

UNIVERSIDAD AUTONOMA DE NUEVO LEON
FACULTAD DE INGENIERIA MECANICA
Y ELECTRICA
DIVISION DE ESTUDIOS DE POSTGRADO



ANALISIS DE ESTABILIDAD DE REDES
NEURONALES DINAMICAS

POR

JOSE PAZ PEREZ PADRON

TESIS

EN OPCION AL GRADO DE
MAESTRO EN CIENCIAS
DE LA INGENIERIA ELECTRICA CON
ESPECIALIDAD EN CONTROL

MONTERREY, N. L.

NOVIEMBRE DE 1998

TM

Z5853

.M2

FIME

1998

P47



1020124501

UNIVERSIDAD AUTONOMA DE NUEVO LEON
FACULTAD DE INGENIERIA MECANICA
Y ELECTRICA
DIVISION DE ESTUDIOS DE POSTGRADO



ANÁLISIS DE ESTABILIDAD DE REDES
NEURONALES DINAMICAS

POR

JOSE PAZ PEREZ PADRON

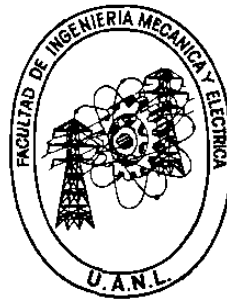
TESIS

EN OPCION AL GRADO DE
MAESTRO EN CIENCIAS
DE LA INGENIERIA ELECTRICA CON
ESPECIALIDAD EN CONTROL

MONTERREY, N. L.

NOVIEMBRE DE 1998

UNIVERSIDAD AUTONOMA DE NUEVO LEON
FACULTAD DE INGENIERIA MECANICA Y ELECTRICA
DIVISION DE ESTUDIOS DE POSTGRADO



ANALISIS DE ESTABILIDAD DE REDES NEURONALES
DINAMICAS

POR

JOSE PAZ PEREZ PADRON

TESIS

EN OPCION AL GRADO DE MAESTRO EN CIENCIAS DE LA
INGENIERIA ELECTRICA CON ESPECIALIDAD EN CONTROL

MONTERREY, N.L.

NOVIEMBRE, 1998

0129-75560

TM
Z5853
M2
FINE
1998
P47



FONDO
TESIS

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
DIVISIÓN DE ESTUDIOS DE POSGRADO

Los miembros del comité de tesis recomendamos que la tesis **ANÁLISIS DE ESTABILIDAD DE REDES NEURONALES DINÁMICAS** realizada por el Lic. José Paz Pérez Padrón sea aceptada para su defensa como opción al Grado de Maestro en Ciencias de la Ingeniería Eléctrica con especialidad en Control.

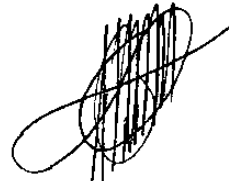
El Comité de Tesis




Asesor
Dr. Edgar N. Sánchez Camperos



Coasesor
Dr. Alexander S. Poznyak



Coasesor
Dr. Oscar L. Chacón Mondragón



Vo.Bo.
M.C. Roberto Villarreal Garza
División de estudios de Posgrado

Monterrey N.L., Noviembre, 1998

© Copyright 1998 por José Paz Pérez Padrón

Reservados todos los derechos

DEDICO ESTE TRABAJO DE TESIS A:

LA MEMORIA DE MI PADRE

José Pérez Ríos (†)

MI MADRE

Sra. Irene Padrón Méndez

MIS HIJOS

José Javier y Jorge Antonio

MI ESPOSA

Sra. M^a Lidia Díaz Enríquez

MIS HERMANOS

Genoveva, Francisco, Feliciano, Joel y Rubén.

TODOS MIS SOBRINOS, PRIMOS, TIOS(AS) y CUÑADOS (AS)

Deseo agradecer de una manera especial

A mi asesor de tesis

Dr. Edgar N. Sánchez C.

A los Drs. A. S. Poznyak y W. Yu

Deseo también agradecer

Al Dir. de la F.I.M.E. de la U.A.N.L. M.C. Cástulo E. Vela Villarreal, Al Subdirector del Postgrado de esta Facultad M.C. Roberto Villarreal Garza, Al M.C. Guadalupe E. Cedillo Garza, así como también Al Director de Estudios de Postgrado de la U.A.N.L. Dr. Ubaldo Ortiz Méndez, por su valiosa ayuda para poder llevar a cabo la defensa de esta tesis, a todos ellos gracias.

A todos los compañeros de trabajo de la F.C.F.M. de la U.A.N.L. en especial para los maestros:

Lic. Martín Martínez Gutiérrez, Lic. Francisco Rodríguez Ramírez, Lic. Santiago Arroyo Garza, Lic. Joel Pérez Padrón, Lic. Pedro E. Rodríguez De La Paz, Lic. Alfredo Alanís Duran y al Lic. Sergio R. Arratia Dávila.

A todos los maestros del programa doctoral en control de la F.I.M.E..

Al Consejo Nacional de Ciencia y Tecnología

Noviembre 19, 1998

RESUMEN

ANÁLISIS DE ESTABILIDAD DE REDES NEURONALES

DINÁMICAS

En este trabajo, en el capítulo 2 y 3 se plantea y se resuelve el problema de la estabilidad de redes neuronales dinámicas (sistemas dinámicos no lineales) desde dos enfoques: a) en el capítulo 2 se analiza la estabilidad de sistemas no lineales desde el punto de vista del análisis clásico de estabilidad de Lyapunov, de donde se obtienen condiciones suficientes en la matriz de peso del sistema que garantizan estabilidad asintótica global de la red neuronal. Esta investigación fue presentada en el congreso internacional del CINVESTAV. b) en el capítulo 3 se analiza la estabilidad de la red neuronal mediante el enfoque entrada estado estable (ISS, por sus siglas en inglés) aquí también, se obtienen condiciones suficientes en la matriz de peso de la red neuronal que garantizan estabilidad entrada estado estable, estas mismas condiciones, también garantizan estabilidad asintótica global desde el punto de vista del análisis clásico de Lyapunov, por lo que este resultado generaliza el análisis de estabilidad desde el punto de vista de Lyapunov, este resultado fue premiado con un segundo lugar en investigación teórica en el congreso internacional Sian'kan 1997 celebrado en Quintana Roo, México.

En el capítulo 4, se plantea y resuelve el problema de la identificación adaptativa robusta de sistemas no lineales y seguimiento de trayectorias usando redes neuronales dinámicas, en donde el sistema nominal es perturbado por un término adicional, representando las incertidumbres del sistema. Por medio del análisis de estabilidad desde el punto de vista de Lyapunov nosotros determinamos condiciones estabilidad para el error de identificación. Luego analizamos la estabilidad del error de seguimiento de las trayectorias para un sistema no lineal previamente identificado; la trayectoria es generada por un modelo no lineal. Para el análisis de identificación usamos una ecuación algebraica de Riccati y para el error de seguimiento una ecuación diferencial de Riccati usando adaptación de parámetros de la red neuronal. Esta investigación fue presentada en la conferencia americana de control (ACC 1997, por sus siglas en inglés).

INDICE

RESUMEN	vi
LISTA DE FIGURAS	x
INTRODUCCIÓN	xii

CAPITULO 1

CONCEPTOS ESENCIALES

1.1	¿Qué es una Red Neuronal?	1
1.2	Modelo de una Neurona	4
1.3	Arquitectura de las Redes Neuronales	7
1.4	Representación del conocimiento	10
1.5	Representación de sistemas en variables de estado	11
1.6	Clasificación de sistemas	13

CAPITULO 2

ANÁLISIS SIMPLIFICADO DE ESTABILIDAD ASINTÓTICA GLOBAL DE REDES NEURONALES DINÁMICAS

2.1	Introducción	15
2.2	Modelo matemático	16
2.3	Análisis de estabilidad asintótica global	18
2.4	Aplicaciones	22
2.5	Conclusiones del capítulo	24

INDICE (continuación)

CAPITULO 3

ANÁLISIS DE ESTABILIDAD ENTRADA-ESTADO PARA REDES NEURONALES DINÁMICAS

3.1	Introducción	25
3.2	Modelo matemático	26
3.3	Punto de equilibrio	27
3.4	Análisis de estabilidad entrada-estado	28
3.5	Aplicaciones	33
3.6	Conclusiones del capítulo	39

CAPITULO 4

IDENTIFICACIÓN ADAPTABLE ROBUSTA DE SISTEMAS NO-LINEALES Y SEGUIMIENTO DE TRAYECTORIAS USANDO REDES NEURONALES DINÁMICAS

4.1	Introducción	40
4.2	Análisis de la estabilidad del error de identificación adaptativa Robusta	43
4.3	Análisis del error de seguimiento de trayectorias	48
4.4	Aplicaciones	53
4.5	Conclusiones del capítulo	61

INDICE (continuación)

CAPITULO 5

CONCLUSIONES Y TRABAJO FUTURO

5.1	Conclusiones	62
5.2	Trabajo futuro	63
Apéndice A. Publicaciones		64
Apendice B. Constancia de un Premio Internacional en Investigación Teórica		90
Referencias		93

LISTA DE FIGURAS

- Fig. 1 Función Umbral
- Fig.2 Función lineal a tramos
- Fig.3 Función Sigmoide
- Fig.4 Red de Prealimentación de una capa
- Fig.5 Red de Prealimentación hacia delante multicapa
- Fig.6 Red Recurrente sin lazos de autoalimentación y sin neuronas ocultas
- Fig.7 Red Reticular de dos dimensiones de 3x3 neuronas
- Fig.8 Plano Fase de $x_1(t)$ contra $x_2(t)$, de un sistema GAS
- Fig.9 Respuesta de $x_1(t)$ de un sistema GAS
- Fig.10 Respuesta de $x_2(t)$ de un sistema GAS
- Fig.11 Plano Fase de $x_1(t)$ contra $x_2(t)$ de un sistema ISS
- Fig.12 Respuesta de $x_1(t)$ de un sistema ISS
- Fig.13 Respuesta de $x_2(t)$ de un sistema ISS
- Fig.14 Plano Fase de $x_1(t)$ contra $x_2(t)$ de un sistema GAS
- Fig.15 Respuesta de $x_1(t)$ de un sistema GAS
- Fig.16 Respuesta de $x_2(t)$ de un sistema GAS
- Fig.17 Esquema de identificación y control
- Fig.18 Esquema de la planta no lineal
- Fig.19 Esquema de la red neuronal
- Fig.20 Resultados de la identificación
- Fig.21 Resultados de la evolución de la matriz de pesos
- Fig.22 Respuesta sin control de retroalimentación
- Fig.23 Esquema para el controlador

LISTA DE FIGURAS (continuación)

Fig.24 Respuesta con control de retroalimentación

Fig.25 Respuesta de la matriz P

Fig.26 Error de seguimiento J_r^A

Fig.27 Entrada al sistema no lineal J_r^u

INTRODUCCIÓN

El objetivo a largo plazo de la Ingeniería de Control Automático es implementar un sistema, que pueda operar independientemente de la intervención humana en un ambiente incierto; un sistema con esas capacidades se denomina autónomo. Una vez implementado, solo requerirá que se le especifiquen las metas deseadas, en lenguaje natural, y él determinará sus objetivos y los mecanismos para lograrlos por medio de una interacción continua con el ambiente. Este tipo de sistemas serán muy útiles cuando la intervención humana sea peligrosa, susceptible a fallas o sencillamente imposible. Sobra decir que un sistema tal aún no ha sido desarrollado; así mismo, las técnicas actuales de análisis, diseño e implementación de sistemas de control automático no son lo suficientemente poderosas y flexibles para construir estos sistemas autónomos.

Los sistemas biológicos constituyen una posible referencia para el diseño de este tipo de sistemas autónomos. Ellos pueden ser fuente de inspiración para el desarrollo de algoritmos de aprendizaje robustos y adaptables, requeridos por los sistemas autónomos. Los sistemas biológicos procesan información en una forma muy diferente a los sistemas de control convencionales; no utilizan modelos matemáticos y son muy exitosos para enfrentar incertidumbres. Aprenden a realizar nuevas tareas y se adaptan fácilmente a medio ambientes cambiantes. Si los principios fundamentales de computación, que el sistema nervioso utiliza, logran ser entendidos, entonces toda una nueva generación de métodos de control automático podrá ser desarrollada; esta nueva generación tendrá capacidades mucho mejores que los sistemas de control actuales, y podrán permitir la implementación de sistemas realmente autónomos.

Se dice que un sistema tiene la capacidad de aprender si éste adquiere información, durante su operación, sobre las características desconocidas de la planta y su medio ambiente, de tal forma que el desempeño global es mejorado de forma continua. Por lo tanto, si se le agrega esta capacidad al sistema de control, entonces se puede expandir el rango global de operación y finalmente obtener un sistema autónomo.

Una clase de modelos, derivados de los sistemas biológicos, que tienen la capacidad para implementar el aprendizaje requerido por los sistemas autónomos, son las redes neuronales artificiales. Aunque el sistema nervioso es muy complejo de analizar, se han desarrollado modelos simplificados que están siendo aplicados en una gran variedad de problemas de ingeniería; en particular en ingeniería de control automático.

Aunque se han reportado muchas aplicaciones de Control Neuronal, como sistemas de control automático basados en redes neuronales artificiales, el análisis de sus propiedades está bastante menos desarrollado. Usualmente en control neuronal se utilizan redes neuronales estáticas; esto es, redes que realizan una transformación no lineal entrada-salida independiente del tiempo. La utilización de redes neuronales dinámicas, representadas por ecuaciones diferenciales no lineales, en control neuronal apenas empieza y ofrece un gran potencial de desarrollo. El análisis de control neuronal, basado en redes neuronales dinámicas, está también en sus etapas iniciales.

Con base en lo anteriormente expuesto, los objetivos del trabajo de investigación, presentados en esta tesis, son:

- Analizar la estabilidad de redes neuronales dinámicas.
- Analizar la robustez y la convergencia del error de un esquema, basado en redes neuronales dinámicas, para seguimiento de trayectoria de sistemas no lineales.

CAPÍTULO 1

CONCEPTOS ESENCIALES

1.1-¿Qué es una red Neuronal?

El esfuerzo por entender el cerebro humano se debe en gran medida a los trabajos pioneros de Ramón y Cajál [1], quien introdujo la idea de la neurona como elemento básico del cerebro. Típicamente, las neuronas son de cinco a seis ordenes de magnitudes más lentas que las compuertas lógicas ('chips'); los eventos en un 'chip' toman del orden de los nanosegundos ((10^{-9}) s, mientras que los eventos en las neuronas toman del orden de los milisegundos ((10^{-3}) s).

El cerebro compensa la relativa lentitud de operación por un sorprendente número de neuronas (células nerviosas) interconectadas masivamente entre ellas; se estima que deben existir del orden de 10 billones de neuronas en la corteza cerebral humana, y 60 trillones de sinapsis o interconexión entre ellas [2]. El resultado neto es que el cerebro es una enorme estructura eficiente. Específicamente, la eficiencia energética en el cerebro es de aproximadamente (10^{-16}) joules(J) por operación por segundo, mientras que el valor correspondiente para la mejor computadora actualmente es de (10^{-6}) joules(J) por operación por segundo [3].

El cerebro es una computadora en paralelo, no lineal y altamente compleja. Tiene la capacidad de organizar las neuronas, así como de ejecutar ciertas operaciones (por ejemplo, reconocimiento de patrones, percepción y control motriz) muchas veces más rápido que la mejor computadora existente.

El uso de redes neuronales ofrece las siguientes ventajas:

1. **Nolinealidad.** Una neurona es básicamente un dispositivo no lineal. Consecuentemente, una red neuronal, constituida de neuronas es ella misma no lineal.
2. **Mapeo entrada-salida.** Un paradigma popular de aprendizaje llamado aprendizaje supervisado involucra la modificación de los pesos sinápticos de una red neuronal aplicando un conjunto de muestras de entrenamiento o ejemplos. Cada ejemplo consiste de una única señal de entrada y la correspondiente respuesta deseada. Así la red aprende de los ejemplos construyendo un mapeo entrada-salida para el problema específico.
3. **Adaptatividad.** Las redes neuronales tienen la capacidad de adaptar sus pesos sinápticos como respuesta a cambios en su medio ambiente. En particular, una red se entrena para operar en un medio ambiente específico y puede ser fácilmente reentrenada al menor cambio en su medio ambiente de operación; una red neuronal puede ser diseñada para cambiar sus pesos sinápticos en tiempo real. La capacidad de las redes neuronales para clasificación de patrones, procesamiento de señales, y aplicaciones de control, junto con su adaptación, hacen de ellas una herramienta ideal para el uso de clasificación de patrones, procesamiento adaptable de señales, y control adaptable.
4. **Validación.** En el contexto de clasificación de patrones, una red neuronal puede ser diseñada para proveer información no únicamente acerca de un patrón particular seleccionado, sino también acerca de la confianza en la decisión tomada. Esta información puede ser usada después para rechazar patrones ambiguos, si es que surgen, y así perfeccionar el desempeño de la red.

5. Tolerancia a fallas. Una red neuronal, implementada en 'hardware', es tolerante a fallas, en el sentido que su desempeño es degradado suavemente bajo condiciones de operación adversas.

6. Implementabilidad en VLSI. La natural masividad paralela de una red neuronal la hace potencialmente rápida para los cálculos de ciertas tareas. La misma característica hace a una red neuronal idealmente apropiada para implementación tecnológica usando una integración de muy alta escala ('VLSI').

7. Analogía Neurobiológica. Los neurobiólogos ven a las redes neuronales (artificiales) como una herramienta para la implementación de fenómenos neurobiológicos. Los ingenieros se inspiran en modelos biológicos para desarrollar nuevos algoritmos.

Una red neuronal artificial: es un procesador masivo distribuido paralelamente que tiene la capacidad de almacenar experiencia y tenerla disponible para su uso. Se asemeja al cerebro en dos aspectos:

1. El conocimiento es adquirido por la red a través de procesos de aprendizaje.
2. Las numerosas interconexiones neuronales conocidas como pesos sinápticos son usadas para almacenar información.

La modificación de los pesos sinápticos provee el método tradicional para el diseño de redes neuronales.

Las redes neuronales son también referidas en la literatura como procesadores neuronales interconectados y distribuidos en paralelo.

1.2-Modelo de una neurona.

Una neurona es una unidad de procesamiento de información la cual es fundamental para la operación de la red neuronal. Se pueden identificar tres elementos básicos en los modelos neuronales, los cuales son descritos enseguida:

1. Un conjunto de sinapsis o líneas de conexión, cada una de los cuales es caracterizada por un peso. Específicamente, una señal x_j en la entrada de la sinapsis j conectada a la neurona k es multiplicada por el peso sináptico w_{kj} .
2. Un elemento para sumar las señales de entrada, pesadas por las respectivas sinapsis de las neuronas.
3. Una función de activación la cual limita la amplitud de las salidas de las neuronas. Las funciones de activación son también referidas en la literatura como función de confinamiento en el sentido que limitan el rango de amplitud de las salidas a un valor finito. Típicamente, el rango de amplitud normalizada de la salida de una neurona está en el intervalo unitario cerrado $[0,1]$ o alternativamente $[-1,1]$.

En algunos modelos de redes neuronales también se incluye un umbral aplicado externamente θ_k , el cual tiene el efecto de desplazar la función de activación.

Tipos de funciones de activación.

Las funciones de activación, denotadas por $\varphi(\cdot)$, definen la salida de la neurona en términos del nivel de actividad en su entrada. Los tres tipos de función de activación más usados son:

1. **Función Umbral.** Para este tipo de función de activación ver (Fig.1), se tiene

$$\varphi(v) = \begin{cases} 1 & \text{si } v \geq 0 \\ 0 & \text{si } v < 0 \end{cases} \quad (1.1)$$

Correspondientemente, la salida de la neurona k empleando tal función Umbral es expresada como

$$y_k = \begin{cases} 1 & \text{si } v_k \geq 0 \\ 0 & \text{si } v_k < 0 \end{cases} \quad (1.2)$$

donde v_k es el nivel de actividad interna de la neurona; esto es,

$$v_k = \sum_{j=1}^p w_{kj} x_j - \theta_k \quad (1.3)$$

Tales neuronas son referidas en la literatura como modelos de McCulloch y Pitts en reconocimiento a los trabajos pioneros hechos por McCulloch y Pitts en (1943).

$\varphi(v)$

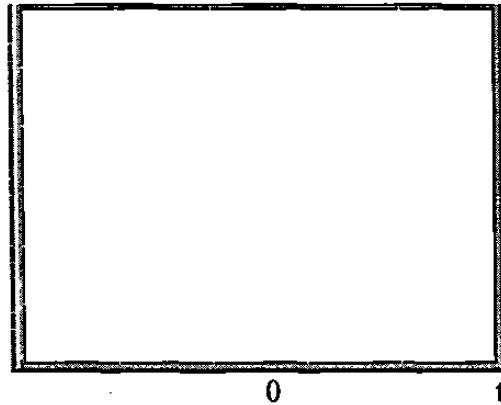


Fig.1 Función Umbral

2. **Función lineal a tramos.** Para la función lineal a tramos, se tiene

$$\varphi(v) = \begin{cases} 1 & \text{si } v \geq 1/2 \\ v & \text{si } -1/2 < v < 1/2 \\ 0 & \text{si } v \leq -1/2 \end{cases} \quad (1.4)$$

donde el factor de amplificación dentro de la región lineal de operación se supone la unidad,(ver Fig.2.)

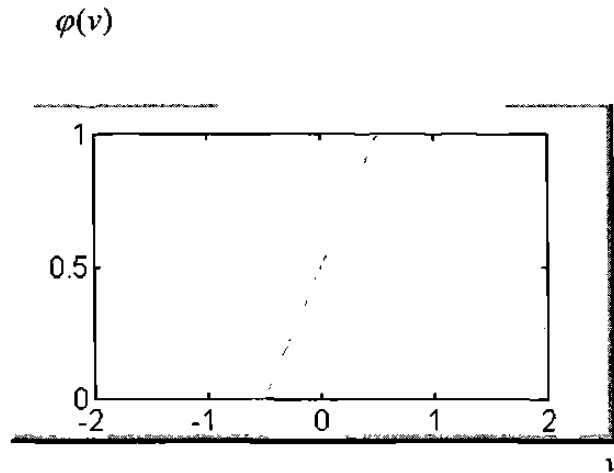


Fig.2 Función lineal a tramos

3. **Función sigmoideal.** La función sigmoide es la función de activación más común usada en la construcción de redes neuronales artificiales . Es una función estrictamente creciente, suave y asintótica, (ver Fig.3). Un ejemplo de la función sigmoide es:

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (1.5)$$

Las funciones de activación anteriormente definidas tienen rango de 0 a 1. Algunas veces es deseable que las funciones de activación tengan rango de -1 a $+1$, Específicamente, la función umbral es redefinida por

$$\varphi(v) = \begin{cases} 1 & \text{si } v > 0 \\ 0 & \text{si } v = 0 \\ -1 & \text{si } v < 0 \end{cases} \quad (1.6)$$

a la cual comúnmente se le llama función signo.

Para una función sigmoide se puede usar la función tangente hiperbólica, la cual se define por:

$$\varphi(v) = \tanh\left(\frac{v}{2}\right) = \frac{1 - \exp(-v)}{1 + \exp(-v)} \quad (1.7)$$

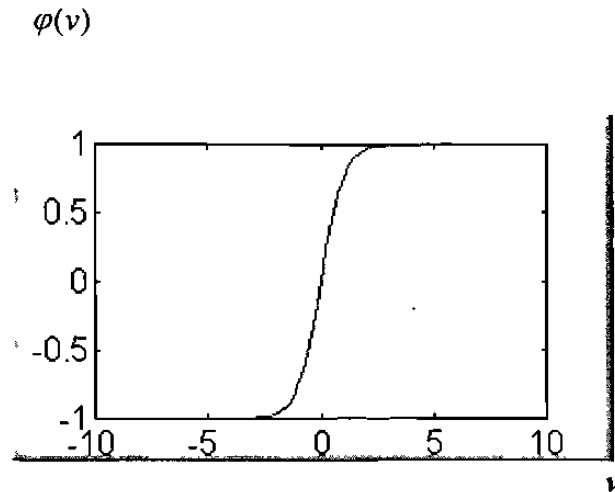


Fig.3 Función sigmoide

Retroalimentación.

Se dice que en un sistema dinámico existe retroalimentación cuando la salida de un elemento en el sistema influencía la entrada a este, al ser aplicado a un elemento particular del mismo sistema.

Claramente la retroalimentación ocurre en casi todas las partes del sistema nervioso de todo animal [4]. Además, juega un mejor papel en el estudio de clases especiales de redes neuronales conocidas como redes neuronales recurrentes.

1.3 Arquitectura de las redes neuronales.

Redes de prealimentación de una capa.

Una red neuronal de capas es una red neuronal organizada en forma de capas. En la forma más simple de una red de capas, se tiene una entrada en la capa de entrada, la cual es transmitida hacia adelante. Esta red es del tipo de prealimentación. Si se tiene una sola capa, entonces se le llama red neuronal de una sola capa, ver (Fig. 4).

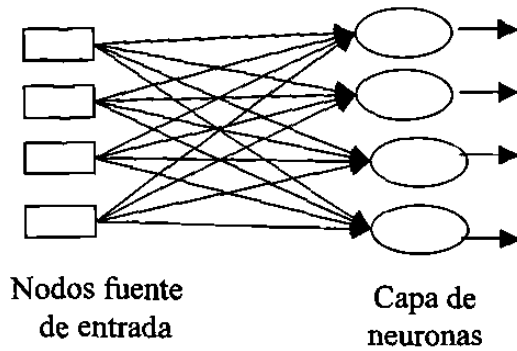


Fig.4 Red de prealimentación de una capa

Redes prealimentadas multicapa.

La segunda clase de redes neuronales prealimentadas se distinguen por la presencia de una o más capas ocultas, cuyos nodos de operación son correspondientemente llamados nodos ocultos o unidades ocultas. Típicamente, las neuronas en cada capa de la red tienen como entradas la señal de salida de la capa que le precede, ver (Fig.5).

Una red neuronal se dice estar completamente conectada en el sentido que todo nodo en cada capa de la red está conectado con todo nodo en la capa siguiente. Si algunas de las conexiones sinápticas no aparecen en la red, se dice que la red es parcialmente conectada.

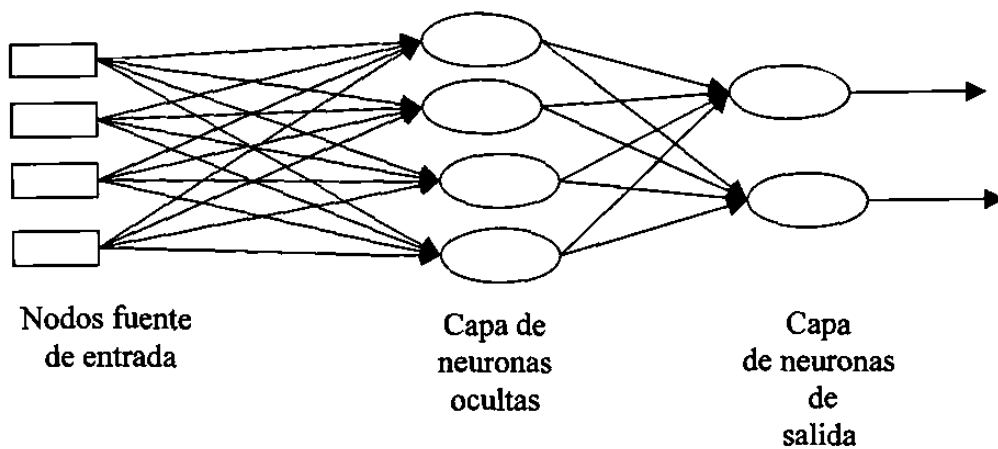


Fig.5 Red de alimentación hacia adelante multicapa.

Redes recurrentes.

Una red neuronal recurrente se distingue de las redes neuronales prealimentadas en que tiene al menos un lazo de retroalimentación, (ver Fig.6).

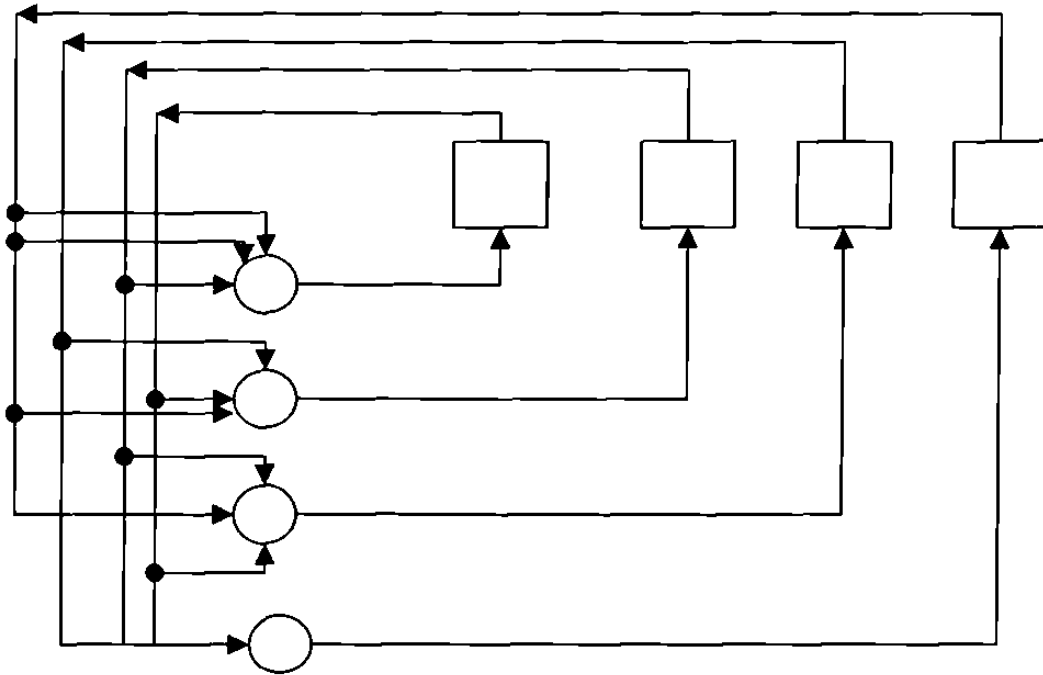


Fig.6 Red recurrente sin lazos de autoalimentación y sin neuronas ocultas

El operador $z = q^{-1}$, retarda en un instante la señal muestreada, por ejemplo:

$$q^{-1}(u(k)) = u(k - 1) \text{ donde } k, \text{ es el } k\text{-ésimo instante de muestreo.}$$

Estructura de retícula.

Una retícula consiste de una dimensión, dos dimensiones, o arreglos de más altas dimensiones de neuronas, con un correspondiente conjunto de nodos fuente que suministran las señales de entrada al arreglo.

Una red de retícula es realmente una red prealimentada con la salida de las neuronas dispuestas en filas y columnas, (ver Fig.7).

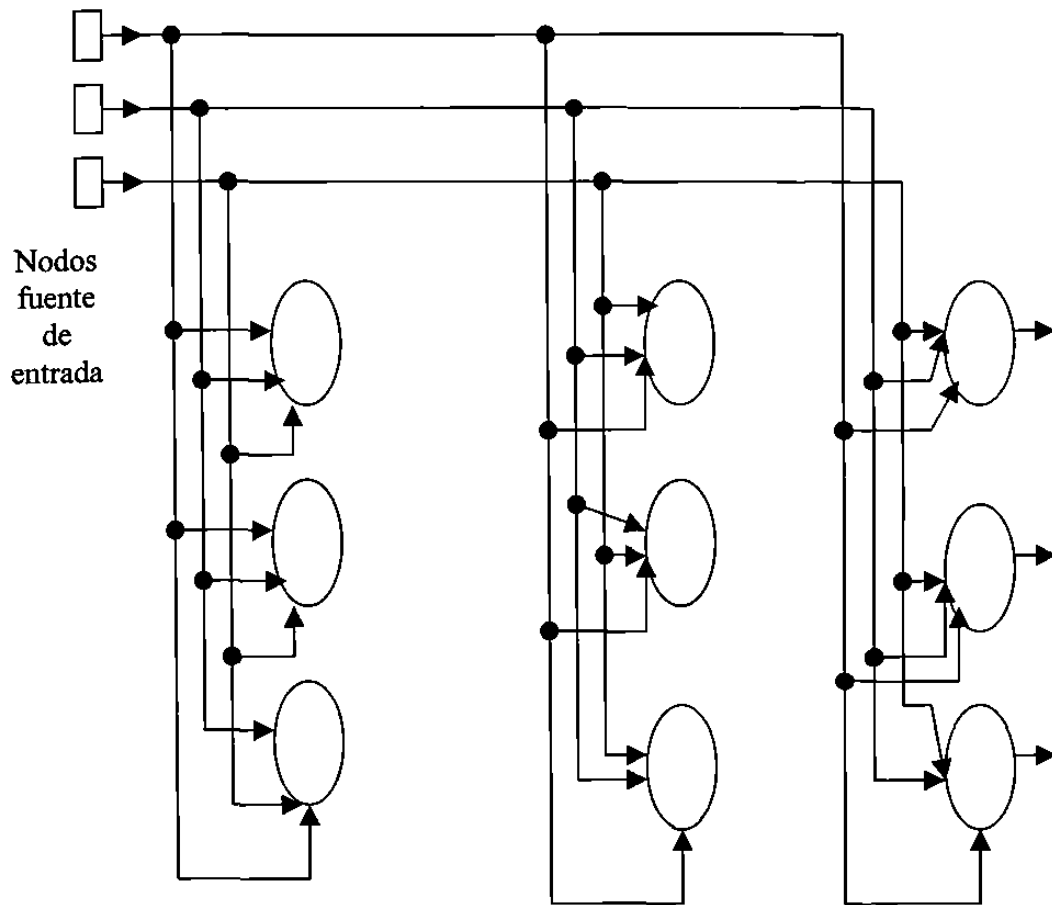


Fig.7 Red reticular de dos dimensiones de 3x3 neuronas.

1.4-Representación del conocimiento.

El conocimiento se refiere a almacenar información o modelos usados por personas o máquinas.

La mayor tarea para una red neuronal es aprender un modelo de un medio ambiente y mantener el modelo lo suficientemente consistente con el mundo real (medio ambiente real) para que efectúe una tarea específica o una aplicación de interés.

El conocimiento del medio ambiente consiste de dos clases de información:

1. El conocimiento del medio ambiente, representado por hechos que fueron o son conocidos. Esta forma de conocimiento es referido como información conocida.
2. Observaciones (mediciones) del medio, obtenidas por sensores.

Así, el conjunto de ejemplos, representan el conocimiento acerca del medio ambiente de interés. Típicamente, el conjunto de ejemplos usados para entrenar a la red consisten de una gran cantidad representativos de la situación real.

No obstante, hay cuatro reglas para la representación del conocimiento:

Regla 1. Entradas similares de clases similares deberán usualmente producir representaciones similares en la red, por lo tanto deberán clasificarse como pertenecientes a la misma categoría.

Hay muchas formas de medidas para determinar la similaridad entre entradas. Una muy comúnmente usada se basa en el concepto de distancia Euclideana.

Regla 2. Partes a ser categorizadas como clases separadas deben tener una representación diferente en la red; esta segunda regla es lo opuesto de la regla 1.

Regla 3. Si una característica particular es importante, entonces deberá haber un gran número de neuronas involucradas en su representación.

Regla 4. La información conocida e invariante se debe tener en cuenta en el diseño de la red neuronal.

1.5-Representación de sistemas en variables de estado

Esta sección se incluye porque en el desarrollo de esta tesis se utiliza la representación en variables de estado.

Por qué usar variables de estado?

El modelo matemático en variables de estado es conveniente puesto que:

1. Las ecuaciones diferenciales son ideales para soluciones digitales o analógicas.

2. El estado provee una herramienta que unifica el estudio de sistemas lineales y no lineales.
3. La forma en variables de estado es invaluable en investigaciones teóricas.
4. El concepto de estado tiene una fuerte motivación física.

Definición de estado de un sistema.

El estado de un sistema es el conjunto de variables $x_1(t), x_2(t), \dots, x_n(t)$ de tal forma que si estas son conocidas en $t = t_0$, entonces se determinan de manera única para $t \geq t_0$ especificando la entrada $u(t)$ del sistema para $t \geq t_0$.

Si $x_1(t), x_2(t), \dots, x_n(t)$ son las variables de estado (o simplemente el estado), y $u_1(t), u_2(t), \dots, u_m(t)$ son las señales de entrada, entonces el sistema puede ser descrito por n ecuaciones diferenciales de primer orden

$$\begin{aligned}
 \dot{x}_1(t) &= f_1(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t) \\
 \dot{x}_2(t) &= f_2(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t) \\
 &\vdots \\
 \dot{x}_n(t) &= f_n(x_1(t), x_2(t), \dots, x_n(t), u_1(t), u_2(t), \dots, u_m(t), t)
 \end{aligned}
 \tag{1.8}$$

Nótese que $\dot{x}_i(t)$ es en general una función no lineal variante en el tiempo que depende, además, del estado y de los controles.

Se define $X(t) = (x_1(t), x_2(t), \dots, x_n(t))^T \in \mathfrak{R}^n$ como el vector de estado de el sistema, y $U(t) = (u_1(t), u_2(t), \dots, u_m(t))^T \in \mathfrak{R}^m$ como el vector de entrada. Las ecuaciones de estado pueden ser entonces escritas como $\dot{X}(t) = F(X(t), U(t), t)$ donde F es una función valuada vectorialmente, con componentes $f_i(x_1(t), \dots, x_n(t), u_1(t), \dots, u_m(t), t), i = 1, 2, \dots, n$.

1.6-Clasificación de sistemas

Los sistemas se clasifican en lineales, no lineales, invariante en el tiempo (estacionario o fijos) y variante en el tiempo. Los sistemas se clasifican de acuerdo a la forma de sus ecuaciones de estado. Por ejemplo, si un sistema es no lineal y variante en el tiempo, las ecuaciones de estado son escritas como:

$$\dot{X}(t) = F(X(t), U(t), t) \quad (1.9)$$

Los sistemas no lineales invariantes en el tiempo son representados por ecuaciones de estado de la forma

$$\dot{X}(t) = F(X(t), U(t)) \quad (1.10)$$

Un sistema es lineal y variante en el tiempo si las ecuaciones de estado son de la forma:

$$\dot{X}(t) = A(t)X(t) + B(t)U(t), \quad (1.11)$$

donde $A(t)$ y $B(t)$ son matrices de dimensiones $n \times n$ y $n \times m$ respectivamente.

Las ecuaciones de estado para sistemas lineales invariantes en el tiempo tienen la forma:

$$\dot{X}(t) = AX(t) + BU(t), \quad (1.12)$$

donde A y B son matrices constantes cuyas dimensiones son igual que en el caso anterior.

Ecuaciones de salida

Las cantidades físicas que pueden ser medidas son llamadas salidas y son denotadas por $y_1(t), y_2(t), \dots, y_q(t)$. Si las salidas son funciones no lineales, variantes en el tiempo dependientes del estado y entrada, escribiremos las ecuaciones de salida como:

$$Y(t) = G(X(t), U(t), t) \quad (1.13)$$

Si la salida es relacionada al estado y al control por medio de relaciones lineales e invariables en el tiempo entonces

$$Y(t) = CX(t) + DU(t) \quad (1.14)$$

Donde C y D son matrices constantes de dimensiones $q \times n$ y $q \times m$ respectivamente.

CAPÍTULO 2

ANÁLISIS SIMPLIFICADO DE ESTABILIDAD ASINTÓTICA GLOBAL DE REDES NEURONALES DINÁMICAS

En este capítulo se presenta el análisis de estabilidad global para redes neuronales dinámicas. Estas condiciones son establecidas mediante una función de Lyapunov cuadrática. A partir de ésta, se derivan las condiciones para que la derivada respectiva en las trayectorias de la red sea estrictamente negativa.

2.1-Introducción

El análisis de la estabilidad de redes neuronales dinámicas se ha desarrollado en gran medida últimamente. Inicialmente se establecieron condiciones para estabilidad asintótica local [5,6]; un problema que se presenta es la existencia de múltiples puntos de equilibrio, que pueden corresponder a mínimos locales. Esta propiedad es importante para memoria asociativa o reconocimiento de patrones. Sin embargo, para aplicaciones en control automático o en optimización, es preferible garantizar estabilidad asintótica global.

El estudio de estabilidad asintótica global de redes neuronales dinámicas es un tema que genera gran interés actualmente. En [7] se destaca su importancia. Aplicando la técnica del mapeo contractante, en [8], se establecen condiciones para estabilidad global. Matsouka [9] generaliza estos resultados haciendo uso de una función de Lyapunov del tipo Lure. Resultados recientes [10,11] se basan en el concepto de matriz diagonalmente estable.

En el presente trabajo se retoma el enfoque empleado en [9], pero se utiliza una función de Lyapunov Cuadrática para establecer condiciones de estabilidad global. La organización del capítulo es como sigue: al inicio se describe el modelo matemático de la

red, después se lleva a cabo el análisis de estabilidad y, finalmente, se establecen las respectivas conclusiones.

2.2-Modelo Matemático

La dinámica de la red neuronal está representada por:

$$\begin{aligned} \tau \frac{dx_i}{dt} &= -x_i + \sum_j w_{ij} y_j + s_i \\ y_i &= g(x_i) \end{aligned} \quad (2.1)$$

donde:

x_i es la variable que representa el estado de cada neurona, $i = 1, 2, 3, \dots, n$.

s_i es una entrada constante.

y_i es la salida de la red.

w_{ij} es el peso de conexión de la unidad j a la unidad i , no necesariamente simétrico.

τ es una constante de tiempo, que sin pérdida de generalidad puede tomarse unitaria.

Se supone que la función de salida g_i es suave, estrictamente creciente y que se satura para $z \rightarrow \infty$.

$\forall z$ entonces:

(G₁) g_i es continuamente diferenciable

(G₂) $0 < g_i'(z) \leq \sup_z \bar{g}_i'(z) = r_i < \infty$

(G₃) g_i es acotada, esto es:

$$\bar{g}_i(z) \leq \sup_z \bar{g}_i(z) < \infty.$$

La función sigmoideal, frecuentemente usada en redes neuronales, $g_i(z) = \frac{1}{1 + e^{-z}}$

, cumple con estas condiciones.

Definiendo:

$$w_{ij} \stackrel{\Delta}{=} w_{ij} r_i, \quad g_i(z) \stackrel{\Delta}{=} \frac{g_i}{r_i}, \quad y \quad \tau = 1 \quad (2.2)$$

entonces (2.1) se puede formular como:

$$\frac{dx_i}{dt} = -x_i + \sum_j w_{ij} g_j(x_j) + s_i \quad (2.3)$$

Por lo que (G2) se expresa como:

$$(G'_2) \quad 0 < g'_i(z) \leq \sup_z g'_i(z) = 1 \quad (2.4)$$

puesto que:

$$g'_i(z) = \frac{\bar{g}'_i(z)}{r_i} \leq \sup_z \left(\frac{\bar{g}'_i(z)}{r_i} \right)$$

$$g'_i(z) \leq \frac{1}{r_i} \sup_z (\bar{g}'_i(z)) = 1$$

$$\therefore 0 < g'_i(z) \leq \sup_z (g'_i(z)) \leq 1$$

Expresando (2.3) en forma matricial, se obtiene:

$$\frac{dX}{dt} = -X + Wg(X) + S \quad (2.5)$$

donde:

$$X = (x_1, x_2, \dots, x_n)^T$$

$$S = (s_1, s_2, \dots, s_n)^T$$

$$g(X) = (g_1(x), g_2(x), \dots, g_n(x))^T$$

W es la matriz cuyos elementos son w_{ij} .

Usualmente la matriz W se denomina matriz de interconexión de pesos.

2.3-Análisis de estabilidad asintótica global

En este trabajo, y de acuerdo a [9], se define que la red neuronal es absolutamente estable cuando el sistema (2.5) tiene un punto de equilibrio globalmente asintóticamente estable, para cualquier entrada constante S .

El objetivo es encontrar condiciones suficientes en la matriz de pesos, W , que aseguren estabilidad absoluta. En la referencias previas, estas condiciones son expresadas en términos de normas o medida de esta matriz. Para establecer dichas condiciones, primero se determina la existencia de al menos un punto de equilibrio y luego se demuestra que es globalmente asintóticamente estable y por lo tanto único.

A. Existencia de un Punto Fijo

Este análisis se toma de [9] y se incluye para facilitar la explicación de los resultados.

Definiendo el mapeo $\Phi: \mathfrak{R}^n \rightarrow \mathfrak{R}^n$ como $\Phi(X) = Wg(X) + S$, entonces, un estado de equilibrio del sistema (2.5) es un punto fijo del mapeo Φ .

En efecto en todo punto de equilibrio se cumple que $X = \Phi(X)$.

LEMA 2.1- Suponiendo g acotada, el mapeo Φ tiene al menos un punto fijo

Prueba:

Sea Ω un hipercubo definido como:

$$\begin{aligned} \Omega &= \{X : \|X - S\|_{\infty} \leq \|W\|_{\infty} M \\ M &= \max_i \sup_z |g_i(z)| \end{aligned} \tag{2.6}$$

entonces:

$$\begin{aligned} \|\Phi(X) - S\|_{\infty} &= \|Wg(X)\|_{\infty} \\ &\leq \|W\|_{\infty} \|g(X)\|_{\infty} \\ &\leq \|W\|_{\infty} M \end{aligned} \tag{2.7}$$

se satisface para todo punto en Ω , implicando que Φ es un mapeo continuo desde un conjunto cerrado convexo acotado a el mismo. Entonces por el teorema del punto fijo de Brower [12], Φ debe tener un punto fijo en Ω , esto es (2.5) tiene un punto de equilibrio.

□□□

Se enuncia el teorema del punto fijo de Brower sin demostración.

Teorema del punto fijo de Brouwer. Sea $b > 0$ y sea $D = \{x \in \mathbb{R}^p : x \leq b\}$. Entonces cualquier función continua con dominio D y contradominio contenido en D tiene al menos un punto fijo.

Sea $X^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ dicho punto de equilibrio del sistema (2.5) y sea la nueva variable $U = (u_1, u_2, \dots, u_n)^T = X - X^*$, entonces (2.5) puede reescribirse en términos de U como:

$$\frac{dU}{dt} = -X + Wg(X) + S \quad (2.8)$$

Teniendo en cuenta que $X = U + X^*$, entonces:

$$\begin{aligned} \frac{dU}{dt} &= -U - X^* + Wg(U + X^*) + S \\ &= -U + W(g(U + X^*) - g(X^*)) \\ &\quad + Wg(X^*) + S - X^* \end{aligned} \quad (2.9)$$

Definiendo la nueva función:

$$f(U) = g(U + X^*) - g(X^*) \quad (2.10)$$

y considerando que

$$Wg(X^*) + S - X^* = \Phi(X^*) - X^* = 0$$

el sistema (2.8) se expresa finalmente en función de la variable U como:

$$\frac{dU}{dt} = -U + Wf(U) \quad (2.11)$$

ahora, puesto que:

$$f_i'(z) = g_i'(z + x_i^*) - g_i'(x_i^*) = g_i'(z + x_i^*) \quad (2.12)$$

por (G_2') se tiene que:

$$f_i'(z) \leq 1, \text{ para } i = 1, 2, \dots, n.$$

Además $f_i(z)$ satisface:

$$\begin{aligned} z \leq f_i(z) < 0, & \text{ si } z < 0 \\ f_i(0) &= 0 \\ 0 < f_i(z) \leq z, & \text{ si } z > 0 \end{aligned} \quad (2.13)$$

Las desigualdades (2.13) se obtienen usando el teorema del valor medio para derivadas.

B. Estabilidad Asintótica:

Una vez establecida la existencia de al menos un punto de equilibrio, se procede a establecer las condiciones para que el origen $U = 0$ (o en forma equivalente $x = x^*$) sea globalmente asintóticamente estable. Esto constituye la aportación principal del presente capítulo, el cual se establece en el siguiente teorema.

TEOREMA 2.1- Si $W_2 < 1$, entonces el sistema (2.8), o en forma equivalente el sistema (2.5), es absolutamente estable.

Prueba-

Definase la siguiente función cuadrática de Lyapunov:

$$V(U) = U^T P U \quad (2.14)$$

donde:

P matriz definida positiva, solución de la ecuación de Lyapunov $PA + A^T P = -I$ con $A = -I$

Derivando (2.14) a lo largo de las trayectorias de (2.8), se obtiene:

$$\frac{dV}{dt} = -2U^T P U + 2U^T P W f(U) \quad (2.15)$$

Aplicando la siguiente desigualdad matricial [13]:

$$X^T Y + Y^T X \leq X^T R X + Y^T R^{-1} Y, \quad (2.16)$$

la cual es válida para cualesquiera matrices $X \in \mathfrak{R}^{n \times k}$, $Y \in \mathfrak{R}^{n \times k}$ y para cualquier matriz definida positiva $R = R^T > 0$, $R \in \mathfrak{R}^{n \times n}$, al último término del lado derecho de (2.15), y tomando $R = P$, entonces:

$$\begin{aligned} 2U^T P W f(U) &\leq U^T P U + (P W f(U))^T P^{-1} (P W f(U)) \\ &= U^T P U + (P W f(U))^T W f(U) \\ &= U^T P U + (W f(U))^T P (W f(U)) \end{aligned} \quad (2.17)$$

por lo tanto:

$$2U^T P W f(U) \leq U^T P U + (W f(U))^T P (W f(U)) \quad (2.18)$$

sustituyendo (2.18) en (2.15) se obtiene:

$$\frac{dV}{dt} \leq -U^T P U + (W f(U))^T P (W f(U)) \quad (3.19)$$

De la igualdad de Lyapunov previamente mencionada, con $P = \frac{1}{2} I$, se tiene lo siguiente:

$$\begin{aligned} \frac{dV}{dt} &\leq -\frac{1}{2} U^T U + \frac{1}{2} (W f(U))^T (W f(U)) \\ &= -\frac{1}{2} U_2^2 + \frac{1}{2} \|W f(U)\|_2^2 \\ &\leq -\frac{1}{2} U_2^2 + \|W\|_2^2 \|f(U)\|_2^2 \end{aligned} \quad (2.20)$$

considerando (2.13), $U_2^2 \geq \|f(U)\|_2^2$, por lo que:

$$\begin{aligned} \frac{dV}{dt} &\leq -\frac{1}{2} U_2^2 + W_2^2 U_2^2 \\ &= \left(-\frac{1}{2} + \frac{1}{2} W_2^2 \right) U_2^2 \end{aligned} \quad (2.21)$$

Para que $\frac{dV}{dt}$ sea estrictamente negativa, se requiere que: $\left(-\frac{1}{2} + \frac{1}{2} W_2^2 \right) < 0$, lo que implica que $W_2 < 1$.

Observación: En esta deducción se usa la norma dos, pero puede usarse cualquier otra norma ya que las normas en \mathfrak{R}^n son equivalentes.

□□□

Este resultado es similar al reportado en [8]. Es evidente que, si en (2.14) se hace $P=I$, la demostración es más simple (fácil); no obstante hemos preferido la anterior presentación, pues permite extender fácilmente el análisis a otros modelos matemáticos de redes neuronales dinámicas.

2.4- Aplicaciones

Para ilustrar los resultados obtenidos se plantea el siguiente problema:

Ejemplo 2.1

Sea la red neuronal:

$$\begin{aligned} \dot{x}_1 &= -x_1 - 0.6701\sigma(x_1) + 0.2673\sigma(x_2) \\ \dot{x}_2 &= -x_2 + 0.2673\sigma(x_1) - 0.6701\sigma(x_2) \end{aligned} \quad (2.22)$$

$$\text{donde, } x \in \mathfrak{R}^2, \sigma(x) = \frac{2}{1 + \exp(-0.5x)}, \quad w = \begin{pmatrix} -0.6701 & 0.2673 \\ 0.2673 & -0.6701 \end{pmatrix}$$

Para este problema, $w = \sqrt{\lambda_{\max}(w^T w)} = 0.9374$, como puede calcularse directamente; entonces, satisface las condiciones del teorema 2.1 y por lo tanto es globalmente asintóticamente estable (GAS). Como puede verse en la figura 8 las trayectorias de la red parten de diferentes estados iniciales y conforme el sistema evoluciona, las trayectorias convergen a un punto (estado) de equilibrio único, lo cual predice dicho teorema.

La respuesta de $x_1(t)$ contra $x_2(t)$, se presentan en la siguiente Fig.8.

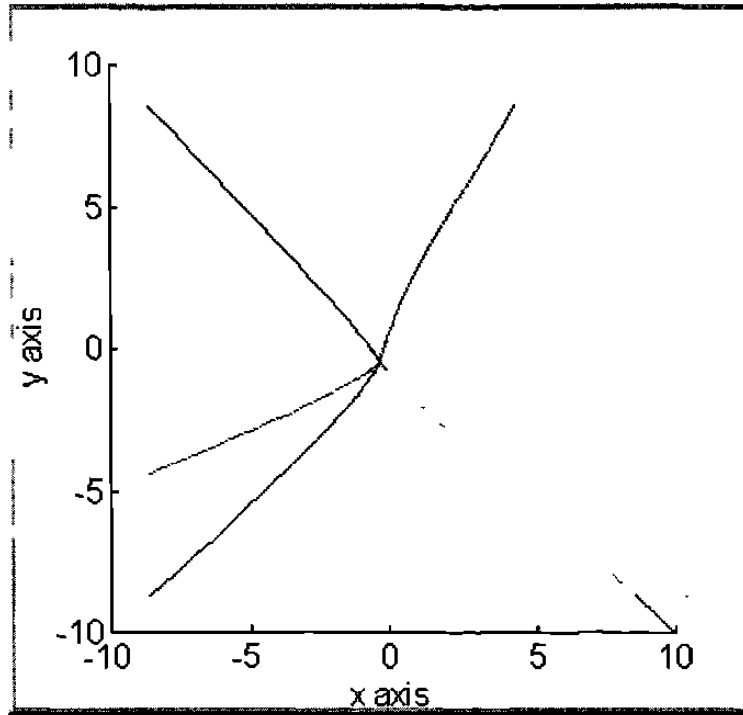


Fig.8 Plano fase de $x_1(t)$ contra $x_2(t)$ de un sistema GAS

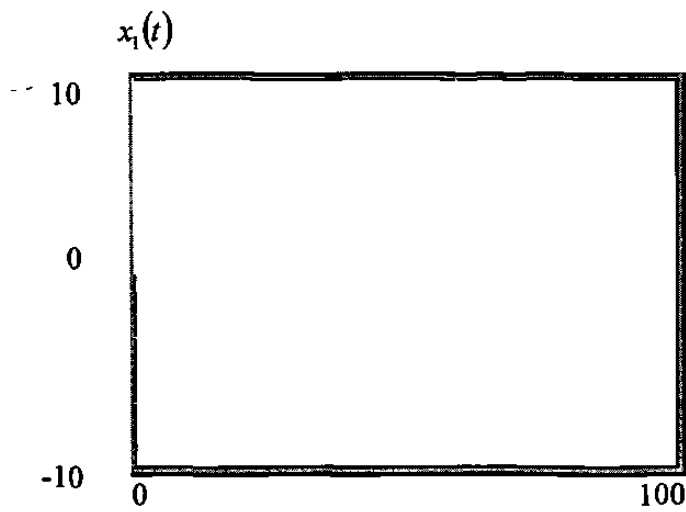


Fig.9 Respuesta de $x_1(t)$ de un sistema GAS

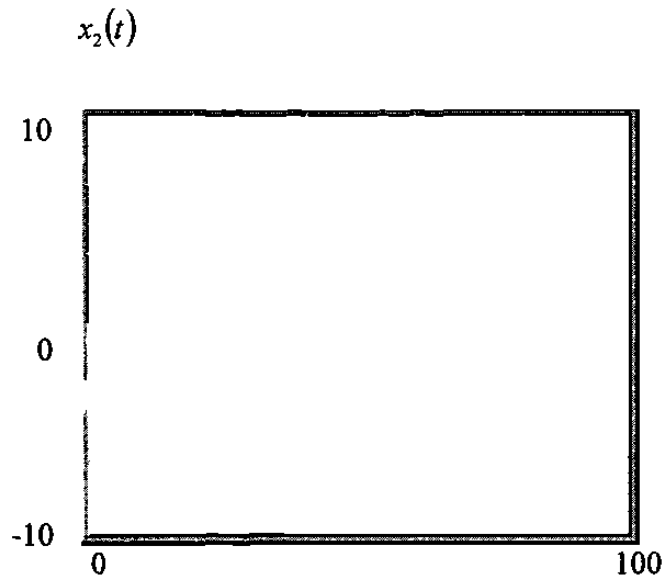


Fig.10 Respuesta de $x_2(t)$ de un sistema GAS

Las fig.9 y Fig.10 muestran las respuestas en el tiempo de $x_1(t)$ contra t y $x_2(t)$ contra t ; como puede observarse estas trayectorias convergen asintóticamente, los valores a los cuales convergen son: $x_1^* = -0.3778$ y $x_2^* = -0.3324$, respectivamente.

Nota: El eje de las abscisas representa el número de iteraciones; estos resultados se simularon con la herramienta SIMULINK de MATLAB.

2.5- Conclusiones

Como contribución principal de este capítulo, se establecen condiciones en la matriz de pesos W de interconexión de redes neuronales dinámicas, que garantizan estabilidad asintótica global. La función de Lyapunov que se utiliza es cuadrática, lo que permite continuar el análisis sobre estabilidad entrada-estado, en el siguiente capítulo.

CAPÍTULO 3

ANÁLISIS DE ESTABILIDAD ENTRADA–ESTADO PARA REDES NEURONALES DINÁMICAS

En este capítulo se presenta el análisis de estabilidad entrada-estado para redes neuronales dinámicas generalizadas. Inicialmente se demuestra la existencia de al menos un punto de equilibrio para la red neuronal, luego se determinan condiciones para satisfacer la propiedad de estabilidad entrada-estado, mediante una función de Lyapunov cuadrática. Se establece también que estas condiciones garantizan estabilidad asintótica global (G.A.S.), en el sentido clásico de estabilidad de Lyapunov.

3.1-Introducción

Existen dos metodologías para analizar la estabilidad de sistemas dinámicos: a) el enfoque entrada-salida (“input-output”) que se basa en técnicas de operadores y cuyo requisito típico es que el operador, que representa al sistema, sea acotado; b) el enfoque de espacio de estado, cuya herramienta fundamental es el análisis por medio de una función de Lyapunov. En contraste con los sistemas lineales, estos dos enfoques no son equivalentes para sistemas no lineales [14]. Un concepto que permite establecer equivalencias entre estos enfoques, aún para sistemas no lineales, es la estabilidad entrada-estado (ISS- “input to state stability”); dicho concepto fue presentado por primera vez en [15] y desde entonces ha sido empleado por diferentes autores para establecer resultados en control de sistemas no lineales [16,17]. La estabilidad entrada-estado puede ser establecida de diferentes maneras equivalentes como disipatividad, márgenes de robustez y en el sentido clásico de Lyapunov [18].

Como continuación al análisis previo sobre estabilidad asintótica global de redes neuronales dinámicas [19], en este capítulo, se presenta el análisis de estabilidad entrada-estado. El análisis se realiza de acuerdo al sentido clásico de estabilidad de Lyapunov.

El capítulo esta organizado como sigue: inicialmente se discute el modelo matemático de la red neuronal y se establece la existencia de al menos un punto de equilibrio, después se derivan condiciones para que la red neuronal previamente descrita sea estable entrada-estado; éstas se establecen en un teorema, que constituye la principal aportación del capítulo, finalmente se discuten las conclusiones relevantes.

3.2-Modelo Matemático

La Red Neuronal Dinámica está descrita por el siguiente modelo matemático.

$$\begin{aligned} \dot{x} &= Ax + W\sigma(x) + C\phi(x)\Gamma(u) \\ x \in \mathfrak{R}^n, u \in \mathfrak{R}^m, \sigma(x) \in \mathfrak{R}^n, \Gamma(u) \in \mathfrak{R}^l \\ A, W, C \in \mathfrak{R}^{n \times n}, \phi(x) \in \mathfrak{R}^{m \times l} \end{aligned} \quad (3.1)$$

donde A es la matriz de estado, W es la matriz de la retroalimentación no lineal, C es la matriz de entrada, y x es el estado de la red neuronal, $\sigma(x)$ y $\phi(x)$ son funciones no lineales.

Este modelo representa redes neuronales dinámicas altamente recurrentes como las que se analizan en [20,21]. Usualmente los elementos de $\sigma(x)$ y $\phi(x)$ se seleccionan sigmoideas, por ejemplo: $\sigma_i(x_i) = \frac{\alpha_i}{1 + e^{-\beta x_i}}$.

Para el análisis de estabilidad entrada-estado, se considera A matriz Hurwitz (todos sus valores característicos tienen parte real estrictamente menor que cero), $\Gamma(0) = 0$, y $\sigma(\cdot)$, $\phi(\cdot)$, $\Gamma(\cdot)$ son globalmente Lipschitz, esto es:

$$\begin{aligned} |\sigma(\Delta x - x) - \sigma(x)| &\leq L_\sigma \Delta x \\ |\phi(\Delta x - x) - \phi(x)| &\leq L_\phi \Delta x \\ |\Gamma(\Delta u - u) - \Gamma(u)| &\leq L_\Gamma \Delta u \\ \forall x, \Delta x \in \mathfrak{R}^n, \forall u, \Delta u \in \mathfrak{R}^m \end{aligned} \quad (3.2)$$

3.3-Punto de equilibrio

Inicialmente se establece que (3.1) presenta al menos un punto de equilibrio aislado.

LEMA 3.1- La red neuronal dinámica descrita por el modelo matemático (3.1), tiene al menos un punto de equilibrio para $u = 0$.

Prueba-

Se desea demostrar que (3.1) tiene al menos un punto de equilibrio, esto es $\exists x^* \in \mathbb{R}^n$, tal que $Ax^* + W\sigma(x^*) = 0$.

Sea $\phi_1(x) = W\sigma(x)$ se quiere probar que $Ax + \phi_1(x) = 0$, o bien $x = A^{-1}\phi_1(x)$ por lo tanto si se toma $\phi_2(x) = -A\phi_1(x)$, o bien: $\phi_2(x) = x$, entonces un punto fijo de $\phi_2(x)$ es un punto de equilibrio de (3.1).

Puesto que $\sigma(x)$ se supone globalmente Lipschitz, entonces es uniformemente acotada [22]; lo que implica que $\phi_2(x)$ también lo es.

A continuación se prueba que $\phi_2(x)$ tiene al menos un punto fijo.

Selecciónese:

$$\Omega = \{x \in \mathbb{R}^n : \|x\|_\infty \leq A^{-1} \|W\|_\infty M\} \quad (3.3)$$

donde $M = \max_i \sup_z \sigma_i(z)$, $\|\cdot\|_\infty$ se define como $\max x_i$ para vectores y como $\max_{i,j} a_{ij}$ para matrices. De (3.3) se tiene:

$$\begin{aligned} \|\phi_2(x)\|_\infty &= \| -A^{-1}W\sigma(x) \|_\infty \\ &\leq A^{-1} \|W\|_\infty \|\sigma(x)\|_\infty \\ &\leq A^{-1} \|W\|_\infty M \end{aligned} \quad (3.4)$$

lo cual implica que para todo $x \in \Omega$, $\phi_2(x) \in \Omega$ por lo tanto $\phi_2(x)$ es un mapeo continuo de un conjunto cerrado acotado y convexo Ω en el mismo; así, nuevamente por el teorema del punto fijo de Brower, $\phi_2(x)$ tiene al menos un punto fijo en Ω , este punto corresponde a un punto de equilibrio del sistema (3.1), con $u = 0$.

Sea este punto fijo x^* entonces:

$$Ax^* + W\sigma(x^*) = 0 \quad (3.5)$$

□□□

3.4-Análisis de estabilidad entrada-estado

En esta sección se analiza en que condiciones el sistema (3.1) tiene la propiedad de estabilidad entrada-estado estable (ISS - "Input to State Stability").

Sea x^* un punto de equilibrio de (3.1) con $u=0$ y definiendo $y = x - x^*$ entonces:

$$\dot{y} = \dot{x} = Ax + W\sigma(x) + C\phi(x)\Gamma(u) \quad (3.6)$$

Sumando y restando $W\sigma(x^*)$ y $C\phi(x^*)\Gamma(u)$ a (3.6) se obtiene :

$$\begin{aligned} \dot{y} = Ay + W(\sigma(y + x^*) - \sigma(x^*)) \\ + C(\phi(y + x^*) - \phi(x^*))\Gamma(u) \\ + C\phi(x^*)\Gamma(u) + Ax^* + W\sigma(x^*) \end{aligned} \quad (3.7)$$

Siendo la suma de los dos últimos términos del lado derecho de (3.7) igual a cero por (3.5), y teniendo en cuenta que $\Gamma(0)=0$, (3.7) se puede escribir como:

$$\begin{aligned} \dot{y} = Ay + W(\sigma(y + x^*) - \sigma(x^*)) \\ + C(\phi(y + x^*) - \phi(x^*))\Gamma(u) \\ + C\phi(x^*)\Gamma(u) - C\phi(x^*)\Gamma(0) \end{aligned} \quad (3.8)$$

La ecuación (3.8) es de la forma $\dot{y} = f(y, u)$ donde el punto de equilibrio es ahora $y = 0, u = 0$; esto es $f(0,0) = 0$, como puede comprobarse directamente.

Antes de realizar el análisis de estabilidad, se introducen los siguientes conceptos:

- a) Función clase K_∞ [6]. Una función continua $f : [0, a) \rightarrow [0, \infty)$ es clase K si es estrictamente creciente y $f(0) = 0$; y es clase

$$K_\infty \text{ si } a = \infty \text{ y } \lim_{r \rightarrow \infty} f(r) = \infty.$$

- b) Estabilidad Asintótica Global (GAS) [6]. Sea $x = 0$ un punto de equilibrio de $\dot{x} = f(x)$ y $D \subset \mathfrak{R}^n$ es la unión de conjuntos abiertos conexos, sin puntos en la frontera, el cual contiene a $x = 0$. Sea $V : D \rightarrow \mathfrak{R}$ una función continuamente diferenciable, tal que $V(0) = 0$, $V(x) > 0$ y $\dot{V} < 0$ en $D - \{0\}$, entonces $x = 0$ es asintóticamente estable. Si la propiedad es satisfecha para $D = \mathfrak{R}^n$, entonces es global asintóticamente estable.
- c) Función ISS Lyapunov[18]. Una función $V(x) : \mathfrak{R}^n \rightarrow \mathfrak{R}^+$ se dice ser una función ISS Lyapunov si existen funciones clase K_∞ $\alpha(r)$, $\theta(r)$ tales que $\dot{V} \leq -\alpha(|x|) + \theta(|u|) \forall x \in \mathfrak{R}^n, \forall u \in \mathfrak{R}^m$.

Se establece ahora un importante teorema acerca de la estabilidad entrada-estado (ISS) de (3.8), o en forma equivalente de (3.1).

Teorema 3.1- El sistema (3.8), o equivalentemente el sistema (3.1), es estable entrada-estado (ISS), si se cumple:

$$W^2 < \frac{\gamma - 3P}{L_\sigma^2 P}, |P| < \frac{\gamma}{3} \quad (3.9)$$

$$\gamma \in \mathfrak{R}^+, P = P^T > 0$$

donde P es la solución de la siguiente ecuación de Lyapunov $PA + A^T P = -\gamma I$. Esta solución existe y es única puesto que la matriz A es Hurwitz.

Prueba-

Para analizar la estabilidad entrada-estado de (3.8) (o en forma equivalente de (3.1)), se considera la siguiente función de Lyapunov candidata:

$$V(y) = y^T P y, P = P^T > 0 \quad (3.10)$$

La derivada con respecto al tiempo t de V en las trayectorias de (3.8) viene dada por:

$$\begin{aligned}
\dot{V} &= y^T \dot{P}y + \dot{y}^T Py \\
&= y^T \dot{P}v + v^T \dot{P}y \\
v &= Ay + W(\sigma(y + x^*) - \sigma(x^*)) \\
&\quad + C(\phi(y + x^*) - \phi(x^*))\Gamma(u) \\
&\quad + C\phi(x^*)(\Gamma(u) - \Gamma(0))
\end{aligned} \tag{3.11}$$

La cual se reduce a:

$$\begin{aligned}
\dot{V} &= -\gamma y^T y + 2y^T P\dot{W}(\sigma(y + x^*) - \sigma(x^*)) \\
&\quad + 2y^T PC(\phi(y + x^*) - \phi(x^*))\Gamma(u) \\
&\quad + 2y^T PC\phi(x^*)(\Gamma(u) - \Gamma(0))
\end{aligned} \tag{3.12}$$

Se analiza ahora el segundo, tercero y cuarto término del lado derecho de (3.12), utilizando nuevamente desigualdad matricial (2.16), con $R = P$:

Para el segundo término de (3.12), se tiene:

$$\begin{aligned}
2y^T (P\dot{W}[\sigma(y + x^*) - \sigma(x^*)]) &\leq \\
& y^T \dot{P}y + (P\dot{W}[\sigma(y + x^*) - \sigma(x^*)])^T P^{-1} (P\dot{W}[\sigma(y + x^*) - \sigma(x^*)]) \\
&= y^T \dot{P}y + (P\dot{W}[\sigma(y + x^*) - \sigma(x^*)])^T (W[\sigma(y + x^*) - \sigma(x^*)]) \\
&= y^T \dot{P}y + (W[\sigma(y + x^*) - \sigma(x^*)])^T P(W[\sigma(y + x^*) - \sigma(x^*)]) \\
&\leq P \|y\|^2 + P \|W[\sigma(y + x^*) - \sigma(x^*)]\|^2 \\
&\leq P \|y\|^2 + L_\sigma^2 \|P W\|^2 \|y\|^2
\end{aligned} \tag{3.13}$$

Por lo tanto:

$$2y^T P\dot{W}[\sigma(y + x^*) - \sigma(x^*)] \leq P \|y\|^2 + L_\sigma^2 \|P W\|^2 \|y\|^2 \tag{3.14}$$

Para el tercer término de (3.12), se obtiene:

$$\begin{aligned}
2y^T (PC[\phi(y+x^*)-\phi(x^*)])\Gamma(u) &\leq \\
&\cdot y^T Py + (PC[\phi(y+x^*)-\phi(x^*)])\Gamma(u))^T P^{-1} (PC[\phi(y+x^*)-\phi(x^*)])\Gamma(u) \\
&= y^T Py + (PC[\phi(y+x^*)-\phi(x^*)])\Gamma(u))^T (C[\phi(y+x^*)-\phi(x^*)])\Gamma(u) \\
&= y^T Py + (C[\phi(y+x^*)-\phi(x^*)])\Gamma(u))^T P(C[\phi(y+x^*)-\phi(x^*)])\Gamma(u)
\end{aligned} \tag{3.15}$$

Siendo $\phi(x)$ globalmente Lipschitz, entonces es uniformemente acotada [22], esto es:

$$\|\phi(y+x^*)-\phi(x^*)\| \leq K_\phi, \quad \forall y \in \mathbb{R}^n \tag{3.16}$$

y entonces el lado derecho de la última igualdad de (3.15) puede ser formulado como:

$$\begin{aligned}
2y^T (PC[\phi(y+x^*)-\phi(x^*)])\Gamma(u) &\leq P\|y\|^2 + PC[\phi(y+x^*)-\phi(x^*)]\Gamma(u)^2 \\
&\leq P\|y\|^2 + PC^2\|\phi(y+x^*)-\phi(x^*)\|^2\Gamma(u)^2 \\
&\leq P\|y\|^2 + K_\phi^2 L_T^2 PC^2\|u\|^2
\end{aligned} \tag{3.17}$$

Finalmente:

$$2y^T (PC[\phi(y+x^*)-\phi(x^*)])\Gamma(u) \leq P\|y\|^2 + K_\phi^2 L_T^2 P\|C\|^2\|u\|^2 \tag{3.18}$$

Para el cuarto término de (3.12), se tiene:

$$\begin{aligned}
2y^T (PC\phi(x^*)[\Gamma(u)-\Gamma(0)]) &\leq \\
&y^T Py + (PC\phi(x^*)[\Gamma(u)-\Gamma(0)])^T P^{-1} (PC\phi(x^*)[\Gamma(u)-\Gamma(0)]) \\
&= y^T Py + (PC\phi(x^*)[\Gamma(u)-\Gamma(0)])^T (C\phi(x^*)[\Gamma(u)-\Gamma(0)]) \\
&= y^T Py + (C\phi(x^*)[\Gamma(u)-\Gamma(0)])^T P(C\phi(x^*)[\Gamma(u)-\Gamma(0)]) \\
&\leq P\|y\|^2 + P\|C\phi(x^*)[\Gamma(u)-\Gamma(0)]\|^2 \\
&\leq P\|y\|^2 + PC^2\|\phi(x^*)\|^2\|\Gamma(u)-\Gamma(0)\|^2
\end{aligned} \tag{3.19}$$

La cual puede ser escrita como:

$$2y^T (PC\phi(x^*)[\Gamma(u)-\Gamma(0)]) \leq P\|y\|^2 + L_T^2 P\|C\|^2\|\phi(x^*)\|^2\|u\|^2 \tag{3.20}$$

Sustituyendo (3.14), (3.18) y (3.20) en (3.12), se obtiene finalmente:

$$\begin{aligned} \dot{V} \leq & -\left(\gamma - 3P - L_\sigma^2 P W^2\right) y^2 \\ & + \left(K_\phi^2 L_\tau^2 P C^2 + L_\tau^2 P C^2 \phi(x^*)^2\right) |u|^2 \end{aligned} \quad (3.21)$$

Definiendo las siguientes funciones:

$$\begin{aligned} \alpha(r) &= \left(\gamma - 3P - L_\sigma^2 P W^2\right) r^2 \\ \theta(r) &= \left(K_\phi^2 L_\tau^2 P C^2 + L_\tau^2 P C^2 \phi(x^*)^2\right) r^2 \end{aligned} \quad (3.22)$$

$r \in \mathfrak{R}$
entonces:

$$\dot{V} \leq -\alpha(x) + \theta(u) \quad (3.23)$$

Como están definidas $\alpha(\cdot)$ y $\theta(\cdot)$, satisfacen las condiciones de clase K_∞ , entonces el sistema (3.8) o en forma equivalente el sistema (3.1) es ISS, si:

$$\left(\gamma - 3P - L_\sigma^2 P W^2\right) > 0 \quad (3.24)$$

lo cual implica que:

$$W^2 < \frac{\gamma - 3P}{L_\sigma^2 P}, \quad P < \frac{\gamma}{3} \quad (3.25)$$

Corolario 3.1-Para que el punto de equilibrio $y=0$ de (3.8) (o en forma equivalente $x = x^*$ de (3.1)) sea globalmente asintóticamente estable (GAS), la desigualdad (3.25) debe ser satisfecha.. □□□

Prueba-

Haciendo en (3.23) $u = 0$ se tiene:

$$\dot{V} \leq -\alpha(x) \quad (3.26)$$

$\dot{V} < 0, \forall x \in \mathfrak{R}^n$, [6] garantiza que (3.8) es globalmente asintóticamente estable (GAS), para cumplir con esta condición se requiere que la desigualdad (3.26) sea satisfecha.

□□□

Una muy interesante aplicación de la propiedad ISS , es la estabilidad de sistemas no lineales interconectados, como se discute en [17].

3.5- Aplicaciones

Para ilustrar los resultados se presenta el siguiente ejemplo en donde se considera primero que la entrada a la red neuronal no es idénticamente cero y posteriormente se considera que la entrada a dicha red es idénticamente cero.

Ejemplo 3.1

Sea la red neuronal:

$$\begin{aligned}
 X &= AX + W\sigma(X) + \Gamma(u), \text{ donde} \\
 A &= -\frac{1}{2}I, \quad W = \begin{pmatrix} -0.3351 & 0.1337 \\ 0.1337 & -0.3351 \end{pmatrix} \quad (3.27) \\
 \sigma(X) &= \begin{pmatrix} 2 / \left(1 + \exp\left(-\frac{1}{2}x\right) \right) \\ 2 / \left(1 + \exp\left(-\frac{1}{2}y\right) \right) \end{pmatrix}, \quad \Gamma(u) = \begin{pmatrix} \text{sen}(u) \\ 1 - \text{cos}(u) \end{pmatrix}
 \end{aligned}$$

$$X \in \mathfrak{R}^2, A, W \in \mathfrak{R}^{2 \times 2}, \sigma(\cdot), \Gamma(\cdot) \in \mathfrak{R}^2$$

Es claro que la norma de la matriz W satisface la desigualdad (3.25),

$$\|W\|^2 < \frac{\gamma - 2\gamma^2}{L_\sigma^2}. \text{ En este ejemplo, } \|W\| = 0.4687, L_\sigma = \frac{1}{4} \text{ y se toma } \gamma = \frac{1}{4}, \text{ claramente la}$$

desigualdad (3.25) es satisfecha.

La gráfica en el plano fase de las respuestas de x_1 contra x_2 de (3.27) se da en la Fig.11.

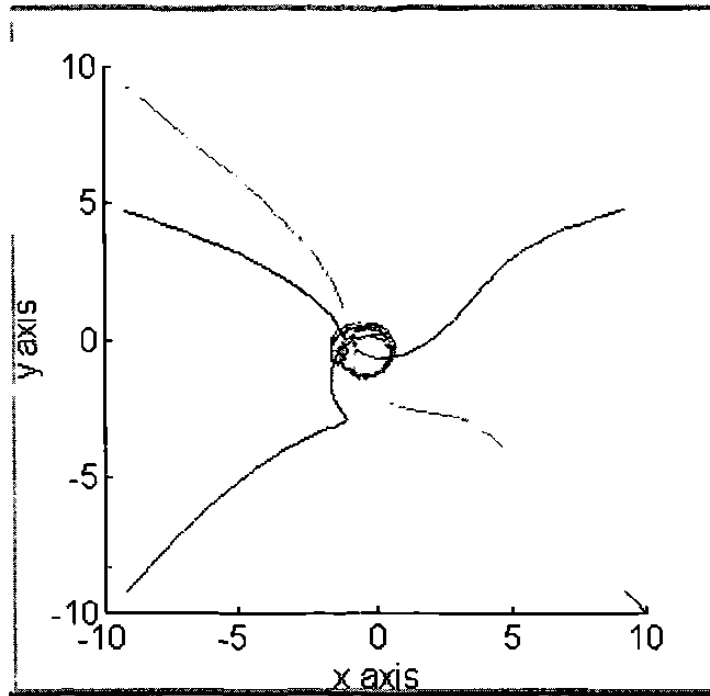


Fig.11 Plano fase de x_1 contra x_2 de un sistema ISS

Como puede observarse, las trayectorias de la red neuronal parten de un estado inicial, y conforme el tiempo transcurre estas permanecen acotadas. Este resultado era de esperarse ya que este sistema es ISS.

Las respuestas de $x_1(t)$ y de $x_2(t)$ con respecto a t están dadas en las Fig.12 y Fig.13.

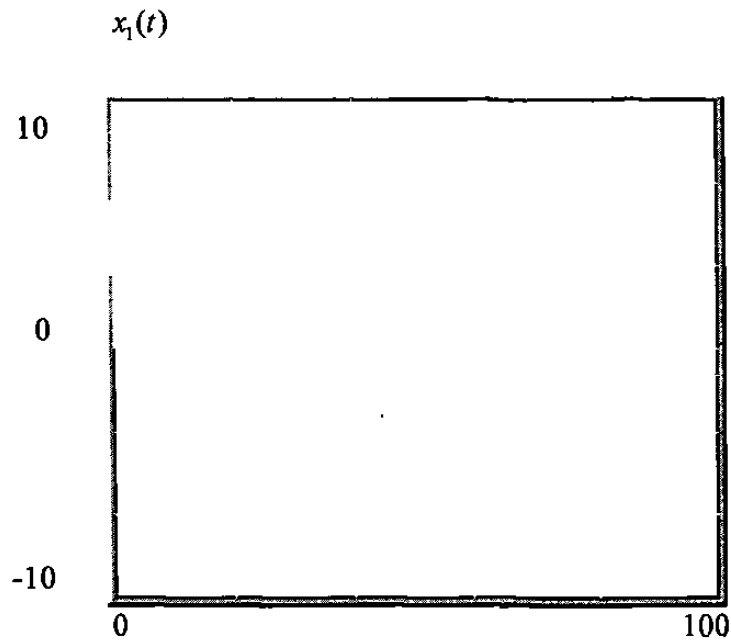


Fig.12 Respuesta de $x_1(t)$ de un sistema ISS

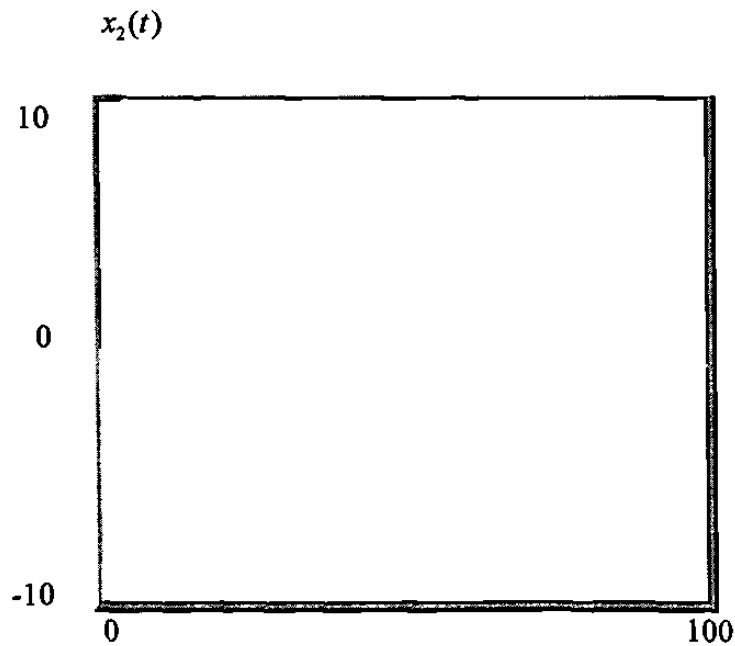


Fig.13 Respuesta $x_2(t)$

Como puede observarse en las Fig.12 y Fig.13 $x_1(t)$ y $x_2(t)$ son acotadas, como era de esperarse ya que el sistema que se presenta en el ejemplo es ISS.

Nota: El eje de las abscisas en la Fig.12 y Fig.13 representa el número de iteraciones.

Ahora consideremos el caso en el cual la entrada de (3.27) es idénticamente cero lo cual lleva al sistema:

$$X = AX + W\sigma(X), \text{ donde} \quad (3.28)$$

$$A = -\frac{1}{2}I, \quad W = \begin{pmatrix} -0.3351 & 0.1337 \\ 0.1337 & -0.3351 \end{pmatrix}$$

$$\sigma(X) = \begin{pmatrix} 2 / \left(1 + \exp\left(-\frac{1}{2}x\right) \right) \\ 2 / \left(1 + \exp\left(-\frac{1}{2}y\right) \right) \end{pmatrix}$$

$$X \in \mathbb{R}^2, \quad A, W \in \mathbb{R}^{2 \times 2}, \quad \sigma(\cdot), \Gamma(\cdot) \in \mathbb{R}^2$$

Nuevamente, como en el caso en el cual la entrada no es cero, es claro que la norma de la matriz W satisface la desigualdad $W^2 < \frac{\gamma - 2\gamma^2}{L_\sigma^2}$; en este caso,

$$W = 0.4687, \quad L_\sigma = \frac{1}{4} \text{ y se toma } \gamma = \frac{1}{4}.$$

La gráfica en el plano fase de las respuestas de x_1 contra x_2 de (3.28) se da en la Fig.14.

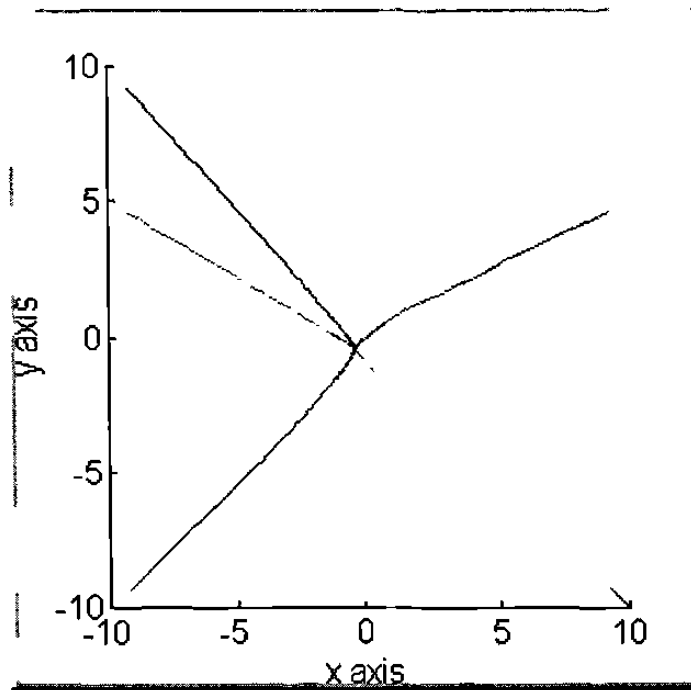


Fig. 14: Plano fase de $x_1(t)$ contra $x_2(t)$ de un sistema GAS

Como puede observarse, las trayectorias del sistema del ejemplo parten de un estado inicial (cualquiera), y conforme el tiempo transcurre estas convergen asintóticamente a un estado de equilibrio en el plano fase. Este resultado era de esperarse ya que el sistema con entrada idénticamente cero es globalmente asintóticamente estable (GAS).

Las respuestas de $x_1(t)$ y de $x_2(t)$ con respecto a t están dadas en la Fig.15 y Fig.16.

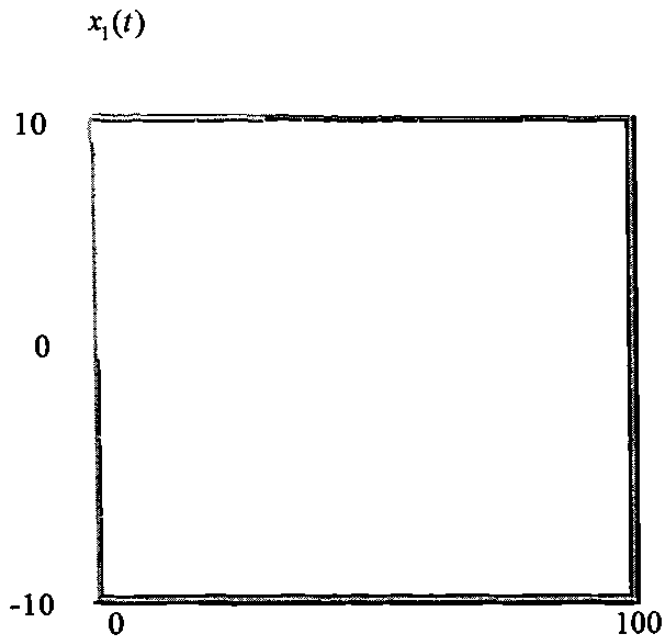


Fig.15 Respuesta de $x_1(t)$ de un sistema GAS

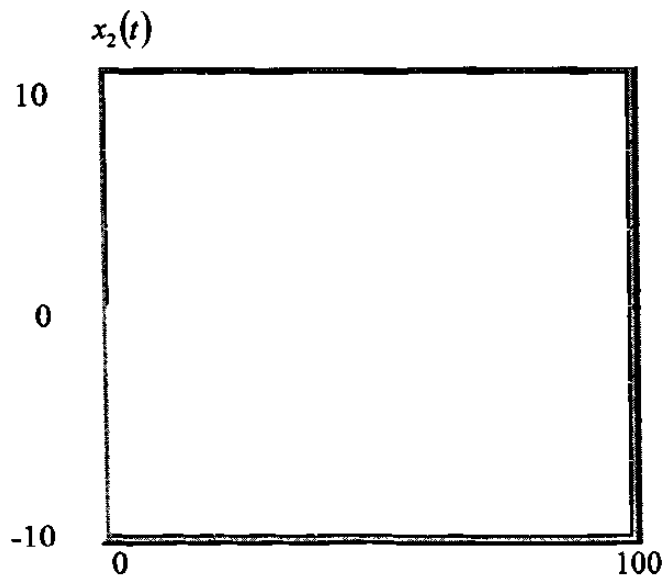


Fig.16 Respuesta de $x_2(t)$ de un sistema GAS

Como puede observarse $x_1(t)$ y $x_2(t)$ en las Fig.15 y Fig.16 convergen asintóticamente, los valores a los cuales convergen son: $x_1^* = -0.3372$ y $x_2^* = -0.4485$, respectivamente, como era de esperarse ya que el sistema (3.28) es GAS.

Nota: El eje de las abscisas de las Fig.15 y Fig.16, representa el número de iteraciones.

3.6- Conclusiones del capítulo

Como contribución principal de este capítulo se establecen condiciones en la matriz de pesos de interconexión W de redes neuronales dinámicas que garantizan estabilidad entrada estado. Como caso particular, para entradas idénticamente iguales a cero ($u = 0$), estas condiciones también garantizan que el punto de equilibrio es globalmente asintóticamente estable.

La propiedad ISS permite la aplicación de redes neuronales dinámicas a control no lineal. Una posible aplicación es desarrollar esquemas de control basados en la técnica del control por modelo (IMC, de su nombre en ingles).

CAPITULO 4

IDENTIFICACIÓN ADAPTABLE ROBUSTA DE SISTEMAS NO LINEALES Y SEGUIMIENTO DE TRAYECTORIAS USANDO REDES NEURONALES DINÁMICAS

En este capítulo, se analiza la identificación adaptable robusta no lineal y seguimiento de trayectorias usando una red neuronal dinámica, con la misma dimensión del espacio de estado que el sistema. Se supone que el espacio de estado del sistema es completamente medible. La dinámica del sistema no lineal nominal es perturbado por un término adicional, representando las incertidumbres del sistema. Por medio del análisis de estabilidad desde el punto de vista de Lyapunov, se determinan condiciones de estabilidad para el error de identificación, luego se analiza la estabilidad del error de seguimiento de las trayectorias para un sistema no lineal previamente identificado; la trayectoria es generada por un modelo no lineal. Para el análisis de identificación se utiliza una ecuación algebraica de Riccati y para el error de seguimiento una ecuación diferencial de Riccati. La estructura del esquema está compuesto en dos partes: el identificador basado en la red neuronal y el controlador de seguimiento. Como aportación principal se establecen dos teoremas: en el primero se da una cota para el error de identificación; y en el segundo, se establece una cota para el error de seguimiento.

4.1 Introducción

Recientemente, existe un gran interés en aplicaciones de redes neuronales en identificación y control de sistemas no lineales. La identificación de sistemas no lineales puede ser abordada como la aproximación del desempeño de sistemas por redes neuronales dinámicas. En esta dirección existen ya algunos resultados. Estos pueden ser clasificados en dos grupos: el primero, como una extensión natural, se basa en propiedades de aproximación de funciones de redes neuronales estáticas [23,24] y está limitado a que el tiempo pertenezca a un conjunto cerrado. En el segundo, se usa la representación por operadores del sistema, para derivar condiciones de la validez de la aproximación por una

red neuronal dinámica. Este segundo enfoque ha sido extensamente analizado por I. W. Sandberg, tanto para tiempo continuo como para tiempo discreto ([25,26] y referencias). La estructura propuesta está constituida por la conexión paralela de neuronas, las cuales no interactúan entre ellas; se requiere que el sistema no lineal satisfaga la condición de memoria aproximadamente-finita. En [27], una red neuronal dinámica, basada en el modelo de Hopfield, fue propuesta para la identificación de sistemas no lineales usando representación operacional; la propiedad de aproximación se propuso como una conjetura. Usando la condición de memoria desvaneciente, esta conjetura fue parcialmente probada en [28]. Ambos conceptos: memoria aproximadamente-finita y memoria desvaneciente, requieren que el sistema no lineal sea estable, en lazo abierto.

Los resultados mencionados arriba dan condiciones para la existencia de una red neuronal dinámica, la cual minimiza el error de aproximación del desempeño del sistema no lineal; sin embargo no se determina el número de neuronas i/o el valor de sus pesos que efectivamente obtienen el error mínimo. Un resultado reciente [29] resuelve el problema del número de neuronas por medio de redes neuronales recursivas de alto orden. Este número se selecciona igual a la dimensión del estado del sistema no lineal, el cual es completamente medible. Esta condición de accesibilidad es simplificada en [20] para sistemas con perturbaciones singulares. Para estos resultados, también se requiere que el tiempo pertenezca a un conjunto cerrado.

A pesar de lo promisorio de las aplicaciones de neuro-control y que, aún para información neuronal, se utilizan funciones de energía para demostrar la convergencia de los valores finales deseados, existen muy pocos resultados sobre estabilidad en neuro-control [30].

Respecto a los sistemas de control no lineales por redes neuronales dinámicas, son muy pocos los resultados publicados. Los más importantes son los de M. A Christodoulo y colaboradores [20,31 y referencias], donde se utiliza una versión particular de redes neuronales recursivas de alto orden. En [20] se identifica un sistema no lineal por medio de redes neuronales dinámicas, luego se calcula la ley de control, basado en el modelo de la

red neuronal, forzando al sistema a seguir un modelo lineal. Los pesos de la red son adaptados en línea para minimizar el error de identificación. La estabilidad de todo el sistema se analiza vía una función de Lyapunov, En [31], se desarrolla un esquema regulador para un sistema no lineal afín en el control; una vez más se analiza estabilidad usando una función de Lyapunov. En ambos artículos, se ilustra la aplicabilidad del esquema propuesto por el control de velocidad de un motor de D.C.. Otros resultados [32,33] utilizan la representación de un sistema no lineal SISO afín en el control, el cual es aproximado por una red neuronal. Esta red es linearizada por un diseño de un lazo interno basado en técnicas de Geometría Diferencial [34], y la ley de control externo es implementada usando un controlador PID.

En [35] se analizan tanto identificación como control de sistemas no lineales. Para control se fuerza al sistema a que siga una señal generada por un modelo no lineal. El análisis de estabilidad del error de identificación y del error de seguimiento es formulado por el método de Lyapunov. Es importante mencionar que la metodología del análisis de estabilidad que aquí se usa es similar a una introducida por A.N. Mitchel and Coworkers [36,37].

En este capítulo se extienden los resultados previos, a sistemas no-lineales perturbados y además se consigue aprendizaje en línea de los parámetros de la red neuronal, los cuales son usados en el control de seguimiento. Este capítulo está organizado como sigue: primero la estabilidad del error de identificación robusta es analizada, luego una cota para el error de seguimiento adaptable es establecida. Finalmente, se ilustra la aplicabilidad de estos resultados con un ejemplo, y se discuten las conclusiones respectivas.

4.2.-Análisis de la estabilidad del error de identificación adaptativa robusta

Se supone la siguiente estructura de la red neuronal, la cual incluye las redes neuronales recursivas de alto orden como en [20]. Se adaptan solo los parámetros de la matriz de retroalimentación no lineal, los cuales son variables en el tiempo. Se adaptan estos parámetros usando los mismos algoritmos que en [20]:

$$\begin{aligned}\dot{x}_{nn} &= A_{nn}x_{nn} + B_{nn}(t)\sigma(x_{nn}) + C_{nn}\phi(x_{nn})\gamma(u_{nn}) \\ x_{nn} &\in \mathfrak{R}^n, u_{nn} \in \mathfrak{R}^m, \sigma(x_{nn}) \in \mathfrak{R}^n, \phi(x_{nn})\gamma(u_{nn}) \in \mathfrak{R}^l \\ A_{nn}, B_{nn} &\in \mathfrak{R}^{n \times n}, C_{nn} \in \mathfrak{R}^{n \times l}\end{aligned}\quad (4.1)$$

donde todas las variables, parámetros, y símbolos son definidos como en [13].

El sistema no lineal a ser identificado esta dado por:

$$\dot{x} = f_0(x, u, t) + f_1(x, u, t), \quad x \in \mathfrak{R}^n, u \in \mathfrak{R}^m \quad (4.2)$$

donde $f_0(x, u, t)$ es el campo vectorial del sistema nominal y $f_1(x, u, t)$ representa las incertidumbres del sistema.

Se supone que el estado del sistema es completamente medible, y que la dimensión del estado y de la entrada iguales para la red neuronal y para el sistema no-lineal. Bajo estas condiciones, se define el error de aproximación como:

$$\Delta = x - x_{nn}, \quad \dot{\Delta} = \dot{x} - \dot{x}_{nn} \quad (4.3)$$

Sumando y restando el término Ax , se obtiene:

$$\begin{aligned}\dot{\Delta} &= A\Delta + [f_0(x, u, t) + f_1(x, u, t) \\ &\quad - A_{nn}x_{nn} - B_{nn}(t)\sigma(x_{nn}) - C_{nn}\phi(x_{nn})\gamma(u) - A\Delta]\end{aligned}\quad (4.4)$$

con A cualquier matriz Hurwitz, y $u_{nn} = u$.

Bajo estas condiciones, la ODE que describe el desempeño del error es:

$$\begin{aligned}\Delta &= A\Delta + h(\Delta, x, u, t) \\ h(\Delta, x, u, t) &= h_0(\Delta, x, u, t) + f_1(x, u, t) \\ h_0(\Delta, x, u, t) &= f_0(x, u, t) - A_{nn}x - (A - A_{nn})\Delta \\ &\quad - B_{nn}(t)\sigma(x - \Delta) - C_{nn}\phi(x - \Delta)\gamma(u)\end{aligned}\quad (4.5)$$

Se supone que las siguientes condiciones son satisfechas:

Suposición No 4.1 – Como en [35], existen matrices definidas positivas H_{n0} y $H_{\Delta 0}$ tal que:

$$h_0^T(\Delta, x, u, t)H_{n0}h_0(\Delta, x, u, t) \leq \varepsilon_{h_0}(x, u, t) + \varepsilon_{h_1}(x, u, t)\Delta^T H_{\Delta 0}\Delta \quad (4.6)$$

para $\varepsilon_{h_0}(\cdot, \cdot, \cdot, \cdot)$ y $\varepsilon_{h_1}(\cdot, \cdot, \cdot, \cdot)$ funciones acotadas positivas, con ε^{h_0} y ε^{h_1} las respectivas cotas, i.e:

$$\begin{aligned} \sup \varepsilon_{h_i}(x, u, t) &= \varepsilon^{h_i} < \infty, i = 0, 1 \\ \forall x \in \mathfrak{R}^n, \forall u \in \mathfrak{R}^m, \forall t \geq 0 \end{aligned} \quad (4.7)$$

Suposición No 4.2 – Existe una matriz positiva definida H_{nf} tal que:

$$f_1^T(x, u, t)H_{nf}f_1(x, u, t) \leq \varepsilon_{f_0}(x, u, t) \quad (4.8)$$

para $\varepsilon_{f_0}(\cdot, \cdot, \cdot, \cdot)$ y $\varepsilon_{f_1}(\cdot, \cdot, \cdot, \cdot)$ funciones positivas acotadas, con ε^{f_0} y ε^{f_1} las respectivas cotas, i.e:

$$\begin{aligned} \sup \varepsilon_{f_i}(x, u, t) &= \varepsilon^{f_i} < \infty, i = 0, 1 \\ \forall x \in \mathfrak{R}^n, \forall u \in \mathfrak{R}^m, \forall t \geq 0 \end{aligned} \quad (4.9)$$

Suposición No 4.3– El campo vectorial nominal (f_0) satisface la siguiente condición de "función en banda" [13]¹:

$$\begin{aligned} \left\| f_0(x, u, t) - A_{n_n}x - (A - A_{n_n}) - B_{n_n}(t)\sigma(x - \Delta) - C_{n_n}\phi(x - \Delta)\gamma(u) \right\| \leq L_f \\ \forall x \in \mathfrak{R}^n, \forall u \in \mathfrak{R}^m, \forall t \geq 0 \end{aligned} \quad (4.10)$$

Suposición No 4.4– Las funciones $\phi(x)$ y $\sigma(x)$ son Lipschitz, p.e.

$$\begin{aligned} \|\phi(x) - \phi(x - \Delta)\| \leq L_\phi \|\Delta\|, \quad \|\sigma(x) - \sigma(x - \Delta)\| \leq L_\sigma \|\Delta\| \\ \forall x, \Delta \in \mathfrak{R}^n \end{aligned} \quad (4.11)$$

Suposición No 4.5– Como es usual en análisis de identificación, $u(t)$ se considera acotada para cualquier tiempo t positivo; esto es:

¹En este trabajo: a) $\|\cdot\|$ denota la norma euclídeana para vectores, y para cualquier matriz A se define como:

$\sqrt{\lambda_{\max}(A^T A)}$, donde $\lambda_{\max}(\cdot)$ es el máximo eigenvalor de la matriz respectiva; b) $\|\cdot\|_w$ es la norma euclídeana pesada.

$$\|\gamma(u)\| \leq L_u \quad (4.12)$$

Suposición No 6 – La matriz $B_{nn}(t)$ es acotada, esto es:

$$\|B_{nn}(t)\| \leq B \quad \forall t \geq 0$$

En [20] se establecen condiciones que validan esta suposición.

Suposición No 4.7– Existe una matriz estrictamente definida positiva Q_0 tal que, para la matriz A arriba definida, la siguiente ecuación matricial de Riccati:

$$\begin{aligned} A^T P_I + P_I A + R P_I + Q &= 0 \\ R &\equiv H_{n0}^{-1} + H_{nf}^{-1}, Q \equiv \varepsilon^{b1} H_{\Delta 0} + Q_0 \end{aligned} \quad (4.13)$$

tiene una solución $P_I = P_I^T$

De acuerdo con [39] tal solución existe si la matriz A es estable, el par $(A, R^{1/2})$ es controlable y el par $(Q^{1/2}, A)$ es observable. Fácilmente estas condiciones son satisfechas si se selecciona la matriz A sea diagonal.

Comentario No 4.1- Todas las matrices $H_{(.)}$ normalizan al vector respectivo a uno con componentes adimensionales.

Comentario No 4.2- Las ecuaciones (4.6) y (4.8) representan restricciones del tipo elipsoidal para los respectivos campos vectoriales.

Lema 4.1- En la desigualdad (4.6), para cualquier valor fijo H_n , $\chi > 0$, y bajo las suposiciones No 2 y No 3, se tiene:

$$\begin{aligned} \varepsilon_{h0}(x, u, t) &= (1 + \chi) \|f_0(x, u, t) \\ &\quad - A_{nn}x - B_{nn}(t)\sigma(x) - C_{nn}\phi(x)\gamma(u)\|_{H_{n0}}^2 \\ \varepsilon^{h0} &\equiv \sup_{x, u, t} \varepsilon_{h0}(x, u, t) \leq (1 + \chi) \|H_{n0}\| L_f^2 \\ \varepsilon_{h1}(x, u, t) &= \varepsilon^{h1} = (1 + \chi^{-1}) (\|B_{nn}(t)\| L_\sigma + \|C_{nn}\| L_\phi L_u \\ &\quad + \|A - A_{nn}\|)^2 \|H_{n0}\| \|H_{\Delta 0}^{-1}\| \end{aligned} \quad (4.14)$$

Prueba- Definiendo $h_0(\Delta, x, u, t)$ como:

$$\begin{aligned} h_0(\Delta, x, u, t) &= h_1(x, u, t) + h_2(\Delta, x, u) \\ h_1(x, u, t) &= f(x, u, t) - A_{nn}x - B_{nn}(t)\sigma(x) - C_{nn}\phi(x)\gamma(u) \\ h_2(\Delta, x, u) &= B_{nn}(t)(\sigma(x) - \sigma(x - \Delta)) - C_{nn}(\phi(x) - \phi(x - \Delta))\gamma(u) \end{aligned} \quad (4.15)$$

y con $h_0 = h_0(\Delta, x, u, t)$, $h_1 = h_1(x, u, t)$, $h_2 = h_2(\Delta, x, u)$, entonces:

$$\|h_0\|_{H_{n0}}^2 = \|h_1 + h_2\|_{H_{n0}}^2 \leq (1 + \chi)\|h_1\|_{H_{n0}}^2 + (1 + \chi^{-1})\|h_2\|_{H_{n0}}^2 \quad (4.16)$$

Así:

$$\varepsilon_{h_0}(x, u, t) = (1 + \chi)\|h_1\|_{H_{n0}}^2 \quad (4.17)$$

La cota ε^{h_0} se obtiene inmediatamente.

Para el término $h_2^2_{H_{n0}}$, nosotros formulamos:

$$\begin{aligned} \|h_2^2_{H_{n0}}\| &\leq B_{nn}(t)(\sigma(x) - \sigma(x - \Delta)) \\ &\quad + C_{nn}(\phi(x) - \phi(x - \Delta))\gamma(u) - (A - A_{nn})\Delta^2 H_{n0} \end{aligned} \quad (4.18)$$

y:

$$\|h_2^2_{H_{n0}}\| \leq (BL_\sigma|\Delta + C_{nn}L_\phi\Delta L_u + A - A_{nn}\Delta)^2 \quad (4.19)$$

de:

$$\|\Delta\|^2 = \Delta^T H_{\Delta 0}^2 H_{\Delta 0}^{-1} H_{\Delta 0}^2 \Delta \leq \lambda_{\max}(H_{\Delta 0}^{-1})\Delta^T H_{\Delta 0} \Delta \leq \|H_{\Delta 0}^{-1}\| \|\Delta\|_{\Delta 0}^2 \quad (4.20)$$

Se sigue inmediatamente la cota ε^{h_1}

□□□

Bajo las suposiciones antes establecidas y el Lema 4.1, se establece el primer teorema.

Teorema 4.1

Para el sistema (4.2) y la red neuronal dada (4.1), las siguientes propiedades son satisfechas:

$$\limsup_{t \rightarrow \infty} \|\Delta(t)\| \leq \Delta^0(P_I) \equiv \sqrt{\frac{\varepsilon^0}{\lambda_{\max}(P_I)\lambda_{\min}(R_p)}} \quad (4.21)$$

donde $R_p = P^{-1/2} Q \sigma^p I^{1/2}$ y $\lambda_{\min}(\cdot)$ es el mínimo valor propio de la matriz respectiva.

Prueba-

Consideremos la siguiente función escalar definida no negativa $V(\Delta) \in \mathfrak{R}^+$,

$$V(\Delta) = \Delta^T P_I \Delta, P_I = P_I^T > 0, P_I \in \mathfrak{R}^{n \times n} \quad (4.22)$$

Calculando la derivada en el tiempo en las trayectorias de la ecuación (4.5), obtenemos:

$$\begin{aligned} \dot{V}(\Delta) &= 2\Delta^T P_I (A\Delta + h) = 2\Delta^T P_I (A\Delta + h^0 + f_1) \\ &= \Delta^T (P_I A + A^T P_I) \Delta + 2\Delta^T P_I h^0 + 2\Delta^T P_I f_1 \end{aligned} \quad (4.23)$$

donde las funciones $h = h(\Delta, x, u, t)$, $h_0 = h_0(\Delta, x, u, t)$, y $f_1 = f_1(x, u, t)$ son definidos en (4.5)

Usando la desigualdad matricial (2.16) se obtiene lo siguiente:

$$\begin{aligned} 2\Delta^T P_I h_0 &\leq h_0^T H_{n0} h_0 + \Delta^T P_I H_{n0}^{-1} P_I \Delta \\ 2\Delta^T P_I f_1 &\leq f_1^T H_{nf} f_1 + \Delta^T P_I H_{nf}^{-1} P_I \Delta \end{aligned} \quad (4.24)$$

Aplicando ahora la suposición No 4.1 y la suposición No 4.2 (desigualdades (4.6), (4.7), (4.8) y (4.9)), se formula:

$$\begin{aligned} 2\Delta P_I h_0 &\leq \Delta^T (P_I H_{n0}^{-1} P_I + \varepsilon_{h1}(x, u, t) H_{\Delta 0}) \Delta + \varepsilon_{h0}(x, u, t) \\ 2\Delta P_I h_0 &\leq \Delta^T (P_I H_{n0}^{-1} P_I + \varepsilon^{h1} H_{\Delta 0}) \Delta + \varepsilon^{h0} \end{aligned} \quad (4.25)$$

y:

$$\begin{aligned} 2\Delta P_I f_1 &\leq \Delta^T (P_I H_{nf}^{-1} P_I) \Delta + \varepsilon_{f0}(x, u, t) \\ 2\Delta P_I f_1 &\leq \Delta^T (P_I H_{nf}^{-1} P_I) \Delta + \varepsilon^{f0} \end{aligned} \quad (4.26)$$

Sustituyendo las desigualdades (4.25), (4.26) en (4.23), entonces:

$$\dot{V}(\Delta) = \Delta^T (P_I A + A^T P_I + P_I (H_{n0}^{-1} + H_{nf}^{-1}) P_I + \varepsilon^{h1} H_{\Delta 0}) \Delta + \varepsilon^0 \quad (4.27)$$

donde $\varepsilon^0 = \varepsilon^{h0} + \varepsilon^{f0}$

Considerando la suposición No 4.5, de (4.27) se obtiene:

$$\begin{aligned} \dot{V}(\Delta) &\leq -\Delta^T Q_0 \Delta + \varepsilon^0 = -\Delta^T P_I^2 \left(P_I^{-2} Q_0 P_I^{-2} \right) P_I^2 \Delta + \varepsilon^0 \\ \dot{V}(\Delta) &\leq -\lambda_{\min}(R_p) V(\Delta) + \varepsilon^0 \end{aligned} \quad (4.28)$$

con $R_p = P_I^{-2} Q_0 P_I^{-2}$

Resolviendo (4.28), obtenemos:

$$V(\Delta(t)) \leq V(\Delta_0)e^{-\lambda_{\min}(R_P)t} + \varepsilon^0 \int_0^t e^{-\lambda_{\min}(R_P)(t-\tau)} d\tau \quad (4.29)$$

la cual puede ser evaluada explícitamente como:

$$V(\Delta(t)) \leq V(\Delta_0)e^{-\lambda_{\min}(R_P)t} + \frac{\varepsilon^0}{\lambda_{\min}(R_P)} (1 - e^{-\lambda_{\min}(R_P)t}) \quad (4.30)$$

Tomando en cuenta la desigualdad:

$$V(\Delta) \geq \lambda_{\min}(P_I) \Delta^2 \quad (4.31)$$

finalmente, se llega a (4.21).

□□

4.3 Análisis del error de seguimiento de trayectorias

Para facilitar el análisis, se supone $\phi(x) = I$. Aún si esta suposición es restrictiva, existe evidencia experimental de la capacidad de esta red para identificación y control de sistemas no lineales [32]. El objetivo del control es forzar el estado del sistema a seguir una señal de referencia, generada por un modelo de referencia no lineal:

$$\dot{x}^* = \varphi(x^*, t) \quad (4.32)$$

Definiendo las siguientes seminormas:

$$\|x\|_Q^2 = \lim_{t \rightarrow \infty} \sup \frac{1}{\tau} \int_0^\tau x^T(t) Q x(t) dt, \quad \|u\|_R^2 = \lim_{t \rightarrow \infty} \sup \frac{1}{\tau} \int_0^\tau u^T(t) R u(t) dt \quad (4.33)$$

$$Q = Q^T > 0, \quad R = R^T > 0$$

la medida del desempeño a optimizar, puede ser formulada como:

$$J_{\min} = \min_{u(t)} J, \quad J = \|x - x^*\|_Q^2 + \|u\|_R^2 \quad (4.34)$$

Así, para cualquier $\eta > 0$, entonces:

$$J \leq (1 + \eta) \|x - x_{nn}\|_Q^2 + (1 + \eta^{-1}) \|x_{nn} - x^*\|_Q^2 + \|u\|_R^2 \quad (4.35)$$

El mínimo del término $\|x - x_{nn}\|_Q^2$ ya ha sido resuelto para el análisis de identificación. Así, el objetivo del control es minimizar $\|x_{nn} - x^*\|_Q^2$.

Se obtiene el error de las trayectorias (error de estado):

$$\Delta = x_m - x^* \quad (4.36)$$

Así, la ecuación diferencial que describe el error de las trayectorias es:

$$\begin{aligned} \dot{\Delta}^* &= A_{nn} \Delta^* + B_{nn}(t) \sigma(\Delta^* + x^*) + C_{nn} \gamma(u) + \xi^*(t) \\ \xi^*(t) &= A_{nn} x^*(t) - \varphi(x^*, t) \end{aligned} \quad (4.37)$$

Para analizar la estabilidad del error de seguimiento, se propone la siguiente función de Lyapunov:

$$V_i(\Delta^*) = \Delta^{*T} P_c(t) \Delta^*, P_c(t) = P_c^T(t) > 0, \forall t \geq 0 \quad (4.38)$$

cuya derivada en el tiempo es:

$$\begin{aligned} \dot{V}_i(\Delta^*) &= 2\Delta^{*T} P_c(t) \dot{\Delta}^* + \Delta^{*T} \dot{P}_c(t) \Delta^* \\ &= 2\Delta^{*T} P_c(t) (A_{nn} \Delta^* + B_{nn}(t) \sigma(\Delta^* + x^*) + C_{nn} \gamma(u) + \xi^*(t)) + \Delta^{*T} \dot{P}_c(t) \Delta^* \end{aligned} \quad (4.39)$$

Primero se toma en cuenta la siguiente identidad:

$$2(C_{nn}^T P_c(t) \Delta^*)^T \gamma(u) = \|C_{nn}^T P_c(t) \Delta^* + \gamma(u)\|^2 - \Delta^{*T} P_c(t) (C_{nn} C_{nn}^T) P_c(t) \Delta^* - \|\gamma(u)\|^2 \quad (4.40)$$

Se supone también que:

$$\begin{aligned} \sigma^T(x) Z \sigma(x) &\leq x^T \Lambda_\sigma x + \lambda_{\sigma\sigma} \\ Z = Z^T &> 0, \Lambda_\sigma = \Lambda_\sigma^T > 0, \lambda_{\sigma\sigma} > 0 \\ \forall x \in \mathfrak{R}^n, \forall t &\geq 0 \end{aligned} \quad (4.41)$$

Se aplica (2.16) a $2(B_{nn}^T(t) P_c(t) \Delta^*)^T \sigma(\Delta^* + x^*)$ y se utiliza (4.41) para obtener:

$$\begin{aligned} 2(B_{nn}^T(t) P_c(t) \Delta^*)^T \sigma(\Delta^* + x^*) &\leq \\ \Delta^{*T} P_c(t) (B_{nn}(t) Z^{-1} B_{nn}^T(t)) P_c(t) \Delta^* &+ (\Delta^* + x^*)^T \Lambda_\sigma (\Delta^* + x^*) + \lambda_{\sigma\sigma} \\ 2(B_{nn}^T(t) P_c(t) \Delta^*)^T \sigma(\Delta^* + x^*) &\leq \\ \Delta^{*T} P_c(t) (B_{nn}(t) Z^{-1} B_{nn}^T(t)) P_c(t) \Delta^* &+ 2\Delta^{*T} \Lambda_\sigma \Delta^* + 2x^{*T} \Lambda_\sigma x^* + \lambda_{\sigma\sigma} \end{aligned} \quad (4.42)$$

Luego, se emplea nuevamente (2.16) a $2(P_c(t) \Delta^*)^T \xi^*(t)$ y así:

$$2(P_c(t) \Delta^*)^T \xi^*(t) \leq \Delta^{*T} P_c(t) \Lambda_\xi^{-1} P_c(t) \Delta^* + \xi^{*T}(t) \Lambda_\xi \xi(t) \quad (4.43)$$

Sustituyendo (4.40), (4.42) y (4.43) en (4.39), y sumando y restando $\Delta^{*T} Q \Delta^*$, se obtiene:

$$\begin{aligned}
\dot{V}(\Delta^*) \leq & \Delta^{*T} \left[P_c(t) A_{nn} + A_{nn}^T P_c(t) + P_c(t) \left(B_{nn}(t) Z^{-1} B_{nn}^T(t) + \Lambda_\xi^{-1} - C_{nn} C_{nn}^T \right) P_c(t) + 2\Lambda_\sigma + Q \right] \Delta \\
& + 2x^{*T} \Lambda_\sigma x^* + \xi^{*T}(t) \Lambda_\xi \xi^*(t) - \gamma(u)^2 + C_{nn}^T P_c(t) \Delta^* + \gamma(u)^2 \\
& - \Delta^{*T} Q \Delta^* + \lambda_{\sigma 0} + \Delta^* \dot{P}_c(t) \Delta^*
\end{aligned} \tag{4.44}$$

Lema 4.2

Sí la ecuación diferencial de Riccati:

$$P_c(t) A_{nn} + A_{nn}^T P_c(t) + P_c(t) \left(B_{nn}(t) Z^{-1} B_{nn}^T(t) + \Lambda_\xi^{-1} - C_{nn} C_{nn}^T \right) P_c(t) + 2\Lambda_\sigma + Q = -\dot{P}_c(t)$$

tiene solución $P_c(t) = P_c^T(t) > 0$, $\forall t \geq 0$, entonces la derivada con respecto al tiempo de la función de Lyapunov $V(\Delta^*)$ satisface la siguiente desigualdad:

$$\begin{aligned}
\dot{V}(\Delta^*) \leq & 2x^{*T} \Lambda_\sigma x^* + \xi^{*T}(t) \Lambda_\xi \xi^*(t) - \|\gamma(u)\|^2 + \|C_{nn}^T P_c(t) \Delta^* + \gamma(u)\|^2 \\
& - \Delta^{*T} Q \Delta^* + \lambda_{\sigma 0} + u^T R u - u^T R u
\end{aligned} \tag{4.45}$$

Prueba-

Se deriva directamente de (4.44) sumando y restando el término $u^T R u$.

□□□

Definiendo la siguiente función $\Psi(u)$ como:

$$\Psi(u) = -\|\gamma(u)\|^2 + \|C_{nn}^T P_c(t) \Delta^* + \gamma(u)\|^2 + u^T R u. \tag{4.46}$$

Entonces, usando la igualdad (4.46), la desigualdad (4.45) puede ser escrita como:

$$V(\Delta^*) \leq 2x^{*T} \Lambda_\sigma x^* + \Psi(u) - \Delta^{*T} Q \Delta^* - u^T R u + \lambda_{\sigma 0} + \xi^{*T}(t) \Lambda_\xi \xi^*(t). \tag{4.47}$$

Para calcular la acción de control $u(t)$, la cual minimiza el error de seguimiento, se debe primero minimizar $\Psi(u)$, por lo que $d\Psi(u)/du$ tiene que ser igual a cero. Para lograr esta minimización, se supone, en un tiempo t (positivo) dado, $x(t)$ y $x_{nn}(t)$ son conocidas y no dependen de $u(t)$.

Comentario No. 4.3- se denomina $u^*(t)$ la señal que minimiza a $\Psi(u)$, como un control óptimo local, puesto que es calculado basandose en información estrictamente "local".

Para realizar el procedimiento de optimización se usa el siguiente esquema del método del gradiente descendiente recursivo:

$$u_k(t) = u_{k-1}(t) - \tau_k d\Psi(u_{k-1}(t))/du, \quad u_0(t) = 0.$$

donde el gradiente $d\Psi(u)/du$ es calculado como:

$$d\Psi(u)/du = 2 \left(\frac{d\gamma(u)}{du} \right)^T C_{nn}^T P_c(t) \Delta^*(t) + 2Ru, \quad (4.48)$$

y la secuencia de parametros escalares τ_k satisface las condiciones:

$$\tau_k \geq 0, \quad \sum_{k=0}^{\infty} \tau_k = \infty, \quad \tau_k \rightarrow 0 \quad (4.49)$$

Por ejemplo, podemos seleccionar $\tau_k = \frac{1}{(1+k)^\tau}$, $\tau \in (0,1]$

Concerniente a $u^*(t)$, establecemos el siguiente Lema.

Lema 4.3

$u^*(t)$ puede ser calculado como el limite de las secuencias $\{u_k(t)\}$, i.e.

$$u_k(t)^* \rightarrow u^*(t), \quad k \rightarrow \infty.$$

Prueba- Se sigue directamente de las propiedades del método del gradiente, y tomando en cuenta (4.32), (4.37) y (4.49).

Corolario 4.1 Si $\gamma(u)$ depende linealmente en u , entonces $\frac{d\gamma(u)}{du} = \Gamma$, y se puede compensar la señal medible ξ^* en (4.37), seleccionando

$$u = u_{comp} + u^*$$

Donde u_{comp} satisface la relación

$$C_{nn} u_{comp}(t) + \xi^*(t) = 0$$

y u^* es seleccionado acorde a la ley de control óptimo cuadrático lineal.

$$u^*(t) = -R^{-1} \Gamma^T C_{nn}^T P_c(t) \Delta^*(t). \quad (4.50)$$

Teorema 4.2

Para el sistema (4.2), dada la red neuronal (4.1) con $\phi(x_m) = I$, el modelo de referencia no lineal (4.32), y la ley de control (4.50), las siguientes propiedades son satisfechas:

$$\|\Delta^*\|_Q^2 + \|u^*\|_{R^*}^2 \leq 2\|u^*\|_{\Lambda_\sigma}^2 + \limsup_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \Psi(u^*(t)) dt + \lambda_{\sigma_0} + \limsup_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \xi^{*T}(t) \Lambda_\xi \xi^*(t) dt \quad (4.51)$$

Prueba- Conociendo el mínimo $\Psi(u^*(t))$, (4.47) puede ser reescrita como:

$$\begin{aligned} V(\Delta^*) &\leq 2x^{*T} \Lambda_\sigma x^* + \Psi(u^*(t)) - \Delta^{*T} Q \Delta^* - u^{*T}(t) R u^*(t) \\ &+ \lambda_{\sigma_0} + \xi^{*T}(t) \Lambda_\xi \xi^*(t). \end{aligned} \quad (4.52)$$

lo cual se reformula como:

$$\begin{aligned} \Delta^{*T} Q \Delta^* + u^{*T}(t) R u^*(t) &\leq 2x^{*T} \Lambda_\sigma x^* + \psi(u^*(t)) \\ &- \dot{V}(\Delta^*) + \lambda_{\sigma_0} + \xi^{*T}(t) \Lambda_\xi \xi^*(t). \end{aligned} \quad (4.53)$$

Entonces, integrando cada término de 0 to τ , y dividiendo cada término por τ , y tomando el limite, para $\tau \rightarrow \infty$, se obtiene:

$$\begin{aligned} \limsup_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \Delta^{*T} Q \Delta^* dt + \limsup_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau u^{*T}(t) R u^*(t) dt &\leq \\ \lim_{\tau \rightarrow \infty} 2 \sup \frac{1}{\tau} \int_0^\tau x^{*T} \Lambda_\sigma x^* dt + \limsup_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \psi(u^*(t)) dt + \limsup_{\tau \rightarrow \infty} \left[-\frac{1}{\tau} \int_0^\tau \dot{V}(\Delta^*(t)) \right] & \quad (4.54) \\ + \limsup_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau [\lambda_{\sigma_0} + \xi^{*T}(t) \Lambda_\xi \xi^*(t)] dt. & \end{aligned}$$

Analizaremos el tercer término, del lado derecho de esta desigualdad, entonces:

$$\begin{aligned} \limsup_{\tau \rightarrow \infty} \left[-\frac{1}{\tau} \int_0^\tau \dot{V}(\Delta^*(t)) \right] &= \limsup_{\tau \rightarrow \infty} \left[-\frac{1}{\tau} (V(\Delta^*(\tau)) - V(\Delta^*(0))) \right] \\ \limsup_{\tau \rightarrow \infty} \left[-\frac{1}{\tau} V(\Delta^*(\tau)) + \frac{1}{\tau} V(\Delta^*(0)) \right] &\leq \limsup_{\tau \rightarrow \infty} \left[\frac{1}{\tau} V(\Delta^*(0)) \right] = 0 \end{aligned} \quad (4.55)$$

Tomando en cuenta (4.55), la desigualdad (4.51) se sigue de (4.54)

□□□

La ecuación (4.52) fija una tolerancia para el error de seguimiento de las trayectorias. La estructura final de identificación y control basado en una red neuronal dinámica es mostrado en la Fig17.

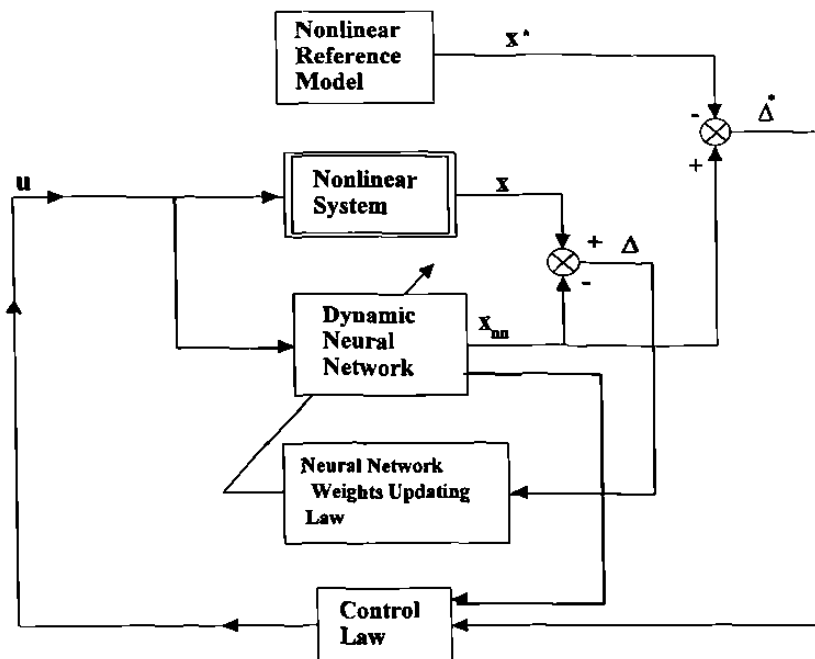


Fig.17: Esquema de identificación y control

4.4-Aplicaciones

Para ilustrar la aplicación del esquema propuesto, se incluye el siguiente ejemplo.

Ejemplo 4.1– Dado el siguiente sistema no lineal:

$$\begin{aligned} \dot{x}_1 &= -\alpha_1 x_1 + \beta_1 \text{sign}(x_2) + u_1 \\ \dot{x}_2 &= -\alpha_2 x_2 + \beta_2 \text{sign}(x_1) + u_2 \end{aligned} \quad (4.56)$$

el cual se presenta en la Fig.18, se selecciona la siguiente red neuronal:

$$\begin{aligned} \dot{x}_{nn,1} &= -a_{nn,1} x_{nn,1} + b_{nn,11} \sigma_{nn,1}(x_{nn,1}) + b_{nn,12} \sigma_{nn,2}(x_{nn,2}) + c_{nn,11} u_1 + c_{nn,12} u_2 \\ \dot{x}_{nn,2} &= -a_{nn,2} x_{nn,2} + b_{nn,21} \sigma_{nn,1}(x_{nn,1}) + b_{nn,22} \sigma_{nn,2}(x_{nn,2}) + c_{nn,21} u_1 + c_{nn,22} u_2 \end{aligned} \quad (4.57)$$

Donde $\sigma(\cdot)$ indica una función sigmoideal. El esquema a bloques de la red neuronal se presenta en la Fig.19.

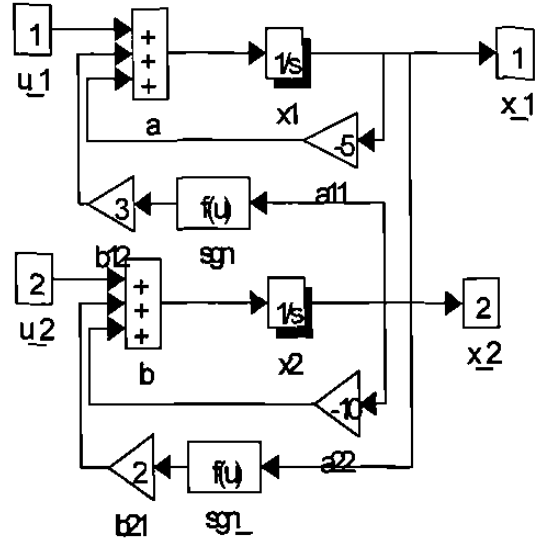


Fig 18. El esquema de la planta no lineal.

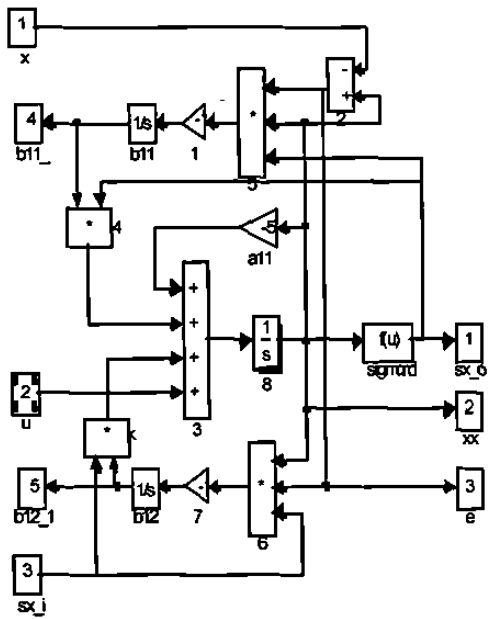


Fig 19. El esquema de la red neuronal

El sistema no-lineal (4.56), aún cuando es sencillo, es interesante; ya que tiene múltiples puntos de equilibrio aislados.

Se considera:

$$a_{m,1} = \alpha_1 = 5, x_1(0) = 10, x_{m,1}(0) = -1, \beta_1 = 3$$

$$a_{m,2} = \alpha_{21} = 10, x_2(0) = -10, x_{m,2}(0) = -5, \beta_2 = 2$$

La función sigmoïdal se selecciona como:

$$y = \frac{2}{1 + e^{-2x}} - 0.5$$

La estructura de la red neuronal es similar a la usada en [20]. Se elijen $Q = Q_0 = H_{n0} = H_{nf} = H_{\Delta 0} = H_{\Delta f} = I, \varepsilon^{h1} = \varepsilon^{f1} = 3$, y se obtiene la solución $P_c(t)$ de la correspondiente ecuación diferencial de Riccati (ver Lema 4.2) iniciando con la condición inicial

$$P_0 = \begin{bmatrix} 0.84 & 0 \\ 0 & 0.36 \end{bmatrix}$$

la cual corresponde a la solución de la ecuación matricial algebraica de Riccati cuando el lado izquierdo de esta ecuación es igual a cero. En nuestro caso $\gamma(u) = u$, así $\Gamma = I$.

Escogiendo $C_{nn} = I$ y $R = \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix}$ y aplicando la ley de control dada por (4.50), se obtienen las curvas de desempeño presentadas posteriormente.

a) Ilustración del proceso de identificación.

Para adaptar en línea los pesos de la red neuronal dinámica se usa el mismo algoritmo de aprendizaje que en [20]. Las acciones de control u_1, u_2 se escogen como senoides (u_1) y como diente de sierra (u_2); los resultados se muestran en la Fig.20. Las líneas sólidas corresponden a la respuesta del estado de el sistema no lineal $x_1(t), x_2(t)$, y las líneas punteadas a la red neuronal $x_{nn,1}^*(t), x_{nn,2}^*(t)$. Los valores de la abscisa corresponden al numero de iteraciones. Se puede ver que la evolución del estado de la red neuronal sigue al sistema no lineal.

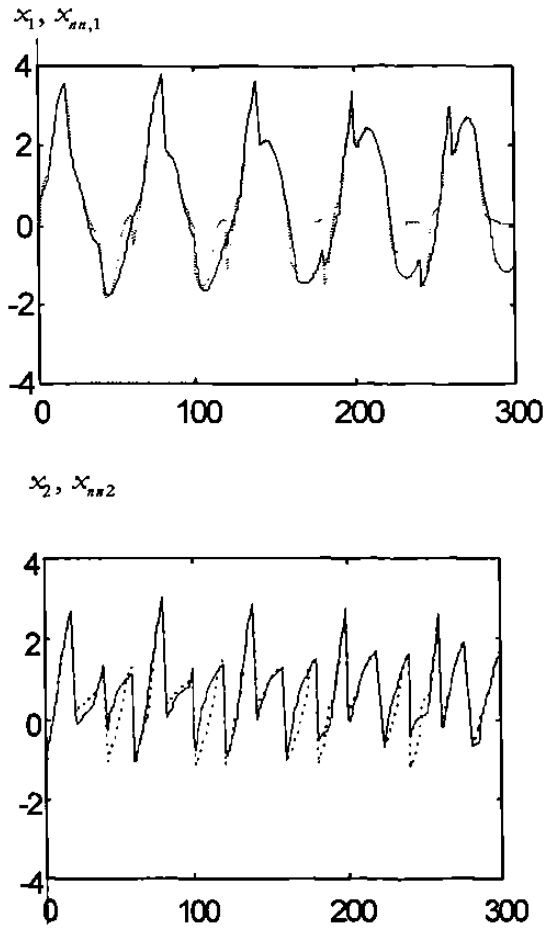


Fig.20 Resultados de la identificación

El resultado de la adaptación de los pesos de la matriz B_{nn} se muestra en la Fig.21.

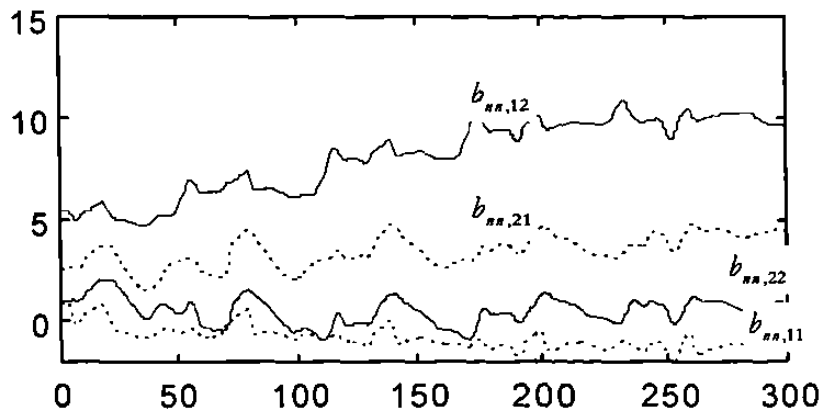


Fig.21 Resultados de la evolución de la matriz de pesos.

b) Ilustración del proceso de control.

Para el desempeño del controlador, primero se simula la respuesta del sistema no lineal con respecto a u_1, u_2 , las cuales son las mismas que antes (sin control de retroalimentación). La Fig.22 presenta las respuestas respectivas $x_1(t), x_2(t)$ y $x_{nn,1}(t), x_{nn,2}(t)$. La línea sólida corresponde a la señal del modelo de referencia, y la línea punteada a la respuesta del sistema no lineal.

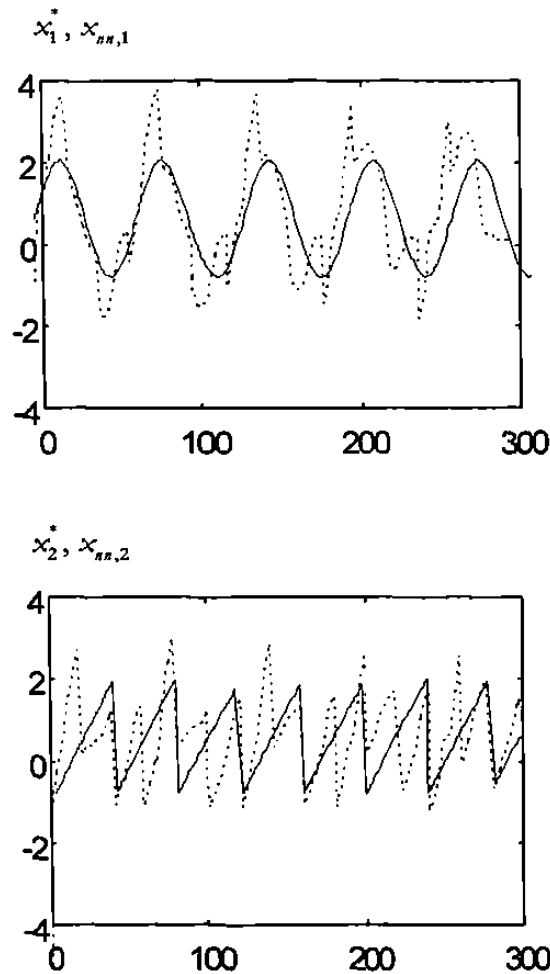


Fig.22 Respuesta sin control de retroalimentación

El esquema del controlador es presentado en la Fig.23. Constituye un control de retroalimentación con una ganancia adaptable en línea.

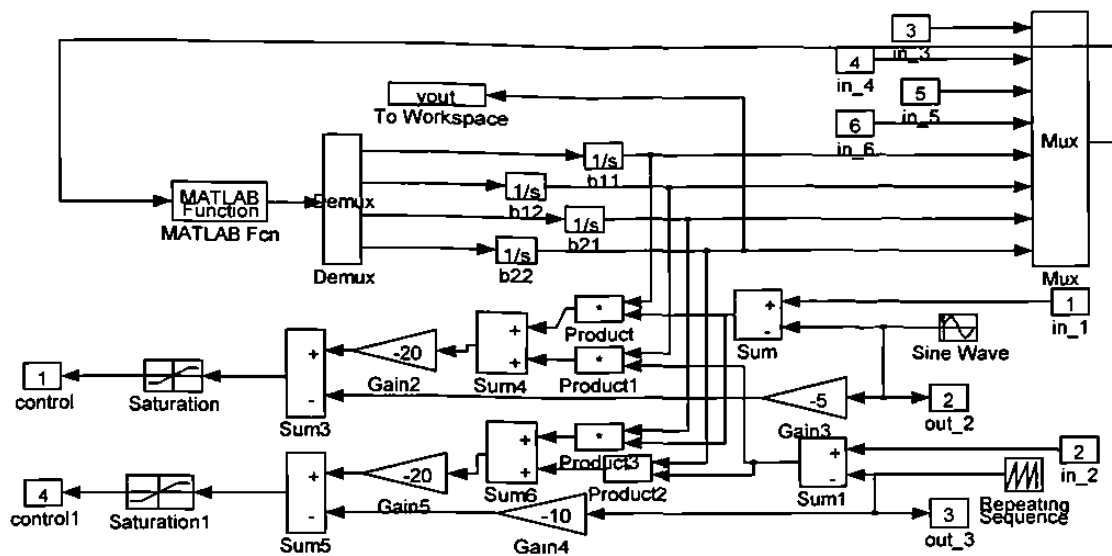


Fig.23 Esquema para el controlador

La Fig.24 presenta las respectivas respuestas, donde la línea sólida corresponde a la señal del modelo de referencia $x_1^*(t)$, $x_2^*(t)$, y la línea punteada a la respuesta del sistema no lineal $x_{nn,1}(t)$, $x_{nn,2}(t)$.

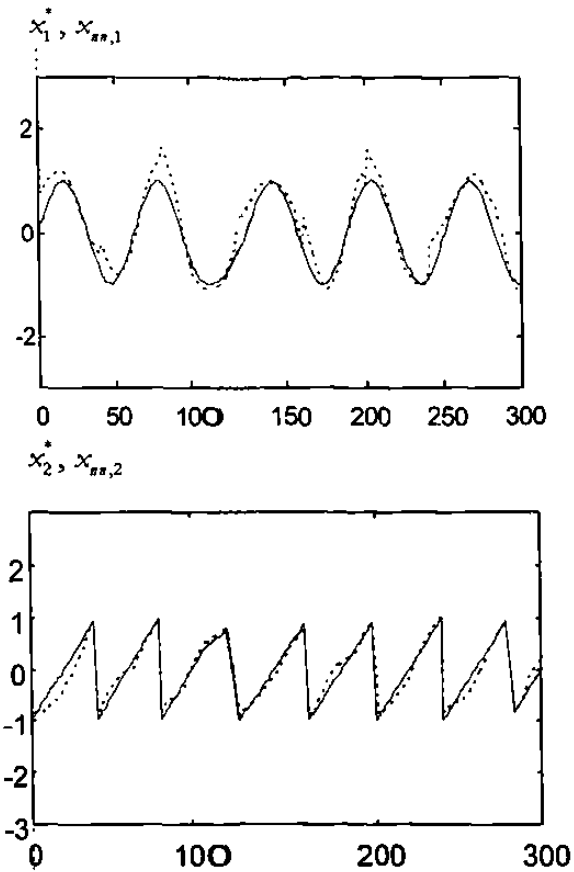


Fig.24 Respuesta con control de retroalimentación

La solución P de la ecuación diferencial es mostrada en Fig.25.

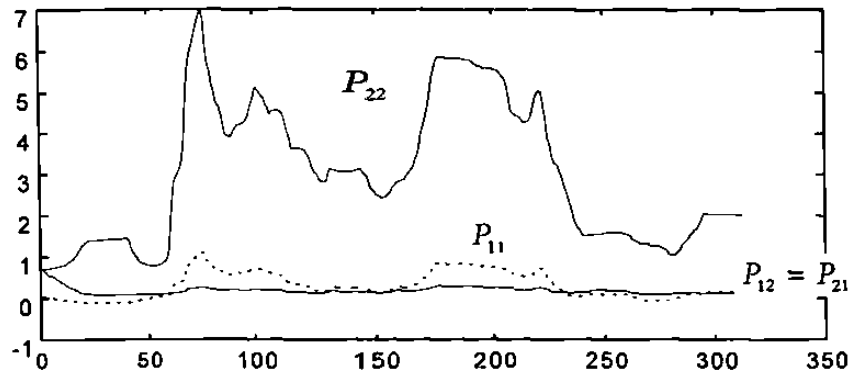


Fig.25 Respuesta de la matriz P

Las evolución de las seminormas $\|\Delta^*\|_q^2 = \limsup_{\tau \rightarrow \infty} J_\tau^\Delta$, $J_\tau^\Delta = \frac{1}{\tau} \int_0^\tau \Delta^{*T}(t) Q \Delta(t) dt$ y

$\|u\|_R^2 = \limsup_{\tau \rightarrow \infty} J_\tau^u$, $J_\tau^u = \frac{1}{\tau} \int_0^\tau u^{*T}(t) R u(t) dt$ pueden verse en las Fig.26 y Fig.27.

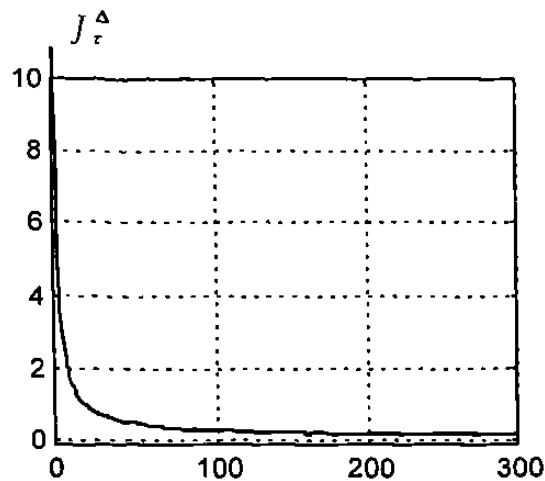


Fig.26 Error de seguimiento J_τ^Δ

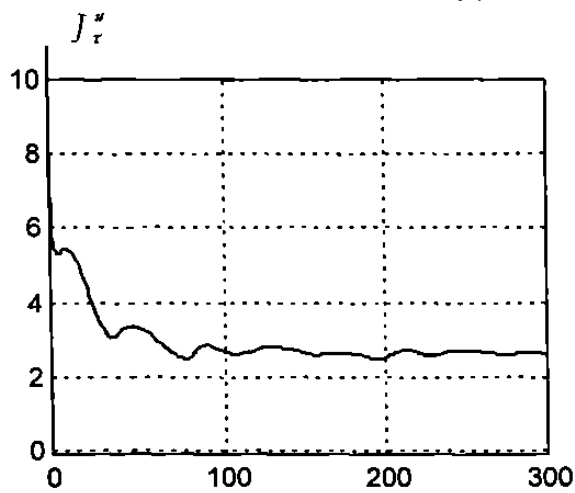


Fig.27: Entrada al sistema no lineal J_τ^u

4.5-Conclusiones del capítulo

En este capítulo se desarrollaron e implementaron un identificador adaptable robusto y un esquema de seguimiento de trayectorias para sistemas no lineales. Este esquema esta compuesto de dos partes: una red neuronal para identificación y un controlador de seguimiento. Se establecen cotas tanto para el error de identificación como para el error de seguimiento.

La aplicabilidad del esquema se ilustra con un sistema no lineal; el cual tiene múltiples puntos de equilibrio y un campo vectorial diferenciable.

CAPITULO V

CONCLUSIONES Y TRABAJO FUTURO

Como parte final de este trabajo de investigación, en primer lugar se presentan las conclusiones y aportaciones, y luego se establecen algunas sugerencias para investigaciones futuras.

5.1 CONCLUSIONES.

La utilización de redes neuronales dinámicas en ingeniería, en particular en sistemas de control automático, es muy promisoria; no obstante, los resultados analíticos son muy escasos. El presente trabajo de investigación está constituido por un conjunto de aportaciones, que enriquece dicho análisis. Específicamente, estas aportaciones son:

a) Estabilidad Global Asintótica

Se establecieron condiciones suficientes sobre la matriz de pesos de una red neuronal tipo Hopfield, para garantizar la estabilidad global asintótica.

El análisis se realizó utilizando el enfoque clásico de Lyapunov; las condiciones obtenidas son menos re restrictivas que aquellas basadas en matrices diagonalmente estables.

b) Estabilidad Entrada-Estado

Para las matrices de pesos de un tipo de redes neuronales recurrentes de alto orden, se establecieron condiciones que garantizan estabilidad entrada –estado.

Esas mismas condiciones garantizan estabilidad asintótica global, extendiendo el resultado previamente obtenido para redes tipo Hopfield, a un tipo de redes neuronales recurrentes de alto orden.

c) Identificación y Control de Sistemas Nolineales

Basándose en un tipo de red neuronal recurrente de alto orden, se desarrolló un esquema de identificación en línea para sistemas nolineales.

Una vez que se obtuvo un modelo de sistemas nolineales por medio de éste identificador neuronal, se desarrolló un esquema para garantizar el seguimiento de trayectorias, aún si éstas son nolineales.

Utilizando el enfoque de Lyapunov, se establecieron cotas tanto para el error de identificación como para el error de seguimiento.

5.2 Trabajos Futuros.

El presente trabajo de investigación puede ser continuado en las siguientes direcciones:

- El análisis de la propiedad de estabilidad entrada-estado puede ser extendido para establecer condiciones de estabilidad entrada-salida para redes neuronales dinámicas.
- Una vez garantizada la estabilidad esntrada-salida, se pueden sintetizar controladores neuronales basados en redes neuronales dinámicas.
- El esquema basado en el identificador neuronal y el controlador para seguimiento de trayectorias, puede ser extendido al caso cuando el estado del sistema no es completamente medible; el identificador neuronal se convierte entonces en un observador neuronal.

APÉNDICE A
PUBLICACIONES

APÉNDICE A

En esta parte se presentan algunos de los resultados presentados en esta Tesis y que fueron aceptados en congresos internacionales y publicados en revista de arbitraje internacional.

- 1) “ Input to State Stability (ISS) Analysis for dynamics Neural Networks” presentado en la 1997 International Conference on Neural networks (ICNN 1997), Houston Texas, USA. June, 1997.
- 2) “Robust Adaptative Nonlinear System Identification and Trayectory Tracking by Dynamic Neural Networks” presentado en la 1997 American Control Conference (ACC), Albuquerque, New Mexico, USA. June, 1997.
- 3) “ Stability Analysis of Dynamic Neural Control”, Artículo publicado en Expert Systems with Applications de la Pergamon Press, ISSN 0957-4174, Volume 14, Number ½, January/ February 1998.

1997 International Conference on Neural Networks (ICNN '97)



H O U S T O N , T E X A S

Westin Galleria Hotel, Houston, Texas, USA

TUTORIALS: June 8, 1997

CONFERENCE: June 9-12, 1997



Proceedings
Volume 2 of 4

INPUT-TO-STATE STABILITY (ISS) ANALYSIS FOR DYNAMIC NEURAL NETWORKS¹

Edgar N. Sanchez^{*} and Jose P. Perez^{**}

^{*}CINVESTAV, Unidad Guadalajara
Apartado Postal 31-438, Plaza La Luna
Guadalajara, Jalisco C.P. 44550, MEXICO
Tel:(52)(3)6841580, Fax:(52)(3)6841708, e-mail:sanchez@gdl.cinvestav.mx

^{**}Electrical Engineering Graduate Program
School of Mechanical and Electrical Engineering
Universidad Autonoma de Nuevo Leon (UANL)
Apartado Postal F-93, C.U.
San Nicolas de los Garza, N.L., C.P. 66450,

Keywords: Neural Network Dynamics, Intelligent Control, Stability, ISS Analysis, Lyapunov Functions.

Abstract: This paper presents the input-to-state (ISS) analysis for dynamic neural networks. We determine, using a Lyapunov function, conditions to guarantee ISS; they also guarantee globally asymptotically stability (GAS).

I. INTRODUCTION

The stability analysis for dynamic neural networks has been intensively developed since late 80's. First, conditions were established for local asymptotically stability [1,2]; however there could exist multiple equilibria, which generate local minima. This property is useful for associative memory or pattern recognition, but to guarantee globally asymptotically stability (GAS) is very important for automatic control and/or optimization. The analysis of GAS for Neural Network is being strongly studied since early 90's; its importance is signaled in [3]. Conditions for GAS are established in [4], applying the contracting mapping approach; Matsuoka [5] generalized these results by means of a Lure type Lyapunov function. Recent results [6,7] are based on the concept of diagonally stable matrices.

There exist two methodologies for the analysis of dynamic systems: a) the input-output approach based on operator theory, which requires the operator, representing the system, to be bounded; b) the state space approach, whose main tool is the analysis by means of a

Lyapunov function. These two approaches are not equivalent for nonlinear systems [8]. The input-to-state (ISS) concept, introduced in [9], allows to establish this equivalence; it has been applied to derive important results concerning nonlinear systems control [10,11]. ISS can be derived by different methods such as: dissipativity, robust margins, or in the classical Lyapunov sense [12].

As a continuation of our previous work related with GAS [13], we present in this paper the ISS analysis for dynamic neural networks; to the best of the authors' awareness, this is the first publication concerning this kind of analysis. In our research, we follow the classical Lyapunov sense. The paper is organized as follows: first we present the mathematical model of the neural network; then we state and prove a theorem, which constitutes our main contribution, about the conditions to guarantee ISS. Finally we establish the relevant conclusions. Definitions of important concepts, such as *class K_∞ function*, *globally asymptotically stability*, and *ISS Lyapunov function*, are included in appendix.

II. MATHEMATICAL MODEL

The dynamic neural network is described by the following mathematical model:

$$\begin{aligned} \dot{x} &= A_{nn}x + B_{nn}\sigma(\tau) + C_{nn}\phi(\tau)\Gamma(u) \\ x &\in \mathfrak{R}^n, u \in \mathfrak{R}^m, \sigma(x) \in \mathfrak{R}^n, \Gamma(u) \in \mathfrak{R}^n \\ A_{nn}, B_{nn}, C_{nn} &\in \mathfrak{R}^{n \times n}, \phi(x) \in \mathfrak{R}^{m \times n} \end{aligned} \quad (2.1)$$

¹This work was supported by CONACYT Project 0652A9506 and partially by the Schools of Mechanical and Electrical Engineering and Mathematics and Physic Sciences of the UANL.

where A_{nn} is the state matrix, B_{nn} is the nonlinear feedback matrix, C_{nn} is the input matrix, x is the neural network state, $\sigma(x)$ and $\phi(x)$ are nonlinear functions, and nn labels the model parameters.

This model describes dynamic neural networks as the ones discussed [14], which is generalization of the Hopfield model. The elements of $\sigma(x)$ and $\phi(x)$ are usually selected as sigmoid, that is

$$\sigma_i(x_i) = \frac{\alpha_i}{1 + e^{-\beta_i x_i}}.$$

We consider A_{nn} Hurwitz (all its eigenvalues have negative real part), $\Gamma(0) = 0$, and $\sigma(\cdot), \phi(\cdot), \Gamma(u)$ globally Lipschitz²:

$$\begin{aligned} \|\sigma(\Delta_v - x) - \sigma(x)\| &\leq L_\sigma \|\Delta_v\| \\ \|\phi(\Delta_v - x) - \phi(x)\| &\leq L_\phi \|\Delta_v\| \\ \|\Gamma(\Delta_u - x) - \Gamma(x)\| &\leq L_\Gamma \|\Delta_u\| \\ \forall x, \Delta_v &\in \mathfrak{R}^n, \forall u, \Delta_u \in \mathfrak{R}^m \end{aligned} \quad (2.2)$$

Selecting A_{nn} Hurwitz is a common assumption for stability analysis of dynamic neural networks [5, 6, 7].

Following the same procedure as in [5], it is trivial to demonstrate that (2.1) presents at least one equilibrium point.

III. ISS ANALYSIS

Given x^* , an equilibrium point of (2.1) for $u=0$, we define $y = x - x^*$ then:

$$\dot{y} = \dot{x} = A_{nn}x + B_{nn}\sigma(x) + C_{nn}\phi(x)\Gamma(u) \quad (3.1)$$

Adding and subtracting $B_{nn}\sigma(x^*) + C_{nn}\phi(x^*)\Gamma(u)$ to (3.1), we obtain:

² In this paper a) $\|\cdot\|$ denotes the Euclidian norm for vectors and, for a matrix A , it is defined as: $\sqrt{\lambda_{\max}(A^T A)}$, with $\lambda_{\max}(\cdot)$ the maximum eigenvalue, b) is A^T the transpose of A

$$\begin{aligned} \dot{y} &= A_{nn}y + B_{nn}[\sigma(y+x^*) - \sigma(x^*)] \\ &+ C_{nn}[\phi(y+x^*) - \phi(x^*)]\Gamma(u) \\ &+ C_{nn}\phi(x^*)\Gamma(u) + A_{nn}x^* + B_{nn}\sigma(x^*) \end{aligned} \quad (3.2)$$

Being the sum of the right part last two terms of (3.2) equal to zero, and taking in account that $\Gamma(0) = 0$, (3.2) can be rewritten as:

$$\begin{aligned} \dot{y} &= A_{nn}y + B_{nn}[\sigma(y+x^*) - \sigma(x^*)] \\ &+ C_{nn}[\phi(y+x^*) - \phi(x^*)]\Gamma(u) \\ &+ C_{nn}\phi(x^*)[\Gamma(u) - \Gamma(0)] \end{aligned} \quad (3.3)$$

At this stage, we establish our main contribution: a theorem about ISS for (3.3), or equivalently for (2.1).

THEOREM - The system (3.1), or equivalently the system (2.1), is ISS if:

$$\begin{aligned} \|B_{nn}\|^2 &< \frac{\gamma - 3\|P\|}{L_\sigma^2\|P\|} \cdot \|P\| < \frac{\gamma}{3} \\ \gamma &\in \mathfrak{R}^+, P = P^T > 0 \end{aligned} \quad (3.4)$$

where P is solution of the Lyapunov equation $A_{nn}^T P + P A_{nn} = -\gamma I$. Such a unique solution exists because A_{nn} is Hurwitz.

Proof - We consider the following function:

$$V(y) = y^T P y, P = P^T > 0 \quad (3.5)$$

Its time derivative along the trajectories of (3.3) is given as:

$$\begin{aligned} \dot{V} &= y^T P \dot{y} + \dot{y}^T P y = y^T P v + v^T P y \\ v &= A y + B_{nn}[\sigma(y+x^*) - \sigma(x^*)] \\ &+ C_{nn}[\phi(y+x^*) - \phi(x^*)]\Gamma(u) \\ &+ C_{nn}\phi(x^*)[\Gamma(u) - \Gamma(0)] \end{aligned} \quad (3.6)$$

which reduces to:

$$\begin{aligned} \sqrt{=} & y^T y + 2y^T PB_{nn} [\sigma(y+x^*) - \sigma(x^*)] \\ & + 2y^T PC_{nn} [\phi(y+x^*) - \phi(x^*)] \Gamma(u) \\ & + 2y^T PC_{nn} \phi(x^*) [\Gamma(u) - \Gamma(0)] \end{aligned} \quad (3.7)$$

We analyze the right part second, third and fourth term of (3.7), using the following matrix inequality [15]:

$$X^T Y + Y^T X \leq X^T \Lambda X + Y^T \Lambda^{-1} Y.$$

This inequality is valid for any matrices $X \in \mathfrak{R}^{n \times k}$, $Y \in \mathfrak{R}^{n \times k}$, and for any matrix $\Lambda = \Lambda^T > 0$, $\Lambda \in \mathfrak{R}^{n \times n}$. In this paper, we select $\Lambda = P$.

For the second term of (3.7), we have:

$$\begin{aligned} 2y^T (PB_{nn} [\sigma(y+x^*) - \sigma(x^*)]) & \leq \\ y^T P y + (PB_{nn} [\sigma(y+x^*) - \sigma(x^*)])^T & \\ P^{-1} (PB_{nn} [\sigma(y+x^*) - \sigma(x^*)]) & \\ \leq & \end{aligned} \quad (3.8)$$

$$\begin{aligned} y^T P y + (PB_{nn} [\sigma(y+x^*) - \sigma(x^*)])^T & \\ (B_{nn} [\sigma(y+x^*) - \sigma(x^*)]) & \end{aligned} \quad (3.13)$$

The right part of (3.8) is equal to:

$$\begin{aligned} & = y^T P y + (B_{nn} [\sigma(y+x^*) - \sigma(x^*)])^T \\ & \quad P (B_{nn} [\sigma(y+x^*) - \sigma(x^*)]) \\ & \leq \|P\| \|y\|^2 + \|P\| \|B_{nn} [\sigma(y+x^*) - \sigma(x^*)]\|^2 \\ & \leq \|P\| \|y\|^2 + \|P\| \|B_{nn}\|^2 \|y\|^2 L_\sigma^2 \end{aligned} \quad (3.9)$$

So:

$$\begin{aligned} 2y^T PB_{nn} [\sigma(y+x^*) - \sigma(x^*)] & \leq \\ \|P\| \|y\|^2 + L_\sigma^2 \|P\| \|B_{nn}\|^2 \|y\|^2 & \end{aligned} \quad (3.10)$$

For the third term of (3.7), we obtain:

$$\begin{aligned} 2y^T (PC_{nn} [\phi(y+x^*) - \phi(x^*)] \Gamma(u)) & \leq \\ y^T P y + (PC_{nn} [\phi(y+x^*) - \phi(x^*)] \Gamma(u))^T & \\ P^{-1} (PC_{nn} [\phi(y+x^*) - \phi(x^*)] \Gamma(u)) & \\ \leq & \\ y^T P y + (C_{nn} [\phi(y+x^*) - \phi(x^*)] \Gamma(u))^T & \\ P (C_{nn} [\phi(y+x^*) - \phi(x^*)] \Gamma(u)) & \end{aligned} \quad (3.11)$$

Being $\phi(x)$ globally Lipschitz, then it is uniformly bounded [16], so:

$$\|\phi(y+x^*) - \phi(x^*)\| \leq K_\phi, \forall y \in \mathfrak{R}^n \quad (3.12)$$

and the right part of (3.11) can be formulated as:

$$\begin{aligned} & = \|P\| \|y\|^2 + \|P\| \|C_{nn} [\phi(y+x^*) - \phi(x^*)] \Gamma(u)\|^2 \leq \\ & \quad \|P\| \|y\|^2 + K_\phi^2 L_\Gamma^2 \|P\| \|C_{nn}\|^2 \|u\|^2 \end{aligned} \quad (3.13)$$

Finally:

$$\begin{aligned} 2y^T (PC_{nn} [\phi(y+x^*) - \phi(x^*)] \Gamma(u)) & \leq \\ \|P\| \|y\|^2 + K_\phi^2 L_\Gamma^2 \|P\| \|C_{nn}\|^2 \|u\|^2 & \end{aligned} \quad (3.14)$$

For the fourth term of (3.7), we have:

$$\begin{aligned} 2y^T (PC_{nn} \phi(x^*) [\Gamma(u) - \Gamma(0)]) & \leq \\ y^T P y + (PC_{nn} \phi(x^*) [\Gamma(u) - \Gamma(0)])^T & \\ P^{-1} (PC_{nn} \phi(x^*) [\Gamma(u) - \Gamma(0)]) & \\ \leq & \\ y^T P y + (C_{nn} \phi(x^*) [\Gamma(u) - \Gamma(0)])^T & \\ P (C_{nn} \phi(x^*) [\Gamma(u) - \Gamma(0)]) & \\ \leq & \\ \|P\| \|y\|^2 + \|P\| \|C_{nn} \phi(x^*) [\Gamma(u) - \Gamma(0)]\|^2 & \end{aligned} \quad (3.15)$$

which can be rewritten as:

$$\begin{aligned}
 2y^T (PC_{nn} \phi(x^*)) [\Gamma(u) - \Gamma(0)] &\leq \\
 \|P\| \|y\|^2 + \|P\| \|C_{nn}\|^2 \|\phi(x^*)\|^2 \|\Gamma(u) - \Gamma(0)\|^2 &\leq \\
 \|P\| \|y\|^2 + L_p^2 \|P\| \|C_{nn}\|^2 \|\phi(x^*)\|^2 \|u\|^2 &
 \end{aligned} \tag{3.16}$$

Substituting (3.10), (3.14) and (3.16) in (3.7), we finally obtain:

$$\begin{aligned}
 \dot{V} &\leq - \left(\gamma - 3\|P\| - L_p^2 \|P\| \|B_{nn}\|^2 \right) \|y\|^2 \\
 &\quad + \left(K_\phi^2 L_p^2 \|P\| \|C_{nn}\|^2 + L_p^2 \|P\| \|C_{nn}\|^2 \|\phi(x^*)\|^2 \right) \|u\|^2
 \end{aligned} \tag{3.17}$$

Defining the following functions:

$$\begin{aligned}
 \alpha(r) &= \left(\gamma - 3\|P\| - L_p^2 \|P\| \|B_{nn}\|^2 \right) r^2 \\
 \theta(r) &= \left(K_\phi^2 L_p^2 \|P\| \|C_{nn}\|^2 + L_p^2 \|P\| \|C_{nn}\|^2 \|\phi(x^*)\|^2 \right) r^2 \\
 r &\in \mathfrak{R}
 \end{aligned} \tag{3.18}$$

then:

$$\dot{V} \leq -\alpha(\|x\|) + \theta(\|u\|) \tag{3.19}$$

V will be a ISS-Lyapunov function [12], if $\alpha(\cdot)$ and $\theta(\cdot)$ are class K_∞ functions [2]. As defined, $\theta(\cdot)$ satisfies this condition; hence, for the system (3.3) (equivalently (2.1)) to be ISS, it is required:

$$\left(\gamma - 3\|P\| - L_p^2 \|P\| \|B_{nn}\|^2 \right) > 0 \tag{3.20}$$

which implies:

$$\|B_{nn}\|^2 < \frac{\gamma - 3\|P\|}{L_p^2 \|P\|}, \quad \|P\| < \frac{\gamma}{3} \tag{3.21}$$

Corollary - For the equilibrium point $y = 0$ of (3.3) (equivalently $x = x^*$ of (2.1)) to be globally asymptotically stable, (3.21) has to be satisfied:

Proof: Taking $u = 0$ in (3.19) then:

$$\dot{V} \leq -\alpha(\|x\|) \tag{3.22}$$

$\dot{V} < 0, \forall x \in \mathfrak{R}^n$ [2] guarantee that (3.3) is globally asymptotically stable. To fulfill this condition, it is required to satisfy (3.22).

See appendix, for definitions of class K_∞ functions, globally asymptotically stability, and ISS-Lyapunov functions.

A very interesting application of the ISS property, to stability of interconnected nonlinear systems, is discussed in [11].

IV. CONCLUSIONS

As our main contribution, we establish conditions for the weights of the interconnection matrix of dynamic neural networks, in order to guarantee input to state stability. For inputs identically equal to zero, these conditions also guarantee globally asymptotically stability.

The ISS property allows the application of dynamic neural network to nonlinear control. Work is in progress to develop control schemes based on the internal model control (IMC) technique.

V. APPENDIX

a) *class K_∞ function*[2]. A continuous function $f: [0, a) \rightarrow [0, \infty)$ is said to belong to class K if it is strictly increasing and $f(0) = 0$; and it is said to belong to class K_∞ if $a = \infty$ and $\lim_{r \rightarrow \infty} f(r) = \infty$.

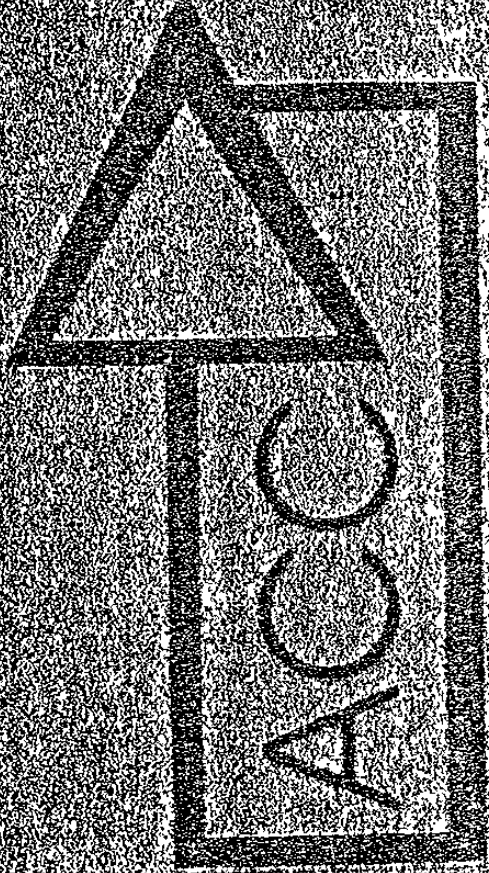
b) *Globally Asymptotically Stability (GAS)* [2]. Let $x=0$ be an equilibrium point of $\dot{x}=f(x)$ and $D \subset \mathfrak{R}^n$ be a union of opened connected set, with none of its boundary points, which contains $x=0$. Let $V: D \rightarrow R$ be a continuously differentiable function, such that $V(0) = 0, V(x) > 0$ and $\dot{V}(x) < 0$ in $D - \{0\}$, then $x=0$ is asymptotically stable. If the property is fulfilled for $D = \mathfrak{R}^n$, then it is global.

c) *ISS Lyapunov function* [12]. A function $V(x): \mathfrak{R}^n \rightarrow \mathfrak{R}^+$ is defined as ISS Lyapunov if there exist $\alpha(r), \theta(r)$ class K_∞ functions such that:

$$\dot{V}(x) \leq -\alpha(\|x\|) + \theta(\|u\|) \quad \forall x \in \mathfrak{R}^n, \forall u \in \mathfrak{R}^m$$

VI. REFERENCES

- [1] A.N.Michel, J.A.Farrel, and W. Porod, "Qualitative analysis of neural networks", *IEEE Trans. on Circuits and Systems*, vol. 36, no. 2, pp 229-243, 1989.
- [2] H.K.Khalil, "Nonlinear Systems", 2nd Edition, Prentice Hall, 1996.
- [3] M.W.Hirsch, "Convergent activation dynamics in continuous time networks", *Neural Networks*, vol. 2, pp. 331-349, 1989.
- [4] D.G.Kelly, "Stability in contractive nonlinear neural networks", *IEEE Trans. on Biomed. Eng.*, vol. 3, no. 3 pp. 241-242, 1990.
- [5] K.Matsouka, "Stability conditions for nonlinear continuous neural networks with asymmetric connections weights", *Neural Networks*, vol. 5, pp. 495-500, 1992.
- [6] E.Kaszakurewics and A.Bhaya, "On a class of globally stable neural circuits", *IEEE Trans. on Circuits and Systems-I*, vol. 41, no. 2, pp 171-174, 1994.
- [7] M Förl et al, "Necessary and sufficient conditions for absolute stability of neural networks", *IEEE Trans. on Circuits and Systems-I*, vol. 41, no. 2, pp 171-174, 1994.
- [8] P.P. Varaiya and R.Liu, "Bounded-input bounded output stability of nonlinear time-varying differential systems", *SIAM Journ. of Control*, vol. 4, pp 698-704, 1966.
- [9] E.D.Sontag, "Smooth stabilization implies coprime factorization", *IEEE Trans. on Automatic Control*, vol. 34, pp 435-443, 1989.
- [10] J. Tziniias, "Sontag's input to state stability condition and global stabilization using state detection", *Systems and Control Letters*, vol. 20, pp 219-226, 1993.
- [11] Z.P. Jiang, A. Teel, and L.Praly, "Small-gain theorem for ISS systems and applications", *Math. of Control, Signals, and Systems*, vol. 7, pp 104-130, 1994.
- [12] E.D.Sontag, "On the input to state stability property", *European Journal of Control*, vol. 1, pp 1-24, 1995.
- [13] José P. Pérez y Edgar N. Sánchez, "Simplified Analysis for GAS Neural Network", in Spanish, *Segunda Conferencia de Ing. Eléctrica*, CINVESTAV-IPN, Mexico City, Mexico, September 1996.
- [14] G.A.Rovithakis, and M.A.Christodoulou, "Adaptive Control of Unknown Plants using Dynamical Neural Networks", *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 24, pp 400-412, 1994.
- [15] A.S.Poznyak and E.N.Sanchez, "Nonlinear systems approximation:error stability analysis", *Intell. Automat. and Soft Compt., an Intl. Journ.*, vol.1, pp 247-258, 1995.
- [16] R.G.Bartle, "The Elements of Real Analysis", 2nd Edition, John Wiley and Sons, New York, 1976.



1997 American Control Conference

UNIVERSITY OF CALIFORNIA
SAN DIEGO, CALIFORNIA

Welcome

Getting Started

Conference Information

Sessions

Authors

Neural Network Applications I

Chair: Michel, Anthony N Univ. of Notre Dame

Co-chair: Fedrowitz, Christian H. Universitaet-GH-Siegen

- Application of dynamic neural networks to approximation and control of nonlinear systems
- Adjustable neural network controller: Application to a large segmented reflector
- A neural network based approach for the identification and optimal control of a cantilever plate
- Commutating switched reluctance motors efficiently via CMAC neural network with learning rate function
- Robust adaptive nonlinear system identification and trajectory tracking by dynamic neural networks
- Estimation of impurity and fouling in batch polymerisation reactors using stacked neural networks

Robust Adaptive Nonlinear System Identification and Trajectory Tracking by Dynamic Neural Networks¹

Alexander S. Poznyak[†], Edgar N. Sanchez[†], Jose P. Perez[‡] and Wen Yu[†]
[†]CINVESTAV-IPN, Seccion de Control Automatico, Av. IPN 2508, A.P. 14-740,
Mexico D.F., 07000, Mexico, e-mail: apoznyak@ctrl.cinvestav.mx.
[‡]Elect. Eng. Grad. Program, FIME, Univ. Aut. de Nuevo Leon (UANL)
A.P. F-93, C.U., San Nicolas de los Garza, N.L., C.P. 66450, Mexico

Abstract

We analyze adaptive nonlinear identification and trajectory tracking using a dynamic neural network, with the same state space dimension as the system. We assume the system space state completely measurable. By means of a Lyapunov-like analysis we determine stability conditions for the identification error. Then we analyze the trajectory tracking error when the adaptive controller is utilized. For the identification analysis we use an algebraic Riccati equation and for the tracking error a differential one using on-line adapted parameters of the neural network. The structure of our scheme is composed by two parts: the neural network identifier and the tracking controller. As our main contributions, we establish two theorems: the first one gives a bound for the identification error and the second one establish a bound for the tracking error.

1. INTRODUCTION

Recently, there has been a big interest in applying neural networks to identification and control of nonlinear systems. Nonlinear system identification can be approached as the approximation of the system behavior by dynamic neural networks. In this direction there exists already some results. They may be classified in two groups: The first one, as a natural extension, is based on the function approximation properties of static neural networks [1] and is limited for time belonging to a closed set. The second one uses the operator representation of the system to derive conditions for the validity of its approximation by a dynamic neural network; it has been extensively analyzed by I. W. Sandberg [2]. The structure proposed is constituted by the parallel connection of neurons, with no interaction between them; it is required that the nonlinear system fulfills the approximately-finite memory condition. In [3], a dynamic neural network, based on the Hopfield model, was proposed for nonlinear systems identification using operator representation based on the fading memory condition [4]. Both of them, the approximately-finite

memory and the fading memory conditions, require the nonlinear system to be stable.

The above mentioned results give just conditions for the existence of a dynamic neural network, which minimizes the approximation error to the nonlinear systems behavior; they do not determine the number of neurons and/or the value of their weights to effectively obtain the minimum error. A recently result [7] solves the problem of the neuron number by means of recursively high-order neural networks. There this number is selected to be equal to the dimension of the nonlinear systems state, which has to be completely measurable. This measurability condition is relaxed in [5] to singular perturbed systems.

There are not many stability analyses in neural control in spite of successful neural control applications reported and that, even for neural information storage application, energy function studies are used for proof of convergence to desired final values [6]. To the best of our knowledge, there are only a few results published regarding nonlinear systems control by dynamic neural networks. In our opinion, the most important are the ones by M.A. Christodoulou and coworkers [5, 7] and references therein. In [5] the neural network weights are on-line adapted to minimize the identification error. Stability of the whole system is performed via a Lyapunov function; as above mentioned their approach can deal with singular perturbed systems. In [7], they develop a direct adaptive regulation scheme and analyze its stability. Other results [8] utilize a SISO affine control representation of the nonlinear system, which is approximated by a dynamic neural network. This neural network is linearized by an inner loop designed using differential geometry techniques, and the outer control law is implemented using a PID controller.

In [9], we analyze both: nonlinear system identification and control. First, the nonlinear system is identified by means of a dynamic neural network; then, we force the identified system to track a signal generated by a nonlinear model using a nonlinear controller. The identification error and tracking error stability analysis is performed by a Lyapunov like method. It is worth mentioning that the

¹The research, reported in this paper, was supported by CONACYT Project 0652A9506 and partially by the Schools of Mechanical and Electrical Engineering and Mathematics and Physics of UANL.

the one introduced by A. N. Mitchel and coworkers for the robustness analysis of neural information storage [10].

This paper extends our previous results in the following ways: we consider on-line learning of the neural network parameters, and a slightly different adaptive controller is used for the trajectory tracking.

2. IDENTIFICATION ERROR STABILITY ANALYSIS

We assume the following neural network structure:

$$\dot{x}_{nn} = A_{nn}x_{nn} + B_{nn}(t)\sigma(x_{nn}) + C_{nn}\gamma(u_{nn}) \quad (1)$$

where $x_{nn} \in \mathbb{R}^n$, $u_{nn} \in \mathbb{R}^m$, $\sigma(x_{nn}) \in \mathbb{R}^n$, $A_{nn}, B_{nn} \in \mathbb{R}^{n \times n}$, $C_{nn} \in \mathbb{R}^{n \times m}$. A_{nn} is a Hurwitz matrix, B_{nn} is the matrix for nonlinear state feedback, C_{nn} is the input matrix, x_{nn} is the state of the neural network,

$\sigma(\cdot)$ is a nonlinear state feedback function, and $\gamma(\cdot)$ is an input function, and nn identifies parameters and variables related to the neural network mathematical model.

The neural network (1) can be classified as a Hopfield-type one. $\sigma(\cdot)$, $\gamma(\cdot)$ are usually represented as sigmoids. We will adapt just the parameters of the nonlinear feedback matrix, so it is time-varying; the parameter adaptation will be done using the same algorithm as in [5].

The nonlinear system to be identified is given as:

$$\dot{x} = f(x, u, t), \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m. \quad (2)$$

We assume the system state completely measurable, and the state and input dimension of the system equal to those of the neural network. Under these conditions, we define the state approximation error as:

$$\Delta = x - x_{nn}, \quad \dot{\Delta} = \dot{x} - \dot{x}_{nn}. \quad (3)$$

Adding and subtracting the term Ax , we obtain:

$$\begin{aligned} \dot{\Delta} = & A\Delta + [f(x, u, t) - A_{nn}x_{nn} \\ & - B_{nn}\sigma(x_{nn}) - C_{nn}\gamma(u) - A\Delta] \end{aligned}$$

with A any Hurwitz matrix, and $u_{nn} = u$.

Under these conditions, the ODE describing the error behavior is:

$$\begin{aligned} \dot{\Delta} = & f(x, u, t) - A_{nn}x_{nn} \\ & - B_{nn}\sigma(x_{nn}) - C_{nn}\gamma(u) - A\Delta. \end{aligned} \quad (4)$$

At this stage, let suppose the following assumptions are fulfilled:

A1 - There exist positive defined matrices H_n and H_Δ such that:

$$\begin{aligned} h^T(\Delta, x, u, t) H_n h(\Delta, x, u, t) \\ \leq \varepsilon_0(x, u, t) + \varepsilon_1(x, u, t) \Delta^T H_\Delta \Delta \end{aligned} \quad (5)$$

for $\varepsilon_0(\cdot, \cdot, \cdot)$ and $\varepsilon_1(\cdot, \cdot, \cdot)$ positive bounded functions, with ε^0 and ε^1 the respective bounds, i.e., $\sup \varepsilon_i(x, u, t) = \varepsilon^i < \infty$, $i = 0, 1$, $\forall x \in \mathbb{R}^n$, $\forall u \in \mathbb{R}^m$, $\forall t \geq 0$.

“strip function” condition [11]¹:

$$\begin{aligned} \|f_0(x, u, t) - A_{nn}x - (A - A_{nn}) \\ - B_{nn}\sigma(x - \Delta) - C_{nn}\gamma(u)\| \leq L_f \end{aligned}$$

for $\forall x \in \mathbb{R}^n$, $\forall u \in \mathbb{R}^m$, $\forall t \geq 0$.

A3 - The functions $\sigma(x)$ and $\phi(x)$ are Lipschitzian with constant L_ϕ and L_σ .

A4 - $\sigma(x)$ satisfies:

$$\sigma^T(x) Z \sigma(x) \leq x^T \Lambda_\sigma x + \lambda_{\sigma\sigma}$$

where $Z = Z^T > 0$, $\Lambda_\sigma = \Lambda_\sigma^T > 0$, $\lambda_{\sigma\sigma} > 0$, $\forall x \in \mathbb{R}^n$.

A5 - There exists a strictly positive defined matrix Q_0 such that, for the matrix A above defined, the following matrix Riccati equation:

$$A^T P + PA + PRP + Q = 0$$

where $R := H_{n0}^{-1} + H_{nf}^{-1}$, $Q := \epsilon^{h1} H_{\Delta 0} + \epsilon^{f1} H_{\Delta f} + Q_0$, has a solution $P_I = P_I^T > 0$. Such solution exists if the matrix A is stable, the pair $(A, R^{\frac{1}{2}})$ is controllable and the pair $(Q, R^{\frac{1}{2}})$ is observable. These conditions are easily fulfilled selecting A diagonal.

Comment 1 - All matrices $H_{(\cdot)}$ normalize the respective vector to one with dimensionless components.

Comment 2 - Equation (5) represents an ellipsoid type restriction for the respective vector fields.

Comment 3 - We use the same learning law as in [5]; therefore matrix $B_{nn}(t)$ is bounded $\|B_{nn}(t)\| \leq B$.

Comment 4 - If $\sigma(x)$ is chosen as a sigmoid function, **A3** and **A4** are fulfilled.

Lemma 1 In inequality (5), for any fixed $H_n, \chi > 0$, and under Assumptions **A1** and **A3**, we have

$$\begin{aligned} \varepsilon_0 = & (1 + \chi) \|f(x, u, t) - A_{nn}x - B_{nn}\sigma(x) - C_{nn}\gamma(u)\|_{H_n}^2, \\ \varepsilon_1 = & (1 + \chi^{-1}) (\|B_{nn}\| L_\sigma + \|A - A_{nn}\|)^2 \|H_n\|, \\ H_\Delta = & I, \varepsilon^0 = \sup_{x, u, t} \varepsilon_0 \leq (1 + \chi) \|H_n\| L_f^2. \end{aligned}$$

Proof: - It directly follows from:

$$\begin{aligned} h(\Delta, x, u, t) = & h_1(x, u, t) + h_2(\Delta, x) \\ h_1(x, u, t) = & f(x, u, t) - A_{nn}x - B_{nn}\sigma(x) - C_{nn}\gamma(u) \\ h_2(\Delta, x) = & B_{nn}(\sigma(x) - \sigma(x - \Delta)) - (A - A_{nn})\Delta \\ \|h\|_{H_n}^2 \leq & (1 + \chi) \|h_1\|_{H_n}^2 + (1 + \chi^{-1}) \|h_2\|_{H_n}^2. \end{aligned}$$

Under the above assumptions and Lemma 1, we establish the first of our main contributions: ■

¹In this article: a) $\|\cdot\|$ denotes the Euclidian norm for vectors, and for any matrix A it is defined as $\sqrt{\lambda_{\max}(A^T A)}$, where $\lambda_{\max}(\cdot)$ is the maximum eigenvalue of the respective matrix; b) $\|\cdot\|_W$ is the weighted euclidian norm.

Theorem 2 For the system (2) and the given initial work (1), the following property holds:

$$\limsup_{t \rightarrow \infty} \|\Delta(t)\| \leq \Delta^0(P) \equiv \sqrt{\frac{\varepsilon^0}{\lambda_{\max}(P)\lambda_{\min}(R_p)}} \quad (6)$$

where $R_p = P_1^{-\frac{1}{2}} Q_0 P_1^{-\frac{1}{2}}$ and $\lambda_{\min}(\cdot)$ is the minimum eigenvalue of the respective matrix.

Proof: - Let consider the nonnegative defined scalar function $V(\Delta) = \Delta^T P_1 \Delta \in \mathfrak{R}_+$. Calculating its time derivative over the trajectories of equation (4), we obtain:

$$\begin{aligned} \dot{V}(\Delta) &= 2\Delta^T P_1 (A\Delta + h) \\ &= \Delta^T (P_1 A + A^T P_1) \Delta + 2\Delta^T P_1 h \end{aligned} \quad (7)$$

where the function $h = h(\Delta, x, u, t)$ is defined in (4). Using the matrix inequality given as:

$$X^T Y + Y^T X \leq X^T \Lambda X + Y^T \Lambda^{-1} Y \quad (8)$$

which is true for all matrices $X \in \mathfrak{R}^{n \times k}$, $Y \in \mathfrak{R}^{n \times k}$, and for any matrix $\Lambda = \Lambda^T > 0$, $\Lambda \in \mathfrak{R}^{n \times n}$, as demonstrated in [11]. Applying **A1** we formulate:

$$2\Delta^T P_1 h \leq \Delta^T (P_1 H_n^{-1} P_1 + \varepsilon^1 H_\Delta) \Delta + \varepsilon^0. \quad (9)$$

Substituting inequalities (9) in (7), then:

$$\dot{V}(\Delta) = \Delta^T (P_1 A + A^T P_1 + P_1 H_n^{-1} P_1 + \varepsilon^1 H_\Delta) \Delta + \varepsilon^0.$$

Considering **A5**, we obtain:

$$\dot{V}(\Delta) \leq -\lambda_{\min}(R_p) V(\Delta) + \varepsilon^0. \quad (10)$$

from which we obtain (6). ■

3. TRACKING ERROR ANALYSIS

The control goal is to force the system states to track a signal, generated by a nonlinear reference model as $\hat{x}^* = \varphi(x^*, t)$. The state trajectory tracking can be formulated as $J_{\min} = \min_{u(t)} J$ where

$$J = \|x - x^*\|_Q^2 + \|u\|_R^2, \|x\|_Q^2 = \overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau x^T(t) Q x(t) dt.$$

So, for any $\eta > 0$, then:

$$J \leq (1 + \eta) \|x - x_{nn}\|_Q^2 + (1 + \eta^{-1}) \|x_{nn} - x^*\|_Q^2 + \|u\|_R^2.$$

The minimum of the term $\|x - x_{nn}\|_Q^2$ has already been solved for the identification analysis. So, the control goal is to minimize $\|x_{nn} - x^*\|_Q^2$. We define the state trajectory error as $\Delta^* = x_{nn} - x^*$. So, the ODE which describes the state trajectory error is:

$$\begin{aligned} \dot{\Delta}^* &= A_{nn} \Delta^* + B_{nn} \sigma(\Delta^* + x^*) + C_{nn} \gamma(u) + \xi^*(t), \\ \xi^*(t) &= A_{nn} x^* - \varphi(x^*, t). \end{aligned} \quad (11)$$

introduce the following Lyapunov function: $V_t(\Delta^*) = \Delta^{*T} P_c(t) \Delta^*$, whose time derivative is:

$$\begin{aligned} \dot{V}_t(\Delta^*) &= 2\Delta^{*T} P_c(t) (A_{nn} \Delta^* + B_{nn} \sigma(\Delta^* + x^*) \\ &\quad + C_{nn} \gamma(u) + \xi^*(t)) + \Delta^{*T} \dot{P}_c(t) \Delta^*. \end{aligned} \quad (12)$$

We take into account the following identity:

$$\begin{aligned} 2(C_{nn}^T P \Delta^*)^T \gamma(u) &= \|C_{nn}^T P \Delta^* + \gamma(u)\|^2 \\ &\quad - \Delta^{*T} P (C_{nn} C_{nn}^T) P \Delta^* - \|\gamma(u)\|^2. \end{aligned} \quad (13)$$

First, we apply (8) to $2(B_{nn}^T(t) P_c(t) \Delta^*)^T \sigma(\Delta^* + x^*)$ and then we utilize **A4** to obtain:

$$\begin{aligned} 2(B_{nn}^T(t) P_c(t) \Delta^*)^T \sigma(\Delta^* + x^*) \\ \leq \Delta^{*T} P_c(t) (B_{nn}(t) Z^{-1} B_{nn}^T(t)) P_c(t) \Delta^* \\ + 2\Delta^{*T} \Lambda_\sigma \Delta^* + 2x^{*T} \Lambda_\sigma x^* + \lambda_{\sigma 0}. \end{aligned}$$

Then, we employ (8) in $2(P_c(t) \Delta^*)^T \xi^*(t)$ so:

$$2(P \Delta^*)^T \xi^*(t) \leq \Delta^{*T} P \Lambda_\xi^{-1} P \Delta^* + \xi^{*T}(t) \Lambda_\xi \xi^*(t). \quad (14)$$

Substituting (13), (14) in (12), adding and subtracting $\Delta^{*T} Q_c \Delta^*$, we formulate:

$$\begin{aligned} \dot{V}(\Delta^*) &\leq \Delta^{*T} [P_c(t) A_{nn} + A_{nn}^T P_c(t) + \\ &\quad P_c(t) (B_{nn}(t) Z^{-1} B_{nn}^T(t) + \Lambda_\xi^{-1} - C_{nn} C_{nn}^T) P_c(t) \\ &\quad + 2\Lambda_\sigma + Q_c + P_c(t)] \Delta^* + 2x^{*T} \Lambda_\sigma x^* + \xi^{*T}(t) \Lambda_\xi \xi^*(t) - \\ &\quad \|\gamma(u)\|^2 + \|C_{nn}^T P_c(t) \Delta^* + \gamma(u)\|^2 - \Delta^{*T} Q_c \Delta^* + \lambda_{\sigma 0}. \end{aligned}$$

Let us choose $P_c(t)$ such that following differential Riccati equation is satisfied:

$$\begin{aligned} P A_{nn} + A_{nn}^T P + P (B_{nn} Z^{-1} B_{nn}^T + \Lambda_\xi^{-1} \\ - C_{nn} C_{nn}^T) P + 2\Lambda_\sigma + Q = -\dot{P}_c(t). \end{aligned} \quad (15)$$

Taking into account *Comment 3* ($\|B_{nn}(t)\| \leq B, \forall t \geq 0$), the term $B_{nn}(t) Z^{-1} B_{nn}^T(t)$ in the left side hand of (15) can be estimated by

$$B_{nn}(t) Z^{-1} B_{nn}^T(t) \leq B^2 \lambda_{\max}(Z^{-1}) I$$

that leads to a differential Riccati equation with constant parameters. Starting with an initial condition $P_c(0) = P_c^T(0) > 0$ close to the equilibrium point of this stationary differential Riccati equation, we can conclude that under conditions analogous to **A5** the original one has a solution $P_c(t) = P_c^T(t) > 0, \forall t \geq 0$. If (15) is satisfied, then the time derivative of the Lyapunov function $V(\Delta^*)$ fulfills the inequality:

$$\begin{aligned} \dot{V}(\Delta^*) &\leq 2x^{*T} \Lambda_\sigma x^* + \xi^{*T}(t) \Lambda_\xi \xi^*(t) - \\ &\quad \|\gamma(u)\|^2 + \|C_{nn} P_c(t) \Delta^* + \gamma(u)\|^2 - \\ &\quad \Delta^{*T} Q_c \Delta^* + \lambda_{\sigma 0} + u^T R_c u - u^T R_c u. \end{aligned} \quad (16)$$

Let define the function $\Psi(u)$ as:

$$\Psi(u) = -\|\gamma(u)\|^2 + \|C_{nn}^T P_c(t) \Delta^* + \gamma(u)\|^2 + u^T R_c u.$$

$$\begin{aligned} \dot{V}(\Delta^*) &\leq 2x^{*T} \Lambda_\sigma x^* + \Psi(u) - \Delta^{*T} Q_C \Delta \\ &- u^T R_C u + \lambda_{\sigma 0} + \xi^{*T}(t) \Lambda_\xi \xi^*(t). \end{aligned} \quad (17)$$

In order to calculate the control action $u(t)$, which minimizes the tracking error, let start by minimizing $\Psi(u)$, so $\frac{d\Psi(u)}{du}$ has to be equal to zero. To perform this minimization, we assume that, at the given t (positive), $x^*(t)$ and $x_{nn}(t)$ are known and do not depend on $u(t)$.

Comment 5 - We name the $u^*(t)$ signal, minimizing $\Psi(u)$, a locally optimal control, because it is calculated based only on "local" information. To solve this optimization problem, let us consider the following recursive gradient scheme:

$$u_k(t) = u_{k-1}(t) - \tau_k \frac{d\Psi(u_{k-1}(t))}{du}, \quad u_0(t) = 0.$$

where the gradient $\frac{d\Psi(u)}{du}$ is calculated as:

$$\frac{d\Psi(u)}{du} = 2 \frac{d\gamma(u)}{du} C_{nn}^T P_c(t) \Delta^*(t) + 2R_C u$$

and the sequence of the scalar parameter τ_k satisfies the condition:

$$\tau_k \geq 0, \quad \sum_{k=0}^{\infty} \tau_k = \infty, \quad \tau_k \rightarrow 0.$$

For example, we can select $\tau_k = \frac{1}{(1+k)^\tau}$, $\tau \in (0, 1]$. Concerning $u^*(t)$, we state the following lemma.

Lemma 3 $u^*(t)$ can be calculated as the limit of the sequence $\{u_k(t)\}$, i.e.,

$$u_k(t) \rightarrow u^*(t), \quad k \rightarrow \infty. \quad (18)$$

Proof: - It directly follows from the properties of the gradient method, taking into account (11) and (18). ■

COROLLARY 1 - If nonlinear input function to the neural network (1) depends linearly on $u(t)$, we can select $\frac{d\gamma(u)}{du} = \Gamma$, and we can compensate the measurable signal $\xi^*(t)$ in (11) by the modified control law:

$$u(t) = u_{comp}(t) + u^*(t)$$

where $u_{comp}(t)$ satisfies the relation:

$$C_{nn} u_{comp}(t) + \xi^*(t) = 0$$

and u^* is selected according to the linear squares optimal control law:

$$u^*(t) = -R_C^{-1} \Gamma^{-1} C_{nn}^T P_c(t) \Delta^*(t). \quad (19)$$

It is worth mentioning the similarity between $u^*(t)$ given by (19) and a linear least squares optimal control law. At this stage, we establish our second main contribution

(1). the nonlinear reference model, and the control law (19), the following property holds:

$$\|\Delta^*\|_Q^2 + \|u^*\|_Q^2 \leq 2\|x^*\|_{\Lambda_\sigma}^2 + \overline{\lim}_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \Psi(u^*(t)) dt + \lambda_{\sigma 0} \quad (20)$$

Proof: - After knowing the minimum $\Psi(u^*(t))$ (17) can be rewritten as:

$$\begin{aligned} \dot{V}(\Delta^*) &\leq 2x^{*T} \Lambda_\sigma x^* + \Psi(u^*(t)) - \Delta^{*T} Q \Delta^* - \\ &u^{*T}(t) R u(t) + \lambda_{\sigma 0} + \lambda_{\sigma 0} + \xi^{*T}(t) \Lambda_\xi \xi^*(t). \end{aligned}$$

Then, integrating each term from 0 to τ , dividing each term by τ , and taking the limit, for $\tau \rightarrow \infty$, of these integrals' supreme and taking into account that

$$\overline{\lim}_{\tau \rightarrow \infty} \left[-\frac{1}{\tau} \int_0^\tau \dot{V}(\Delta^*(t)) \right] \leq \overline{\lim}_{\tau \rightarrow \infty} \left[\frac{1}{\tau} V(\Delta^*(0)) \right] = 0 \quad (21)$$

we obtain (20). ■

Equation (20) fixes a tolerance level for the trajectory tracking error.

4. APPLICATIONS

To illustrate the applicability of our approach, we include the following example. Given the nonlinear system:

$$\dot{x}_i = -\alpha_i x_i + \beta_i \text{sign}(x_{j \neq i}) + u_i, \quad i, j = 1, 2. \quad (22)$$

Let select the neural network as follows:

$$\begin{aligned} \dot{x}_{nn,i} &= -a_{nn,i} x_{nn,i} + b_{nn,i1} \sigma_{nn,1}(x_{nn,1}) \\ &+ b_{nn,i2} \sigma_{nn,2}(x_{nn,2}) + c_{nn,i1} u_1 + c_{nn,i2} u_2, \quad i = 1, 2. \end{aligned}$$

where $\sigma(\cdot)$ indicates a sigmoidal function. The given nonlinear system (22), even simple, is interesting enough; it has multiple isolated equilibrium. We consider:

$$\begin{aligned} a_{nn,1} &= \alpha_1 = 5, x_1(0) = 10, x_{nn,1}(0) = -1, \beta_1 = 3, \\ a_{nn,2} &= \alpha_2 = 10, x_2(0) = -10, x_{nn,2}(0) = -5, \beta_2 = 2. \end{aligned}$$

The sigmoid function is chosen as: $\sigma(x) = 2(1 + e^{-2x})^{-1} - 0.5$. The structure of the neural network is similar to the one used in [5]. We select $Q_C = Q_0 = H_{n0} = H_{nf} = H_{\Delta 0} = H_{\Delta f} = I$, $\varepsilon^{h1} = \varepsilon^{f1} = 3$, and obtain the solution $P_c(t)$ of the corresponding differential Riccati Equation (15) starting from the initial conditions: $P_0 = \begin{bmatrix} 0.84 & 0 \\ 0 & 0.84 \end{bmatrix}$, that corresponds to the solution of the algebraic matrix Riccati equation when the left side of this equation is equal to zero. Choosing $C_{nn} = I$ and $R_C = \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix}$, and applying the control law given by (19), we obtain the results explained below.

a) *Nonlinear identification*

To adapt on-line the dynamic neural network weights we use the same learning algorithm as in [5]. The input signal u_1, u_2 are chosen as sine wave and saw-tooth function; the results are shown as Fig.2. and Fig.3. The solid lines correspond to nonlinear system state responses $x_1(t), x_2(t)$, and the dashed line to neural network ones $x_{nn,1}(t), x_{nn,2}(t)$. The abscissa values correspond to the number of iterations. It can be seen that the neural network state time evolution follows the ones of nonlinear system.

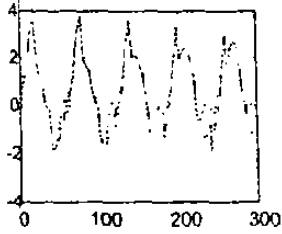


Figure 2: x_1

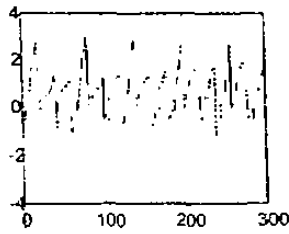


Figure 3: x_2

b) Trajectory Tracking.

We implement the control law given by equation (19); it constitutes a feedback control with an on-line adaptive gain. Fig.4 and Fig.5 present the respective responses, where the solid line corresponds to reference model signals $x_1^*(t), x_2^*(t)$, and the dashed line to the nonlinear system responses $x_1(t), x_2(t)$

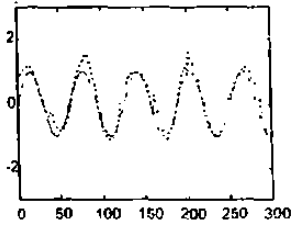


Figure 4: x_1

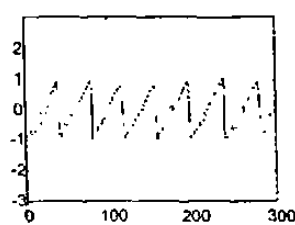


Figure 5: x_2

The time evolution for the seminorms

$$\|\Delta^*\|_Q^2 = \limsup_{\tau \rightarrow \infty} J_\tau^\Delta, \quad J_\tau^\Delta = \frac{1}{\tau} \int_0^\tau \Delta^{*T}(t) Q \Delta(t) dt.$$

can be seen in Fig.6. The time evolution for the entries of P_c in (15) is shown as Fig 7.

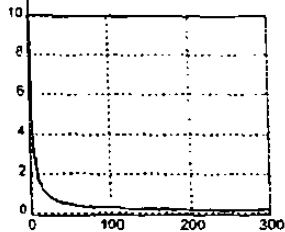


Figure 6: J_τ^Δ

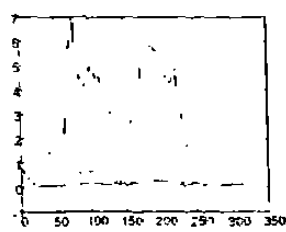


Figure 7: P_c

5. CONCLUSIONS

In this paper we propose a new adaptive neural controller, which is composed of two parts: a neural network identifier and a tracking controller. We establish bounds for both the identification and the tracking error. It is

worth mentioning that the reference model is nonlinear; so the developed structure allows the implementation of many applications.

We tested the performance with an interest system; it has multiple equilibrium and the vector field is not differentiable. In the future we will develop the implementation of this scheme for nonlinear systems presenting phenomena such as: saturation, friction, and hysteresis.

As a continuation of our research, we were able to develop and implement an adaptive identification and trajectory tracking scheme for nonlinear systems. Work is in progress to relax the condition of complete measurability of the system state.

References

- [1] E.Sontag, "Neural nets as systems models and controllers", in Proc. 7th Yale Workshop on Adaptive and Learning Systems, pp 73-79, Yale University, 1992.
- [2] I.W. Sandberg, "Approximation theorems for discrete-time systems", IEEE Trans. on Circ. and Syst., Vol. 38, pp 564-566, 1991.
- [3] E.Alcorta and E.N.Sanchez, "Nonlinear identification via neural networks", in Proc.4th IFAC, Intl. Symp. on Adapt.Syst. in Cont. and Sign. Process., pp 675-679, June 1992.
- [4] E.N.Sanchez, "Dynamic neural network for nonlinear systems identification", in Proc.33rd IEEE CDC, pp 2480-2481, December 1994.
- [5] G.A.Rovithakis, and M.A.Christodoulou, "Adaptive control of unknown plants using dynamical neural networks", IEEE Trans. on Syst., Man and Cybern., vol. 24, No 3, pp 400-412, March 1994.
- [6] F.L.Lewis and T.Parisini, "Neural network control with guaranteed stability", proposal for invited session, 34 IEEE CDC, February 1995.
- [7] A.Rovithakis, and M.A.Christodoulou, "Adaptive control of unknown plants using dynamical neural networks", IEEE Trans. on Syst., Man and Cybern., vol. 25, No 12, pp 1578-1594, December 1995.
- [8] A.Delgado, C.Kambahmpati, and K.Warwick, "Dynamic recurrent neural network for systems identification and control", IEE Proc.-Cont. Theo. Appl., Vol. 142, No 4, July 1995.
- [9] A.Poznyak, and E.N. Sanchez, "Nonlinear system identification and trajectory tracking using dynamic neural networks", accepted for 35th IEEE CDC, 1996.
- [10] K.Wang, and A.N.Mitchel, "Robustness and perturbation analysis of a class of artificial neural networks", Neural Networks, Vol. 7, No 2, pp 251-257, 1994.
- [11] A.S.Poznyak, and E.N.Sanchez, "Nonlinear system approximation by neural networks: error stability analysis", Intl. Journ. of Intell. Autom. and Soft Comput., Vol. 1, No 3, pp 247-258, 1995.

ISSN 0957-4174
Volume 14, Number 1/2, January/February 1998

Expert Systems with Applications

An International
Journal

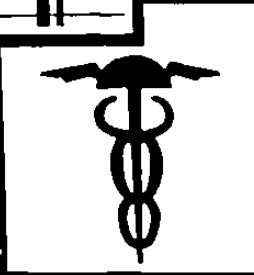
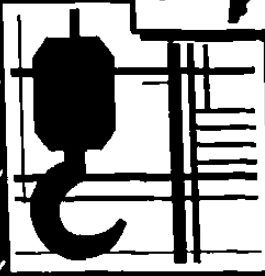
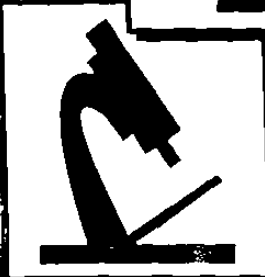
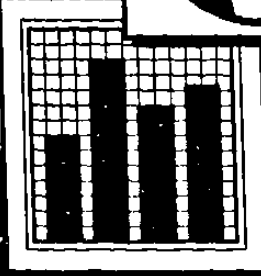
Editor-in-Chief
Jay Liebowitz

Special Issue:
ARTIFICIAL INTELLIGENCE IN MEXICO

Guest Editors:
FRANCISCO J. GANTU
CARLOS ZOZAYA



Pergamon



Stability analysis of dynamic neural control

Alexander S. Poznyak* and Wen Yu

CINVESTAV-IPN, Sección de Control Automático, Av. IPN 2508, A.P. 14 740, Mexico D.F., 07000, México

Edgar N. Sanchez

CINVESTAV, Unidad Guadalajara, Apartado Postal 31-438, Plaza La Luna, Guadalajara, Jalisco, 44550, México

Jose P. Perez

Mathematics and Physics Sciences School, Universidad Autónoma de Nuevo Leon, Calle Pedro de Alba s/n, Ciudad Universitaria, San Nicolas de los Garza, N.L. C.P. 66450, México

Abstract

In this paper, the authors summarize their research related to dynamic neural control. In particular, results on nonlinear system identification, nonlinear trajectory tracking, and input-to-state stability (ISS) of dynamic neural networks are presented. The main analysis tool utilized is the Lyapunov approach. References for the detailed demonstrations are given. We illustrate the applicability of the results by means of examples. © 1998 Elsevier Science Ltd. All rights reserved.

1. INTRODUCTION

In spite of reported successful neural control application (Agarwal, 1997), and that even for neural information storage application, energy function studies are used for proof of convergence to desired values, there are not many stability analysis in neural control (Lewis & Parissini, 1995). Regarding nonlinear system control by dynamic neural networks, to the best of our knowledge, there are only a few results published. In our opinion, the most important are Rovithakis and Christodoulou (1994, 1995), and references therein. They utilize a particular version of recursively high-order neural networks. In Rovithakis and Christodoulou (1994), they identify a nonlinear system by means of dynamic neural networks, then calculate the control law, based on the neural network model, to force the system to follow a linear model. The neural network weights are on-line adapted to minimize the identification error. Stability of the whole system is analyzed via a Lyapunov function; their scheme can deal with singular perturbed systems. In Rovithakis and Christodoulou (1995), they develop a

direct adaptive regulation scheme for affine in the control nonlinear systems; again they analyze stability using a Lyapunov function. In both papers, they illustrate the applicability of the respective approach by a D.C. motor speed control. Other results (Delgado et al., 1995) utilize a SISO affine control representation of the nonlinear system, which is approximated by a dynamic neural network. This neural network is linearized by an inner loop designed based on differential geometry techniques; the outer control law is implemented using a PID controller.

There exist two methodologies for the stability analysis of nonlinear dynamic systems: (a) the input-output approach based on operator theory, which requires the operator, representing the system, to be bounded; (b) the state space approach, whose main tool the Lyapunov approach. These two methods are not equivalent for nonlinear systems (Varaiya & Liu, 1966). The input-to-state stability (ISS) concept, introduced in Sontag (1989), allows us to establish this equivalence; it has been applied to derive important new results concerning nonlinear control (Jiang et al., 1994; Tziniyas, 1993). ISS can be derived by different ways, such as: dissipativity, robust margins, or in the classical Lyapunov sense

* Author for correspondence

(Sontag, 1995).

In this paper, we discuss both nonlinear system identification and control. For control, we force the identified system to track a signal generated by a nonlinear model. The identification error and tracking error stability analysis is performed via a Lyapunov-like method. We establish, as theorems, the bounds for these errors. We also present the ISS analysis for dynamic neural networks; to the best of our knowledge, this is the first time such an analysis is performed.

The paper outline is as follows: in Section 2, nonlinear identification is discussed, then in Section 3 we present the nonlinear tracking scheme. In Section 4, we establish ISS conditions for dynamic neural networks; then in Section 5 we include examples which illustrate the applicability of the results. Finally, in Section 6, we state the relevant conclusions.

2. NONLINEAR SYSTEMS IDENTIFICATION

In this section, we develop the analysis of nonlinear system identification by means of dynamic neural networks. We assume the following structure for the neural network:

$$\begin{aligned} \dot{x}_{nn} &= A_{nn}x_{nn} + B_{nn}\sigma(x_{nn}) + C_{nn}\phi(x_{nn})\gamma(u_{nn}) \quad (1) \\ x_{nn} &\in \mathbb{R}^n, u_{nn} \in \mathbb{R}^m, \sigma(x_{nn}) \in \mathbb{R}^n, \phi(x_{nn})\gamma(u_{nn}) \in \mathbb{R}^n \\ A_{nn}, B_{nn} &\in \mathbb{R}^{n \times n}, C_{nn} \in \mathbb{R}^{n \times m} \end{aligned}$$

where: x_{nn} is the state of the neural network, u_{nn} is the neural network input, A_{nn} is the state feedback matrix, B_{nn} is the nonlinear feedback matrix, C_{nn} is the input matrix, $\sigma(\cdot)$, $\phi(\cdot)$ are sigmoidal functions, $\gamma(\cdot)$ is a bounded nonlinear function, and nn labels the neural network parameters.

The nonlinear system is defined as:

$$\dot{x} = f(x, u, t), \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m \quad (2)$$

where: x is the system state, u is the input, f is a vector valued function, and t is time.

We assume the system state completely measurable, and the neural network state dimension equal to the nonlinear system one. The identification error is defined as:

$$\Delta = x - x_{nn}, \quad \dot{\Delta} = \dot{x} - \dot{x}_{nn} \quad (3)$$

Adding and subtracting $A\tau$, we obtain:

$$\begin{aligned} \dot{\Delta} &= A\Delta + [f(x, u, t) - A_{nn}x_{nn} - B_{nn}\sigma(x_{nn}) - C_{nn}\phi(x_{nn})\gamma(u) \\ &\quad - A\Delta] \end{aligned} \quad (4)$$

with A any Hurwitz matrix and $u_{nn} = u$.

The ordinary differential equation describing the error

dynamic is given as:

$$\dot{\Delta} = A\Delta + h(\Delta, x, u, t) \quad (5)$$

$$\begin{aligned} h(\Delta, x, u, t) &= f(x, u, t) - A_{nn}x - (A - A_{nn})\Delta - B_{nn}\sigma(x - \Delta) \\ &\quad - C_{nn}\phi(x - \Delta)\gamma(u) \end{aligned}$$

We assume that the following assumptions are fulfilled.

Assumption 2.1. As in Poznyak and Sanchez (1995), there exist matrices H_n, H_j positive definite, such:

$$h^T(\Delta, x, u, t)H_n h(\Delta, x, u, t) \leq \epsilon_0(x, u, t) + \epsilon_1(x, u, t)\Delta^T H_j \Delta \quad (6)$$

for $\epsilon_0(\dots), \epsilon_1(\dots)$ bounded positive functions, with ϵ^0, ϵ^1 the respective bounds, i.e.:

$$\sup \epsilon_i(x, u, t) = \epsilon^i < \infty, \quad i = 0, 1 \quad (7)$$

$$\forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^m, \forall t \geq 0$$

Assumption 2.2. f satisfies¹

$$\|f(x, u, t) - A_{nn}x - B_{nn}\sigma(x) - C_{nn}\phi(x)\gamma(u)\| \leq L_f \quad (8)$$

$$\forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^m, t \geq 0$$

Assumption 2.3. $\phi(x)$ and $\sigma(x)$ are Lipschitz, i.e.:

$$\|\phi(x) - \phi(x - \Delta)\| \leq L_\phi \|\Delta\|, \|\sigma(x) - \sigma(x - \Delta)\| \leq L_\sigma \|\Delta\| \quad (9)$$

$$\forall x, \Delta \in \mathbb{R}^n$$

Assumption 2.4. As usual in identification, $u(t)$ is bounded for $\forall t > 0$, so:

$$\|\gamma(u)\| \leq L_u \quad (10)$$

Assumption 2.5 There exists Q_0 strictly positive definite matrix such, for the previous defined A matrix, the following matrix Riccati equation:

$$A^T P + PA + PRP + Q = 0 \quad (11)$$

$$R = H_n^{-1}, Q = \epsilon^1 H_d + Q_0$$

has a solution $P = P^T > 0$

This solution exists if the matrix A is stable, $(A, R^{1/2})$ is controllable, and $(Q^{1/2}, A)$ is observable (Willems, 1971). These conditions are easily fulfilled, selecting A diagonal.

At this stage, we establish the following lemma.

Lemma 2.1. In inequality equation (6), for any H_n ,

¹In this paper (a) $\|\cdot\|$ denotes the Euclidian norm for vectors, and for any matrix A is defined as $\sqrt{\lambda_{\max}(A^T A)}$, with $\lambda_{\max}(\cdot)$ its maximum eigenvalue. (b) $\|\cdot\|_w$ is the weighted Euclidian norm

$\chi > 0$ fixed, and under assumptions 2.2 and 2.3, then:

$$\varepsilon_0(x, u, t) = (1 + \chi) \|f(x, u, t) - A_{nn}x - B_{nn}\sigma(x) - C_{nn}\phi(x)\chi(u)\|_{H_n}^2 \quad (12)$$

$$\varepsilon_1(x, u, t) = (1 + \chi^{-1})(\|B_{nn}\|L_\sigma + \|C_{nn}\|L_\phi L_u + \|A - A_{nn}\|)^2 \|H_n\| H_\Delta = I$$

$$\varepsilon^0 = (1 + \chi) \|H_n\| L_f^2$$

Proof. It follows directly from:

$$h(x, u, \Delta, t) = h_1(x, u, t) + h_2(x, u, \Delta) \quad (13)$$

$$h_1(x, u, t) = f(x, u, t) - A_{nn}x - B_{nn}\sigma(x) - C_{nn}\phi(x)\chi(u)$$

$$h_2(x, u, \Delta) = B_{nn}(\sigma(x) - \sigma(x - \Delta)) + C_{nn}(\phi(x) - \phi(x - \Delta))\chi(u) - (A - A_{nn})\Delta$$

$$\|h\|_{H_n}^2 = \|h_1 + h_2\|_{H_n}^2 \leq (1 + \chi) \|h_1\|_{H_n}^2 + (1 + \chi^{-1}) \|h_2\|_{H_n}^2$$

Under the above mentioned assumptions and Lemma 2.1, we establish the following theorem.

Theorem 2.1. For the nonlinear system equation (2) and the neural network equation (1), the identification error is bounded by:

$$\limsup_{t \rightarrow \infty} \|\Delta(t)\| \leq \Delta^0(P) \equiv \sqrt{\frac{\varepsilon^0}{\lambda_{\min}(P)\lambda_{\max}(R_p)}} \quad (14)$$

where $R_p = P^{-1/2} Q_0 P^{1/2}$ and $\lambda_{\min}(\cdot)$ is the minimum eigenvalue of the respective matrix.

Proof. Let consider the non-negative defined scalar function $V(\Delta) \in \mathfrak{R}$, as follows:

$$V(\Delta) = \Delta^T P \Delta, P = P^T > 0, P \in \mathfrak{R}^{n \times n} \quad (15)$$

Calculating its time derivative along the trajectories of equation (5), we obtain:

$$\dot{V}(\Delta) = 2\Delta^T P(A\Delta + h) = \Delta^T (PA + A^T P)\Delta + 2\Delta^T Ph \quad (16)$$

where the function $h = h(\Delta, x, u, t)$ is defined in equation (5).

Using the matrix inequality given as:

$$X^T Y + Y^T \leq X^T \Lambda X + Y^T \Lambda^{-1} Y \quad (17)$$

which is true for all matrices $X \in \mathfrak{R}^{m \times 1}$, $Y \in \mathfrak{R}^{n \times 1}$, and for any matrix $\Lambda = \Lambda^T > 0$, $\Lambda \in \mathfrak{R}^{m \times m}$, as demonstrated in [11], we derive:

$$2\Delta^T Ph \leq h^T H_n h + \Delta^T P H_n^{-1} P \Delta \quad (18)$$

Applying assumptions equation (1) equation (5), we

obtain:

$$V(\Delta) \leq -\Delta^T Q_0 \Delta + \varepsilon^0 = -\Delta^T P^{-1/2} \left(P^{1/2} Q_0 P^{-1/2} \right) P^{1/2} \Delta + \varepsilon^0 \quad (19)$$

$$\dot{V}(\Delta) \leq -\lambda_{\min}(R_p) V(\Delta) + \varepsilon^0$$

with $R_p = P^{-1/2} Q_0 P^{-1/2}$

Solving equation (19), we derive:

$$V(\Delta(t)) \leq V(\Delta_0) e^{-\lambda_{\min}(R_p)t} + \varepsilon^0 \int_0^t e^{-\lambda_{\min}(R_p)(t-\tau)} d\tau \quad (20)$$

which can be explicitly evaluated as:

$$V(\Delta(t)) \leq V(\Delta_0) e^{-\lambda_{\min}(R_p)t} + \frac{\varepsilon^0}{\lambda_{\min}(R_p)} (1 - e^{-\lambda_{\min}(R_p)t}) \quad (21)$$

Taking into account the inequality:

$$V(\Delta) \geq \lambda_{\min}(P) \|\Delta\|^2 \quad (22)$$

finally, we obtain equation (15). ■

In Poznyak and Sanchez (1996a), we discuss the stability analysis for the identification error, when the components of B_{nn} are adapted on-line. In this case, because we use the same learning law as in [3], so $B_{nn}(t)$ is bounded, i.e. $\|B_{nn}(t)\| \leq B, \forall t \geq 0$.

Regarding the above mentioned results, we state the following comments.

Comment 2.1. equation (14) sets a bound for the maximum of the identification error.

Comment 2.2. equation (14) is valid even if the nonlinear system is unstable, the initial conditions are different for the nonlinear system and the neural network, or the time goes to infinity. This is a big improvement over existing results.

3. NONLINEAR SYSTEMS TRAJECTORY TRACKING

We consider $\phi(x) = I$, for the neural network equation (1). The goal is to force the system states of the nonlinear system equation (2) to track a reference signal, which is generated by a nonlinear reference model

$$\dot{x} = \varphi(x^*, t) \quad (23)$$

Defining the following semi-norms:

$$\|x\|_Q^2 = \limsup_{\tau \rightarrow \infty} \sup_{t \geq \tau} \int_t^\infty x^T(t) Q x(t) dt, \quad (24)$$

$$\|u\|_R^2 = \limsup_{\tau \rightarrow \infty} \sup_{t \geq \tau} \int_t^\infty u^T(t) R u(t) dt$$

$$Q = Q^T > 0, R = R^T > 0$$

the state trajectory tracking is formulated as:

$$J_{\min} = \min_{u(t)} J, J = \|x - x^*\|_Q^2 + \|u\|_R^2 \quad (25)$$

For any $\eta > 0$, then:

$$J \leq (1 + \eta) \|x - x_{\min}\|_Q^2 + (1 + \eta^{-1}) \|x_{\min} - x^*\|_Q^2 + \|u\|_R^2 \quad (26)$$

The minimum of $\|x - x_{\min}\|_Q^2$ has been obtained for the identification analysis. So the control goal is to minimize $\|x_{\min} - x^*\|_Q^2$; we define the state trajectory error as $\Delta^* = x_{\min} - x^*$.

The ordinary differential equation which describes this error dynamics is:

$$\dot{\Delta}^* = A_{nn} \Delta^* + B_{nn} \sigma(\Delta^* + x^*) + C_{nn} \gamma(u) + \xi^*(t) \quad (27)$$

$$\xi^*(t) = A_{nn} x^* - \varphi(x^*, t)$$

We define the following Lyapunov function:

$$V(\Delta^*) = \Delta^{*T} P \Delta^*, P = P^T > 0 \quad (28)$$

whose time derivative, along the tracking error trajectory, is:

$$\begin{aligned} \dot{V}(\Delta^*) &= \Delta^{*T} P \dot{\Delta}^* \\ &= 2\Delta^{*T} P (A_{nn} \Delta^* + B_{nn} \sigma(\Delta^* + x^*) + C_{nn} \gamma(u) + \xi(t)) \end{aligned} \quad (29)$$

$$\begin{aligned} \dot{V}(\Delta^*) &= \Delta^{*T} (P A_{nn} + A_{nn}^T P) \Delta^* + 2(B_{nn}^T P \Delta^*)^T \sigma(\Delta^* + x^*) \\ &\quad + 2(C_{nn}^T P \Delta^*) \gamma(u) + 2(P \Delta^*)^T \xi^*(t) \end{aligned}$$

Considering the following inequality

$$\begin{aligned} 2(C_{nn}^T P \Delta^*)^T \gamma(u) &= \|C_{nn}^T P \Delta^* + \gamma(u)\|^2 \\ &\quad - \Delta^{*T} P (C_{nn} C_{nn}^T) P \Delta^* - \|\gamma(u)\|^2 \end{aligned} \quad (30)$$

assuming that:

$$\sigma^T(x) Z \sigma(x) \leq x^T \Lambda_\sigma x + \lambda_{\sigma 0} \quad (31)$$

$$Z = z^T > 0, \Lambda_\sigma = \Lambda_\sigma^T > 0, \lambda_{\sigma 0} > 0,$$

$$\forall x \in \mathbb{R}^n, t \geq 0$$

and using the matrix inequality equation (17), we obtain:

$$\begin{aligned} \dot{V}(\Delta^*) &\leq \Delta^{*T} (P A_{nn} + A_{nn}^T P + 2\Lambda_\sigma + Q P (B_{nn} Z^{-1} B_{nn}^T \\ &\quad + \Lambda_\xi^{-1} - C_{nn} C_{nn}^T) P) \Delta^* \end{aligned} \quad (32)$$

$$\begin{aligned} &+ 2x^{*T} \Lambda_\sigma x^* + \xi^{*T}(t) \Lambda_\xi \xi^* - \|\gamma(u)\|^2 + \|C_{nn} P \Delta^* + \gamma(u)\|^2 \\ &\quad - \Delta^{*T} Q \Delta^* + \lambda_{\sigma 0} \end{aligned}$$

Related to this Lyapunov function derivative, we establish the following Lemma

Lemma 3.1. The time derivative of the Lyapunov function equation (28) satisfies the inequality:

$$\begin{aligned} \dot{V}(\Delta^*) &\leq 2x^{*T} \Lambda_\sigma x^* + \xi^{*T}(t) \Lambda_\xi \xi^*(t) - \|\gamma(u)\|^2 + \|C_{nn} P \Delta^* \\ &\quad + \gamma(u)\|^2 - \Delta^{*T} Q \Delta^* + u^T R u - u^T R u + \lambda_{\sigma 0} \end{aligned} \quad (33)$$

Proof. It directly follows if the matrix Riccati equation:

$$P A_{nn} + A_{nn}^T P + P (B_{nn} Z^{-1} B_{nn}^T + \Lambda_\xi^{-1} - C_{nn} C_{nn}^T) P + 2\Lambda_\sigma + Q = 0 \quad (34)$$

has a solution $P = P^T > 0$, and adding and subtracting $u^T R u$. ■

Defining the function $\Psi(u)$ as:

$$\Psi(u) = \xi^{*T}(t) \Lambda_\xi \xi^*(t) - \|\gamma(u)\|^2 + \|C_{nn} P \Delta^* + \gamma(u)\|^2 + u^T R u \quad (35)$$

then:

$$\dot{V}(\Delta^*) \leq 2x^{*T} \Lambda_\sigma x^* + \Psi(u) - \Delta^{*T} Q \Delta^* - u^T R u \quad (36)$$

To determine the control action $u(t)$, which minimizes the tracking error; we start by minimizing $\Psi(u)$, so $(d\Psi(u)/du) = 0$. We assume that for a fixed t (positive), $x(t)$ and $x_{\min}(t)$ are known and do not depend on $u(t)$.

Comment 3.1. We name $u^*(t)$, minimizing $\Psi(u)$, a local optimal control, because it is derived based only on 'local' information.

$d\Psi(u)/du$ is calculated as:

$$\begin{aligned} &2 \left(\frac{\partial \xi^*(t)}{\partial u} \right)^T \Lambda_\xi \xi^*(t) - 2 \left(\frac{d\gamma(u)}{du} \right)^T \gamma(u) \\ &\quad + 2 \left(\frac{d\gamma(u)}{du} \right)^T (C_{nn} P \Delta^* + \gamma(u)) + 2R u \end{aligned} \quad (37)$$

Related to $u^*(t)$, we state the following

Lemma 3.2. $u^*(t)$ is given by

$$u^*(t) = -R^{-1} \left(\frac{d\gamma(u)}{du} \right)^T C_{nn} P \Delta^* \quad (38)$$

Proof. It directly follows from $\xi^*(t) = A_{nn} x^* - \varphi(x^*, t)$. ■

Corollary 3.1. If $\gamma(u)$ depends linearly on u , then:

$$\frac{d\gamma(u)}{du} = \Gamma \quad (39)$$

$$u^*(t) = -R^{-1} \Gamma^T C_{nn}^T P \Delta^*$$

We establish, now, a theorem about the tracking error.

Theorem 3.1. For the nonlinear system equation (2), the neural network equation (1), with $\phi(x_{nn})=I$, the nonlinear reference model equation (23) and the control law equation (38), the following property holds:

$$\|\Delta^*\|_Q^2 + \|u^*(t)\|_R^2 \leq 2\|x^*\|_{\lambda_c}^2 + \limsup_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \Psi(u^*(t)) dt + \lambda_{\sigma 0} \quad (40)$$

Proof. After knowing the minimum $\Psi(u^*(t))$, equation (37) can be rewritten as:

$$\dot{V}(\Delta^*) \leq 2x^{*T} A_c x^* + \Psi(u^*(t)) - \Delta^{*T} Q \Delta^* - u^{*T} R u^*(t) + \lambda_{\sigma 0} \quad (41)$$

which can be reformulated as:

$$\Delta^{*T} Q \Delta^* + u^{*T} R u^*(t) \leq 2x^{*T} A_c x^* + \Psi(u^*(t)) - \dot{V}(\Delta^*) + \lambda_{\sigma 0} \quad (42)$$

Then, integrating each of these term, from 0 to τ , dividing them by τ , and calculating the limit, for $\tau \rightarrow \infty$, of the integrals suprema, we obtain equation (40). ■

The global structure for the dynamic neural identifier and controller is shown in Fig. 1.

The detailed demonstrations are included in Poznyak and Sanchez (1996b).

At this stage, we extend the previous analysis to consider nonlinear systems with uncertainties modelled as:

$$\dot{x} = f_0(x, u, t) + f_1(x, u, t), \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m \quad (43)$$

where $f_0(\dots)$ is the nominal system and $f_1(\dots)$ represents the uncertainties.

To develop the respective analysis, which is very similar to the one above explained, Assumption 2.1 has

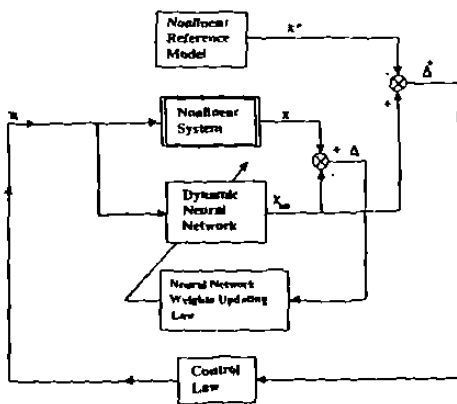


FIGURE 1. Identification and control scheme.

to be modified as follows.

Assumption 3.1. As in Poznyak and Sanchez (1995), there exist matrices $H_{n0}, H_{\Delta 0}$ positive definite such:

$$h_0^T(\Delta, x, u, t) H_{n0} h_0(\Delta, x, u, t) \leq \epsilon_{n0}(x, u, t) + \epsilon_{h1}(x, u, t) \Delta^T H_{\Delta 0} \Delta \quad (44)$$

for $\epsilon_{n0}(\dots), \epsilon_{h1}(\dots)$ bounded positive functions, with $\epsilon^{n0}, \epsilon^{h1}$, i.e:

$$\sup \epsilon_{ni}(x, u, t) = \epsilon^{ni} < \infty, \quad i=0,1 \quad (45)$$

$$\forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^m, \forall t \geq 0$$

Additionally, we assume the following:

Assumptions 3.2. There exist matrices $H_{n1}, H_{\Delta 1}$ positive defined such:

$$f_1^T(x, u, t) H_{n1} f_1(x, u, t) \leq \epsilon_{n1}(x, u, t) + \epsilon_{h1}(x, u, t) \Delta^T H_{\Delta 1} \Delta \quad (46)$$

for $\epsilon_{n1}(\dots), \epsilon_{h1}(\dots)$ bounded positive functions, with $\epsilon^{n1}, \epsilon^{h1}$, i.e:

$$\sup \epsilon_{ni}(x, u, t) = \epsilon^{ni} < \infty, \quad i=0,1 \quad (47)$$

$$\forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^m, \forall t \geq 0$$

From this point the analysis follows as the one above discussed. The detailed demonstrations are presented in Poznyak et al. (1997).

The analysis, when B_{nn} is on-line adapted, is again very similar. The most important difference is the matrix Ricatti equation (35), now becomes a differential matrix Ricatti equation as follows:

$$PA_{nn} + A_{nn}^T P + P(B_{nn} Z^{-1} B_{nn}^T + \Lambda_c^{-1} - C_{nn} C_{nn}^T) P + 2\Lambda_c + Q = -\dot{P} \quad (48)$$

For this case, we establish the following corollary.

Corollary 3.2. If, for neural network equation (1),

$\gamma(u)$ depends linearly on u , then $\frac{d\gamma(u)}{du} = \Gamma$, and the

signal $\xi^*(t)$ is compensated by the modified input $u(t) = u_{comp}(t) + u^*(t)$, where $u_{comp}(t)$ fulfills $C_{nn} u_{comp}(t) + \xi^*(t) = 0$ and $u^*(t) = -R_c^{-1} \Gamma^{-1} C_{nn}^T P_c(t) \Delta^*(t)$.

In Poznyak et al. (1996c), all the demonstrations are included.

Comment 3.2. The scheme shown in Fig. 1 is very innovative; it allows to track a reference signal generated by a nonlinear model, and is robust in presence of uncertainties.

4. INPUT-TO-STATE STABILITY (ISS) ANALYSIS

The dynamic neural network is described by equation (1), with A_{nn} Hurwitz, $\Gamma(0)=0, \gamma\sigma(\cdot), \phi(\cdot), \Gamma u$ globally Lipschitz:

$$\|\sigma(\Delta_u - x) - \sigma(x)\| \leq L_\sigma \|\Delta_u\| \tag{49}$$

$$\|\phi(\Delta_u - x) - \phi(x)\| \leq L_\phi \|\Delta_u\|$$

$$\|\Gamma(\Delta_u - x) - \Gamma(u)\| \leq L_\Gamma \|\Delta_u\|$$

$$\forall x, \Delta_u \in \mathbb{R}^n, \forall u, \Delta_u \in \mathbb{R}^m$$

Following the same procedure as in Matsouka (1992), it is trivial to demonstrate that equation (1) presents at least one equilibrium point (Perez & Sanchez, 1997).

Given x^* an equilibrium point of equation (1) with $u=0$, we define $y=x-x^*$ then:

$$\dot{y} = \dot{x} = A_{nn}x + B_{nn}\sigma(x) + C_{nn}\phi(x)\Gamma(u) \tag{50}$$

Adding and subtracting $B_{nn}\sigma(x^*) + C_{nn}\phi(x^*)\Gamma(u)$ to equation (50), and taking in account that $y=x-x^*$, we obtain:

$$\begin{aligned} \dot{y} = & A_{nn}y + B_{nn}[\sigma(y+x^*) - \sigma(x^*)] \\ & + C_{nn}[\phi(y+x^*) - \phi(x^*)]\Gamma(u) + C_{nn}\phi(x^*)[\Gamma(u) - \Gamma(0)] \end{aligned} \tag{51}$$

Once this model is defined, we proceed with the ISS analysis; we use the following definitions.

Definition 4.1 (Class K_∞ function) (Khalil, 1996) A continuous function $f: [0, a) \rightarrow [0, \infty)$ is said to belong to class K if it is strictly increasing and $f(0)=0$; and it is said to belong to class K_∞ if $a=\infty$ and $\lim_{r \rightarrow \infty} f(r) = \infty$.

Definition 4.2 (ISS Lyapunov function) (Sontag, 1995). A function $V(x): \mathbb{R}^n \rightarrow \mathbb{R}^+$ is defined as ISS Lyapunov if there exist $\alpha(r), \theta(r)$ class K_∞ functions such that:

$$\dot{V}(x) \leq -\alpha(\|x\|) + \theta(\|u\|), \forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^m$$

We establish the following theorem about ISS dynamic neural networks.

Theorem 4.1. The system equation (50), or equivalently the system equation (1), is ISS if:

$$\|B_{nn}\|^2 < \frac{\gamma - 3\|P\|}{L_\phi^2\|P\|}, \|P\| < \frac{\gamma}{3}, \gamma \in \mathbb{R}^+, P = P^T > 0. \tag{52}$$

where P is solution of the Lyapunov equation $A_{nn}^T P + PA_{nn} = -\gamma I$. Such a unique solution exists because A_{nn} is Hurwitz.

Proof. We consider the following function:

$$V(y) = y^T P y, P = P^T > 0 \tag{53}$$

Its time derivative along the trajectories of equation (51) is given as:

$$\begin{aligned} \dot{V} = & -\gamma y^T y + 2y^T P B_{nn}[\sigma(y+x^*) - \sigma(x^*)] \\ & + 2y^T P C_{nn}[\phi(y+x^*) - \phi(x^*)]\Gamma(u) \\ & + 2y^T P C_{nn}\phi(x^*)[\Gamma(u) - \Gamma(0)] \end{aligned} \tag{54}$$

We analyze the right part, the second, third and fourth term, of equation (54) using the matrix inequality equation (7), with $A=P$. Taking in account that $\phi(x)$ is globally Lipschitz, then it is uniformly bounded, i.e: $\|\phi(y+x^*) - \phi(x^*)\| \leq K_\phi, \forall y \in \mathbb{R}^n$, we finally obtain:

$$\begin{aligned} V \leq & -(\gamma - 3\|P\| - L_\phi^2\|P\|\|B_{nn}\|^2)\|y\|^2 \\ & + (K_\phi^2 L_\phi^2\|P\|\|C_{nn}\|^2 \\ & + L_\phi^2\|P\|\|C_{nn}\|^2\|\phi(x^*)\|^2)\|u\|^2 \end{aligned} \tag{55}$$

Defining the following functions:

$$\alpha(r) = (\gamma - 3\|P\| - L_\phi^2\|P\|\|B_{nn}\|^2)r^2 \tag{56}$$

$$\theta(r) = (K_\phi^2 L_\phi^2\|P\|\|C_{nn}\|^2 + L_\phi^2\|P\|\|C_{nn}\|^2\|\phi(x^*)\|^2)r^2$$

$$r \in \mathbb{R}$$

then:

$$\dot{V} \leq -\alpha(\|x\|) + \theta(\|u\|) \tag{57}$$

V will be an ISS-Lyapunov function, if $\alpha(\cdot)$ and $\theta(\cdot)$ are class K_∞ functions. As defined $\theta(\cdot)$ satisfies this condition; hence, for the system equation (50) (equivalently equation (1)) to be ISS, it is required:

$$(\gamma - 3\|P\| - L_\phi^2\|P\|\|B_{nn}\|^2) > 0 \tag{58}$$

which implies:

$$\|B_{nn}\|^2 < \frac{\gamma - 3\|P\|}{L_\phi^2\|P\|}, \|P\| < \frac{\gamma}{3} \tag{59}$$

Corollary 4.1. For the equilibrium point $y=0$ of equation (50) (equivalently $x=x^*$ of equation (1)) to be globally asymptotically stable (GAS), equation (59) has to be satisfied.

Proof. It directly follows from equation (59), with $u=0$.

The detailed demonstrations are discussed in Perez and Sanchez (1997), and Sanchez and Perez (1997).

Comment 4.1. The ISS result for dynamic neural networks is very important; it allows to utilize this type of neural networks for new schemes of nonlinear control [8,9].

5. EXAMPLES

To illustrate the applicability of our results, we include the following examples.

5.1. Identification and Trajectory Tracking

Given the nonlinear system:

$$\dot{x}_1 = -\alpha_1 x_1 + \beta_1 \text{sign}(x_2) + u_1 \tag{60}$$

$$\dot{x}_2 = -\alpha_2 x_2 + \beta_2 \text{sign}(x_1) + u_2$$

Let us select the neural network as follows:

$$\dot{x}_{nn,1} = -a_{nn,1} x_{nn,1} + b_{nn,11} \sigma_{nn,1}(x_{nn,1}) + b_{nn,12} \sigma_{nn,2}(x_{nn,2}) + c_{nn,11} u_1 + c_{nn,12} u_2 \tag{61}$$

$$\dot{x}_{nn,2} = -a_{nn,2} x_{nn,2} + b_{nn,21} \sigma_{nn,1}(x_{nn,1}) + b_{nn,22} \sigma_{nn,2}(x_{nn,2}) + c_{nn,21} u_1 + c_{nn,22} u_2$$

where $\sigma(\cdot)$ indicates a sigmoidal function.

The given nonlinear system equation (23), even simple, is interesting enough; it has multiple isolated equilibrium, and the right part of the differential equation is not smooth.

We consider:

$$a_{nn,1} = \alpha_1 = 5, x_1(0) = 10, x_{nn,1}(0) = -1, \beta_1 = 3$$

$$a_{nn,2} = \alpha_2 = 10, x_2(0) = -10, x_{nn,2}(0) = -5, \beta_2 = 2$$

$$c_{nn,11} = c_{nn,22} = 1, c_{nn,12} = c_{nn,21} = 0$$

The sigmoid function is chosen as:

$$\sigma(x) = \frac{2}{1 + e^{-2x}} - 0.5$$

We select $Q_0 = H_n = H_d = I, \alpha^1 = 3$. To adapt on-line the

dynamic neural network weights we use the same learning algorithm as in [11]. The control signals u_1, u_2 are chosen as sine wave and saw-tooth function; the results are shown as Fig. 2. The solid lines correspond to nonlinear system state responses $x_1(t), x_2(t)$, and the dashed line to neural network ones $x_{nn,1}(t), x_{nn,2}(t)$. The abscissa values correspond to the number of iterations. It can be seen that the neural network state time evolution follows the ones of nonlinear system.

The time evolution for the entries of the nonlinear feedback matrix B_{nn} is shown at Fig. 3.

Then, we implement the control law given by equation (40), with an on-line adaptive gain. Figure 4 presents the respective responses, where the solid line corresponds to reference model signals $x_1^*(t), x_2^*(t)$, and the dashed line to the nonlinear system responses $x_1(t), x_2(t)$.

The time evolution for the seminorms $\|\Delta\|_q^2 = \limsup_{\tau \rightarrow \infty} J_{\tau}^{\Delta}, J_{\tau}^{\Delta} = \frac{1}{\tau} \int_0^{\tau} \Delta^T(t) Q \Delta(t) dt$ and $\|u\|_R^2 = \limsup_{\tau \rightarrow \infty} J_{\tau}^u, J_{\tau}^u = \frac{1}{\tau} \int_0^{\tau} u^T(t) R u(t) dt$ can be seen in Fig. 5.

Comment 5.1. This example has an unstable equilibrium point. Unstable equilibria could exist in industrial applications such as chemical processes or electric power systems.

5.2. ISS Analysis

We present two cases: the first one fulfills equation (59), and the second one violates this condition.

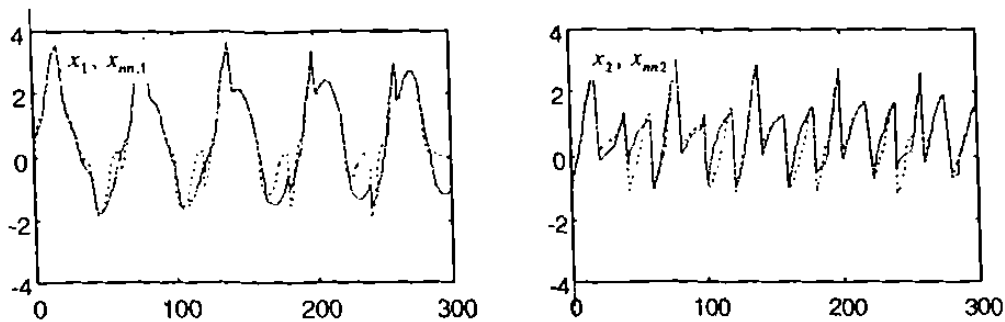


FIGURE 2. Identification results.

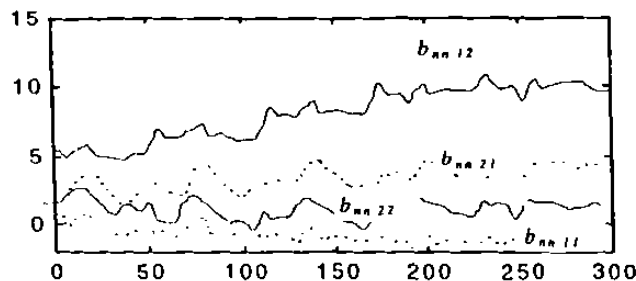


FIGURE 3. Time evolution of the nonlinear feedback matrix entries.

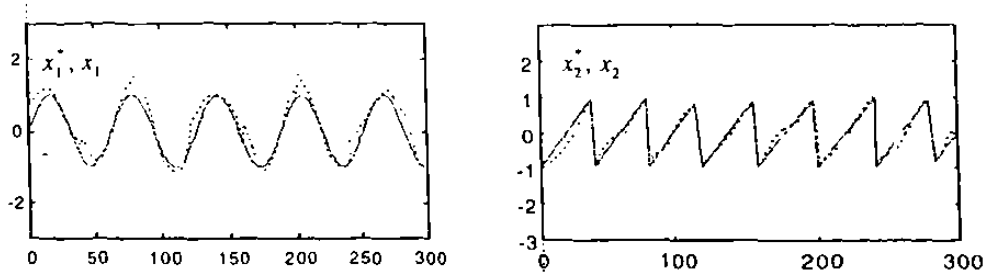


FIGURE 4. Response with feedback control.

Case 1.

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0.3 & 0.8 \\ 0.4 & 0.3 \end{pmatrix} \begin{pmatrix} \tanh(x_1) \\ \tanh(x_2) \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\gamma = 1, P = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \Rightarrow \|B_{nn}\| < 1, \lambda(B_{nn}) = \begin{pmatrix} 0.866 \\ 0.266 \end{pmatrix}, \sigma(B_{nn}) = \begin{pmatrix} 0.960 \\ 0.239 \end{pmatrix}$$

where $\lambda(\cdot)$ are the eigenvalues of the respective matrix, and $\sigma(\cdot)$ are the corresponding singular values.

This neural network fulfills the ISS conditions. Figure 6 shows the respective phase plane and the input to state time response of this neuron.

Case 2.

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} \tanh(x_1) \\ \tanh(x_2) \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

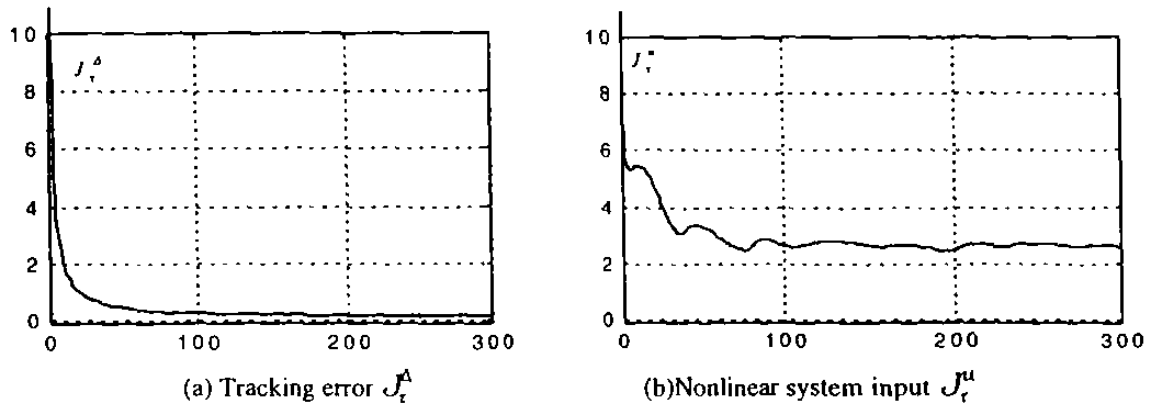


FIGURE 5. Time evolution of the semi-norms: (a) tracking error J_r^A ; (b) nonlinear system input J_r^u .

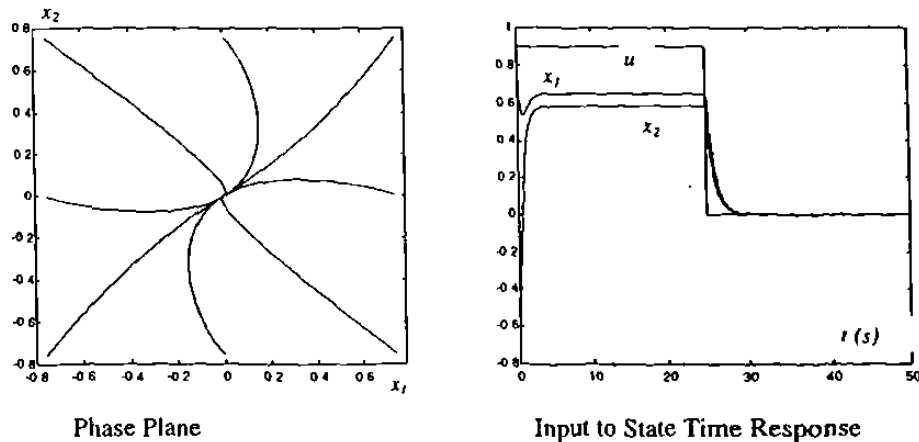


FIGURE 6. ISS neural network responses: (a) phase plane; input to state time response.

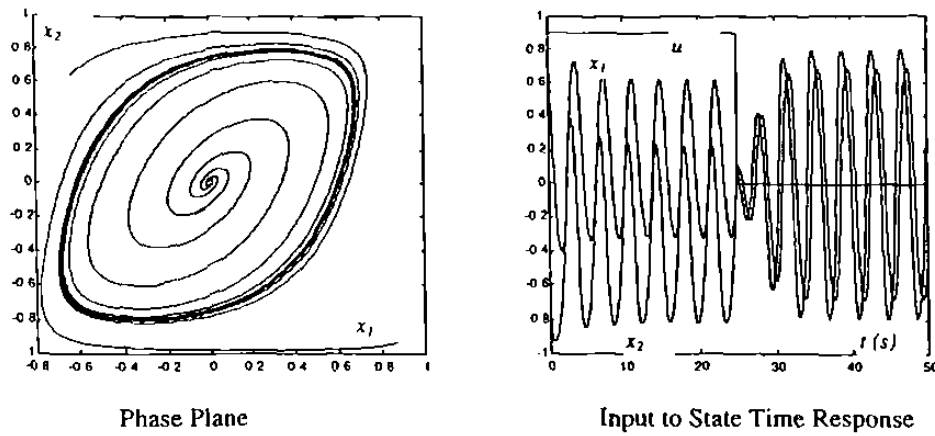


FIGURE 7. Oscillatory neural network: (a) phase plane; input to state time response.

$$\gamma=1, P=\begin{pmatrix} 1 & 0 \\ 4 & 1 \\ 0 & 4 \end{pmatrix} \Rightarrow \|B_m\| < \lambda(B_m) = \begin{pmatrix} 2.5+j1.941 \\ 2.5-j1.941 \end{pmatrix}, \sigma(B_m) = \begin{pmatrix} 5.12 \\ 1.95 \end{pmatrix}$$

This neural network does not satisfy the ISS conditions. Figure 7 shows the respective phase plane and the input to state time response.

6. CONCLUSIONS

As conclusion, we state the following.

We are able to implement nonlinear system identification by means of a dynamic neural network; a bound for the identification error is established. We do not require neither the system to be stable nor the time to be constrained to a closed interval. This is a big improvement over existing results, which ask for one of these two requirements to be satisfied. We only need the control actions to be bounded and the right part of the differential equation, describing the behavior of the nonlinear system, to satisfy a kind of sector condition.

Regarding trajectory tracking, we derived a control law to guarantee a bound for the respective error, under the assumptions that the trajectories generated by the reference model are bounded. To establish the identification error bound and the trajectory tracking error bound, we utilize a Lyapunov like analysis. In both cases, it is required that a matrix Riccati equation has a solution, which can be easily solved with existing software. Our approach includes time varying systems.

In relations to ISS, we establish conditions for the weights of the interconnection matrix of dynamic neural networks, in order to guarantee this property. For inputs identically equal to zero, these conditions also guarantee globally asymptotically stability. The ISS property allows the application of dynamic neural network to nonlinear control schemes such as internal model control (IMC).

Acknowledgements—The authors thank the support of CONACYT

project 0652A9506. They also thank the anonymous reviewers for their evaluation and comments which contributed to improve this article.

REFERENCES

- Agarwal, M. (1997) A systematic classification of neural network-based control. *IEEE Control Systems Magazine*, 17, 75–93
- Delgado, A., Kambhampati, C., & Warwick, K. (1995) Dynamic recurrent neural network for systems identification and control. In *IEE Proceedings Control Theory Applications*, vol. 142
- Jiang, Z. P., Teel, A., & Praly, L. (1994) Small gain theorem for ISS systems and applications. *Mathematics of Control, Signals, and Systems*, 7, 104–130
- Khalil, H. K. (1996) *Nonlinear Systems*, (2nd ed.), Prentice Hall
- Lewis, F. L., & Parisini, T. (1995) Neural network control with guaranteed stability, proposal for invited session. In *34th IEEE Conference on Decision and Control*
- Matsouka, K. (1992) Stability conditions for nonlinear continuous neural networks with asymmetric connections weights. *Neural Networks*, 5, 495–500
- Perez, J. P., and Sanchez, E. N. (1997) Dynamic neural networks input-to-state stability analysis, (in Spanish). In *2nd Conference of the Colombian Automation Association*, Bucaramanga, Colombia
- Poznyak, A. S., & Sanchez, E. N. (1995) Nonlinear system approximation by neural networks: error stability analysis. *Intelligent Automation and Soft Computing. An International Journal*, 1, 247–258
- Poznyak, A. S., & Sanchez, E. N. (1996a) Non linear system on-line identification using dynamic neural networks. *Intelligent Automation and Soft Computing. An International Journal*, submitted
- Poznyak, A. S., & Sanchez, E. N. (1996b) Nonlinear systems identification and tracking using dynamic neural networks. In *Proceedings of 35th IEEE Conference on Decision and Control*, Kobe, Japan.
- Poznyak, A. S., Sanchez, E. N., Perez, J. P., and Yu, W. (1996c) Nonlinear adaptive trajectory tracking using dynamic neural networks. *IEEE Transactions on Neural Networks*, submitted
- Poznyak, A. S., Sanchez, E. N., Perez, J. P., and Yu, W. (1997) Robust nonlinear system identification and trajectory tracking by dynamic neural networks. In *Proceedings of 1997 American Control Conference*, Albuquerque NM.
- Rovithakis, G. A., & Christodoulou, M. A. (1994) Adaptive control of unknown plants using dynamical neural networks. *IEEE Transactions on Systems, Man and Cybernetics*, 24, 400–412
- Rovithakis, G. A., & Christodoulou, M. A. (1995) Adaptive control of unknown plants using dynamical neural networks. *IEEE Transactions on Systems, Man and Cybernetics*, 25, 15728–15794

- Sanchez, E. N., and Perez, J. P. (1997). Input to-state stability (ISS) analysis for dynamic neural networks. In *Proceedings 1997 Intl Joint Conf on Neural Networks*, Houston TX.
- Sontag, E. D. (1989). Smooth stabilization implies coprime factorization. *IEEE Transactions on Automatic Control*, **34**, 435–443.
- Sontag, E. D. (1995). On the input to state stability property. *European Journal of Control*, **1**, 1–24.
- Tziniias, J. (1993). Sontag's input to state stability condition and global stabilization using state detection. *Systems and Control Letters*, **20**, 219–226.
- Varaiya, P. P., & Liu, R. (1966). Bounded-input bounded output stability of nonlinear time-varying differential systems. *SIAM Journal of Control*, **4**, 698–704.
- Willems, J. C. (1971). Least squares optimal control and algebraic Riccati equations. *IEEE Transactions on Automatic Control*, **16**, 621–634.

APÉNDICE B

PREMIO INTERNACIONAL EN INVESTIGACIÓN TEÓRICA

APENDICE B

En esta parte se incluye la costancia de un premio internacional en investigación teórica que se obtuvo con la presentación de este trabajo de investigación.

Sian Ka'an '97
**International
Conference**
Otorga el presente
DIPLOMA

a:

A. Poznyak, E. Sánchez, J. P. Pérez, W. Yu

por *Segundo Lugar en la categoría de trabajo Teórico*
con el trabajo

Dynamic Neural Control: Stability Analysis

en el Segundo Taller Internacional México-Estados Unidos en Redes Neuronales y
Neurocontrol celebrado en Playa del Carmen, Quintana Roo,
México, del 19 al 29 de agosto de 1997.

Bernard Widrow

Felipe Lara

George Vendaris

Dejan Sobajic

Comité Científico

Patrocinado por

National Science Foundation • Academia Mexicana de Ciencias • Sociedad Mexicana de
Instrumentación • Florida Institute of Technology • Consejo Nacional de Ciencias y
Tecnología • Universidad Autónoma de México • Centro de Instrumentos • Coordinación de la
Investigación Científica • Programa Universitario de Energía Científica • Programa
Universitario de Investigación y Desarrollo Espacial • Dirección General de
Posgrado • Facultad de Ingeniería • Universidad de Quintana Roo • Gobierno del Estado de
Quintana Roo • Municipio de Cancún • Municipio de Solidaridad • IBM de México S.A. de
C.V. • Sun Microsystems de México S.A. de C.V. • Apple Computer de México S.A. de C.V.

REFERENCIAS BIBLIOGRÁFICAS

1. Ramón y Cajal, S., 1991. *Histologie du Système Nerveux de l'homme et des Vertébrés*. Paris: maloine; Edition Francaise Revue: Tome I , 1952; Tome II, 1955; Madrid: Consejo Superior de Investigaciones Cientificas.
2. Shepherd, G. M., and C. Koch, 1990. ' Introduction to Synaptic Circuits'. In the *Synaptic Organization of the Brian* (G. M. Shepherd, ed.), pp.3-31. New york : Oxford University Press.
3. Faggin, F., 1991. 'VLSI Implementation of neural network', Tutorial Notes. International Joint Conference on Neural Network, Seattle, WA.
4. Freeman, W. J., 1975. *Mass Action in the Nervous System*. New York: academic Press.
5. A.N.Michel, J.A.Farrel, and W. Porod, "Qualitative analysis of neural networks", *IEEE Trans. on Circuits and Systems*, vol. 36, no. 2, pp 229-243, 1989.
6. H.K.Khalil, "Nonlinear Systems", Ed. Maxwell Macmillan, 1992.
7. M.W.Hirsch, "Convergent activation dynamics in continuous time networks", *Neural Networks*, vol. 2, pp. 331-349, 1989.
8. D.G. Kelly, "Stability in contractive nonlinear neural networks", *IEEE Trans. on Biomed. Eng.*, vol. 3, no. 3 pp. 241-242, 1990.
9. K.Matsouka, "Stability conditions for nonliner continuous neural networks with asymmetric connections weigthts", *Neural Networks*, vol. 5, pp. 495-500, 1992.
10. E. Kaszkurewics and A.Bhaya, "On a class of globally stable neural circuits", *IEEE Trans. on Circuits and Systems-I*, vol. 41, no. 2, pp 171-174, 1994.
11. M Forti et al, "Necessary and sufficient conditions for absolute stability of neural networks", *IEEE Trans. On Circuits and Systems-I*, vol. 41, no. 2, pp 171-174, 1994.
12. K.Hale, "Ordinary differential equations", Ed. Krieger, 1980.
13. A.S.Poznyak and E. N. Sánchez, "Nonlinear systems approximation by neural networks: error stability analysis", *Intell. Automat. and Soft Compt., an Intl. Journ.*, Vol. 1, No. 3, pp 247-258, 1995.

14. P. P. Varaiya and R. Liu, "Bounded -input Bounded-ouput stability of nonlinear time-varying differential systems", *SIAM Journ. Of control*, vol. 4, pp. 698-704, 1966.
15. E. D. Sontag, "Smooth stabilization implies coprime factorization", *IEEE Trans. On Automatic Control*, vol.34, pp 435-443, 1989.
16. J. Tzinias, "Sontag's input to state stability condition and global stabilization using state detection", *Systems and Control Letters*, vol. 20, pp 219-226, 1993.
17. Z.P. Jiang, A. Teel, and L. Praly, "Small-gain theorem for ISS Systems and applications", *Math. Of Control, Signals and Systems*, vol. 7, pp 104-130, 1994.
18. E. D. Sontag, "O the input state stability property," *Europen Journal of Control*, pp 1-24, 1995.
19. José P. Pérez Padrón y Edgar N. Sánchez C, "Simplified Analysis for GAS Neural Network", in Spanish, Segunda conferencia de Ing. Eléctrica, CINVESTAV-IPN, Mexico City, Mexico, September 1996.
20. G. A.Rovithakis, and M.A. Christodoulou, "Adaptative Control of Unknown plants using Dynamical Neural Networks", *IEEE Trans. On Systems, Man, and Cybernetics*, vol. 24 pp 400- 412, 1994.
21. E.B. Kosmatopoulous et al, "High order neural network structures for identification of dynamical systems", *IEEE Trans. on Neural Networks* vol. 6 pp 422-430, 1995.
22. R. G. Bartle, "The Elements of Real Analysis", 2nd Edition, John Wiley and Sons New York ,1976.
23. E.Sontag, "Neural nets as systems models and controllers", in *Proc. 7th Yale Workshop on Adaptive and Learning Systems*, pp 73-79, Yale University, 1992.
24. K.Funahashi, and Y.Nakamura, "Approximation of dynamical systems by continuous time recurrent neural networks", *Neural Networks*, Vol. 6, pp 801-806, 1993.
25. I.W. Sandberg, "Approximation theorems for discrete-time systems", *IEEE trans. on Circ. and Syst.*, Vol. 38, pp 564-566, 1991.
26. I.W.Sandberg, "Uniform Approximation and the Circle Criterion", *IEEE Trans. on Automatic Control*, Vol. 38, pp 1450-1458., 1992.
27. E.Alcorta and E.N.Sánchez, "Nonlinear Identification via Neural Networks", in *Proc.4th IFAC, Intl. Symp. on Adapt.Syst. in Cont. and Sign. Proc.*, June 1992.

28. E.N.Sánchez, "Dynamic Neural Network for Nonlinear Systems Identification", in Proc.33rd IEEE CDC, December 1994.
29. E.B.Kosmatopoulos et al, "High order neural network structures for identification of dynamical systems", IEEE Trans. on Neural Networks, Vol. 6, No 2, pp 422-431, March 1995.
30. F.L.Lewis and T.Parisini, "Neural network control with guaranteed stability", proposal for invited session, 34 IEEE CDC, February 1995.
31. A.Rovithakis, and M.A.Christodoulou, "Adaptive control of unknown plants using dynamical neural networks", IEEE Trans. on Syst., Man and Cybern., vol. 25, No 12, pp 1578-1594, December 1995.
32. M.Nikolaou, and V.Hanagandi, "Control of nonlinear dynamical systems modeled by recurrent neural networks", AIChE Journal, Vol. 39, No 11, pp 1890-1894, November 1993.
33. A.Delgado, C.Kambahmpati, and K.Warwick, "Dynamic recurrent neural network for systems identification and control", IEE Proc.-Cont. Theo. Appl., Vol. 142, No 4, July 1995.
34. A.Isidori, "Nonlinear control systems", 2nd edition, Ed. Springer Verlag, 1989.
35. A.Poznyak, E. N. Sánchez, and G. Acosta, "Nonlinear System Identification and Trajectory Tracking using Dynamic Neural Networks", Proc. IEEE CDC 1996, Kobe, Japón, December 1996.
36. K.Wang, and A.N.Mitchel, "Robustness and perturbation analysis of a class of nonlinear systems with applications to neural networks", IEEE Trans. on Circuits and systems, part 1, Vol.41, No1, pp 24-32 January 1994.
37. K.Wang, and A.N.Mitchel, "Robustness and perturbation analysis of a class of artificial neural networks", Neural Networks, Vol. 7, No 2, pp 251-257, 1994.
38. M.M.Polycarpou,"Stable adaptive neural control scheme for nonlinear systems", IEEE Trans. on Automatic Control, Vol. 41, No 3, pp 447-450, 1996.
39. J.C.Willems, "Least squares optimal control and algebraic Riccati equations", IEEE Trans. on Automatic Control, Vol. 16, No 6, pp 621-634, 1971.

