los diferentes algoritmos para el sistema lineal 2

En este caso, el mejor algoritmo resulta ser FCLOE.

3.2. Planta 2

La siguiente planta que se estudiará cuenta con las siguientes características:

$$\begin{aligned}
 A(q^{-1}) &= 6,233 \times 10^{-5} + 6,229 \times 10^{-5}q^{-1} \\
 B(q^{-1}) &= 1 - 1,973q^{-1} + 0,998q^{-2} \\
 R(q^{-1}) &= 1 - 0,85q^{-1} \\
 S(q^{-1}) &= 1 \\
 T(q^{-1}) &= 1 - 0,85q^{-1}
 \end{aligned}
 \tag{46}$$

La excitación y salida del sistema se muestran en la figura 3. Se uso la misma entrada que en la señal anterior.

A continuación se corren los algoritmos de identificación para este sistema. Los resultados obtenidos se muestran en la tabla 2.

Para esta planta los algoritmos no encuentran una planta del todo satisfactoria. Sin embargo, la aproximación que realizan puede considerarse bajo muchos aspectos suficiente.

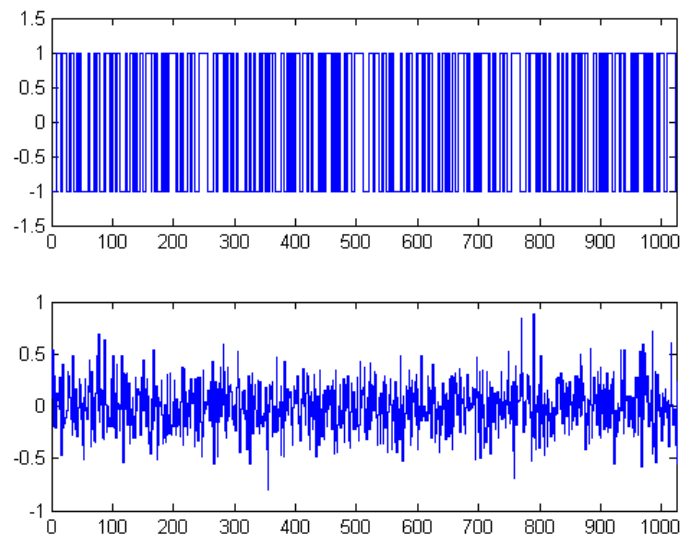


Figura 5: Señales de excitación (superior) y salida del modelo 2 (inferior)

3.3. Planta 3

Características de la planta: En el proceso de identificación el primer paso corresponde a realizar experimentos con la planta. Se simulará entonces un modelo con controlador RST y con los siguientes parámetros. La salida de la planta está contaminada con ruido blanco con potencia de 0.05.

$$A(q^{-1}) = 1 - 1,3528q^{-1} + 1,5502q^{-2} - 1,2798q^{-3} + 0,9115q^{-4} \quad B(q^{-1}) = 0,4116q^{-3} + 0,524q^{-4} \quad (47)$$

Los parámetros del controlador son los siguientes:

$$\begin{aligned} R(q^{-1}) &= 0,4526 - 0,4564q^{-1} - 0,6857q^{-2} + 1,0955q^{-3} - 0,1449q^{-4} \\ S(q^{-1}) &= 1 + 0,2345q^{-1} - 0,8704q^{-2} - 0,4474q^{-3} + 0,0833q^{-4} \\ T(q^{-1}) &= 0,2612 \end{aligned} \quad (48)$$

Al simular la planta con estas características se obtiene la siguiente figura:

En esta prueba se supondrá que no existe ningún conocimiento previo sobre el modelo original. Para estimar el número de parámetros previos de los cuales puede depender el modelo, se utilizará el método sugerido en [X. He, H. Asada: .^A

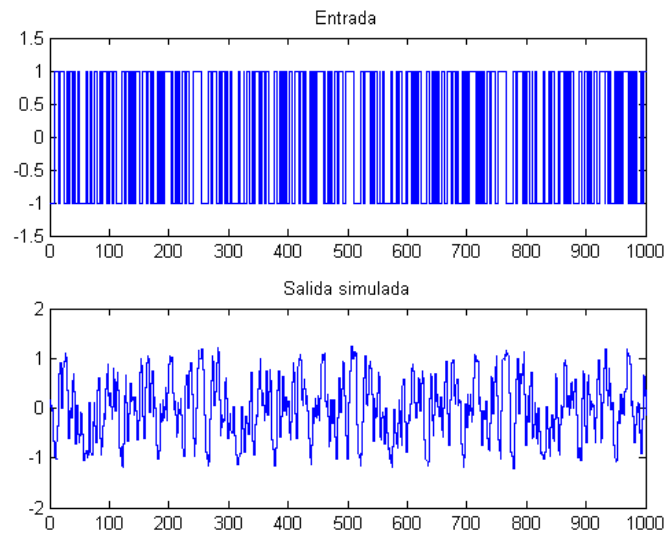


Figura 6: Señales de excitación (superior) y salida del modelo 2 (inferior)

Algoritmo	A Estimada	B Estimada	Intervalo de Vinnicombe
CLOE	[1 -1.3473 1.4549 -1.1332 0.8230]	[0.4390 0.4704]	0.2771
FCLOE	[1 -1.3354 1.5330 -1.2770 0.9195]	[0.4130 0.5406]	0.0551
AFCLOE	[1 -1.3357 1.5424 -1.2908 0.9275]	[0.4104 0.5479]	0.0707
XCLOE	[1 -1.2581 1.1765 -0.8005 0.6629]	[0.4790 0.4179]	0.658

Tabla 3: Resultados obtenidos con los diferentes algoritmos para el sistema lineal 3

New Method for Identifying Orders of Input-Output Models for Nonlinear Dynamic Systems”]. .

De esta gráfica se puede observar que después de cuatro parámetros de entrada y salida anteriores, las dos muestras no están muy correlacionadas. Se selecciona entonces un modelo de cuarto orden ara identificar el sistema.

A continuación se procede a estimar los valores mediante los algoritmos propuestos por Landau. Es decir por CLOE, FCLOE y XCLOE. Los modelos estimados son mostrados en la tabla 3. .

En este caso se observa que la mejor identificación es realizada por FCLOE.

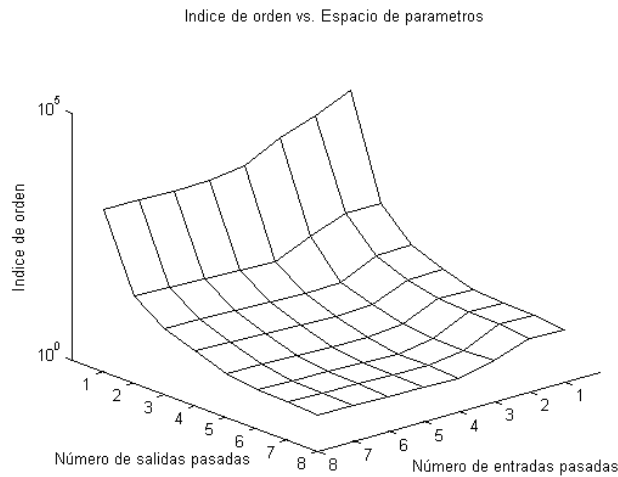


Figura 7: Determinación del orden del sistema

3.4. Planta 4

Características de la planta: Considérese la planta descrita por las siguientes ecuaciones:

$$A(q^{-1}) = 1 - 2q^{-1} + 1,45q^{-2} - 0,35q^{-3} \quad B(q^{-1}) = q^{-1} - 0,5q^{-2} \quad (49)$$

La planta se encuentra controlada por un sistema RTS en lazo cerrado con las siguientes características:

$$\begin{aligned} R(q^{-1}) &= 0,8659 - 1,2763q^{-1} - 0,5204q^{-2} \\ S(q^{-1}) &= 1 - 0,6283q^{-1} - 0,3717q^{-2} \\ T(q^{-1}) &= 1 \end{aligned} \quad (50)$$

El sistema se simulara, agregando un ruido blanco con potencia 0.01. La respuesta de entrada-salida obtenida durante la simulación se muestra en la figura 8.

A continuación se procede a calcular el modelo estimado usando cada uno de los algoritmos propuestos. Los resultados se muestran en la tabla 4 .

Se muestra que el método encuentra unas funciones de transferencia que se aproximan bastante a las originales. En este caso, la mejor aproximación se obtiene con el algoritmo CLOE.

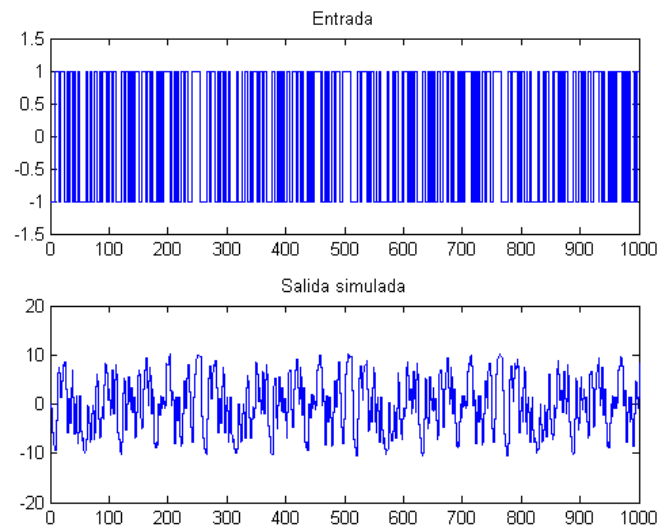


Figura 8: Señales de excitación (superior) y salida del modelo 2 (inferior)

Algoritmo	A Estimada	B Estimada	Intervalo de Vinnicombe
CLOE	[1 -2.0037 1.4574 -0.3546]	[1.0023 0.4942]	0.0014
FCLOE	[1 -1.9970 1.4440 -0.3466]	[1.0009 0.5051]	0.0016
AFCLOE	[1 -1.9914 1.4348 -0.3414]	[0.9975 0.5140]	0.0033
XCLOE	[1 -1.9966 1.4433 -0.3463]	[1.0029 0.5052]	0.0025

Tabla 4: Resultados obtenidos con los diferentes algoritmos para el sistema lineal 4

Capítulo IV

MODELAMIENTO E IDENTIFICACIÓN DIFUSOS

Las técnicas y algoritmos para el reconocimiento de los parámetros de un modelo de dato se conocen como identificación difusa. Parámetros de una estructura difusa como funciones de pertenencia y peso de las reglas son sintonizadas usando los datos de entrada y salida. En este proceso no se hace necesario conocimiento a priori, en lugar de ello se espera que las funciones de pertenencia y reglas extraídas proporcionen una interpretación del comportamiento del sistema. La identificación es vista como una descomposición de un sistema no lineal, lo cual genera un balance entre la complejidad (no uniforme) y la exactitud del modelo.

El modelamiento difuso reúne el procesamiento lógico de información con estructuras matemáticas capaces de representar mapeos no lineales complejos. La estructura basada en reglas de los sistemas difusos, contribuye a los modelos difusos generados por tratamiento de datos, ya que generan una descripción cualitativa, que combinada con el conocimiento de expertos ayuda a comprender, validar y/o simplificar el modelo.

4.1. Modelos Takagi Sugeno Kang

Los modelos TSK lineales son modelos difusos que constan de reglas con la siguiente estructura:

$$\text{Si } x(k) \text{ es } A_i \text{ Entonces } y_i = a_i'x(k) + b_i \quad i = 1, 2, \dots, c \quad (51)$$

Donde c es el número de reglas, x es el vector de entrada, A_i y y_i son el conjunto

difuso (multidimensional) y la función lineal de la i -ésima regla respectivamente. Así mismo, a_i y b_i son el vector paramétrico y el desplazamiento escalar de la i -ésima función lineal. La proposición de entrada x es A_i puede ser expresada como una combinación lógica de proposiciones unidimensionales definidas en cada componente del vector de entrada x .

$$\text{Si } x_i(k) \text{ es } A_{i1} \text{ y } x_p(k) \text{ es } A_{ip} \text{ Entonces } y_i = \mathbf{a}'_i \mathbf{x}(k) + b_i \quad (52)$$

Tal como en los modelos difusos Mamdani el grado de cumplimiento de la regla esta dado por:

$$\beta_i(x) = \mu_{A_{i1}}(x_1) \wedge \mu_{A_{i2}}(x_2) \wedge \dots \wedge \mu_{A_{ip}}(x_p) \quad (53)$$

Donde \wedge representa una T-norma [2] [19]. La salida del sistema TSK es obtenida combinando las reglas de la siguiente forma:

$$y = \frac{\sum_{i=1}^c \beta_i(\mathbf{x})(\mathbf{a}'_i \mathbf{x} + b_i)}{\sum_{i=1}^c \beta_i(\mathbf{x})} \quad (54)$$

Tomando la normalización del grado de cumplimiento como:

$$\varphi_i(x) = \frac{\beta_i(x)}{\sum_{i=1}^c \beta_i(x)} \quad (55)$$

El modelo TSK con una estructura común de consecuente puede ser expresado como un modelo pseudolineal con parámetros dependientes de las entradas. Entonces la ecuación (54) se puede reescribir como:

$$y = \sum_{i=1}^c \varphi_i(\mathbf{x})(\mathbf{a}'_i \mathbf{x} + b_i) \quad (56)$$

$$y = \left(\sum_{i=1}^c \varphi_i(\mathbf{x}) \mathbf{a}'_i \right) \mathbf{x} + \sum_{i=1}^c \varphi_i(\mathbf{x}) b_i = \mathbf{a}^T(\mathbf{x}) \mathbf{x} + b(\mathbf{x}) \quad (57)$$

La anterior expresión muestra que los modelos TSK pueden cumplir el papel de regresores de funciones, es decir, pueden aproximar con cierto margen de exactitud cualquier función $y = f(\mathbf{x})$.

4.2. Técnicas de Clustering

Las técnicas de clustering surgieron dentro de la disciplina de patrones de reconocimiento. Esta técnica busca hallar subgrupos con cierto grado de similaridad dentro de una colección de datos ya sean cuantitativos (numéricos), cualitativos (categorías) o una combinación de éstos. Además de identificar cada subgrupo, determina parámetros representativos de cada uno. Originalmente se planteó la idea de que cada dato solo podría pertenecer a un solo subgrupo, lo cual es considerado como "hard clustering" (teoría clásica de conjuntos). Mas adelante se tomo la idea de clustering difuso en el cual particionamiento de datos se hace de forma tal que la transición entre conjuntos es gradual en lugar de abrupta como lo haría el "hard clustering" [2] [10]. Para el estudio, la similaridad es definida como una distancia, la cual puede ser medida entre vectores que contienen los datos y un elemento prototipo del cluster, normalmente el centro [2].

Los datos son típicamente observación de un proceso físico, entonces cada observación consiste de n variables medidas, agrupadas en un vector columna n -dimensional $\mathbf{z}_k = [z_{1k}, \dots, z_{nk}]^T$, un conjunto de N observaciones es denotado por $\mathbf{Z} = [z_k | K = 1, 2, \dots, N]$ y está representado como una matriz $n \times N$:

$$\mathbf{Z} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1N} \\ z_{21} & z_{22} & \dots & z_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nN} \end{bmatrix},$$

Cuando son aplicadas las técnicas de clustering al modelamiento e identificación de sistemas dinámicos, las columnas de \mathbf{Z} contienen muestras de las señales en el tiempo y las columnas son en consecuencia, variables físicas observadas en el sistema. Para representar la dinámica del sistema son incluidos también típicamente valores pasados de las variables en la matriz \mathbf{Z} .

Existen varios algoritmos de clustering los cuales tienen en cuenta: la forma geométrica, densidad y relación espacial de cada cluster, además de la distancia entre ellos [2] [10]. Se encuentran algoritmos iterativos basados en la optimización de una función objetivo diferenciable, con la cual se mide la conveniencia de la

partición. Una de las primeras técnicas de agrupamiento planteada que utiliza una función objetivo es conocida como Fuzzy C-means [2] [20] [11]. Debido a la medida de similaridad de esta técnica, este algoritmo solo es capaz de detectar cluster de forma circular. Ya que lo que se busca es generar submodelos lineales, esta técnica no es apropiada.

Debido a lo anterior, como punto de partida para este trabajo se tomaron como referencia cuatro de algoritmos que buscan hallar tendencias lineales: Gustafson-Kessel, Maximum Likelihood, la modificación a la versión simplificada del Maximum Likelihood y Fuzzy c-Regression Models [10] [8] [1] [11]. Los dos primeros proponen una distancia adaptiva con el fin de detectar clusters de diferentes formas geométricas y orientación, es decir, generan clusters hiperelipsoidales que pueden ser aproximados a hiperplanos. La desventaja de estos algoritmos es que al derivar los antecedentes (funciones de pertenencia) que forman el modelo TSK, se produce errores por proyección o poca interpretabilidad cuando se corrige este error. El tercer algoritmo maneja clusters de forma hiperelipsoidal pero paralelos a los ejes de las variables de entrada. Debido a esto no se producen errores por descomposición al derivar los antecedentes teniendo una alta interpretabilidad; la desventaja es que se pierde predictividad de modelo. El algoritmo FCRM busca generar los parámetros de prototipos funcionales que se desempeñen como regresores locales y cuya combinación lineal cumpla el papel de regresor global. Una opción es escoger subfunciones lineales de tal forma que el algoritmo identifique clusters hiperplanares, donde los parámetros de cada submodelo se obtienen directamente del algoritmo. En este caso las funciones de pertenencia realizan la labor de etiquetas sobre los datos, las cuales determinan con cual de los submodelos de regresión se está más asociado. Este algoritmo ha sido propuesto por Hathaway y Bezdek en [2] [11].

4.2.1. Algoritmo Gustafson-Kessel

Este algoritmo se basa en la optimización de la siguiente función objetivo [2] [10] [4]:

$$J = (Z; U, V, A) = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m d_{ij}^2 \quad (58)$$

Siendo $U = [\mu_{ik}]$, $V = [v_1, v_2, \dots, v_c]$, $A = [A_1, A_2, \dots, A_c]$, c es el número de clusters, N es el número de datos que se toman de la dispersión, μ_{ik} es la función de pertenencia del k -ésimo dato al i -ésimo cluster, m es un índice que indica la fusividad (fuzzyness) entre las fronteras de cada cluster; entre mas alto sea m más difusas son las fronteras de cada cluster y d_{ij}^2 es una distancia (norma) entre puntos denominada norma Mahalanobis, la cual es de la siguiente forma:

$$d_{ij}^2 = (\mathbf{z}_k - \mathbf{v}_i)^T \mathbf{A}_i (\mathbf{z}_k - \mathbf{v}_i) \quad (59)$$

donde \mathbf{z}_k , es un vector columna que contiene el k -ésimo dato tomado, cuya dimensión n es idéntica a la dimensión del espacio producto de entrada-salida, v_i es un vector columna prototipo de dimensión n denominado centro del cluster y \mathbf{A}_i es una matriz simétrica cuyo tamaño es $(n \times n)$, esta matriz indica la orientación de cada cluster. En este caso la norma Mahalanobis mide la distancia entre el k -ésimo dato \mathbf{z}_k y el centro del i -ésimo cluster \mathbf{v}_i .

En el momento de minimizar se toman como variables μ_{ik} , \mathbf{A}_i y \mathbf{v}_i , teniendo en cuenta las siguientes restricciones [2] [4]:

$$\sum_{i=1}^c \mu_{ik} = 1 \quad 1 \leq k \leq N \quad (60)$$

$$|\mathbf{A}_i| = \rho_i \quad \rho_i > 0 \quad \forall_i \quad (61)$$

donde ρ_i es denominado el volumen de cada cluster. Para minimizar la ecuación(58) con las restricciones y (60) se utilizan multiplicadores de Lagrange. Sin las restricciones la minimización conduciría a respuestas triviales $\mu_{ik} = 0$ y $\mathbf{A}_i = 0$. La primera restricción implica que la suma de los grados de pertenencia de cada dato a los clusters debe ser igual a uno, lo cual es acorde con la teoría de probabilidad. La segunda limita el volumen de cada cluster a un valor determinado.

El proceso de minimización se realiza de forma alternante, es decir, se minimiza una variable suponiendo las otras constantes y se iguala a cero. Como resultado

la minimización se obtiene las siguientes expresiones, sobre las cuales se realiza el proceso iterativo.

$$\mathbf{v}_i^{(l)} = \frac{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m \mathbf{z}_k}{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m} \quad (62)$$

$$\mathbf{F}_i^{(l)} = \frac{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m (\mathbf{z}_k - \mathbf{v}_i^{(l)}) (\mathbf{z}_k - \mathbf{v}_i^{(l)})^T}{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m} \quad (63)$$

$$\mathbf{A}_i^{(l)} = |\rho_i \det(\mathbf{F}_i^{(l)})^{1/n} (\mathbf{F}_i^{(l)})^{-1}| \quad (64)$$

$$(d_{ij}^2)^{(l)} = (\mathbf{z}_k - \mathbf{v}_i^{(l)})^T \mathbf{A}_i^{(l)} (\mathbf{z}_k - \mathbf{v}_i^{(l)}) \quad (65)$$

$$\mu_{ik}^{(l)} = \frac{1}{\sum_{j=1}^c (d_{ik}^{(l)} / d_{jk}^{(l)})^{2/m-l}} \quad (66)$$

donde l indica el número de la iteración. Las iteraciones inician determinando el número de cluster c , el exponente m (en la literatura generalmente se toma un valor de 2) determinando los hipervolumenes de cada cluster ρ_i y seleccionando de manera aleatoria la matriz de partición difusa \mathbf{U} de tamaño $(c \times N)$ la cual contiene los grados de pertenencia de los datos y cumple con (60). Entonces \mathbf{U} se encuentra organizada así:

$$\mathbf{U} = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,N} \\ \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{c,1} & \mu_{c,2} & \cdots & \mu_{c,N} \end{bmatrix},$$

Posteriormente con la ecuación (63) es calculado el valor del i -ésimo centro v_i . Teniendo el valor de v_i es calculada la matriz de covarianza difusa \mathbf{F}_i . A partir de (64) se calcula la matriz \mathbf{A}_i necesaria para hallar la distancia de la expresión (64). La ecuación ((61)) describe una hiperelipse rotada. La dirección y dimensión de cada uno de los ejes rotados están establecidas por los autovectores y autovalores de la matriz \mathbf{A}_i respectivamente. Lo anterior fija la geometría de los clusters. La actualización de \mathbf{U} se logra mediante (80). El proceso concluye cuando la diferencia entre \mathbf{U} actual y anterior es menor que el límite de tolerancia ε (0.01 ó 0.001), es decir:

$$\mathbf{U} = \|\mathbf{U}^l - \mathbf{U}^{l-1}\| < \varepsilon \quad (67)$$

Este algoritmo tiene la ventaja de ser poco sensible a los parámetros de inicialización. Presenta limitaciones en cuanto al volumen de cada cluster, ya que es necesario predeterminarlo debido a la restricción en (60). Cuando no se tiene conocimiento previo del sistema, el hipervolumen de cada cluster es fijado en uno, lo cual impide que este algoritmo sea capaz identificar clusters con diferentes volúmenes. Además, si el sistema presenta comportamientos estrictamente lineales o la cantidad de datos es muy pequeña, A_i puede presentar problemas de singularidad, interrumpiendo el proceso iterativo. En [4] se presentan dos técnicas para corregir los problemas matemáticos que pueden surgir en el cálculo de la matriz de covarianza A_i debido a linealidades del sistema.

4.3. Algoritmo RoFeR

El funcional propuesto a minimizar en el algoritmo GK es el de la ecuación (58) con la distancia descrita por la ecuación (59), sujeta a las restricciones (60) y (61).

Por ser A_{is} simétrica, la funcional dada por la ecuación (58), describe clusters de forma hipereplisoidal rotadas. Por tanto, si lo que se desea es obtener clusters paralelos a los ejes, se debe modificar la matriz A_{is} por una matriz diagonal [9].

$$J(Z; U, V, A) = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m (\mathbf{z}_k - \mathbf{v}_i)^T \mathbf{A}_{id} (\mathbf{z}_k - \mathbf{v}_i) \quad (68)$$

De nuevo sujeta a las condiciones (60) y (61). Siendo A_{id} una matriz diagonal, la funcional descrita por la ecuación (68) se puede describir como:

$$J(Z; U, V, A) = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m [(\mathbf{x}_k - \mathbf{v}_{ix})^T \mathbf{A}_{ixd} (\mathbf{x}_k - \mathbf{v}_{ix}) + \sigma_i (y_k - \mathbf{v}_{iy})^2] \quad (69)$$

Donde x_k y y_k son un vector y una constante que contienen las componentes de entrada y la salida del k -ésimo dato respectivamente, v_{ix} y v_{iy} son un vector y una constante, con las componentes de entrada y la salida del centro del i -ésimo cluster, A_{ixd} es una matriz diagonal, donde cada elemento corresponde a cada una de las variables de entrada de la matriz A_{id} y σ_i corresponde a la componente relacionada con la variable de salida de la matriz A_{id} . Tomando como base la funcional dada en

la ecuación (70), donde se producen clusters de ejes paralelos respecto a la variables de entrada, se modifica el término relacionado con la variable de salida con el fin de producir submodelos lineales de la siguiente forma:

$$J(Z; U, V, A) = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m [(\mathbf{x}_k - \mathbf{v}_{ix})^T \mathbf{A}_{ixd} (\mathbf{x}_k - \mathbf{v}_{ix}) + \sigma_i (y_k - f(\mathbf{x}_k, \Theta_i))^2] \quad (70)$$

Donde $f(\mathbf{x}_k, \Theta_i) = a_i^T x_k + b_i$ representa cada uno de los i -ésimos submodelos lineales relacionados con el i -ésimo cluster. También se debe resaltar que al igual que en el algoritmo GK donde existe una restricción extra respecto al hipervolumen de cada cluster en este algoritmo se debe colocar la siguiente restricción.

$$|\mathbf{A}_{ixd}| \sigma_i = \eta_i \quad \eta_i > 0 \quad \forall_i \quad (71)$$

De la ecuación (71) se puede notar que el término relacionado con la variable de salida es prácticamente igual a la funcional que se propone para el algoritmo de FCRM [11]. Obviamente esta funcional también debe estar sujeta a la restricción . Para minimizar (70) sujeta a (60) y (71) se aplican multiplicadores de Lagrange. Combinado los desarrollos hechos en [10] y [11], se llega a las siguientes expresiones:

$$\mathbf{v}_{ix}^{(l)} = \frac{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m \mathbf{x}_k}{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m} \quad (72)$$

$$\Phi_i = \begin{bmatrix} (\mu_{i1}^{l-1})^m & 0 & \dots & 0 \\ 0 & (\mu_{i2}^{l-1})^m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\mu_{iN}^{l-1})^m \end{bmatrix} \quad (73)$$

$$\mathbf{X}_e = \begin{bmatrix} (x_1^T) & 1 \\ (x_2^T) & 1 \\ \vdots & \vdots \\ (x_N^T) & 1 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (74)$$

$$\Theta_i = (\mathbf{X}_e^T \Phi_i \mathbf{X}_e)^{-1} (\mathbf{X}_e^T \Phi_i \mathbf{y}) \quad (75)$$

$$f_{iy}^{(l)} = \frac{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m (\mathbf{y}_k - f(\mathbf{x}_k - \Theta_i))^2}{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m} \quad (76)$$

$$\mathbf{A}_{ixd}^{(l)} = \left[\eta_i f_{iy}^{(l)} |\mathbf{F}_{ixd}| \right]^{1/n} \mathbf{F}_{ixd}^{-1} \quad (77)$$

$$\sigma_i = \left[\eta_i f_{iy}^{(l)} |\mathbf{F}_{ixd}| \right]^{1/n} (f_{iy}^{(l)})^{-1} \quad (78)$$

$$(d_{ij}^2)^{(l)} = (\mathbf{x}_k - \mathbf{v}_{ix})^T \mathbf{A}_{ixd} (\mathbf{x}_k - \mathbf{v}_{ix}) + \sigma_i (y_k - f(\mathbf{x}_k, \Theta_i))^2 \quad (79)$$

$$\mu_{ik}^{(l)} = \frac{1}{\sum_{j=1}^c (d_{ik}^{(l)} / d_{jk}^{(l)})^{2/m-l}} \quad (80)$$

donde l indica el número de la iteración.

4.3.1. Formulación Robusta y de Aglomeración Competitiva

Teniendo en cuenta las técnicas desarrolladas en (RCA) y (FCRA) se propone la siguiente funcional [6] [7]:

$$J = (Z; U, V, A) = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m \rho_i d_{ij}^2 - \alpha \sum_{i=1}^c \left[\sum_{k=1}^N w_{ik} \mu_{ik} \right]^2 \quad (81)$$

Con d_{ij} descrita por la ecuación (79) y sujeta a las restricciones (60) y (61). En la ecuación (81) ρ_i es la función robusta de pérdida asociada al i -ésimo cluster y $w_{ik} = \partial \rho_i(d_{ij}^2) / \partial d_{ij}^2$ es la función de peso que representa la tipicidad del dato x_k respecto al i -ésimo cluster. El segundo término es la suma negativa de los cuadrados de las cardinalidades de cada cluster y es minimizada cuando la cardinalidad de alguno de los clusters es igual a N . La función de pérdida en el primer término reduce el efecto del ruido y la función de peso descarta los datos con ruido al calcular la cardinalidad en el segundo. Seleccionado α correctamente el funcional J puede ser utilizada para encontrar clusters compactos mientras que los datos son particionados en un mínimo número de clusters. En un comienzo α debe ser pequeño para permitir la formación de pequeños clusters. Después debe aumentar gradualmente para fomentar el proceso de aglomeración. Finalmente cuando se halla encontrado un número óptimo de clusters debe reducirse con el fin de detener el proceso de aglomeración para permitir que el algoritmo converja. Una apropiada elección de α

se encuentra en [7]

$$\alpha_k = \lambda(k) \frac{\sum_{i=1}^c \sum_{k=1}^N (\mu_{ik}^{(l-1)})^m \rho_i \mathbf{d}_{ij}^2}{\sum_{i=1}^c \left[\sum_{k=1}^N w_{ik} \mu_{ik} \right]^{m2}} \quad (82)$$

$$\lambda_k = \begin{cases} \lambda_0 e^{-|k_0-k|/\tau} & \text{si } k > 0 \\ \lambda_0 e^{-|k_0-k|/\tau} & \text{si } k = 0 \end{cases} \quad (83)$$

En este caso λ y α son funciones del número de iteración para cumplir con el comportamiento descrito anteriormente. Para minimizar la ecuación (81) con respecto a cada uno de los parámetros del cluster se mantiene constante μ_{ik} , y se llega a la siguiente ecuación

$$\sum_{k=1}^N (w_{ik} \mu_{ik}^{(m)}) \frac{\partial d_{ij}^2}{\partial \beta_i} \quad (84)$$

Donde β_i denota los parámetros de i -ésimo cluster. Como resultado de la ecuación (84) se llega a ecuaciones similares a las formuladas en (72)-(80) presentando como única diferencia el término , el cual es reemplazado por μ_{ik}^m [9].

4.4. Derivación de Modelos TSK a Partir de los Parámetros Obtenidos en los Métodos de Clustering

El problema de derivar los modelos TSK se reduce a hallar los parámetros de la función lineal en los consecuentes y las funciones de pertenencia en los antecedentes. Como resultado de los algoritmos de clustering se obtienen particiones difusas junto con los parámetros que identifican a cada cluster. Tomando estos parámetros como base se deben hallar los parámetros de las funciones en los consecuentes y las funciones de pertenencia de los antecedentes.

4.4.1. Derivación de Consecuentes

Para derivar los consecuentes los dos métodos usados más comúnmente son: mínimos cuadrados totales y mínimos cuadrados ponderados:

4.4.1.1. Mínimos cuadrados totales (TLS)

En los algoritmos GK y RoFeR se obtienen clusters de forma hiperelipsoidal. Los parámetros que describen cada cluster son la matriz de covarianza y la media difusa,

los cuales, indican la orientación y el punto medio de cada cluster respectivamente. A partir de la interpretación geométrica de los autovectores de la matriz de covarianza y de la media difusa se pueden formar ecuaciones de hiperplanos. Observe que la orientación y longitud de cada eje de los hiperelipses formados están determinadas por los autovectores y autovalores de la matriz de covarianza respectivamente. Las subregiones lineales se pueden aproximar por hiperplanos achatados en alguno de los ejes, de tal forma que el eje más pequeño del hiperelipse debe tender a ser perpendicular al hiperplano que se quiere aproximar. El autovalor más pequeño indica cuál de los ejes del hiperelipse es más pequeño y por tanto determina un posible vector perpendicular al hiperplano. Además, el hiperplano debe pasar por un punto determinado, en este caso la media de cada cluster. Esta aproximación se muestra en la figura 9.

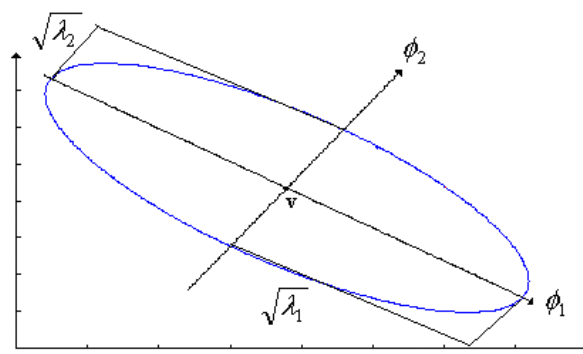


Figura 9: La estructura propia de la matriz de covarianza difusa provee información acerca de la forma y orientación del cluster

λ indica los valores propios y ϕ los vectores propios asociados). En la figura ϕ_2 representa el eje más pequeño del hiperelipse el tiene dirección normal a la línea que se quiere aproximar. Por otro lado ϕ_1 indica el eje más grande del hiperelipse representando la línea a aproximar. Finalmente \mathbf{v} es el centro del cluster. Basándose en esta interpretación geométrica se puede plantear la siguiente ecuación:

$$\mathbf{U} = \|\mathbf{U}^l - \mathbf{U}^{l-1}\| < \varepsilon \quad (85)$$

$$\phi_i^T (\mathbf{z} - \mathbf{v}_i) = 0 \quad (86)$$

Donde ϕ_i^T es el autovector correspondiente al autovalor más pequeño de la i -ésima matriz de covarianza y v_i es el i -ésimo centro. Ahora el autovector ϕ_i^T y el centro v_i se pueden describir descomponiéndolos en sus componentes de entrada y salida de la siguiente forma:

$$\phi_i^T = [(\phi_i^x)^T, \phi_i^y] \quad v_i^T = [(v_i^x)^T, v_i^y] \quad (87)$$

Donde ϕ_i^x y $(v_i^x)^T$ son las componentes correspondientes a las variables de entrada de los vectores y respectivamente. Mientras que ϕ_i^y y $(v_i^y)^T$ son las constantes que corresponden a la salida. Teniendo en cuenta esto la ecuación (1.60) se puede describir de la siguiente forma:

$$[(\phi_i^x)^T, \phi_i^y] \left([x^T, y] - [(v_i^x)^T, v_i^y]^T \right) = 0 \quad (88)$$

Mediante manipulación algebraica se obtiene:

$$y = \frac{-1}{\phi_i^y} (\phi_i^x)^T \mathbf{x} + \frac{1}{\phi_i^y} (\phi_i^y)^T \mathbf{v}_i \quad (89)$$

Donde

$$\mathbf{a}_i = \frac{-1}{\phi_i^y} (\phi_i^x)^T \quad (90)$$

$$b_i = \frac{1}{\phi_i^y} (\phi_i^y)^T \mathbf{v}_i \quad (91)$$

Siendo los consecuentes de la i -ésima regla. Como se dijo anteriormente esta solución esta basada en la interpretación geométrica de los cluster, pero también esta demostrado que los parámetros pueden ser obtenidos como una solución de TLS.

4.5. Derivación de Antecedentes

Es necesario establecer las funciones de pertenencia de manera tal que se permita el cálculo de los grados de pertenencia para cualquier dato de entrada, esto con el fin de obtener un modelo apropiado. Lo anterior es logrado al proyectar la

matriz de partición difusa U sobre las variables de entrada, obteniendo las funciones de pertenencia punto a punto de los datos tratados. El procedimiento es proyectar los conjuntos difusos multidimensionales establecidos en U sobre las variables individuales de la parte del antecedente. Es difícil la interpretación e implementación multidimensional de conjuntos difusos, por lo tanto los antecedentes en (51) deben ser reescritos como una combinación de proposiciones simples con conjuntos difusos unidimensionales. Para este fin se pueden aplicar dos estrategias: en una de ellas, las variables pueden ser las variables originales de regresión en cuyo caso la proyección es justamente una proyección ortogonal de los datos. O pueden obtenerse variables transformadas del antecedente por medio de la proyección del autovector usando los primeros p autovectores de las matrices de covarianza de cluster. Esta proyección del autovector es útil para clusters que son rotados respecto a los ejes del espacio de regresión, y no pueden ser representados con precisión por proyección ortogonal del eje. En ambos casos las funciones de pertenencia proyectadas (definidas punto a punto) son aproximadas por algunas funciones paramétricas convenientes

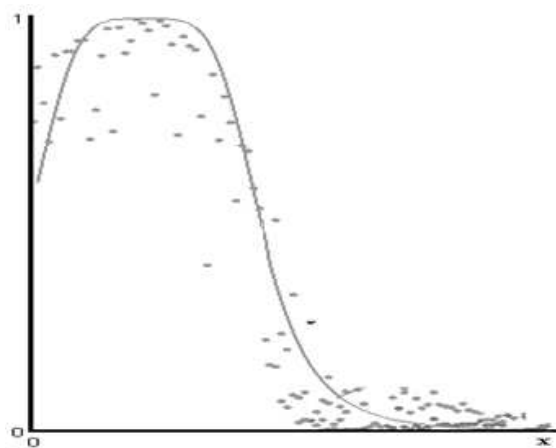


Figura 10: Aproximación mediante una ecuación paramétrica de datos proyectados

4.5.0.2. Proyección Ortogonal

Proyecta la función de pertenencia multivariable sobre cada uno de los ejes correspondientes a las diferentes variables de entrada de la siguiente forma:

$$\mu_{A_i}(x_j) = \text{proj}_j \mathbf{U}(i) \quad (92)$$

Donde $\mathbf{U}(i)$ denota la i -ésima fila de la matriz de partición difusa. Con la expresión anterior se obtiene las reglas TSK de la forma conjuntiva presentada en (51) y (52). En este tipo de proyección se puede perder información como resultado de un error de descomposición. Es decir, los cluster pueden presentar rotación del espacio de tal manera que no queden paralelos a los ejes de las variables. O en otros casos, gracias a la naturaleza de la distribución de datos, los cluster puedan tomar formas que no puedan ser generadas por una intersección de conjuntos unidimensionales. En la figura 11 se muestra el caso donde un cluster de forma elipsoidal y paralelo a los ejes se ha obtenido y como después de proyectado y reconstruido se obtiene un cluster de forma circular paralelo a los ejes, de esta forma se ilustra el error por descomposición.

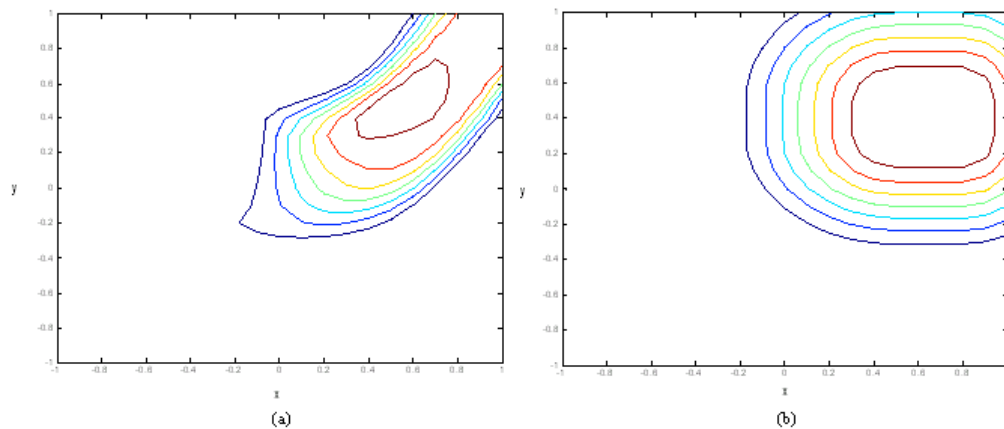


Figura 11: Error por descomposición. (a) Cluster obtenido (b) Cluster reconstruido después de la proyección

4.5.0.3. Proyección del Autovector

Como se dijo anteriormente, los autovectores contienen información acerca de orientación de los clusters en el espacio de regresión, es decir la dirección de los ejes rotados de las hiperelipses, por tanto se puede utilizar esta propiedad para formar una matriz de rotación y luego proyectar ortogonalmente. La matriz de regresión \mathbf{X} es proyectada de la siguiente manera:

$$\mathbf{X}'_i = \mathbf{X}\mathbf{H}_i \quad (93)$$

donde \mathbf{H}_i es la matriz de proyección conteniendo en sus columnas el autovector unitario de la i -ésima matriz de covarianza. Es decir, cada cluster tiene su propia matriz de proyección. Entonces, las nuevas entradas se convierten en combinaciones lineales de las entradas originales (por el proceso de rotación de ejes a través de la matriz). Aunque este tipo de proyección evita el error por descomposición presente en la proyección ortogonal, exhibe problemas de interpretación lingüística de los antecedentes de la regla, debido a que la transformación de las variables puede no ser relevante a la dinámica del sistema.

4.6. Modelos de Entrada - Salida

La representación entrada-salida es aplicada frecuentemente para construir modelos difusos a partir de datos generados por sistemas dinámicos entendidos pobremente. En lugar de usar un vector de estado relevante desde el punto de vista físico, el estado del sistema es representado por un número finito de entradas y salidas anteriores del sistema. La estructura más común es el modelo NARX (Non-linear AutoRegressive with eXogenous input), la cual puede representar los modos observables y controlables de una gran variedad de sistemas no lineales de tiempo discreto. El modelo NARX establece una relación entre la recopilación de datos pasados entrada-salida y las salidas predichas.

$$\hat{y}(k+1) = F(y(k), \dots, y(k-n_y+1), u(k), \dots, u(k-n_u+1)) \quad (94)$$

donde k denota las muestras en tiempo discreto, n_y y n_u son enteros relacionados con el orden del sistema y F denota un modelo difuso. En el modelo NARX, el vector

de regresión es una colección de un finito número de entradas y salidas pasadas:

$$\mathbf{x}(k) = [(y(k), \dots, y(k - n_y + 1), u(k - 1), \dots, u(k - n_u + 1))]^T \quad (95)$$

El regresor es la salida predicha $\hat{y}(k + 1)$. En términos de reglas, el modelo es dado por:

$$\begin{aligned} \text{Si } y(k) \text{ es } A_{i,2} \text{ y } \dots \text{ y } y(k - n_y + 1) \text{ es } A_{i,n_y} \text{ y} \\ u(k) \text{ es } B_{i,1} \text{ y } u(k - 1) \text{ es } B_{i,2} \text{ y } \dots \text{ y } u(k - n_u + 1) \text{ es } B_{i,n_u} \\ \text{Entonces } y(k + 1) = \sum_{j=1}^n a_{ij} y(k - j + 1) + \sum_{j=1}^m b_{ij} u(k - j + 1) + c_i \end{aligned} \quad (96)$$

Otros modelos comunes entrada-salida son los de error de salida no lineal (NOE, Nonlinear Output Error), que incluyen las salidas pasadas del modelo (predicciones) en lugar de las salidas medidas del proceso y las llamadas formas innovadoras (NARMAX), que incluyen los valores pasados de la predicción de error en el vector de regresión. Estas estructuras no son usadas para identificación basada en clustering, ya que el vector de regresión no puede ser construido directamente a partir de datos y el problema de regresión correspondiente debe ser resuelto de manera recurrente [2] [20].

Capítulo V

MODELAMIENTO DIFUSO DE PLANTAS NO LINEALES EN LAZO CERRADO

El desarrollo de modelos de sistemas reales es una importante temática en muchos campos de la ingeniería. Dichos modelos pueden ser usados para simulación, análisis del comportamiento del sistema, mejor entendimiento de los mecanismos principales, diseño de nuevos procesos y control del sistema. El desarrollo de un modelo que represente adecuadamente la realidad del problema es una tarea importante. Si el modelo no es suficientemente exacto los pasos subsecuentes de análisis, predicción, control, entre otros, pueden no ser exitosos. El modelamiento y la identificación son importantes pasos en el diseño de control, supervisión y detección de fallas de sistemas. Muchos sistemas no son tratables mediante aproximaciones de modelamiento en lazo abierto convencional debido a la dificultad de aislar la planta del controlador o la inestabilidad del sistema una vez se haya eliminado la realimentación.

La aproximación en lazo cerrado, considerada como difícil, ha sido el objeto de un creciente interés en los últimos años con el desarrollo de nuevos algoritmos y una mejor comprensión del tema. De igual manera las técnicas de identificación y modelamiento difuso han tenido desarrollos en formulaciones menos empíricas y más formales de sistemas basados en datos de entrada y salida.

Tiempo atrás el control difuso había sido implementado a partir de modelos empíricos; con el desarrollo de las técnicas de clustering difuso, se han encontrado justificaciones teóricas y matemáticas en el planteamiento de modelos difusos. En la actualidad no se encuentran suficientes registros donde los investigadores hayan

adaptado satisfactoriamente la identificación en lazo cerrado a los algoritmos de clustering. Esto impide que muchas aplicaciones puedan utilizar clustering difuso como un método de modelamiento, debido a comportamientos inestables en lazo abierto, comportamientos integradores o que la planta no pueda ser aislada del resto del sistema. Así mismo, es necesario realizar un análisis que indique si la implementación en lazo cerrado permitirá la manipulación de sistemas que no son tratables mediante aproximaciones de modelamiento convencional debido a la falta de precisión, conocimiento formal acerca del sistema, fuerte comportamiento no lineal, alto grado de incertidumbre, características de variación en el tiempo; características que para algunas aproximaciones el modelamiento difuso ha presentado resultados aceptables frente a las técnicas clásicas.

Es por eso que en este trabajo se realiza el planteamiento de una metodología para el tratamiento de sistemas en lazo cerrado a partir de la información obtenida por técnicas de clustering y su posterior enfoque en modelos TSK, evaluando su desempeño, restricciones presentes en el desarrollo y en la posible solución al tipo de identificación indirecta.

A continuación se describe la metodología planteada y los resultados obtenidos:

5.1. Metodología Planteada

La figura 12 muestra la representación básica de un sistema en lazo cerrado. La caja punteada indica la representación de un sistema de *caja negra*.

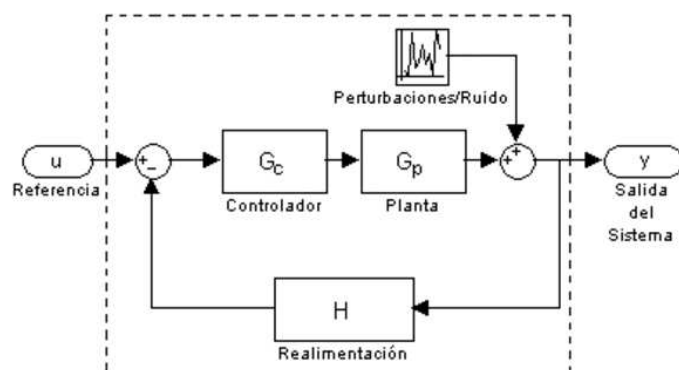


Figura 12: Representación clásica de un sistema de lazo cerrado

Con éstas dos representaciones y basados en el tipo de aproximación de la sección 2.1.2, en este trabajo se intenta encontrar una aproximación en lazo abierto a partir de la aproximación indirecta en lazo cerrado, de lo cual se plantea la metodología de la figura 13.

Para está identificación se utilizan las técnicas implementadas por Gustafson y Kessel en [10] conocida como GK y la técnica desarrollada por González y Barato en [9] la cual se referenciará en este trabajo como RoFer, técnicas basadas en datos de entrada-salida. Está identificación se lleva a cabo con el toolbox FMID desarrollado por Robert Babuska [5] para uso con Matlab para GK y la implementación realizada por Victor Grisales [9] en dicho toolbox para RoFer. A diferencia de GK, RoFer permite la obtención automática de un número adecuado de clusters. El índice de desempeño utilizado para determinar el grado de validez de la aproximación es el VAF (Variance Accounted For) el cual entrega en porcentaje la medida de la varianza entre dos señales.

Como se describe en las ecuaciones (51) y (52) los consecuentes de los modelos TSK son representados por funciones lineales y con estructuras como las de la ecuación (94), lo que permite que un sistema multientrada sea expresado como en la ecuación (96).

Ahora bien una vez realizado el clustering y obtenido el modelo difuso los consecuentes se encontrarán expresados de la forma:

$$y_{i,j}(k) = \sum_{j=1}^{ny} a_{ij}y(k-j) + \sum_{j=1}^{nu} b_{ij}u(k-j) + c_i \quad (97)$$

donde i indica la i -ésima regla y estará ligada al número de clusters que representen de manera adecuada la dinámica del sistema, es decir, $i = 1, 2, \dots, C$, j representa el número de entradas y salidas del sistema (multivariable). La ecuación (97) puede ser representada en forma de *función de transferencia* de un sistema discreto de la forma general:

$$\frac{y}{u} = \frac{b_1z^{-1} + b_2z^{-2} + \dots + b_{nu}z^{-nu}}{1 - a_1z^{-1} + a_2z^{-2} + \dots + a_{ny}z^{-ny}} \quad (98)$$

y teniendo en cuenta la ecuación (97) para un sistema discreto con la estructura

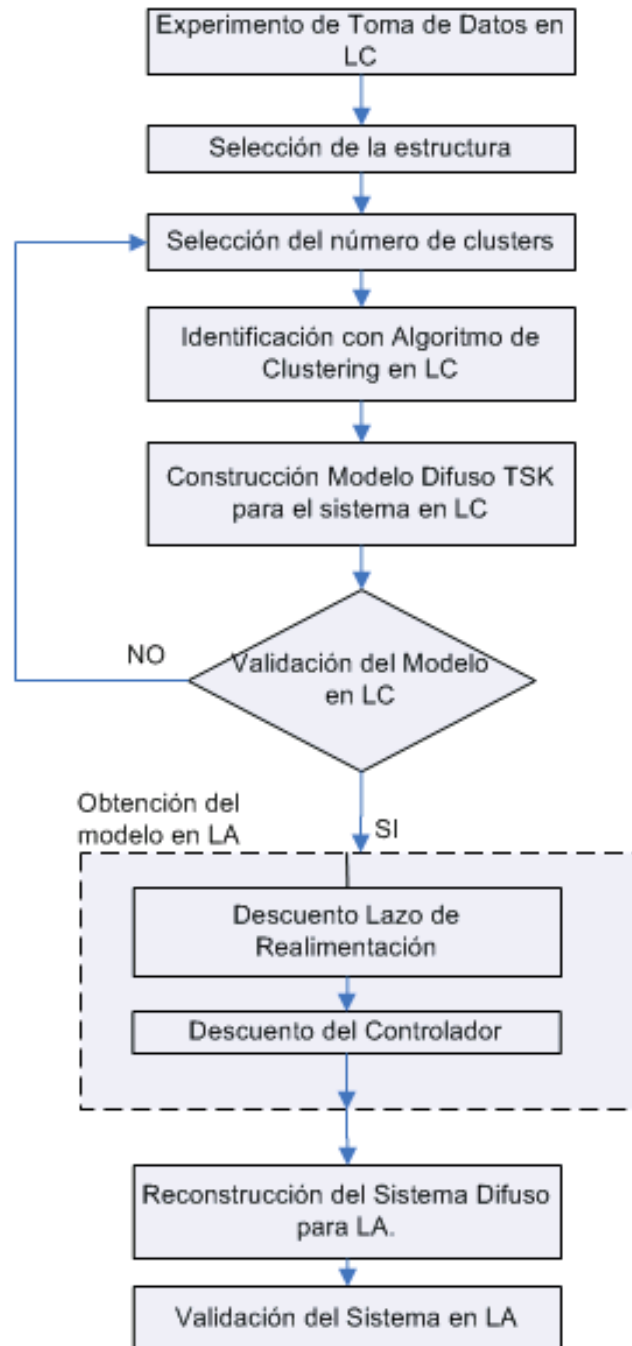


Figura 13: Diagrama que representa la metodología propuesta

del sistema de la figura 12 se tiene que:

$$G_{lc}(z^{-1}) = \frac{G_{ld}(z^{-1})}{1 + G_{ld}(z^{-1})H(z^{-1})} \quad (99)$$

donde

$$G_{ld}(z^{-1}) = G_c(z^{-1})G_p(z^{-1}) \quad (100)$$

es la ganancia del lazo directo. Si se asume la estructura general donde $H(z^{-1}) = 1$, indicando una realimentación unitaria y dado que para el tipo de aproximación indirecta es necesario el conocimiento del controlador (obteniendo su función de transferencia en el dominio discreto) se plantea la obtención de los consecuentes del modelo difuso de la planta en lazo abierto despejando G_p de las ecuaciones (99) y (100).

$$G_{ld}(z^{-1}) = \frac{G_{lc}(z^{-1})}{[1 - G_{lc}(z^{-1})]} \quad (101)$$

Para descontar el controlador, se aprovecha que éste es lineal para establecer un modelo difuso TSK que comparta el universo de entrada (Antecedentes) y las reglas con el modelo de lazo directo, para operar posteriormente los submodelos lineales correspondientes. Ya que el controlador es lineal, es posible obtener un modelo difuso adecuado para cualquier subdivisión del universo de entrada, es decir, para cualquier conjunto de antecedentes y de reglas.

Para la supresión del controlador se aprovecharía el hecho de que el controlador fuese lineal. Al ser un sistema lineal es posible establecer un modelo difuso TSK que lo represente adecuadamente para cualquier universo de entrada y conjunto de reglas. La idea es que una vez se suprima el lazo de realimentación del modelo, se obtenga un modelo difuso TSK del controlador que tenga el mismo universo de entrada (antecedentes) y las mismas reglas que dicho modelo de lazo directo. Para posteriormente descontar los submodelos lineales correspondientes.

5.2. Aplicación

El contexto de aplicación es el siguiente: Se pretende realizar la identificación de una planta no lineal que cuenta con un controlador lineal. Sin embargo no es posible realizar pruebas de lazo abierto, es decir, desconectando el controlador, por

restricciones físicas (resulta un procedimiento muy complicado) o de estabilidad. Para emular esta situación se tomó el sistema MIMO usado por Babuska en [3] [5] como ejemplo para identificación en lazo abierto y se le diseñó un controlador clásico PID. Dicho sistema está conformado por cuatro tanques interconectados.

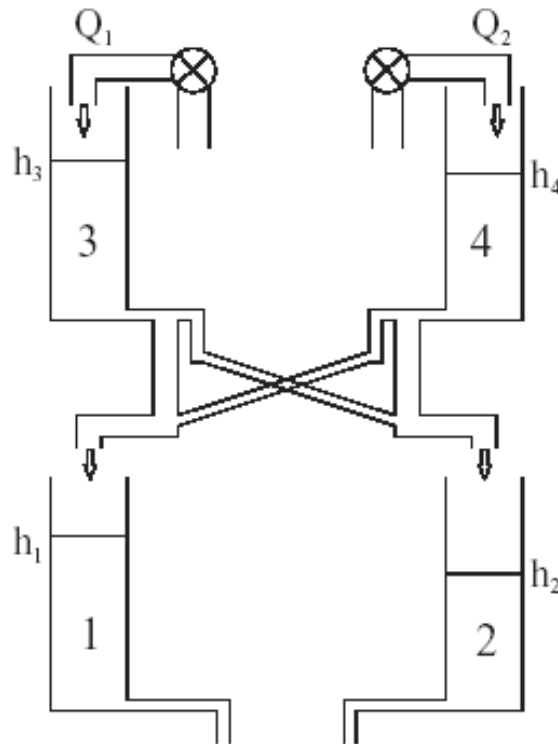


Figura 14: Cuatro Tanques en Cascada

para el cual las entradas son dos tasas de flujo $u = [Q_1, Q_2]^T$ y las salidas son los cuatro niveles de los tanques $y = [h_1, h_2, h_1, h_2]^T$. Este sistema está discreto por la ecuación diferencial no lineal:

$$\dot{h} = A\sqrt{h} + BQ \quad (102)$$

donde,

$$A = \begin{bmatrix} -\frac{S_{2,1}}{S_{1,1}}\sqrt{2g} & 0 & r_{3,1}\frac{S_{2,3}}{S_{1,1}}\sqrt{2g} & r_{4,1}\frac{S_{2,4}}{S_{1,1}}\sqrt{2g} \\ 0 & -\frac{S_{2,2}}{S_{1,2}}\sqrt{2g} & r_{3,2}\frac{S_{2,3}}{S_{1,2}}\sqrt{2g} & r_{4,2}\frac{S_{2,4}}{S_{1,2}}\sqrt{2g} \\ 0 & 0 & -\frac{S_{2,3}}{S_{1,3}}\sqrt{2g} & 0 \\ 0 & 0 & 0 & -\frac{S_{2,4}}{S_{1,4}}\sqrt{2g} \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{S_{1,2}} & 0 \\ 0 & \frac{1}{S_{1,4}} \end{bmatrix},$$

donde $r_{3,1} + r_{3,2} = r_{4,1} + r_{4,2} = 1$, el subíndice j denota el tanque j -ésimo, h es el nivel del líquido en el tanque, S_1 es el área del tanque ($10^{-3}m^2$), S_2 es el área de la salida del tanque ($10^{-6}m^2$), $r_{i,j}$ es la restricción del tanque i al tanque j ($r_{3,1} = r_{4,2} = 0,8$) y g es la gravedad.

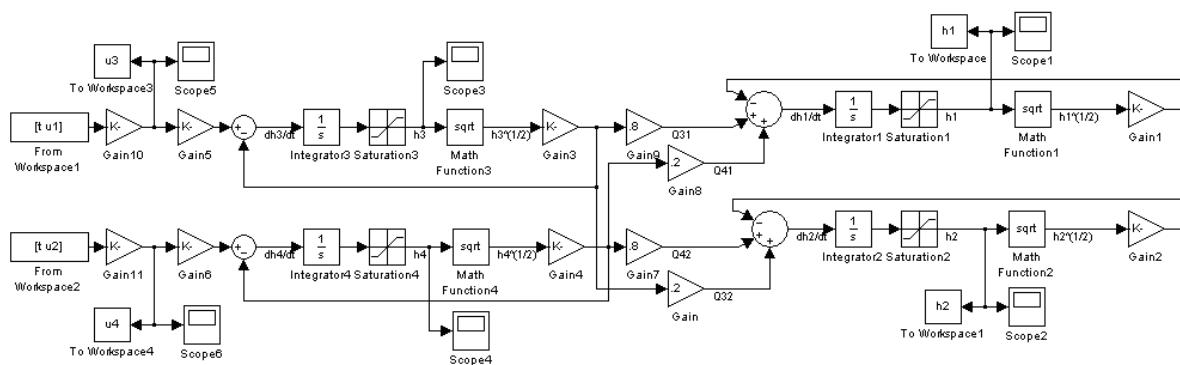


Figura 15: Sistema de Cuatro Tanques Interconectados

La restricción que se tiene sobre el controlador es que debe ser lineal. Como la planta es no lineal, dicho controlador no puede funcionar adecuadamente para cualquier estado del sistema. Por esto, para el diseño del controlador se debió escoger un punto de operación, en el cual, se realizó la linealización de la planta. El controlador obtenido funciona adecuadamente en un rango cercano a dicho punto de operación. Además, como las restricciones del problema no implican un comportamiento óptimo del sistema del control, se aprovecho que el sistema MIMO se podía aproximar mediante un par de sistemas SISO simétricos para simplificar el problema de diseño del controlador. Por lo tanto se reducía el sistema a 2 entradas y 2 salidas, estableciendo el acople de los tanques como perturbaciones del sistema.

Se tomó uno de estos sistemas SISO, se linealizó al rededor de un punto de operación y sobre este sistema se sintonizó un controlador PID usando el segundo

método de Ziegler-Nicholson [18]. Se estableció una acción de control proporcional y K_p se incrementó desde 0 hasta el valor donde la salida se aproximó a presentar oscilaciones sostenidas. En este punto se registraron la ganancia que producía este efecto (K_{cr}) y el periodo de las oscilaciones (P_{cr}). Entonces los valores de K_p , T_i y T_d del controlador se obtuvieron como $0,6K_{cr}$, $0,5P_{cr}$ y $0,125P_{cr}$ respectivamente. El controlador obtenido de esta forma, presentó una respuesta a una entrada escalón aceptable alrededor del punto de operación, como se evidencia en la figura que representa un lazo sencillo de dos tanques 16.

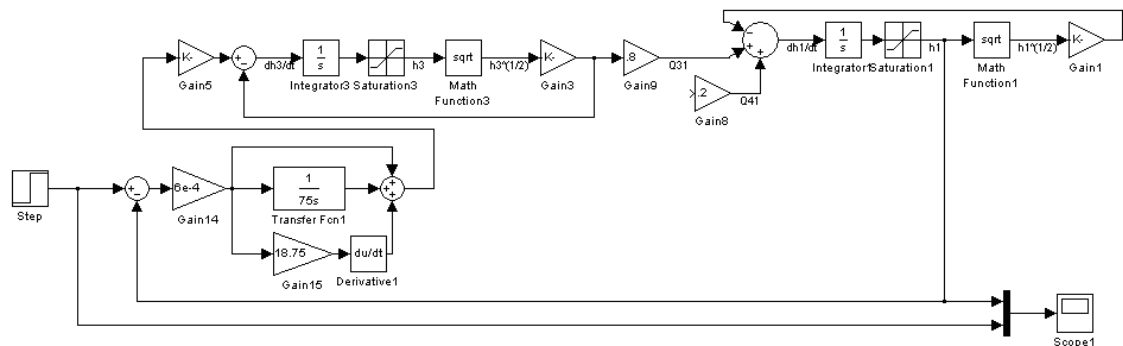


Figura 16: Sistema de Dos Tanques - Controlador PID

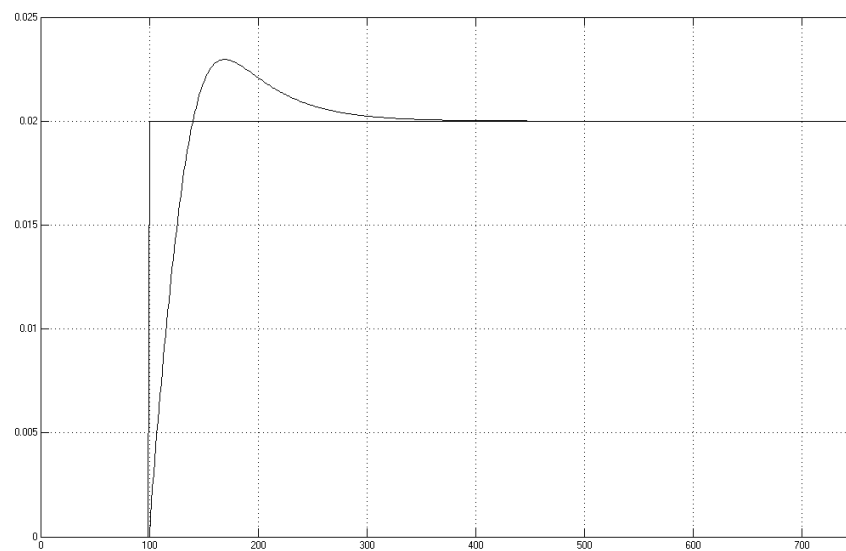


Figura 17: Respuesta a una entrada escalón para diseño del controlador PID

El sistema con cuatro tanques, dos entradas, dos salidas junto con el par de controladores PID, sería identificado como caja negra, presentado los siguientes resultados:

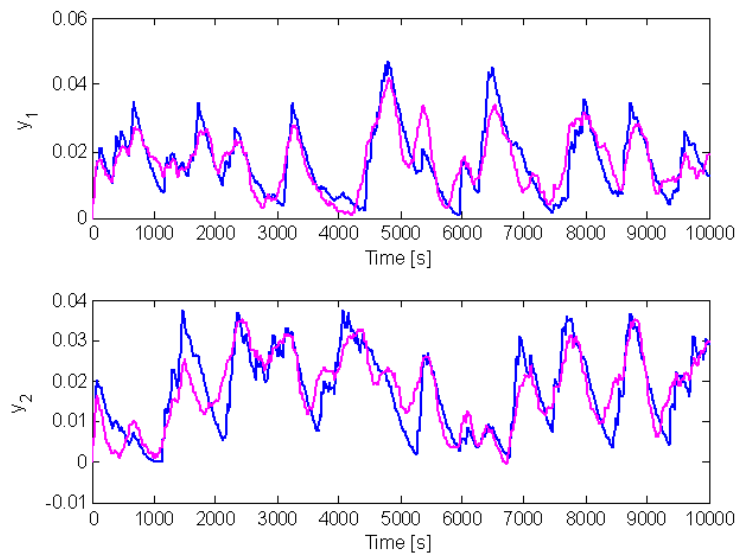


Figura 18: Identificación del sistema de 4 tanques interconectados, con el algoritmo de Gustafson-Kessel. Toolbox FMID. La señal azul indica la salida real del sistema y la señal discontinua roja la aproximación GK.

Arrojando un VAF (Variance Accounted For) de 72,5348 % y 74,4026 %, para las salidas y_1 y y_2 respectivamente.

Considerando que la entrada u es de orden 4, la salida y es de orden 2; el número de clusters que brindo el mejor resultado fue 2. Un tiempo de muestreo igual al del controlador PID de 10s.

1. **If** $y_1(k-1)$ is A_{11} **and** $y_1(k-2)$ is A_{12} **and** $y_2(k-1)$ is A_{13} **and** $y_2(k-2)$ is A_{14} **and** $u_1(k-1)$ is A_{15} **and** $u_1(k-2)$ is A_{16} **and** $u_1(k-3)$ is A_{17} **and** $u_1(k-4)$ is A_{18} **and** $u_2(k-1)$ is A_{19} **and** $u_2(k-2)$ is A_{110} **and** $u_2(k-3)$ is A_{111} **and** $u_2(k-4)$ is A_{112} **then**

$$y_1(k) = y_1(k-2) + y_2(k-1) + 1.27 \cdot 10^{-1} y_2(k-2) + 1.22 \cdot 10^{-1} u_1(k-1) - 9.32 \cdot 10^{-3} u_1(k-2) - 4.22 \cdot 10^{-2} u_1(k-3) - 3.54 \cdot 10^{-2} u_1(k-4) + 2.77 \cdot 10^{-2} u_2(k-1) + 3.92 \cdot 10^{-2} u_2(k-2) - 5.45 \cdot 10^{-2} u_2(k-3) - 5.43 \cdot 10^{-3} u_2(k-4) - 1.20 \cdot 10^{-4}$$

Figura 19: Ejemplo de las reglas que arroja el modelo difuso obtenido para el sistema de 4 tanques interconectados con el Toolbox de R. Babuska

Para el sistema de dos tanques con controlador PID, como se muestra en la figura 16, se obtuvieron los siguientes resultados en la identificación: un VAF de 98.3815 % teniendo en cuenta que a la entrada y la salida les fue asignado un orden de 4. Se corrió el algoritmo con 3 clusters, con los que se obtuvo mejor respuesta de la aproximación.

Los resultados de simulación son evidenciados en las gráficas 20, 21:

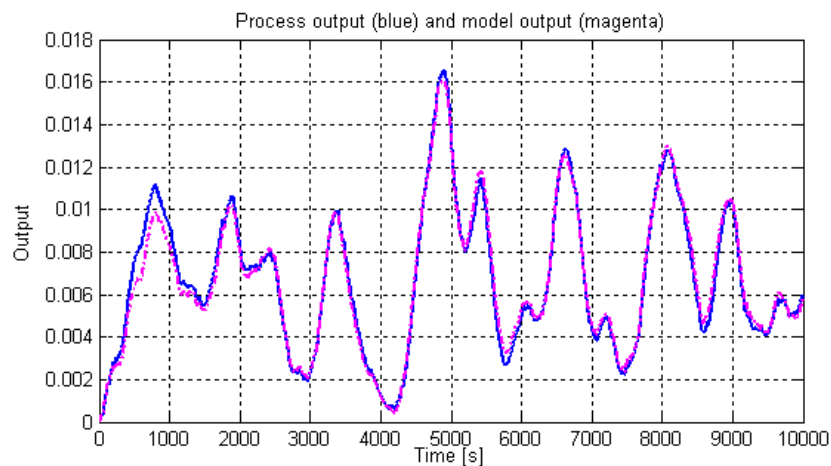


Figura 20: Salida del Modelo Difuso obtenido para el sistema de dos tanques con controlador

PID. En azul salida real del proceso y en magenta aproximación del modelo.

Pretendiendo realizar el descuento del controlador como se explica en la sección 5.1 en las ecuaciones (99) y (101) se obtuvieron resultados no acordes con las señales esperadas. Por lo tanto, se realizaron dos pruebas para observar si los resultados obtenidos reflejaban restricciones inherentes al sistema y a los métodos utilizados para su identificación y modelamiento.

5.2.1. Análisis de la Realimentación del Sistema Difuso

El procedimiento para obtener el modelo de una planta a partir del modelo difuso TSK de un sistema de control que incluye dicha planta, conociendo el controlador, puede dividirse en dos etapas, supresión del lazo de realimentación y eliminación del controlador.

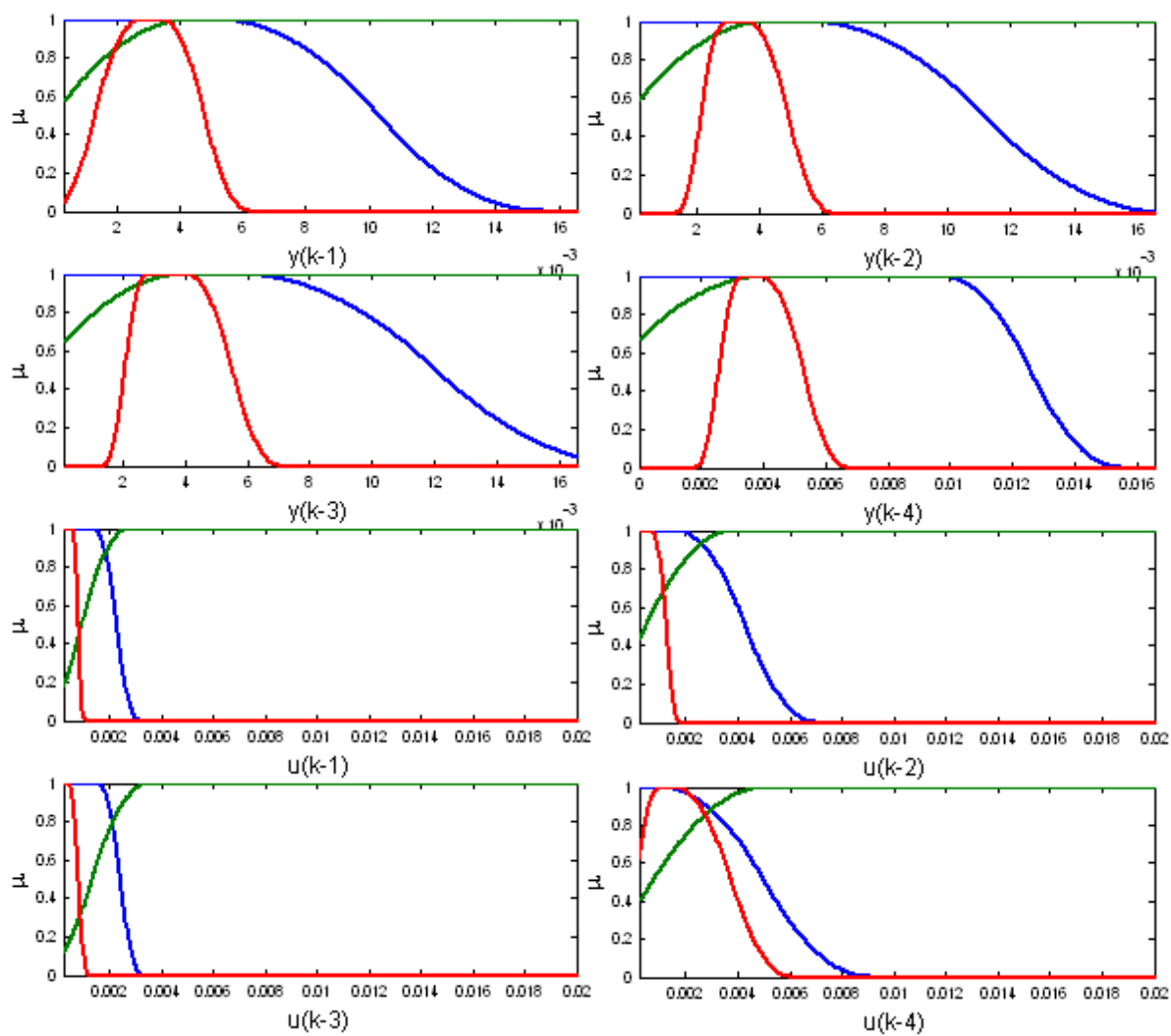


Figura 21: Funciones de pertenencia de los antecedentes del modelo obtenido.

Se tomó el caso de un bloque de dos tanques sin controlador, como el caso más sencillo para realizar la prueba de descuento del lazo de realimentación. La figura 22 indica la implementación de dos tanques con un lazo de realimentación negativa. La figura 23 muestra de antemano cual es la respuesta del bloque de tanques sin ser aplicada dicha realimentación. Se observa como para una entrada de amplitud pequeña, el sistema se torna inestable.

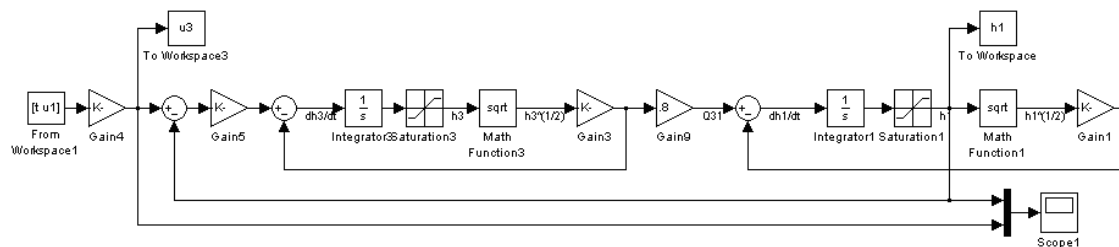


Figura 22: Diagrama de bloques para el sistema realimentado

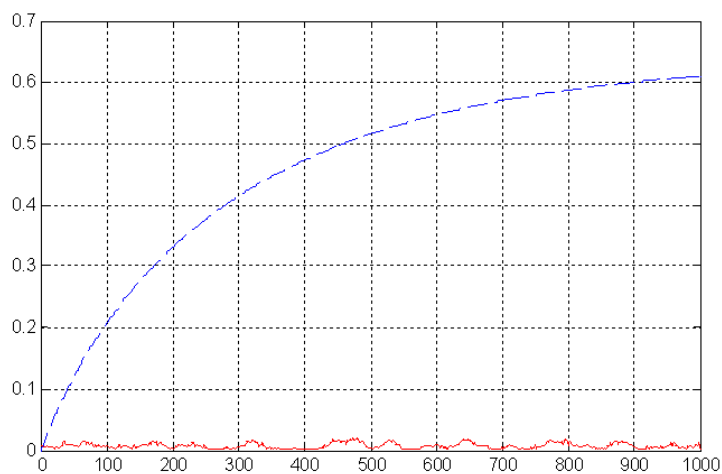


Figura 23: Respuesta en lazo abierto; Entrada (Rojo Señal Continua) y Salida (Azul Señal Discontinua)

En la figura 24 se gráfica la señal de entrada aplicada a par de tanques realimentados y la señal de salida obtenida. Con esta pareja de datos de entrada-salida se realiza la identificación de la planta mediante las técnicas de clustering para luego

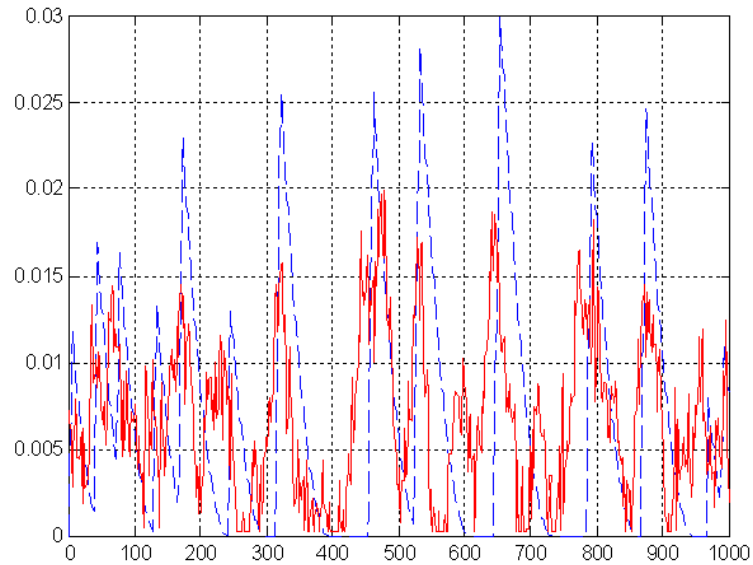


Figura 24: Respuesta en lazo cerrado; Entrada (Rojo Señal Continua) y Salida (Azul Señal Discontinua)

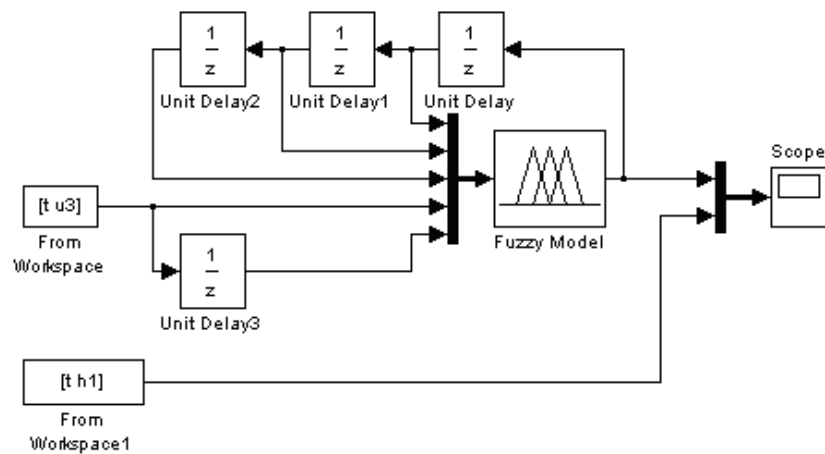


Figura 25: Diagrama de bloques para la aproximación difusa con lazo de realimentación

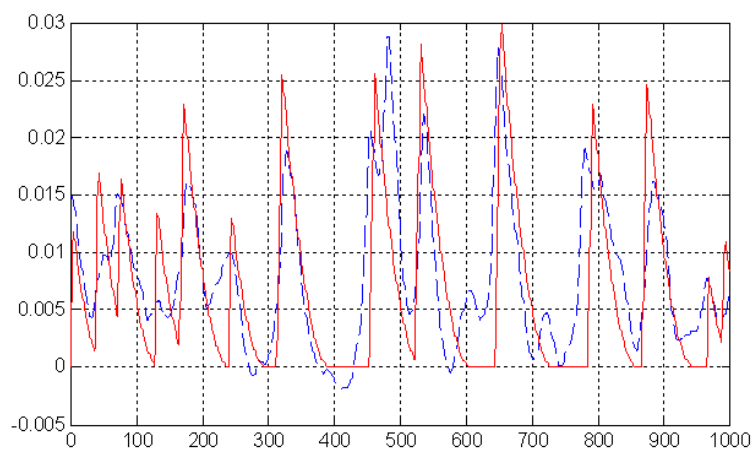


Figura 26: Comparación de la salida del sistema real realimentado (Rojo Continua) y Salida del modelo difuso obtenido en lazo cerrado (Azul Discontinua).

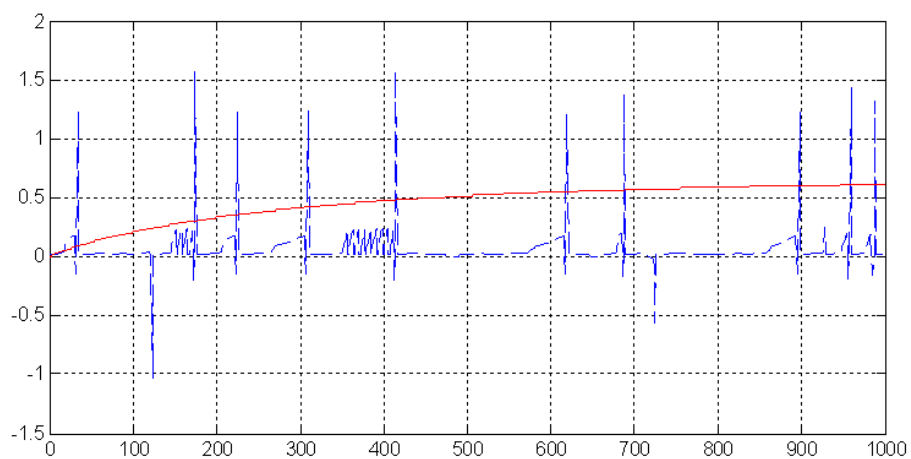


Figura 27: Salida del Sistema (Rojo) y Salida del Modelo Difuso (Azul), en lazo abierto.

obtener una representación TSK, es decir, un modelo difuso, figura 25. La aproximación obtenida con el modelo difuso es comparada la respuesta del sistema real en lazo cerrado, en la figura 26. Allí se observa que el modelo TSK entrega una respuesta cercana a la deseada. Por lo tanto se acepta el modelo obtenido. Los consecuentes que se encuentran expresados como ecuaciones de diferencia se expresan como funciones de transferencia para poderlos manipular aplicando la ecuación (101), eliminando el término afine.

Una vez realizado el descuento de la realimentación, la función de transferencia G_{ld} es reescrita como ecuaciones de diferencia para ser reemplazadas en el modelo obtenido como los consecuentes del sistema en lazo abierto. El término afine es retomado, con el mismo valor del encontrado en lazo cerrado. La figura 25 muestra la implementación del sistema difuso como un bloque FIS de simulink en Matlab® y la comparación con la respuesta deseada.

La figura 27 gráfica dicha comparación y evidencia como la aproximación del sistema difuso no representa la verdadera dinámica de la planta en lazo abierto

5.2.2. Aplicación de Realimentación Positiva

El método más directo para eliminar un lazo de realimentación negativa, a nivel computacional, es un lazo equivalente de realimentación positiva puesto en paralelo al primero, es decir, exactamente entre los mismos extremos y en la misma dirección, ya que de esta forma el lazo positivo sustrae lo que el lazo negativo adiciona y viceversa. Esto se verifica tanto para sistemas lineales como no lineales.

Para evidenciar esto observe la gráfica 28, allí se indica la respuesta de la planta en lazo abierto. A continuación, en la figura 29 se muestra la conexión de los dos lazos de realimentación en un sistema de implementación clásica. La siguiente figura 30 entrega el mismo resultado que el obtenido en lazo abierto de la figura 29.

Luego el modelo difuso TSK del sistema realimentado es obtenido, en la figura 32 se evidencia la buena aproximación que entrega el modelo. La figura 31 indica como es aplicado un lazo de realimentación positiva, imitando el tratamiento hecho para el sistema de implementación clásica. La respuesta obtenida evidencia que el comportamiento de dicho modelo no corresponde con el comportamiento del sistema

sin el lazo de realimentación, figura 33.

Entonces para realizar una identificación adecuada del sistema de lazo cerrado fue necesario usar una señal de amplitud grande, para obtener variaciones considerables a la salida ya que el sistema realimentado presenta cierta estabilidad. Dicha señal no se puede utilizar para caracterizar el sistema en lazo abierto ya que este es menos estable y con una señal de mayor amplitud este se satura.

Para caracterizar el sistema en lazo abierto fue necesario usar una señal pequeña, para evitar la saturación del sistema y hacer que la salida represente su dinámica. Dicha señal no se puede utilizar para realizar una identificación adecuada del sistema en lazo cerrado ya que este es más estable y los cambios en la salida son mínimos con una señal de entrada pequeña.

La realimentación positiva no funciona en el modelo difuso TSK, ya que si bien este funciona en la escala de la señal con que fue identificado, no representa adecuadamente la dinámica del sistema en una escala diferente.

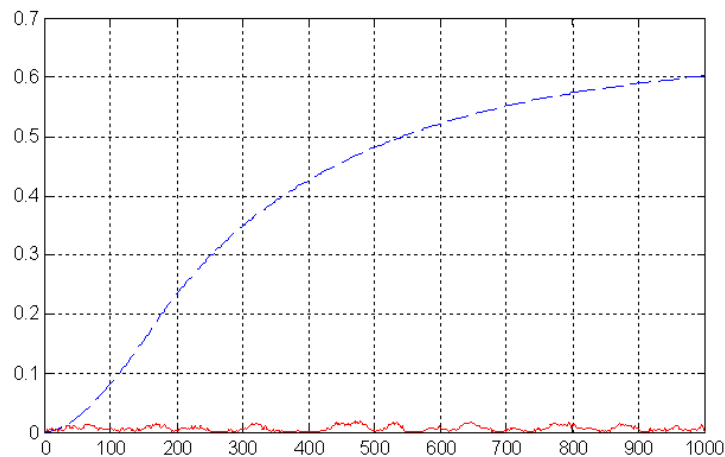


Figura 28: Entrada (Rojo) y Salida (Azul Discontinua) del sistema de lazo directo

5.2.3. Sensibilidad del Modelo a la Escala

Se realizaron pruebas para ver la respuesta del sistema cuando la señal de entrada no se encuentra en el rango de identificación. Al escalar la entrada del sistema y del modelo por un factor de 10, en la salida se obtiene respectivamente:

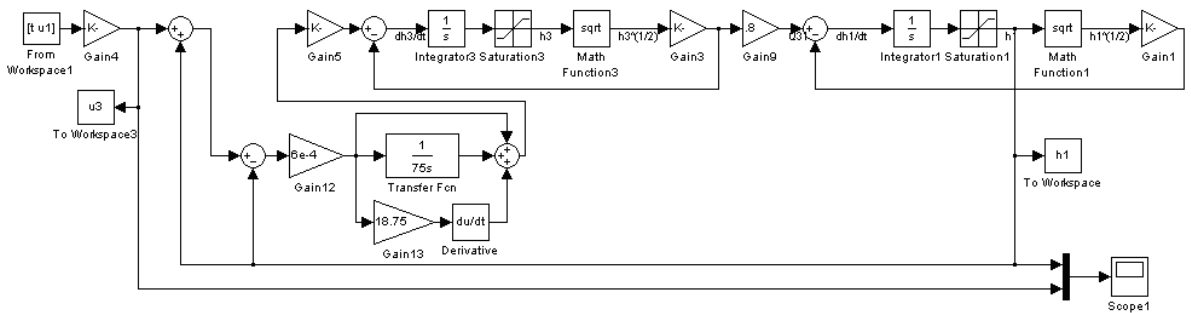


Figura 29: Diagrama de bloques con realimentación positiva

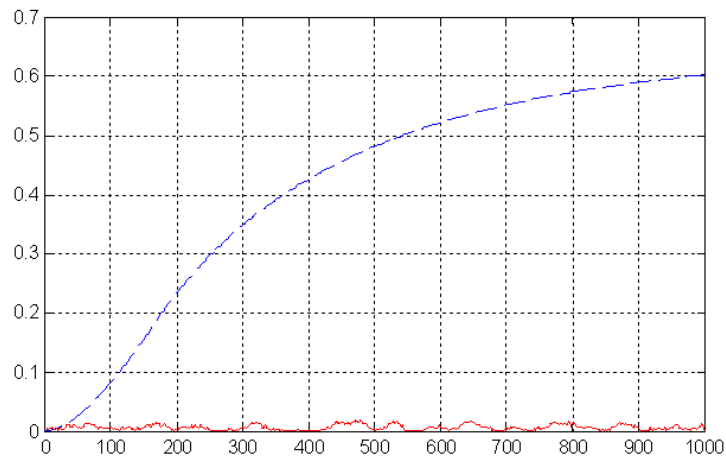


Figura 30: Entrada y Salida del sistema con los dos lazos de realimentación

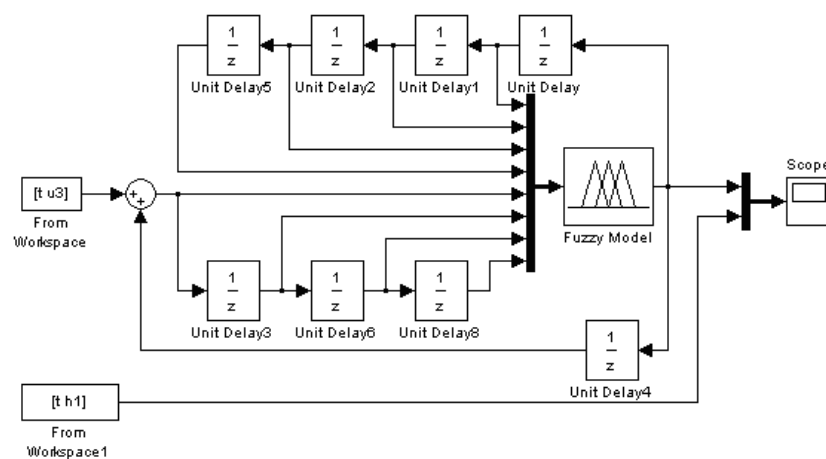


Figura 31: Diagrama de bloques para el sistema Difuso con realimentación positiva

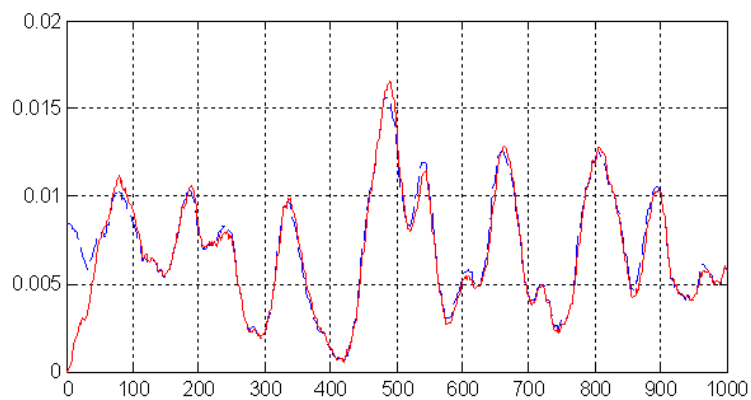


Figura 32: Salida del sistema(Azul) y Salida del modelo difuso (Rojo) en lazo directo

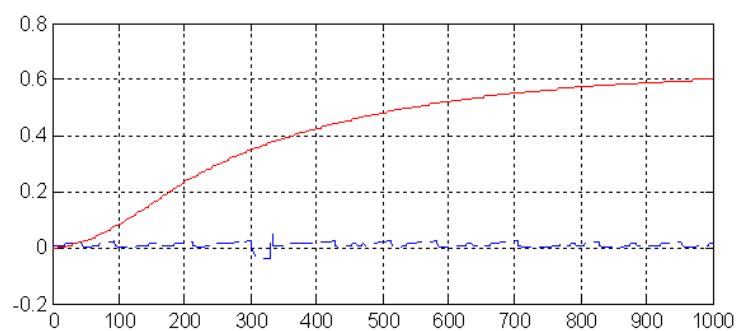


Figura 33: Salida del sistema(Azul) y Salida del modelo difuso (Rojo) con los dos lazos de realimentación

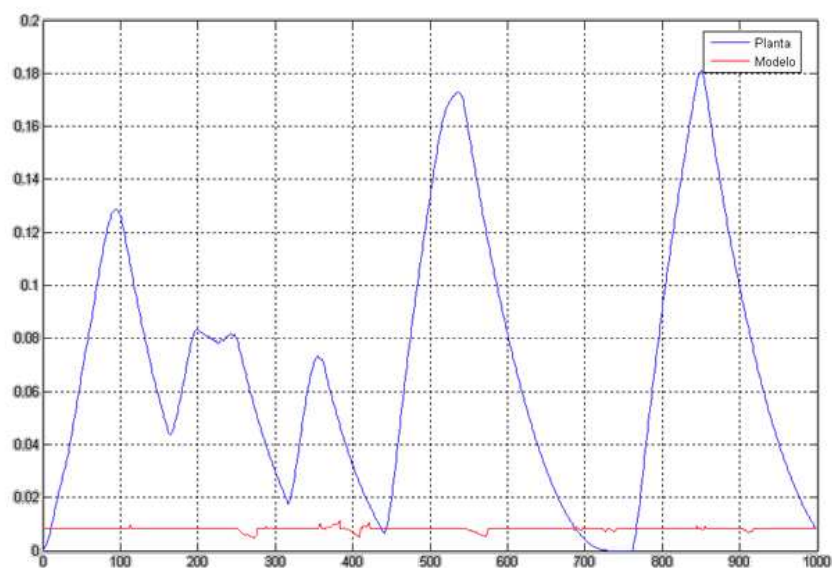


Figura 34: Salida del Modelo (Rojo) y Salida de la Planta (Azul)

Al escalar la entrada del sistema y del modelo por un factor de 0.1, en la salida se obtiene la gráfica de la figura 35, para esta aproximación el VAF obtenido es de -10.9504 %.

Y al escalar la entrada del sistema y del modelo por un factor de 0.01, en la salida se obtiene la gráfica de la figura 36:

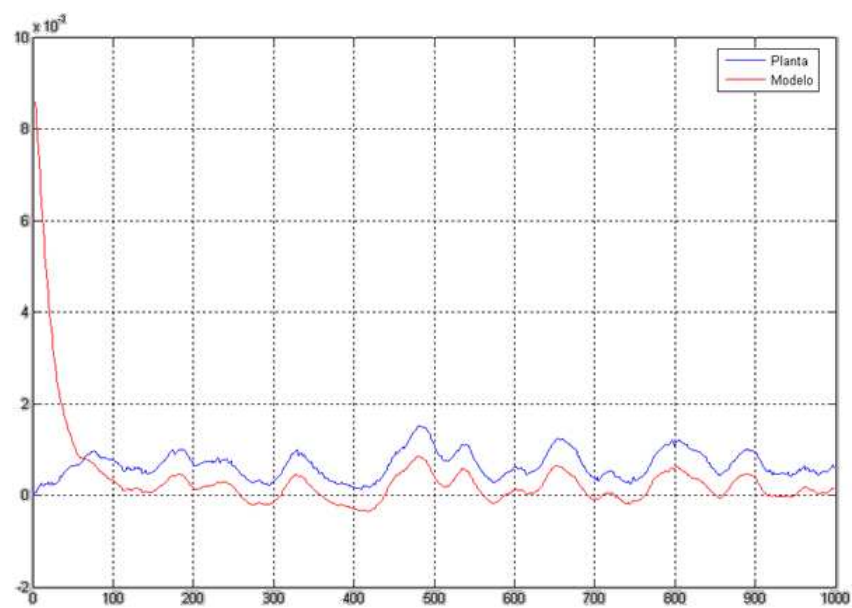


Figura 35: Salida del Modelo (Rojo) y Salida de la Planta (Azul)

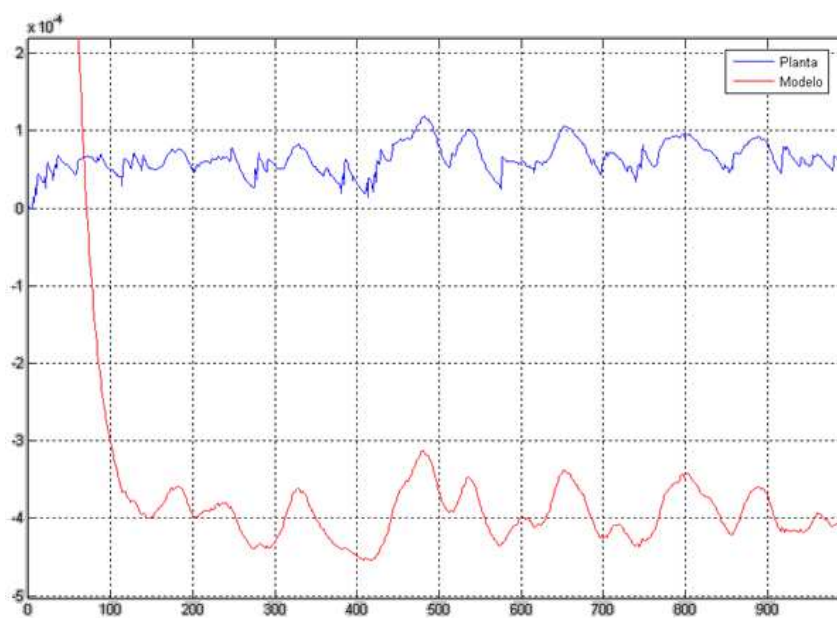


Figura 36: Salida del Modelo (Rojo) y Salida de la Planta (Azul)

Capítulo VI

Conclusiones

Los algoritmos de identificación recursiva basados en la minimización del error de salida presentaron un buen desempeño ante sistemas lineales en lazo cerrado bajo simulación. Es necesario el conocimiento amplio del controlador y del modelo de ruido de los sistemas. La identificación de 4 sistemas específicos se realizó mediante los algoritmos CLOE, FCLOE y XCLOE desarrollados por I.D.Landau en [15]. El mejor desempeño de cada uno de ellos dependerá de la aplicación. Es por eso que se evalúan por medio del índice de desempeño del intervalo de Vinnicombe.

Las técnicas de identificación difusa se basan en la clasificación de datos de entrada y salida, permitiendo la creación de modelos difusos TSK, los cuales permiten la representación de sistemas no lineales como un conjunto submodelos lineales. Por tal motivo se aplicaron para la identificación de sistemas no lineales en lazo cerrado.

Se evaluó la identificación del sistema completo Controlador-Planta-Realimentación a partir de la aplicación simulada y se obtuvo una aproximación satisfactoria medida a partir del índice de desempeño VAF (Variance Accounted For) el cual entrega en porcentaje la medida de la varianza entre dos señales. Una vez obtenida la identificación tipo caja negra al sistema total, se aplicó la metodología propuesta, encontrando restricciones por escala en las señales de identificación del sistema total y de la entrada de la planta.

Se analizaron las respuestas para el procedimiento de descuento del lazo de realimentación planteado en la metodología y no se obtuvieron respuestas adecuadas para la aplicación lo cual limitó el análisis de los pasos posteriores.

Los modelos de los sistemas no lineales son muy sensibles a la escala de las

señales implementadas en el experimento de obtención de datos. Cuando se obtiene un modelo adecuado de un sistema no lineal a partir de una determinada señal, en general, dicho modelo no representa adecuadamente la dinámica del sistema para señales de ordenes de magnitud diferentes, ni siquiera si se utiliza la señal de identificación escalada. Debido a que, una vez se halla realizado la excitación el universo de discurso estará limitado al rango de la señal de entrenamiento aplicada, de tal manera que si se introduce posteriormente una señal fuera de ese rango, el sistema de inferencia no podrá evaluar ese valor, generando desviaciones en el modelo.

Se deben realizar pruebas adicionales para determinar si existen otras características que degraden la construcción del modelo o de la aproximación en lazo abierto en cada uno de los pasos planteados en la metodología.

Se debe analizar si es adecuado el procedimiento de eliminación del término afine de los submodelos lineales obtenidos y su posterior inserción en el modelo en lazo abierto, sin la necesidad de un tratamiento algebraico específico.

Referencias

- [1] J. Abonyi, R. Babuska, F Szeifert, Modified Gath-Geva Fuzzy Clustering for Identification of Takagi-Sugeno Fuzzy Models”, IEEE Trans. on Systems, Man and Cybernetic, Part B, October, 2002
- [2] R. Babuska, Fuzzy Modeling for Control, Kluwer Academic Publisher. 1998. United States of América.
- [3] R. Babuska, J. Roubos, H. Verbruggen, Idetification of MIMO systems by input-output TS Fuzzy Models, Delft University of Technology. Department of Electrical Engineering, Control Laboratory Melkeweg 4.
- [4] R. Babuska, P.J. van der Veen, U Kaymak, Improved Covariance Estimation for Gustafson-Kessel Clustering, Proc. Of the 2002 IEEE International Conference on Fuzzy Systems,vol. 2, pp. 1081-1085, 2002.
- [5] R. Babuska, Fuzzy Modeling and Identification Toolbox User’s Guide, august 1998, Copyright.
- [6] Ch.Chuang, Sh.Su, S.Chen, Robust TSK Fuzzy Modeling for Function Approximation With Outliers, EE Trans.Fuzzy Systems, Vol.9,No.6, 2001, pp. 810-821
- [7] H.Frigui, R.Krishnapuram, A Robust Competitive Clustering Algorithm With Applications in Computer Vision , IEEE Trans, Pattern Analysis and Machine Intelligence., Vol.21,No.5, 1999, pp. 450-465.
- [8] I. Gath, A. B. Geva, Unsupervised Optimal Fuzzy Clustering, IEEE Trans. On Pattern Analisis and Machine Intelligence, vol. 11, No. 7, pp. 773-781, 1989.

- [9] M. González, S. Barato, J. Soriano and V. Grisales, Robust Agglomerative Clustering Algorithm for Fuzzy Modeling Purposes. ACC04 American Control Conference 2004, Boston USA, Junio 30 - Julio 2 de 2004.
- [10] D. E. Gustafson, W. C. Kessel, "Fuzzy Clustering with a Fuzzy Covariance", Proc.IEEE CDC, pp 761-766, San Diego, CA, USA, 1979.
- [11] R.Hathaway, J.Bezdek, Switching Regression Models and Fuzzy Clustering, IEEE Trans .Fuzzy Systems, Vol.1,No.3, 2002, pp. 195-204.
- [12] E.Kim, M.Park, S.Kim,M.Park, A Transformed Input-Output Approach to Fuzzy Modeling, IEEE Trans. Fuzzy Systems, Vol.6,No.4, 1998, pp. 596-604.
- [13] B. Kosko, Neuronal-Networks and Fuzzy Systems: a dynamical systems approach to machine intelligence, Prentice Hall, 1992, ISBN 0-13-611435-0.
- [14] I. D. Landau, Identification et commandes des systèmes, Hermes, Paris 1988
- [15] I. D. Landau and R. Loranzo, Control Adaptive, Ed. Springer-Verlag, 1998.
- [16] L. Ljung , System Identification: Theory for the user, Ed. Prentice Hall, September 1998
- [17] L. Ljung and Söderström T., Theory and Practice of Recursive Identification, Ed. MIT Press, 1987
- [18] K.Ogata, Ingeniería de Control Moderna, Prentice Hall, Tercera Edición, 1998, Mexico.
- [19] T.Takagi, M.Sugeno, Fuzzy Identification of Systems and Its Applications to Modeling and Control, IEEE Trans. Syst, Man and Cybernetics, Vol.15, No.1, 1985, pp. 116-132
- [20] L.X.Wang, A Course in Fuzzy Systems and Control. Prentice Hall, 1997
- [21] H. Wang et al., A new approach to closed-loop system identification via a vector autoregressive model, Biomedical Engineering Society, vol 30, pp. 1204-1214, 2002