

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL
INSTITUTO POLITÉCNICO NACIONAL
DEPARTAMENTO DE CONTROL AUTOMÁTICO

Redes neuronales dinámicas con diferentes escalas de tiempo para identificación de sistemas no lineales.

TESIS QUE PRESENTA EL:

M. en C. Alejandro Cruz Sandoval¹

PARA OBTENER EL GRADO DE
Doctor en Ciencias en la Especialidad de Control Automático

DIRECTOR DE TESIS:
Dr. Wen Yu Liu

Becario de CONACYT¹

México, D.F. Agosto del 2006.

AGRADECIMIENTOS

A mi asesor:

Dr. Wen Yu Liu

Por su valiosa asesoría, paciencia y consejos otorgados para concluir con éxito esta tesis, que prácticamente es un proyecto de vida.

A los investigadores:

Dr. Ieroham Barouch
Dr. Moisés Bonilla Estrada
Dr. Marco A. Moreno Armendariz
Dr. Andrés Ferreira Ramírez

Por sus comentarios y observaciones para complementar este trabajo, así como para aceptar ser parte del jurado.

A la

Srita. Lucero M. Fernández Campos

Por su gran calidad humana y apoyo secretarial.

Al CONACYT

Por su apoyo económico, ya que de no ser por este no sería posible consolidar este proyecto.

Y finalmente, a mis amigos, compañeros(a) y profesores por su grata compañía, apoyo y ayuda brindada durante todo este proceso de formación.

DEDICATORIA

A ustedes, mis queridos padres por su comprensión y apoyo incondicional:

Emilio Cruz Martínez (q.e.p.d.)
Gloria Sandoval García

A ustedes hermanos:

Fernando, Emilio, Gerardo, Gloria,
Esther, Guadalupe, Lorena y Pablo.

Y a todos los profesores e investigadores del departamento de Control Automático que con su valiosa participación contribuyen a la formación de individuos útiles a la sociedad.

Índice general

| | |
|--|-----------|
| 1. Introducción | 1 |
| 1.1. Motivación | 3 |
| 1.2. Objetivos | 5 |
| 1.3. Estructura | 6 |
| 2. Preliminares | 9 |
| 2.1. Pasividad y estabilidad de entrada-estado (ISS) | 9 |
| 2.2. Perturbaciones singulares | 14 |
| 2.2.1. Sistemas lineales | 15 |
| 2.2.2. Sistemas no lineales | 17 |
| 3. Redes neuronales con diferentes escalas de tiempo y propiedades | 23 |
| 3.1. Estructura de redes neuronales con diferentes escalas de tiempo | 24 |
| 3.2. Propiedades de redes neuronales de escalas de dos tiempos | 31 |
| 3.2.1. Sin control ($u = 0$) | 31 |
| 3.2.2. Con control ($u \neq 0$) | 33 |
| 3.2.3. Perturbación singular en el intervalo infinito. | 39 |
| 3.2.4. Múltiples lentos y rápidos. | 41 |
| 4. Estabilidad de redes neuronales con diferentes escalas de tiempo | 45 |
| 4.1. Pasividad de redes neuronales de escalas de dos tiempos | 46 |
| 4.2. Estabilidad exponencial | 52 |

| | | |
|-----------|--|------------|
| 4.3. | Estabilidad robusta | 56 |
| 4.4. | Simulaciones | 63 |
| 4.4.1. | Pasividad de redes neuronales con diferentes escalas de tiempo | 63 |
| 4.4.2. | Estabilidad exponencial de redes neuronales con diferentes escalas de tiempo | 66 |
| 4.4.3. | Estabilidad robusta de redes neuronales con diferentes escalas de tiempo | 69 |
| 4.5. | Conclusión del capítulo | 70 |
| 5. | Identificación vía redes neuronales con diferentes escalas de tiempo | 73 |
| 5.1. | Redes neuronales de capa simple | 75 |
| 5.2. | Redes Neuronales multicapa | 83 |
| 5.3. | Simulación | 93 |
| 5.4. | Conclusión del capítulo | 95 |
| 6. | Aplicación a Robots de unión flexible | 99 |
| 6.1. | Introducción | 99 |
| 6.1.1. | Componentes y estructura de robots | 100 |
| 6.2. | Modelo de un robot de unión flexible | 106 |
| 6.2.1. | Modelado dinámico de robots | 106 |
| 6.2.2. | Modelo de un robot de unión flexible | 109 |
| 6.2.3. | Dinámicas de robots de eslabón flexible | 115 |
| 6.3. | Modelado e identificación vía redes neuronales con diferentes escalas de tiempo | 120 |
| 6.4. | Simulación | 121 |
| 6.5. | Conclusión del capítulo | 125 |
| 7. | Conclusiones y trabajos futuros | 127 |
| 8. | Apéndice: artículos publicados | 141 |

Índice de figuras

| | |
|--|----|
| 1.1. Sistema de identificación y control para el brazo de robot de eslabón flexible. | 2 |
| 2.1. La consistencia graficada requerida. | 16 |
| 3.1. Red Neuronal Diferencial Recurrente. | 25 |
| 3.2. Red Neuronal Diferencial Recurrente sin entrada. | 25 |
| 3.3. Red Neuronal Diferencial Recurrente con capas de neuronas ocultas. | 27 |
| 3.4. Red Neuronal Recurrente Dinámica. | 27 |
| 3.5. Motor de C.D. | 28 |
| 4.1. Respuesta a las entradas acotadas. | 64 |
| 4.2. Respuestas a entrada cero. | 65 |
| 4.3. Respuesta a la salida retroalimentada. | 65 |
| 4.4. Para $T = 0,1$ es inestable. | 66 |
| 4.5. Estabilidad de la red neuronal diferencial (Ejemplo 1) | 67 |
| 4.6. Circuito RC con fuente de voltaje dependiente. | 69 |
| 4.7. Estabilidad de la red neuronal diferencial (Ejemplo 2) | 70 |
| 4.8. Estabilidad de la red neuronal diferencial (Ejemplo 3) | 71 |
| 5.1. Red Neuronal Dinámica con dos escalas de tiempo. | 76 |
| 5.2. Red Neuronal Dinámica con dos escalas de tiempo. | 84 |
| 5.3. DC motor | 94 |

| | |
|--|-----|
| 5.4. Identificación de ω_r | 95 |
| 5.5. Identificación de \dot{i}_r | 96 |
| 5.6. Identificación de ω_r | 96 |
| 5.7. Identificación de \dot{i}_r | 97 |
| 6.1. Un robot de unión flexible | 122 |
| 6.2. Identificación de x_1 | 123 |
| 6.3. Identificación de x_2 | 123 |
| 6.4. Identificación de x_1 | 124 |
| 6.5. Identificación de x_2 | 124 |

Capítulo 1

Introducción

Una diversa clase de sistemas físicos en ciencias e ingeniería contienen fenómenos dinámicos lentos y rápidos que ocurren en escalas de tiempo diferentes (Fig.1.1). Recientes resultados muestran que las técnicas de redes neuronales dinámicas al parecer pueden ser muy efectivas para modelar, identificar y controlar una muy variada clase de sistemas no lineales complejos, cuando no se tiene información completa del modelo físico, o cuando se considera un sistema complejo como una “caja negra” [33]. Esta aproximación de modelo libre puede utilizar las agradables características de redes neuronales artificiales con diferentes escalas de tiempo, pero la carencia de un modelo para el sistema a modelar y controlar provoca dificultades en la obtención de resultados teóricos sobre la estabilidad del mismo. Algunas de las aplicaciones de redes neuronales artificiales requieren que los puntos de equilibrio de la red neuronal diseñada sean estables [37]. De tal manera, que resulta muy importante estudiar la estabilidad de redes neuronales artificiales dinámicas. Existe una clase de redes neuronales diferenciales [66] con escalas de tiempo diferentes que pueden modelar las dinámicas de ambos niveles de actividad neuronal; estos pueden ser los términos de memoria corta (STM) y los de memoria larga (LTM). Su capacidad de almacenar patrones deseados como puntos de equilibrio requiere criterios de estabilidad que incluyen la mutua interacción entre modificaciones sinápticas y dinámicas de aprendizaje. Las dinámicas de redes neuronales artificiales diferenciales con diferentes escalas de tiempo pueden ser extremadamente complejas, exhi-

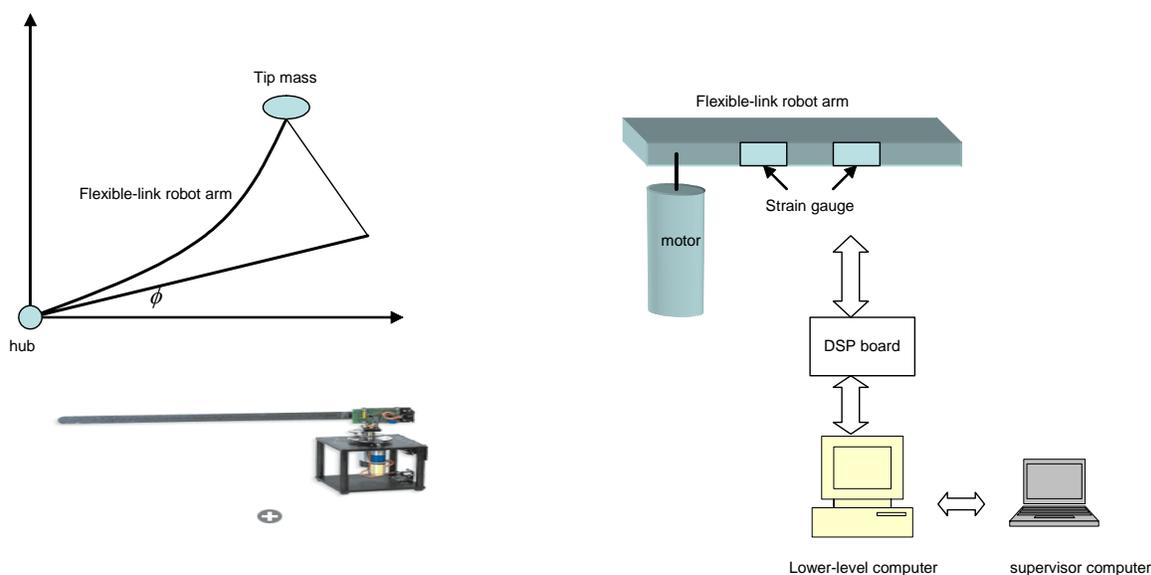


Figura 1.1: Sistema de identificación y control para el brazo de robot de eslabón flexible.

biendo convergencia para puntos atractores y atractores periódicos [1]. Las redes neuronales artificiales donde ambos términos de memoria corta y larga, son variables dinámicas que no pueden ser colocadas en esta forma debido a las ecuaciones que las modelan, es decir, las ecuaciones de Cohen-Grossberg no pueden modelar dinámicas sinápticas [22]. Sin embargo, una diversas clase de sistemas neuronales competitivos han sido considerados.

No existen muchos resultados sobre el análisis de estabilidad de redes neuronales a pesar de sus exitosas aplicaciones. Por ejemplo, la estabilidad asintótica global (GAS) de redes neuronales ha sido desarrollada durante los últimos años, ya que una función semidefinida negativa de la interconexión matricial puede hacer que el circuito neuronal de Hopfield-Tank sea (GAS) [15]; la estabilidad de neuro-circuitos fue establecida por el concepto de estabilidad diagonal [38]. Dentro del marco de trabajo de los sistemas de Lur, la estabilidad de perceptrones multicapa (MPL) y las redes neuronales recurrentes fueron propuestas en [87] y [53]. Por otra parte, el método de análisis de Estabilidad de Entrada Estado (ISS) [40] es una efectiva herramienta para redes neuronales dinámicas, y en [75] esta planteado que si los

pesos son suficientemente pequeños, las redes neuronales dinámicas son (ISS) y (GAS) con entrada cero.

La estabilidad de los errores de identificación y seguimiento están siendo investigados. Los autores [35] y [46] estudiaron las condiciones de estabilidad cuando los perceptrones son usados para identificar y controlar sistemas no lineales.

El análisis por medio del enfoque de Lyapunov es una herramienta popular para probar la estabilidad. [72] y [102] discutieron la estabilidad de redes neuronales dinámicas de una capa. Para el caso de redes neuronales dinámicas de alto orden y redes multicapas los resultados de estabilidad pueden ser encontrados en [61] y [66].

La teoría de pasividad [80] es otra efectiva herramienta para analizar la estabilidad de sistemas no lineales, ya que es posible tratar con sistemas no lineales usando solo las características generales de las dinámicas de entrada-salida, y ofrecer soluciones elegantes para la prueba de estabilidad absoluta. El marco de trabajo de pasividad es un enfoque promisorio para el análisis de estabilidad de redes neuronales dinámicas, porque puede producir conclusiones sobre la estabilidad usando solo características de entrada-salida del sistema. Las propiedades de perceptrones multicapa fueron ya examinados por [46] y [10]. Por medio del análisis de interconexión de modelos de error, ellos derivaron la relación entre pasividad y estabilidad de lazo cerrado. Pero la ventaja del conocimiento adquirido hasta ahora en esta área, es el hecho de que el análisis en lazo cerrado basado en los métodos de pasividad para redes neuronales dinámicas con diferentes escalas de tiempo todavía no han sido establecidos en la literatura especializada.

1.1. Motivación

En esta tesis se pretende utilizar las técnicas de perturbaciones singulares y escalas de tiempo implementadas con redes neuronales artificiales.

El control adaptivo de sistemas no lineales ha sido una área bastante activa en años recientes. Es una dificultad el control de plantas desconocidas. Un enfoque común para tratar este problema es de utilizar técnicas de identificación simultánea. Las redes neuronales arti-

ficiales han sido empleadas en la identificación y control de sistemas no lineales desconocidos debido a su masivo paralelismo, rápida adaptación y capacidad de aprendizaje. Una red neuronal basada naturalmente en control lleva a problemas en control no lineal y control adaptivo no lineal. La pasada década ha presenciado una gran actividad en este campo, con un aumento de conciencia por parte de los investigadores en que los problemas pueden ser diereccionados en el marco de trabajo de la teoría de control matemático.

Aunque muchos resultados y descubrimientos importantes han sido hechos en control neuronal, un número de preguntas abiertas permanecen sin respuesta.

1. Varios problemas físicos que están caracterizados por la presencia de un pequeña perturbación que está activa en un largo período de tiempo, tendrá un efecto acumulativo nada insignificante. Métodos de perturbación pueden ser usados para obtener una solución aproximada en la forma de una expansión en un parámetro pequeño. Los métodos especiales son llamados métodos de perturbación. Estos pueden usar escalas múltiples de tiempo en problemas de valor inicial y extensión de coordenadas en regiones de cambio brusco en problemas con valores en la frontera. Un requerimiento básico de los métodos de perturbación es que el sistema no lineal debe ser completamente conocido. Si usamos las propiedades de aproximación universal de redes neuronales artificiales, será posible aplicar los métodos de perturbación para sistemas no lineales desconocidos ?.

2. Muchos sistemas de larga escala o escala grande, tales como sistemas de potencia, pueden ser descompuestos en áreas coherentes lentas y subsistemas rápidos en la formación tipo *Lure* por la propiedad de escala de tiempo. Algunos sistemas mecánicos pueden también ser dividido en subsistemas rápidos y lentos, por ejemplo la dinámica de un robot de eslabón flexible puede ser dividido en dos partes en escalas de tiempo separadas. Debido a que el control neuronal normal utiliza únicamente escalas de tiempo, algunos artículos usan dos redes neuronales para el control de subsistemas rápidos y lentos independientemente. Aquí los métodos de perturbación son usados para separar la planta. Los controladores son únicamente normales en escalas de tiempo. Será posible diseñar redes neuronales multiescala de tiempo para controlar plantas multiescala de tiempo directamente ?

3. Las redes neuronales operan en dos modos: *cómputo y apendizaje*. Por ejemplo, las

redes neuronales recurrentes tipo Hopfield, en donde la operación de cómputo definida por la ecuación del sistema, es un evento sináptico rápido asociado con una constante de tiempo pequeña (RC), mientras el aprendizaje (cambio de peso sináptico) puede ser tomado en cuenta como un proceso lento con una constante de tiempo grande $\frac{1}{\lambda}$, λ es el rango de aprendizaje. Casi todo el análisis de control neuronal considera estos dos modos de operación en la misma escala de tiempo. Si se utilizan escalas multi- tiempo, será posible diseñar ambas partes de cómputo y aprendizaje de redes neuronales para asegurar la convergencia de aprendizaje y estabilidad de lazo cerrado ?

4. Muchos sistemas prácticos involucran sensores que proporcionan señales en rangos de muestreo lento. El controlador y el sensor de salida tienen diferentes escalas de tiempo. Algunos sistemas de control tienen diferentes períodos de control, por ejemplo, el sistema visual-servoing de unión servoing es más rápido que el control basado en imagen. Sin embargo, lo mejor de cierto conocimiento ya adquirido, es que las técnicas de control neuronal multirango no han sido todavía establecidas en la literatura especializada. Entonces será posible usar redes neuronales para aplicar control multirango?

1.2. Objetivos

En esta tesis se pretende usar redes neuronales con múltiple escalas de tiempo para tratar de resolver los cuatro problemas planteados anteriormente. Estos nuevos tipos de modelos toman beneficios de redes neuronales artificiales y escalas múltiple de tiempo. Sus capacidades de aprendizaje son multirango, así que se puede hacer identificación y control adaptivo para sistemas de escalas de tiempo múltiple. Y debido a que estas tienen formas universales para clases largas de sistemas de escalas en el tiempo múltiple, es posible desarrollar una prueba universal aproximada para superar la difícil validación en sistemas completos de escala grande. Para mostrar la efectividad de redes neuronales adaptivas con múltiples escalas de tiempo, se implementarán dos sistemas no lineales típicos multi-escalas de tiempo, el brazo de robot con eslabón flexible y un sistema electromecánico, para verificar los resultados teóricos.

Básicamente son dos los objetivos fundamentales que se contemplan en esta tesis, y que son los siguientes:

1. Analizar y separar en diferentes escalas de tiempo sistemas no lineales, utilizando la técnica de perturbaciones singulares.
2. Modelar e identificar sistemas sometidos a perturbaciones singulares con redes neuronales artificiales con diferentes escalas de tiempo.

1.3. Estructura

Esta tesis esta estructurada de la siguiente manera:

En la introducción se indica lo que básicamente se abordará durante la tesis, mencionando algunas ideas y logros alcanzados en tareas de identificación y control de redes neuronales, así como posibles técnicas para mejorar lo hecho actualmente, en campos de control con redes con diferentes escalas de tiempo.

En el capítulo II se introduce el concepto de pasividad, como uno de los conceptos básicos y fundamentales para el análisis de estabilidad de sistemas no lineales. En este mismo capítulo, también se aborda el concepto de estabilidad de *entrada-estado* con la finalidad de analizar la estabilidad de sistemas no lineales. Estos tópicos toman solo en consideración únicamente las señales de *entrada -estado-salida* de un sistema o una planta.

También se abordan los conceptos básicos de la teoría de perturbaciones singulares, como una de las técnicas para el análisis y modelado de sistemas dinámicos con diferentes escalas de tiempo, tanto para sistemas lineales y sistemas no-lineales.

En el capítulo III se plantean tópicos relativamente nuevos de redes neuronales dinámicas, estructuradas con diferentes escalas de tiempo. Básicamente se parte de los conceptos de perturbaciones singulares del libro de Hassan K. Khalil, pero dándole el enfoque con redes neuronales dinámicas. Se intenta unificar tanto los criterios de redes neuronales artificiales con los de perturbaciones singulares.

En el capítulo IV y con base en los conceptos del capítulo II, se deducen criterios de estabilidad para redes neuronales con escalas de dos tiempos, validando los resultados obtenidos

con una serie de simulaciones.

En el capítulo V se inicia una de los objetivos fundamentales de esta tesis, que es el de identificación y modelado de sistemas no lineales, con redes neuronales con diferentes escalas de tiempo, realizando también una serie de simulaciones para validar su efectividad.

El objetivo y logro fundamental de este trabajo, es el de utilizar redes neuronales dinámicas con diferentes escalas de tiempo, para modelar e identificar sistemas de complejidad diversa.

Capítulo 2

Preliminares

2.1. Pasividad y estabilidad de entrada-estado (ISS)

La pasividad es una de las importantes propiedades de sistemas dinámicos que proporciona una relación especial entre la entrada y salida de un sistema y es comúnmente usado en el análisis de estabilidad y estabilización de una diversa clase de sistemas no lineales. Hablando brúscamente, si un sistema no lineal es pasivo puede ser estabilizado por cualquier retroalimentación negativa lineal constante ante la carencia de una descripción a detalle de su modelo matemático. Esta propiedad al parecer es muy atractiva en diferentes aplicaciones físicas. En vista de esto, el siguiente enfoque para diseñar un controlador retroalimentado para sistemas no lineales es diversamente usado: primero, una retroalimentación no lineal interna especial es introducida para pasivar el sistema no lineal dado; segundo, una simple retroalimentación lineal negativa externa es introducida para proporcionar una propiedad de estabilidad para el sistema de lazo cerrado obtenido.

En general, la teoría de pasividad hace frente a sistemas controlados cuyas propiedades no lineales están mal definidas (generalmente por medio de cotas sectoriales). Sin embargo, ofrece una elegante solución para el problema de estabilidad absoluta de tales sistemas. El marco de trabajo de pasividad puede conducir a conclusiones generales sobre la estabilidad de diversa clase de sistemas de control no lineal, usando solo algunas características generales

de las dinámicas de *entrada-salida* del sistema controlado y el mapeo *entrada-salida* del controlador. Por ejemplo, si el sistema es pasivo y su estado cero detectable, cualquier salida retroalimentada estabiliza el equilibrio del sistema no lineal.

Cuando los sistemas dinámicos son totalmente o parcialmente desconocidos, la pasividad retroalimentada equivalente produce un importante problema. Esta propiedad puede ser probada por un diseño especial de controladores pasivados robustos. Pero todos ellos requieren más conocimiento detallado sobre las dinámicas del sistema. De tal manera, para ser realizado exitosamente, un control pasivado adaptivo necesita la estructura del sistema bajo consideración además los parámetros desconocidos deben ser lineales. Si se trata con el control pasivado no adaptivo, la parte nominal (sin perturbaciones externas) del sistema se asume ser completamente conocida.

Si el sistema es considerado como una caja negra (solo algunas propiedades generales son asumidas para ser verificadas para garantizar la existencia de la solución de los correspondientes modelos de ecuaciones diferenciales ordinarias), el modelado y control basado en el aprendizaje usando redes neuronales artificiales ha emergido como una herramienta viable. Este enfoque de modelo libre es presentado como una agradable característica de redes neuronales artificiales, pero la falta del modelo para la planta controlada hace difícil obtener resultados teóricos sobre la estabilidad y realización de un sistema no lineal cerrado por un diseño de un sistema neuronal. En la práctica ingenieril, es muy importante tener cualquier garantía teórica de que el neurocontrolador pueda estabilizar un sistema dado antes de su aplicación para una planta real industrial o mecánica. Esto es porque el neurocontrolador diseñado puede ser considerado como un cambio para la comunidad de control moderno. La mayoría de publicaciones en la identificación y control de sistemas no lineales usa redes neuronales estáticas, por ejemplo, perceptrones multicapa, que son implementados por la aproximación de una función no lineal en el lado derecho de las ecuaciones del modelo dinámico. La principal desventaja de estas redes neuronales es que la actualización de los pesos no usa cualquier información sobre una estructura de datos local y la aproximación de la función aplicada es sensible a los datos de entrenamiento. Las redes neuronales dinámicas recurrentes pueden exitosamente mejorar esta desventaja además de proporcionar una ade-

cuado comportamiento en la presencia de dinámicas no modeladas, por sus estructuras de retroalimentación incorporadas. Estas tienen poderosas capacidades de representación. Una de las mejores redes neuronales dinámicas fué introducida por John J. Hopfield.

Por esta razón el marco de trabajo de redes neuronales artificiales es muy conveniente para la pasivación de sistemas no lineales desconocidos. Como antes se planteó, existen muchas desventajas en usar redes neuronales estáticas para el control de sistemas dinámicos.

En este capítulo se usarán redes neuronales dinámicas recurrentes para pasivar un sistema no lineal desconocido. Una función de almacenamiento especial está definida de tal manera que el objetivo de pasivación e identificación puede ser alcanzada simultáneamente.

El objetivo de esta sección es para introducir el concepto de pasividad y presentar algunos resultados de estabilidad que pueden ser obtenidos usando las herramientas de redes neuronales recurrentes dinámicas y perturbaciones singulares. Básicamente se enfocará en la definición clásica de *entrada-salida* de un sistema.

El principal interés de esta sección es entender algunos conceptos de pasividad y estabilidad de entrada estado (ISS). Considérese una clase de sistemas no lineales indicados por:

$$\begin{aligned}\dot{x}_t &= f(x_t, u_t) \\ y_t &= h(x_t, u_t)\end{aligned}\tag{2.1}$$

donde $x_t \in \mathfrak{R}^n$ es el estado, $u_t \in \mathfrak{R}^m$ es el vector de entrada, $y_t \in \mathfrak{R}^m$ es el vector de salida. $f : \mathfrak{R}^n \times \mathfrak{R}^m \rightarrow \mathfrak{R}^n$ es localmente Lipschitz, $h : \mathfrak{R}^n \times \mathfrak{R}^m \rightarrow \mathfrak{R}^m$ es continua. También se asume que para cualquier $x^0 = x_0 \in \mathfrak{R}^n$, la salida $y_t = h(\Phi(t, x^0, u))$ del sistema (2.1) es tal que $\int_0^t |u_s^T y_s| ds < \infty$, para todo $t \geq 0$, *i.e.*, la *energía almacenada en el sistema* (2.1) está acotada.

Recordando algunas propiedades de pasividad y estabilidad de sistemas pasivos (ver [6] y [29]).

Definición 2.1 *Un sistema (2.1) se dice que es pasivo de la entrada u_t a la salida y_t , si existe una C^r función no negativa $S(x_t) : \mathfrak{R}^n \rightarrow \mathfrak{R}$, llamada función de almacenamiento, tal*

que, para todo u_t , todas las condiciones iniciales x^0 y toda $t \geq 0$ la siguiente desigualdad se cumple:

$$\dot{S}(x_t) \leq u_t^T y_t - \varepsilon u_t^T u_t - \delta y_t^T y_t - \rho \psi(x_t), \quad (x_t, u_t) \in \mathfrak{R}^n \times \mathfrak{R}^m.$$

donde ε , δ , ρ son constantes no negativas, $\psi(x_t)$ es una función semidefinida positiva de x_t tal que $\psi(0) = 0$. La función $\rho\psi(x_t)$ es llamada razón de disipación del estado. Lo que es más, el sistema se dice que es:

- de *menor pérdida* si $\varepsilon = \delta = \rho = 0$ y $\dot{S}(x_t) = u_t^T y_t$; es decir, no existe disipación de energía en el sistema.
- estrictamente pasivo a la entrada si $\varepsilon > 0$
- estrictamente pasivo a la salida si $\delta > 0$
- estrictamente pasivo de estado si $\rho > 0$
- estrictamente pasivo si existe una función definida positiva $V(x_t) : \mathfrak{R}^n \rightarrow \mathfrak{R}$ tal que $\dot{S}(x_t) \leq u_t^T y_t - V(x_t)$

Propiedad 1. Si la función de almacenamiento $S(x_t)$ es diferenciable y el sistema dinámico es pasivo, la función de almacenamiento $S(x_t)$ satisface:

$$\dot{S}(x_t) \leq u_t^T y_t$$

Así como existen dos diferentes nociones de estabilidad: 1) *la estabilidad en el sentido de Lyapunov* 2) *la estabilidad de entrada-salida*. Estos dos enfoques son opuestos en toda la gama de conceptos. Por una parte, la estabilidad de Lyapunov se aplica para puntos de equilibrio de realizaciones de espacio de estado. Y por otra parte, la estabilidad entrada-salida trata con sistemas que son mapeados entre entradas y salidas, ignorando la descripción del estado interno, que puede ser o no ser dada por una realización de espacio de estado. Por lo tanto, en esta parte, se trata de cerrar este hueco entre estas dos nociones y se introduce el

concepto de estabilidad de entrada estado (ISS). Se asume que los sistemas son descritos por una realización de espacio de estado que incluye una función variable a la entrada.

Definición 2.2 *Un sistema (2.1) se dice que es globalmente estable de entrada estado (ISS) si existe una \mathcal{K} -función $\gamma(s)$ (continua y estrictamente creciente $\gamma(0) = 0$) y una \mathcal{KL} -función $\beta(s, t)$ (\mathcal{K} -función y para cada fijo $s_0 \geq 0$, $\lim_{t \rightarrow \infty} \beta(s_0, t) = 0$), tal que, para cada entrada $u_t \in L_\infty$ ($\sup \{\|u_t\|, t \geq 0\} < \infty$) y cada estado inicial $x^0 \in \mathbb{R}^n$, se cumple que:*

$$\|x(t, x^0, u_t)\| \leq \beta[\|x^0\|, t] + \gamma[\|u_t\|].$$

para cada $t \geq 0$ [85].

Propiedad 2. Si un sistema es estable de entrada estado (ISS), el comportamiento del sistema permanecerá acotado cuando sus entradas están acotadas.

Definición 2.3 *Un sistema no lineal (2.1) se dice que es pasivo si existe una C^r función nonegativa $S(x) : \mathbb{R}^n \rightarrow \mathbb{R}$, llamada función de almacenamiento, tal que, para todo u , todas las condiciones iniciales x^0, z^0 y toda $t \geq 0$ se cumplen las siguientes desigualdades:*

$$\begin{aligned} \dot{S}(x, z) &\leq u^T y - \epsilon u^T - \delta y^T y - \rho \psi(x) \\ (x, z, u) &\in \mathbb{R}^n \times \mathbb{R}^m \end{aligned}$$

donde: ϵ, δ y ρ son constantes no negativas, $\psi(x)$ es una función semidefinida positiva de x . La función $\rho\psi(x)$ es llamada la razón de disipación de estado. Un sistema se dice ser estrictamente pasivo si existe una función definida positiva

$V(x, z) : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que:

$$\dot{S}(x, z) \leq u^T y - V(x, z)$$

Propiedad. Si la función de almacenamiento $S(x, z)$ es diferenciable y el sistema dinámico es pasivo, la función de almacenamiento $S(x, z)$ satisface $\dot{S}(x, z) \leq u^T y$.

Comentario 2.1 *La definición anterior prácticamente es la misma que la mencionada inicialmente, solo que ahora se hace uso de las dos variables de estado que generalmente intervienen en sistemas perturbados singularmente, con la finalidad de ser separados en dos subsistemas dinámicos con diferentes escalas de tiempo.*

2.2. Perturbaciones singulares

Un problema fundamental en la teoría de sistemas y control, es el modelado matemático de un sistema físico. La representación real de los sistemas se realiza mediante modelos matemáticos descritos por ecuaciones diferenciales o en diferencias. La presencia de algunos parámetros parásitos, tales como pequeñas constantes de tiempo, masas, momentos de inercia, resistencias, inductancias y capacitancias pequeñas, es a menudo la causa del incremento del orden y rigidez de estos sistemas. La rigidez, atribuida a la ocurrencia simultánea de fenómenos o dinámicas lentas y rápidas, da lugar a escalas de tiempo. Los sistemas en que la supresión de un pequeño parámetro es la degeneración o degradación de la dimensión del sistema son etiquetados como sistemas *perturbados singularmente*, que son una representación especial de la clase general de sistemas en escalas de tiempo. La contrariedad de acoplamientos dimensionales con rigidez posee formidables complejidades computacionales para el análisis y control de estos sistemas. Los métodos de *perturbaciones singulares y escalas de tiempo*, dotados con las dos características correctivas de reducción dimensional y alivio de rigidez, son considerados como un *boom* por los ingenieros de control. Estas técnicas han alcanzado un cierto nivel de madurez en la teoría de sistemas de control continuo descritos por ecuaciones diferenciales. La metodología tiene un impresionante record de aplicaciones en una diversa gama de campos tales como circuitos eléctricos, redes, sistemas de potencia, electromagnetismo, mecánica de fluidos, mecánica estructural, mecánica celeste, geofísica, dinámica de reactores nucleares, oceanografía, biología y ecología. Los sistemas discretos son muy imperantes en ciencias e ingeniería. Existen principalmente tres fuentes de modelos discretos definidos por ecuaciones en diferencias conteniendo diversos parámetros pequeños. La primer fuente es la simulación digital de sistemas continuos donde las ecuaciones diferenciales

son aproximadas por las correspondientes ecuaciones en diferencias. El estudio de sistemas con datos muestreados y sistemas de control computacional conduce de manera natural a otra fuente de modelos discretos. Finalmente, muchos sistemas económicos, biológicos y sociales, son representados por modelos discretos. A pesar del hecho de que el control de un sistema físico con una computadora digital es de vital importancia, el campo de *perturbaciones singulares y escalas de tiempo* en sistemas de control en tiempo discreto es de reciente origen. Recientemente existe considerable interés en este campo incluyendo control óptimo y adaptivo. El rápido crecimiento de actividad en la investigación en *perturbaciones singulares y escalas de tiempo* se ha visto enriquecido por la publicación de excelentes artículos, libros, monografías, reportes y conferencias especiales.

2.2.1. Sistemas lineales

Considérese un sistema lineal perturbado singularmente, de la forma:

$$\begin{aligned}\dot{x} &= A_{11}x + A_{12}z \\ \epsilon\dot{z} &= A_{21}x + A_{22}z\end{aligned}\tag{2.2}$$

Posicionando $\epsilon = 0$ se obtiene el sistema lineal perturbado singularmente:

$$\begin{aligned}\dot{x} &= A_{11}x + A_{12}z \\ 0 &= A_{21}x + A_{22}z\end{aligned}\tag{2.3}$$

Si A_{22} es no singular, entonces la ecuación algebraica lineal puede ser resuelta para que se obtenga el sistema reducido:

$$\dot{x} = (A_{11} - A_{12}A_{22}^{-1}A_{21})x\tag{2.4}$$

de la manera siguiente:

$$z = -A_{22}^{-1}A_{21}x$$

La configuración del múltiple del espacio es $M = \{(x, z) : z = -A_{22}^{-1}A_{21}x\}$. También, el sistema de escala de tiempo rápido en la escala de tiempo rápida es:

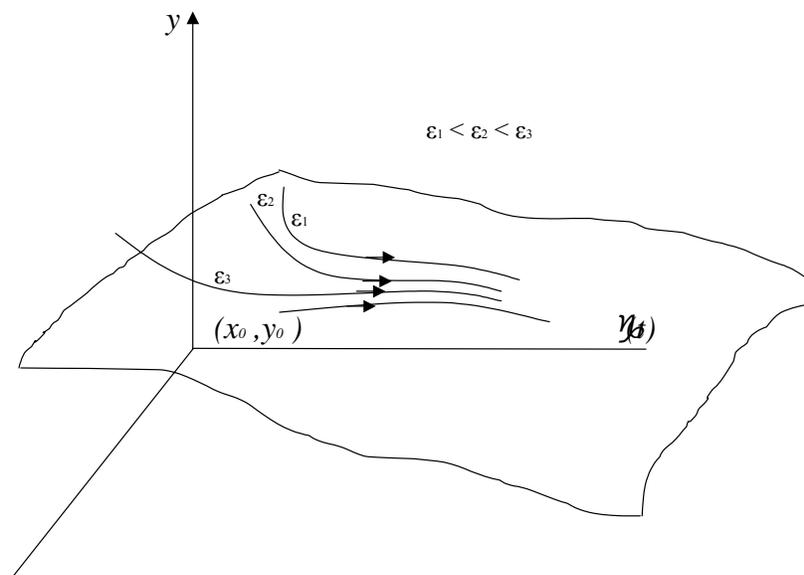


Figura 2.1: La consistencia graficada requerida.

$$\frac{\partial z}{\partial t} = A_{21}x + A_{22}z \quad (2.5)$$

Por lo tanto la estabilidad de la matriz A_{22} determina cuando el múltiple M es atractivo o no para la velocidad dinámica. Se generalizará esta noción para sistemas no lineales.

Se define:

$$M_1 = \{(x, z) : g(x, z) = 0, \det D_2 g(x, z) \neq 0\}.$$

Sea $(x_0, z_0) \in M_1$. Entonces por el teorema de la función implícita existe una integral única de la curva de (2.7), $\gamma(t) = (x(t), z(t))$ hasta (x_0, z_0) definido en $[0, \alpha)$ para algún $\alpha > 0$. Algunas trayectorias del sistema original (2.6), originándose arbitrariamente de $(x(0), z(0)) \in \mathbb{R}^{n+m}$ para todo $\epsilon > 0$ tiende uniformemente (consistentemente) para $\gamma(t)$ cuando $\epsilon \rightarrow 0$ en un intervalo cerrado de $(0, \alpha)$ como se ve en la figura(2.1).

Comentario 2.2 *Planteamos estos teoremas sin pruebas, pero versiones locales de estas*

fueron probadas por [45], [59] para intervalos de tiempo acotados, y [30] para intervalos de tiempo infinitos. Una buena versión de las pruebas están dadas en [40]. Finalmente, una versión geométrica moderna de estas pruebas usando el teorema del Múltiple Central fué dado por [14].

2.2.2. Sistemas no lineales

Como ya se indicó anteriormente, un problema básico en sistemas de control es el modelado matemático de un sistema físico y la representación real de muchos sistemas es por medio de ecuaciones diferenciales de orden alto. La presencia de algunos parámetros parásitos tales como pequeñas constantes de tiempo, masas, momentos de inercia, resistencias, inductancias y capacitancias es muy a menudo la causa del incremento de orden de estos sistemas. Generalmente hablando, un sistema en que la supresión de pequeños parámetros es responsable para la degeneración o degradación en la dimensión es un término expresado como sistema *perturbado singularmente*. Alternativamente, un problema descrito por una ecuación diferencial que involucra un pequeño parámetro ϵ es llamado un problema perturbado singularmente si el orden de la ecuación diferencial disminuye para $\epsilon = 0$ en lugar de $\epsilon \neq 0$. Obviamente, el parámetro pequeño ϵ multiplica a la derivada de mayor orden de la ecuación diferencial.

- Por lo contrario, en un problema de perturbación normal o regular, el parámetro pequeño ϵ no es responsable de la reducción de orden del problema, ya que no multiplica a la derivada de mayor orden de un sistema. Es posible utilizar la siguiente representación matemática de un sistema no lineal perturbado singularmente, el cual también es conocido como modelo de perturbación singular standard:

$$\begin{aligned} \dot{x} &= f(t, x, z, \epsilon), & x(0) &= x_0 \\ \epsilon \dot{z} &= g(t, x, z, \epsilon), & z(0) &= z_0 \end{aligned} \tag{2.6}$$

- Donde la variable estado x se mueve lentamente comparado con z .

Para ϵ pequeño y positivo, la primera ecuación describe las dinámicas lentas, mientras que la segunda ecuación define las dinámicas rápidas. La idea principal será aproximar a x con la solución del problema reducido posicionando $\epsilon = 0$, es decir:

$$\begin{aligned}\dot{x} &= f(t, x, z, \epsilon), & x(0) &= x_0 \\ 0 &= g(t, x, z, \epsilon), & z(0) &= z_0\end{aligned}$$

$$\dot{\bar{x}} = f(t, x, h(t, \bar{x}), 0) \quad \bar{x}(0) = x_0 \quad (2.7)$$

donde: $h(t, \bar{x})$ esta definido por la ecuación: $0 = g(t, x, h(t, x), 0)$.

- El problema de reducción de orden o degradado:

$$\frac{dx^{(0)}}{dt} + x^{(0)} = 0$$

- obtenido por la supresión del pequeño parámetro ϵ en la ecuación diferencial original, tendrá una condición en la frontera $x^{(0)}(t=1) = x_f$ si ϵ tiende a 0_+ y $x^{(0)}(t=0) = x_i$ si ϵ tiende a 0_- . En cada caso, una condición en la frontera será sacrificada en el proceso de degradación:

$$\begin{aligned}\lim_{t \rightarrow 0} \lim_{\epsilon \rightarrow 0} [x(t, \epsilon)] &= x^{(0)}(0) = x_f \\ \lim_{t \rightarrow 0} \lim_{\epsilon \rightarrow 0} [x(t, \epsilon)] &= x_i\end{aligned}$$

- De las ecuaciones (2.6)y(2.7), es claro, que para pequeños valores de ϵ , una raíz es grande comparada con la otra raíz. Esto significa que la solución consiste de componentes (*modos*) lento y rápido correspondiendo a valores propios pequeños y grandes, respectivamente.

Se revelan claramente las siguientes características importantes de perturbaciones singulares:

- La perturbación es de $\epsilon = 0$ a $\epsilon \neq 0$ y por lo tanto el problema (2.8) ó (2.12) es llamado un problema perturbado singularmente (*perturbación singular*), descrito por una ecuación diferencial con el parámetro pequeño ϵ multiplicado por la derivada de mayor orden.
- La solución tendrá una convergencia no uniforme.
- Existe un nivel o capa frontera donde la solución cambia rápidamente.
- Los procesos de perturbación (de $\epsilon = 0$ a $\epsilon \neq 0$) y proceso de degradación (de $\epsilon \neq 0$ a $\epsilon = 0$) son opuestos en efecto y acción.
- El problema degenerado o degradado, también llamado el problema no perturbado, es de orden reducido y no puede satisfacer todas las condiciones en la frontera dadas para el original, completo, o problema perturbado.
- El problema perturbado singularmente posee una propiedad de escala de dos tiempos debido a la presencia simultánea de fenómenos lentos y rápidos y el problema se dice ser rígido.

En una diversa variedad de aplicaciones prácticas, las variables de estado de sistemas no lineales involucra escalas de tiempo diferentes, es decir, escalas de tiempo rápidas y lentas. Este fenómeno es causado por una variedad de factores diferentes, el más común es la existencia de elementos pequeños o parásitos en los sistemas dinámicos. Esta situación es modelada por el siguiente conjunto de ecuaciones diferenciales:

$$\begin{aligned} \dot{x} &= f(x(t), z(t), \epsilon) \\ \epsilon \dot{z} &= g(x(t), z(t), \epsilon) \end{aligned} \tag{2.8}$$

donde: $x \in \mathbb{R}^n$, $z \in \mathbb{R}^m$. En la ecuación (2.8), el parámetro pequeño ϵ modela el hecho de que la variable z evoluciona más rápido que la variable x . Para $\epsilon \neq 0$ el sistema (2.8) tiene un espacio de estado de dimensión $n + m$. Sin embargo en $\epsilon = 0$ las ecuaciones degeneran en :

$$\begin{aligned}\dot{x} &= f(x(t), z(t), 0) \\ 0 &= g(x(t), z(t), 0)\end{aligned}\tag{2.9}$$

El sistema (2.9), consiste de m ecuaciones algebraicas y n ecuaciones diferenciales. Una descripción geométrica más completa de estas esta dada notando que si 0 es un valor regular de $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, entonces el conjunto $M = \{(x, z) : g(x, y) = 0\}$ es un múltiple $n - dimensional$. Las ecuaciones (2.9) necesitan ser interpretadas. La interpretación simple es que el sistema (2.9), representa un sistema dinámico en M en la siguiente manera: Si la segunda ecuación $g(x, y) = 0$ puede ser resuelta para z unicamente local como una función de x , se dice que $z = \psi(x)$ (Por el teorema de la función aimplícita, esto es, si es posible que $D_2g(x, z)$ es invertible para $(x, z) \in M$) entonces el múltiple M puede ser localmente parametrizado por x , y la dinámica del sistema en M puede ser descrito en términos de:

$$\dot{x} = f(x(t), \psi(x))\tag{2.10}$$

Observar que la ecuación (2.10) representa una ecuación $n - dimensional$, que es referida como el sistema reducido. Es obvia la importancia para deteminar cuándo las trayectorias del sistema original (2.8) converge a este sistema reducido de orden. El comportamiento de la variable z del sistema reducido de orden puede ser también descrito por una ecuación diferencial:

$$\dot{z} = \frac{\partial \psi(x)}{\partial x} f(x, \psi(x))\tag{2.11}$$

El sistema reducido de orden es también referido como el sistema de escala de tiempo lento. El sistema de escala de tiempo rápido es el sistema obtenido por reescalamiento del tiempo t para $\tau = \frac{t}{\epsilon}$ para obtener:

$$\begin{aligned}\frac{\partial x}{\partial \tau} &= \epsilon f(x, z) \\ \frac{\partial z}{\partial \tau} &= g(x, z)\end{aligned}\tag{2.12}$$

cuando $\epsilon \rightarrow 0$ las ecuaciones se tornarán:

$$\gamma \begin{cases} \frac{\partial x}{\partial \tau} = 0 \\ \frac{\partial z}{\partial \tau} = g(x, z) \end{cases} \quad (2.13)$$

Las dinámicas del sistema aceleran γ en la escala de tiempo rápida. τ muestra que la variable x es *fijada*. El conjunto de equilibrio de la aceleración del sistema γ es precisamente M , la configuración del sistema múltiple. Las dinámicas rápidas de:

$$\frac{\partial z}{\partial \tau} = g(x, z)$$

describen cuando el múltiple M es atractivo o repulsivo para las dinámicas parásitas.

Capítulo 3

Redes neuronales con diferentes escalas de tiempo y propiedades

Por el buen comportamiento de las redes neuronales recurrentes dinámicas y como aproximadores universales de funciones complejas no lineales, es posible la representación y modelado de un sistema no lineal utilizando redes neuronales artificiales. La representación real de muchos sistemas son indicados por ecuaciones diferenciales de alto orden, de la misma manera, una red neuronal recurrente dinámica, puede ser modelada por medio de ecuaciones diferenciales de cierta complejidad, lo que les da el nombre de redes neuronales diferenciales.

Las técnicas de redes neuronales artificiales, al parecer son bastante efectivas y flexibles, para modelar, identificar y controlar diversos tipos de sistemas no lineales complejos. Las redes neuronales han sido empleadas en la identificación y control de sistemas no lineales desconocidos debido a su masivo paralelismo, rápida adaptación y capacidad de aprendizaje con bastante éxito. Por lo tanto se pretende unificar los criterios y propiedades, tanto de las técnicas de perturbaciones singulares, como las de redes neuronales recurrentes o dinámicas, con la finalidad de utilizarlas en conjunto y así poder modelar, identificar y en posteriores investigaciones, crear leyes de control, para diversos sistemas físicos de naturaleza no lineal.

3.1. Estructura de redes neuronales con diferentes escalas de tiempo

Las ecuaciones de redes neuronales diferenciales que describen la dinámica lenta y rápida de un sistema perturbado singularmente o en escalas de dos tiempos, en forma de ecuación de estado son:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(z) + W_3\phi_1(x)u \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(x) + W_4\phi_2(z)u\end{aligned}\tag{3.1}$$

donde: $x \in R^n, z \in R^n$ son las variables de estado, $A \in R^{n \times n}, B \in R^{n \times n}$ son las matrices de estabilidad (Hurwitz). $W_1 \in R^{n \times 2n}, W_2 \in R^{n \times 2n}, \sigma_1(x, z) \in R^{2n}, \sigma_k(x, z) = [\sigma_k(x_1) \cdots \sigma_k(x_n), \sigma_k(z_1) \cdots \sigma_k(z_n)]^T \in R^{2n}$, $\phi(\cdot)$ es $R^{2n \times 2n}$ matriz diagonal, $\phi_k(x, z) = \text{diag}[\phi_k(x_1) \cdots \phi_k(x_n), \phi_k(z_1) \cdots \phi_k(z_n)]$, $k = 1, 2$. En la figura (3.1), se muestra la estructura de red neuronal diferencial recurrente, que modela la ecuación de red neuronal con perturbación singular (3.1). Y también se muestra en la figura (3.2) la misma estructura de red neuronal diferencial recurrente, pero sin entrada u y puede tener o no, capas de neuronas ocultas.

Comentario 3.1 *La red neuronal dinámica (3.1) ya ha sido discutida por muchos autores, por ejemplo, [12],[8],[10] y [17]. Se puede observar que el modelo de Hopfield es un caso especial de esta red neuronal con $A = \text{diag}\{a_i\}$, $a_i := -\frac{1}{R_i C_i}$, $R_i > 0$ y $C_i > 0$. R_i y C_i son la resistencia y capacitancia en el i -ésimo nodo de la red, respectivamente.*

Comentario 3.2 *Un caso general, de redes neuronales diferenciales con entrada de control con y sin capas de neuronas ocultas son:*

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(x, z) + W_3\phi_1(x, z)u \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(x, z) + W_4\phi_2(x, z)u\end{aligned}\tag{3.2}$$

donde $u(k) = [u_1, u_2 \cdots u_m, 0, \cdots 0]^T \in R^{2n}$, $W_1 \in R^{n \times 2n}, W_2 \in R^{n \times 2n}, W_3 \in R^{n \times 2n}$ y $W_4 \in R^{n \times 2n}$ son las matrices de pesos, $\sigma_k(x, z) = [\sigma_k(x_1) \cdots \sigma_k(x_n), \sigma_k(z_1) \cdots \sigma_k(z_n)]^T \in$

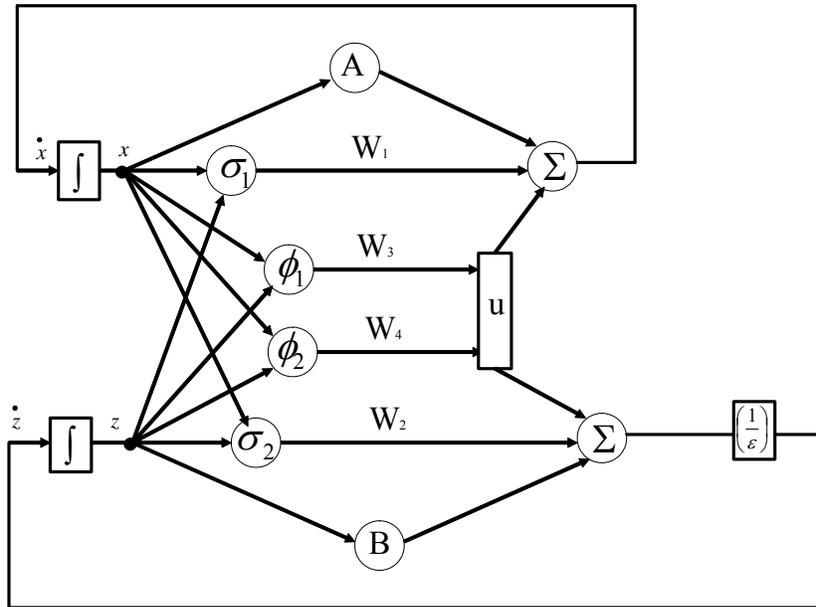


Figura 3.1: Red Neuronal Diferencial Recurrente.

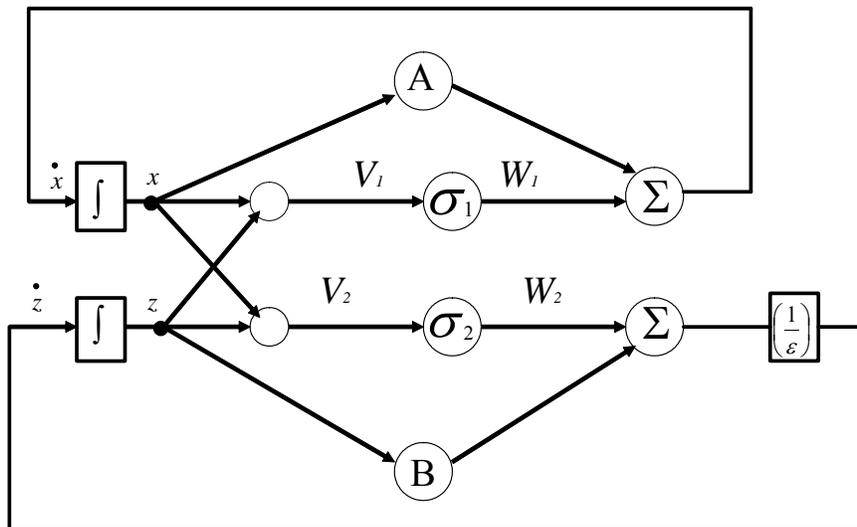


Figura 3.2: Red Neuronal Diferencial Recurrente sin entrada.

R^{2n} , $\phi(\cdot)$ es $R^{2n \times 2n}$ matriz diagonal, $\phi_k(x, z) = \text{diag}[\phi_k(x_1) \cdots \phi_k(x_n), \phi_k(z_1) \cdots \phi_k(z_n)]$, $k = 1, 2$. El análisis de estabilidad de estas redes neuronales diferenciales pueden fácilmente ser extendidas de los resultados de investigaciones realizadas en esta tesis.

Pero también es posible considerar una muy semejante red neuronal dinámica de Hopfield con capas de neuronas ocultas:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(V_1z) + W_3\phi_1(V_3x)u \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(V_2x) + W_4\phi_2(V_4z)u\end{aligned}\quad (3.3)$$

donde: x, z son las variables de estado de la red neuronal, A, B son las matrices de estabilidad Hurwitz, W_1, W_3 son las matrices de salida de los pesos de las capas externas, W_2, W_4 son las matrices de salida de los pesos de las capas ocultas, V_1, V_2, V_3, V_4 son las matrices de los pesos de las capas ocultas, $\sigma_1, \sigma_2, \phi_1, \phi_2$ son las funciones de activación generalment no lineales y u es la entrada de la red. Pero también es posible modelarla, agregando las variables de estado involucradas en un sistema sometido a perturbaciones singulares:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(V_1[x, z]^T) + W_3\phi_1(V_3[x, z]^T)u \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(V_2[x, z]^T) + W_4\phi_2(V_4[x, z]^T)u\end{aligned}\quad (3.4)$$

En la figura (3.3) se muestra la estructura de red neuronal diferencial recurrente con perturbación singular, con capas de neuronas ocultas, la cual esta modelada por la ecuación (3.4).

El modelado de sistemas físicos no lineales en la forma de perturbaciones singulares con redes neuronales dinámicas puede no ser fácil. No siempre es claro como escoger los parámetros a ser considerados como pequeños. Afortunadamente, en muchas aplicaciones, el conocimiento de procesos físicos y el conjunto de componentes de un sistema permite usarlos de manera correcta. Se tratará de hacerlo mediante el ejemplo clásico de un motor de corriente directa, para que de alguna manera relacionar sus respectivas variables con las de la red neuronal perturbada singularmente. También es posible utilizar un modelo de red neuronal recurrente dinámica general, para poder utilizarla como herramienta en el modelado de sistemas físicos con perturbaciones singulares. La red propuesta es la que se bosqueja en la figura (3.4).

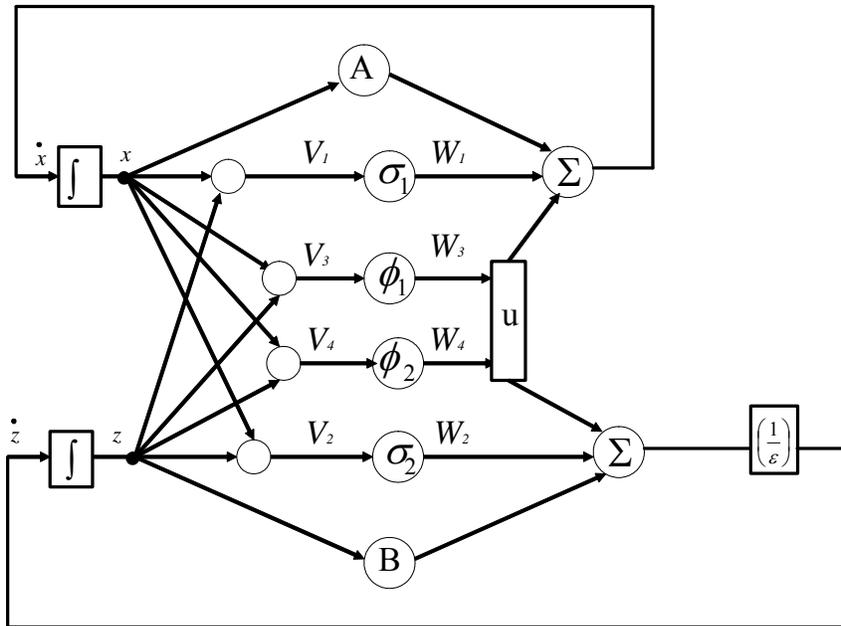


Figura 3.3: Red Neuronal Diferencial Recurrente con capas de neuronas ocultas.

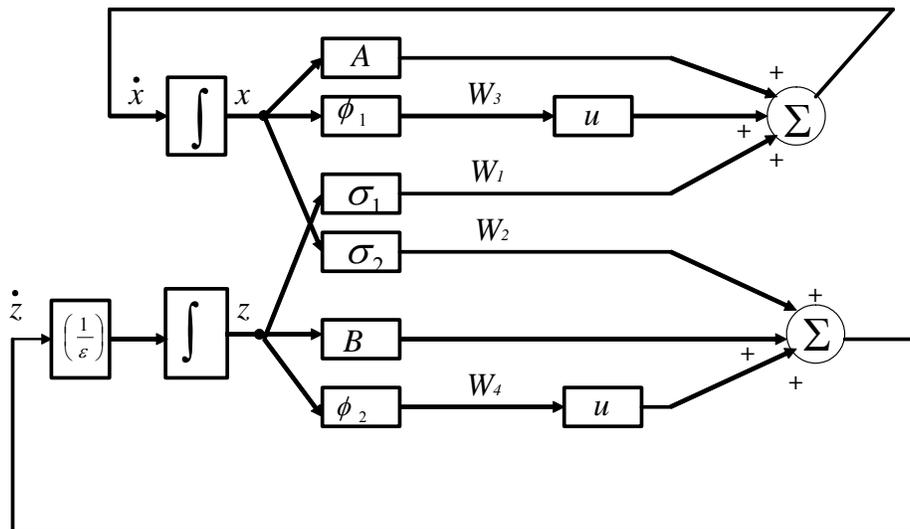


Figura 3.4: Red Neuronal Recurrente Dinámica.

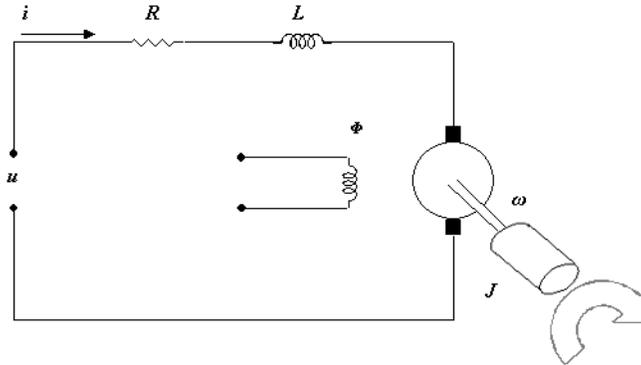


Figura 3.5: Motor de C.D.

Comentario 3.3 La técnica de perturbaciones singulares es usada para el análisis de un sistema completo. La ventaja de este enfoque es que se puede dividir el problema original o completo en dos subsistemas: el subsistema lento o sistema de estado cuasi-continuo y el subsistema rápido o sistema con nivel en la frontera o capa frontera. Ambos sistemas pueden ser estudiados independientemente.

Ejemplo 3.1 Un motor de C.D., como se muestra en la siguiente Fig.3.5, controlado por armadura puede ser modelado por una ecuación de estado de segundo orden; se intenta unificar el criterio anterior, con redes neuronales perturbadas singularmente.

$$J \frac{d\omega}{dt} = ki$$

$$L \frac{di}{dt} = -k\omega - Ri + u$$

donde i, u, R, L son la corriente de armadura, voltaje, resistencia, e inductancia, J es el momento de inercia, ω es la velocidad angular, ki y $k\omega$ son, respectivamente, el par y la fuerza electromotriz (*e.m.f.*) desarrollada con flujo de excitación constante. La primera ecuación de estado es una ecuación de par mecánico, y la segunda es una ecuación para el transitorio eléctrico en el circuito de armadura. Típicamente, L es pequeño y puede jugar el papel del parámetro ϵ . En forma de red neuronal con diferentes escalas de tiempo (3.3):

$$\begin{aligned} J\dot{\omega} &= A\omega + W_1\sigma_1(V_1i) + W_2\phi_1(V_2\omega)u \\ Li &= Bi + W_3\sigma_2(V_3\omega) + W_4\phi_2(V_4\omega)u \end{aligned}$$

Esto significa que, con $\omega = x$ y $i = z$, el modelo del motor esta en la forma estandard de (3.1)-(3.2) cuando $R \neq 0$. Despreciando L , y resolviendo el sistema se obtiene:

$$0 = -k\omega - Ri + u$$

o equivalentemente:

$$0 = Bi + W_3\sigma_2(V_3\omega) + W_4\phi_2(V_4i)u$$

para obtener la raíz única:

$$i = \frac{u - k\omega}{R}$$

que es equivalente a lo siguiente:

$$i = [-(W_3\sigma_2(V_3i) + W_4\phi_2(V_4i)u)]B^{-1}$$

y sustituyendo dentro la ecuación del par. El modelo resultante es:

$$J\dot{\omega} = -\frac{k^2}{R}\omega + \frac{k}{R}u$$

cuya equivalencia deducida es:

$$J\dot{\omega} = A\omega + W_1\sigma_1(V_1[-(W_3\sigma_2(V_3i) + W_4\phi_2(V_4i)u)]B^{-1}) + W_2\phi_1(V_2\omega)u$$

el comúnmente usado modelo de primer orden del motor de D.C.

Este sistema esta sujeto para consideraciones de análisis respecto a sus puntos de equilibrio, es decir:

$$\dot{x} = \dot{z} = 0$$

Comentario 3.4 *Aunque el modelo realizado anteriormente para el motor de C.D., básicamente correspondería a un sistema lineal, la aplicación de redes neuronales con diferentes escalas de tiempo es factible hacerlo al mismo sistema solamente agregando la no linealidad, como por ejemplo acoplando el eje de rotación del motor a una posible carga, lo que repercutiría en la velocidad angular del mismo, por la fuerza de fricción como fuerza de perturbación externa no lineal.*

Los métodos de perturbaciones singulares representan en ingeniería simplificaciones de modelos dinámicos, y son usados como un método de aproximación para el análisis y simplificación de sistemas no lineales. Estos revelan estructuras de escalas multi tiempo, inherentes en muchos problemas prácticos. La solución de la ecuación de estado a menudo exhibe el fenómeno, de que algunas variables se mueven mucho más rápido con respecto a otras, lo que conduce a la clasificación de variables *lentas* y *rápidas*. En ejemplos prácticos que no conducen a sistemas dinámicos, es necesario realizar un modelo de dos etapas por la separación de escalas de tiempo. Se obtendrá un modelo reducido que representa el fenómeno lento, mientras que el modelo con nivel o capa frontera involucra una escala de tiempo más rápida y representa desviaciones para predecir el comportamiento lento del sistema. La teoría de perturbaciones singulares abraza una diversa variedad de fenómenos dinámicos que poseen las dinámicas lenta y rápida de un sistema completo, y están presentes en muchos problemas neurodinámicos. Las redes neuronales diferenciales con dinámicas de escalas de dos tiempos pueden ser formuladas de una manera más general e interpretadas como sistemas perturbados singularmente. Se adoptarán las notaciones de [54] y [87] en la introducción y representación de un sistema neuronal perturbado singularmente. Las ecuaciones (3.6) y (3.7) se puede considerar que representan o modelan la dinámica lenta y rápida del sistema no lineal. Ambas ecuaciones son representaciones generales de las ecuaciones (3.1) y (3.2). Este enfoque de escala de tiempo es asintótico, es decir exacto, en el límite con un radio ϵ de la velocidad de la dinámica lenta contra la dinámica rápida cuando $\epsilon \rightarrow 0$. Cuando ϵ es pequeño y es posible cancelarlo, las aproximaciones son obtenidas del sistema reducido al ser separado en escalas de tiempo:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(x, z) + W_3\phi_1(x, z)u \\ 0 &= Bz + W_2\sigma_2(x, z) + W_4\phi_2(x, z)u\end{aligned}\tag{3.5}$$

$$\dot{x} = Ax + W_1\sigma_1(x, z) + W_3\phi_1(x, z)u \quad \text{dinámicas lentas}\tag{3.6}$$

$$\frac{dz}{d\tau} = Bz + W_2\sigma_2(x, z) + W_4\phi_2(x, z)u \quad \text{dinámicas rápidas}\tag{3.7}$$

$$\dot{x} = Ax + W_1\sigma_1(x, [W_2\sigma_2(x, z) + W_4\phi_2(x, z)u]B^{-1}) + W_3\phi_1(x, z)u \quad \text{Si stema Reducido}\tag{3.8}$$

Donde $\tau = \frac{t}{\epsilon}$ es una escala de tiempo flexible y el vector $x \in \mathbb{R}^n$ es tratado como un parámetro desconocido fijo. Básicamente, el objetivo fundamental de la teoría de perturbaciones singulares es bosquejar conclusiones acerca del comportamiento de un sistema original o completo, basándose únicamente en estudiar y analizar el sistema de orden reducido o simplificado obtenido de (3.8), posicionando $\epsilon = 0$, es decir, $0 = Bz + W_2\sigma_2(z) + W_4\phi_2(x, z)u$, para posteriormente sustituirla en la primera ecuación.

3.2. Propiedades de redes neuronales de escalas de dos tiempos

3.2.1. Sin control ($u = 0$)

Se analizará el llamado modelo de perturbación singular estandar de la forma (3.1), $u = 0$

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(V_1z) \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(V_2x)\end{aligned}$$

donde posicionando $\epsilon = 0$ se produce un fundamental y abrupto cambio en las propiedades dinámicas del sistema, cuando la red neuronal diferencial:

$$\epsilon \dot{z} = Bz + W_2 \sigma_2(V_2 x)$$

degenera o se degrada en la ecuación algebraica o trascendental siguiente:

$$\begin{aligned} 0 &= Bz + W_2 \sigma_2(V_2 x) & u &= 0 \\ z &= B^{-1}(-W_2 \sigma_2(V_2 \hat{x}_t)) & u &= 0 \end{aligned}$$

Comentario 3.5 *La esencia de la teoría desarrollada aquí es que la discontinuidad de las soluciones causada por perturbaciones singulares puede ser evitada si se analizan en escalas de tiempo separadas. Esta aproximación de escalas de tiempo es una característica fundamental del método de perturbaciones singulares. Esto surge del hecho de que al degradarse el problema completo, se pierde cierta información respecto a la parte de la dinámica rápida del sistema, sobre todo porque existen condiciones iniciales que al ser despreciado el parámetro ϵ , una de las condiciones iniciales no se toma en cuenta.*

El modelo de perturbación singular de un sistema dinámico con redes neuronales es un modelo de estado donde las derivadas de algunos de los estados esta multiplicada por un pequeño parámetro positivo ϵ , es decir:

$$\dot{x} = Ax + W_1 \sigma_1(V_1 z) \tag{3.9}$$

$$\epsilon \dot{z} = Bz + W_2 \sigma_2(V_2 x) \tag{3.10}$$

Asumiendo que las funciones: σ_1, σ_2 son continuamente diferenciables en sus argumentos para $[(V_1 x), (V_2 x)] \in [0, t_1] \times D_x \times D_z \times [0, \epsilon_0]$,

donde: $D_x \subset \mathbb{R}^n$ y $D_z \subset \mathbb{R}^m$ son conjuntos conexos abiertos. Cuando colocamos $\epsilon = 0$ en la ecuación anterior (3.9), la dimensión de la ecuación de estado se reduce de $(n + m)$ a (n) porque la red neuronal diferencial (3.10) degenera en la ecuación:

$$0 = Bz + W_2 \sigma_2(V_2 x) \tag{3.11}$$

Es posible afirmar que el modelo (3.9) y (3.10) esta en la forma estandard si (3.11) tiene $k \geq 1$ raíces reales aisladas:

$$z = h_i(t, x), \quad i = 1, 2, \dots, k \quad (3.12)$$

para cada $(t, x, u) \in [0, t_1] \times D_x$. Estas consideraciones aseguran que un bien definido modelo reducido $n - \text{dimensional}$ corresponderá para cada raíz del modelo reducido. Para obtener el $i - \text{ésimo}$ modelo reducido, sustituyendo (3.12) en (3.11), en $\epsilon = 0$, para obtener:

$$\dot{x} = Ax + W_1\sigma_1(V_1h_i(t, x)) \quad (3.13)$$

Será claro dentro el contexto qué raíz de (3.11) es necesario usar. Este modelo es algunas veces llamado un modelo de estado cuasi-continuo porque z , cuya velocidad $\dot{z} = \frac{\sigma_2}{\epsilon}$ puede ser grande cuando ϵ es pequeño y $\sigma_2 \neq 0$, puede rápidamente converger a la raíz de (3.12) que es el equilibrio de (3.11). El modelo (3.13) es también conocido como el modelo o dinámica lenta del sistema.

3.2.2. Con control ($u \neq 0$)

Las perturbaciones singulares causan un comportamiento de escalas multitiempo de sistemas dinámicos caracterizados por la presencia de transitorios lentos y rápidos en la respuesta del sistema a estímulos externos. Hablando vagamente, la respuesta lenta o estado cuasi-continuo esta aproximada por el modelo reducido (3.13), mientras la desigualdad entre la respuesta del modelo reducido y el del modelo completo (3.11) y (3.12) es el transitorio rápido. Para ver este punto, se considera el problema de resolver la ecuación de estado:

$$\begin{aligned} \dot{x} &= Ax + W_1\sigma_1(V_1z) + W_3\phi_1(V_3x)u & x(t_0) &= \xi(\epsilon) \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(V_2x) + W_4\phi_2(V_4x)u & z(t_0) &= \eta(\epsilon) \end{aligned} \quad (3.14)$$

donde $\xi(\epsilon)$ y $\eta(\epsilon)$ dependen suavemente sobre ϵ y $t_0 \in [0, t_1)$. Sean $x(t, \epsilon)$ y $z(t, \epsilon)$ que denotan la solución del problema completo de (3.5). Cuando se define el correspondiente problema para el modelo reducido (3.13), se pueden solo especificar n condiciones iniciales,

ya que el modelo es de $n - \acute{e}simo$ orden. Naturalmente, reteniendo el estado inicial de x para obtener el problema reducido que representa la dinámica lenta del sistema:

$$\dot{x} = Ax + W_1\sigma_1(V_1z) + W_3\phi_1(V_3x)u \quad x(t_0) = \xi_0 \stackrel{def}{=} \xi(0) \quad (3.15)$$

Se denota la solución de (3.15) por $\bar{x}(t)$. Porque la variable z ha sido excluida del modelo reducido y sustituida por su estado cuasi-continuo $h(t, x)$, la información solo es posible obtenerla acerca de z resolviendo (3.15) para calcular:

$$\bar{z}(t) \stackrel{def}{=} h(t, \bar{x}(t))$$

que describe el comportamiento cuasi-continuo de estado de z cuando $x = \bar{x}$. Por contraste para la variable original z comenzando en t_0 de $\eta(\epsilon)$ prescrito, el estado cuasi-continuo \bar{z} no es libre para empezar de un valor prescrito, y puede haber una gran desigualdad entre su valor inicial $\bar{z}(t_0) = h(t_0, \xi_0)$ y el estado inicial prescrito $\eta(\epsilon)$. Por lo tanto, $\bar{z}(t)$ no puede ser una aproximación uniforme de $z(t, \epsilon)$. Lo mejor que se puede esperar es que el estimativo:

$$z(t, \epsilon) - \bar{z}(t) = O(\epsilon)$$

se mantendrá sobre un intervalo excluyendo t_0 , es decir, para $t \in [t_b, t_1]$, donde $t_b > t_0$. Por otra parte, es razonable esperar la estimación:

$$x(t, \epsilon) - \bar{x}(t) = O(\epsilon)$$

se mantendrá uniforme para todo $t \in [t_0, t_1]$, ya que:

$$x(t_0, \epsilon) - \bar{x}(t_0) = \xi(\epsilon) - \xi(0) = O(\epsilon)$$

Si el error $z(t, \epsilon) - \bar{z}(t)$ es realmente $O(\epsilon)$ sobre $[t_b, t_1]$, entonces debe ser verdad que durante el intervalo inicial (nivel o capa frontera) $[t_0, t_b]$, la variable z se aproxima a \bar{z} . Tomando en cuenta que la velocidad de z puede ser alta, porque $\dot{z} = \frac{\sigma^2}{\epsilon}$. En efecto, colocando $\epsilon = 0$ en (3.14), se hace el transitorio de z instantáneo cuando $\sigma \neq 0$. Del estudio previo de

la estabilidad de los puntos de equilibrio, será claro que no se puede esperar que z converja a su estado cuasi-continuo \bar{z} , bajo la certeza que las condiciones de estabilidad son satisfechas. Tales condiciones resultarán del análisis disponible. Que es más conveniente en el análisis la realización del cambio de variables:

$$y = z - h(t, x) \quad (3.16)$$

donde y es la nueva variable en la que se puede establecer el nuevo origen, para alcanzar el equilibrio; entonces se puede utilizar lo siguiente:

$$z = y + h(t, x) \quad (3.17)$$

que cambia el estado cuasi-continuo de z en el origen. En las nuevas variables (x, y) , el problema completo es:

$$\dot{x} = Ax + W_1\sigma_1(V_1z) + W_3\phi_1(V_3x)u \quad x(t_0) = \xi(\epsilon) \quad (3.18)$$

$$\begin{aligned} \epsilon\dot{y} &= By + W_2\sigma_2(V_2, x) + W_4\phi_2(V_4, x)u - \epsilon\frac{\partial h}{\partial t} - \epsilon\frac{\partial h}{\partial t} By + \\ &\quad + W_2\sigma_2(V_2, x) + W_4\phi_2(V_4, x)u \\ y(t_0) &= \eta(\epsilon) - h(t_0, \xi(\epsilon)) \end{aligned} \quad (3.19)$$

$$\dot{x} = f(t, x, y + h(t, x), \epsilon), \quad x(t_0) = \xi(\epsilon)$$

$$\begin{aligned} \epsilon\dot{y} &= g(t, x, y + h(t, x), \epsilon) - \epsilon\frac{\partial h}{\partial t} \\ &\quad - \epsilon\frac{\partial h}{\partial x}f(t, x, y + h(t, x), \epsilon), \quad y(t_0) = \eta(\epsilon) - h(t_0, \xi(\epsilon)) \end{aligned}$$

El estado cuasi-continuo de (3.19) es ahora $y = 0$, que cuando se sustituye en (3.18) resulta en el modelo reducido. Para analizar (3.19), se observa que $\epsilon\dot{y}$ puede permanecer finito cuando ϵ tiende a cero y \dot{y} tiende a infinito. Entonces se establece que:

$$\epsilon\frac{dy}{dt} = \frac{dy}{d\tau}; \text{ por lo tanto, } \frac{d\tau}{dt} = \frac{1}{\epsilon}$$

y usando $\tau = 0$ como el valor inicial en $t = t_0$. La nueva variable $\tau = \frac{(t-t_0)}{\epsilon}$ es flexible, es decir si ϵ tiende a cero, τ tiende a infinito

para t finito solo para ligeramente más grande que t_0 por una diferencia fija (independiente de ϵ). En la escala de tiempo τ , esta representada por:

$$\begin{aligned} \frac{dy}{d\tau} &= By + W_2\sigma_2(V_2x) + W_4\phi_2(V_4x)u - \epsilon\frac{\partial h}{\partial t} - \epsilon\frac{\partial h}{\partial t} = \\ &= By + W_2\sigma_2(V_2x) + W_4\phi_2(V_4x)u \\ y(t_0) &= \eta(\epsilon) - h(t_0, \xi(\epsilon)) \end{aligned} \quad (3.20)$$

Las variables t y x en la ecuación externa variarán lentamente, en la escala de tiempo τ , y estan dadas por:

$$t = t_0 + \epsilon\tau, \quad x = x(t_0 + \epsilon\tau, \epsilon)$$

Posicionando $\epsilon = 0$ se fijan estas variables en $t = t_0$ y $x = \xi_0$, y se reduce al sistema autónomo:

$$\begin{aligned} \frac{dy}{d\tau} &= By + W_2\sigma_2(V_2x) + W_4\phi_2(V_4x)u - \epsilon\frac{\partial h}{\partial t} - \epsilon\frac{\partial h}{\partial t} = \\ &= By + W_2\sigma_2(V_2x) + W_4\phi_2(V_4x)u \\ y(0) &= \eta(0) - h(t_0, \xi_0) \stackrel{def}{=} \eta_0 - h(t_0, \xi_0) \end{aligned} \quad (3.21)$$

que es el equilibrio en $y = 0$. Si este punto de equilibrio es asintóticamente estable y $y(0)$ pertenece a su región de atracción, es razonable esperar que la solución de (3.21) se alcanzará en la vecindad de $O(\epsilon)$ del origen durante el intervalo en el límite en la frontera. Más allá de este intervalo, es necesario la propiedad de estabilidad que garantice que $y(\tau)$ permanecerá cerca de cero, mientras los parámetros (t, x) que varían lentamente se mueven de sus valores iniciales (t_0, ξ_0) . Para analizar esta situación, se permite que los parámetros previamente fijos tomen los valores en la región de la variación lenta de los parámetros (t, x) . Se asume que la solución $\bar{x}(t)$ del problema reducido esta definido para $t \in [0, t_1]$ y $\bar{x}(t) \in D_x \subset \mathbb{R}^n$, para algún dominio D_x . Entonces reescribiendo (3.21) como:

$$\frac{dy}{d\tau} = Ay + W\sigma(y + h(t, x)) \quad (3.22)$$

donde $(t, x) \in [0, t_1] \times D_x$ son tratados como parámetros fijos. Se referirá a (3.22) como el modelo con límite en la frontera o sistema con nivel o capa frontera. Algunas veces, también se referirá a (3.21) como el modelo con nivel en la frontera. Esto no causará confusión, porque (3.21) es una evaluación de (3.22) para un tiempo inicial dado y un estado inicial. La crucial propiedad de estabilidad que se necesita para (3.22) es la estabilidad exponencial en su origen, uniformemente en los parámetros fijos, como se plantea en la siguiente definición.

Definición 3.1 *El punto de equilibrio $y = 0$ del sistema con nivel o capa frontera (3.22) es exponencialmente estable, uniformemente en $(t, x) \in [0, t_1] \times D_x$, si existen constantes k, γ, ρ_0 positivas, tales que las soluciones de (3.22) satisfacen:*

$$\|y(\tau)\| \leq k \|y(0)\| \exp(-\gamma\tau), \quad \forall \|y(0)\| < \rho_0, \quad \forall (t, x) \in [0, t_1] \times D_x, \quad \forall \tau \geq 0 \quad (3.23)$$

Aparte del caso trivial donde la solución del modelo con nivel en la frontera puede ser conocido de manera cerrada, la verificación de la estabilidad exponencial del origen se podrá hacer por linealización o vía análisis de Lyapunov. Se puede mostrar que si la matriz Jacobiana $\left[\frac{\partial\sigma}{\partial y}\right]$ satisface la condición del valor propio:

$$\operatorname{Re} \left[\lambda \left\{ \frac{\partial\sigma}{\partial y}(t, x, h(t, x), 0) \right\} \right] \leq -c < 0, \quad \forall (t, x) \in [0, t_1] \times D_x \quad (3.24)$$

entonces existen constantes k, γ, ρ_0 para que (3.23) sea satisfecha. Esto, naturalmente, es un resultado local; es decir, la constante ρ_0 puede ser muy pequeña. Alternativamente, se puede mostrar que si existe una función de Lyapunov $V(t, x, y)$ que satisface:

$$c_1 \|y\|^2 \leq V(t, x, y) \leq c_2 \|y\|^2 \quad (3.25)$$

$$\frac{\partial V}{\partial y} \phi(t, x, y + h(t, x), 0) \leq -c_3 \|y\|^2 \quad (3.26)$$

para $(t, x, y) \in [0, t_1] \times D_x \times D_y$, donde $D_y \subset \mathbb{R}^m$ es un dominio que contiene el origen, entonces (3.23) se satisface con las estimaciones:

$$\rho_0 = \rho \sqrt{\frac{c_1}{c_2}}, \quad k = \sqrt{\frac{c_2}{c_1}}, \quad \gamma = \frac{c_3}{2c_2}$$

en donde $B_\rho \subset D_y$.

Teorema 3.1 *Considérese el problema de perturbación singular de (3.14) y sea $z = h(t, x)$ una raíz aislada de (2.9). Se asume que las siguientes condiciones son satisfechas para todo:*

$$[t, x, z - h(t, x), \epsilon] \in [0, t_1] \times D_x \times D_y \times [0, \epsilon_0]$$

para algunos dominios $D_x \subset \mathbb{R}^n$ y $D_y \subset \mathbb{R}^m$, en que D_x es convexo y D_y contiene el origen:

- Las funciones σ, ϕ , sus primeras derivadas parciales con respecto a (x, z, ϵ) , y la primera derivada parcial de ϕ con respecto a t son continuas; la función $h(t, x)$ y el Jacobiano $\left[\frac{\partial g(t, x, z, 0)}{\partial z} \right]$ tienen primera derivada parcial continua con respecto a sus argumentos; el dato inicial $\xi(\epsilon)$ y $\eta(\epsilon)$ son funciones suaves de ϵ .

- El problema reducido (3.18) tendrá una única solución $\bar{x} \in S$, para $t \in [t_0, t_1]$, donde S es un subconjunto compacto de D_x .

- El origen es un punto de equilibrio exponencialmente estable del modelo (3.22), uniformemente en (t, x) ; sea $R_y \subset D_y$ la región de atracción de (3.21) y sea Ω_y un subconjunto compacto de R_y .

Entonces, existe una constante positiva ϵ^* tal que para todo $\eta_0 - h(t_0, \xi_0) \in \Omega_y$ y $0 < \epsilon < \epsilon^*$, el problema de perturbación singular de (3.14) tendrán una solución única $x(t, \epsilon), z(t, \epsilon)$ sobre $[t_0, t_1]$, y:

$$x(t, \epsilon) - \bar{x}(t) = O(\epsilon) \tag{3.27}$$

$$z(t, \epsilon) - h(t, \bar{x}(t)) - y\left(\frac{t}{\epsilon}\right) = O(\epsilon) \tag{3.28}$$

se mantendrá uniforme para $t \in [t_0, t_1]$, donde $y(\tau)$ es la solución del modelo (3.27) con límite o capa frontera. Más aún, dado cualquier $t_b > t_0$, existe un $\epsilon^{**} \leq \epsilon^*$ tal que:

$$z(t, \epsilon) - h(t, \bar{x}(t)) = O(\epsilon) \quad (3.29)$$

se mantendrá uniforme para $t \in [t_b, t_1]$ cuando $\epsilon < \epsilon^{**}$.

Cabe mencionar y aclarar que el término $O(\epsilon)$ se refiere como una aproximación o estimación y se le conoce comúnmente, como símbolo de orden o de Landu.

Comentario 3.6 *Este teorema es conocido como el teorema de Tikhonov. Su prueba usa las propiedades de estabilidad del modelo con nivel o capa frontera para mostrar que:*

$$\|y(t, \epsilon)\| \leq \exp\left[\frac{-\alpha(t-t_0)}{\epsilon}\right] + \epsilon\delta$$

Comentario 3.7 *La cota precedente es usada en (3.18) para probar (3.27), que es admisible, debido a que $\int_0^t \exp(-\frac{\alpha s}{\epsilon}) ds$ es $O(\epsilon)$. La prueba finaliza con el análisis del error de (3.19) en la escala de tiempo τ para probar (3.28) y (3.29). Ver C.17 de [40]*

3.2.3. Perturbación singular en el intervalo infinito.

El teorema anterior es válido solo en intervalos de tiempo $O(1)$. Este hecho puede ser fácilmente visto de la prueba del teorema [40]. En particular, se establece que:

$$\|x(t, \epsilon) - \bar{x}(t)\| \leq \epsilon k [1 + t_1 - t_0] \exp[L(t_1 - t_0)]$$

donde k, L son constantes positivas y que prueba el error estimado para x .

Para cualquier t_1 finito, el estimado externo es $O(\epsilon)$, pero $O(\epsilon)$ no es uniforme en t para todo $t \geq t_0$. Para que el planteamiento anterior se cumpla, se necesita mostrar que:

$$\|x(t, \epsilon) - \bar{x}(t)\| \leq \epsilon k, \quad \forall t \in [t_0, \infty)$$

Esto puede ser hecho bajo algunas condiciones de estabilidad adicional. En el siguiente teorema, se requiere que el sistema reducido (2.9) tendrá un punto de equilibrio exponencialmente estable en el origen y usa la función de Lyapunov para estimar su región de atracción.

Teorema 3.2 [40] *Considérese el problema de la red neuronal recurrente perturbada singularmente de (3.14) y sea $z = h(t, x)$ una raíz aislada de (2.9). Se asume que las siguientes se satisfacen para todo:*

$$[t, x, z - h(t, x), \epsilon] \in [0, \infty) \times D_x \times D_y \times [0, \epsilon_0]$$

para algunos dominios $D_x \subset \mathbb{R}^n$ y $D_y \subset \mathbb{R}^m$, que contiene sus respectivos orígenes:

Teorema 3.3 • *En cualquier subconjunto compacto de $D_x \times D_y$, las funciones f, g , sus primeras derivadas parciales con respecto a (x, z, ϵ) , y la primera derivada parcial de g con respecto a t son continuas y acotadas, $h(t, x)$ y $\left[\frac{\partial g(t, x, z, 0)}{\partial z}\right]$ tienen primeras derivadas parciales con respecto a sus argumentos, y $\left[\frac{\partial f(t, x, h(t, x), 0)}{\partial z}\right]$ es Lipschitz en x , uniformemente en t ; el dato inicial $\xi(\epsilon)$ y $\eta(\epsilon)$ son funciones suaves de ϵ ;*

- *el origen es un punto de equilibrio exponencialmente estable del sistema reducido (2.9); existe una función de Lyapunov $V(t, x)$ que satisface las condiciones del teorema [40] para (2.9) para $(t, x) \in [0, \infty) \times D_x$ y $\{W_1(x) \leq c\}$ es un subconjunto compacto de D_x ;*

- *el origen es un punto de equilibrio exponencialmente estable del sistema con límite en la frontera (3.22), uniforme en (t, x) ; sea $\mathfrak{R}_y \subset D_y$ la región de atracción de (3.22) y sea Ω_y un subconjunto compacto de \mathfrak{R}_y .*

Entonces para cada conjunto compacto $\Omega_x \subset \{W_2(x) \leq \rho c, 0 < \rho < 1\}$ existe una constante positiva ϵ^ tal que para todo $t_0 \geq 0$, $\xi_0 \in \Omega_x$, $\eta_0 - h(t_0, \xi_0) \in \Omega_y$, y $0 < \epsilon < \epsilon^*$, el problema de perturbación singular de (3.14) tendrá una solución única $x(t, \epsilon)$, $z(t, \epsilon)$ en $[t_0, \infty)$, y :*

$$x(t, \epsilon) - \bar{x}(t) = O(\epsilon) \tag{3.30}$$

$$z(t, \epsilon) - h\left(t, \bar{x}(t)\right) - y\left(\frac{t}{\epsilon}\right) = O(\epsilon) \tag{3.31}$$

se mantiene uniforme para $t \in [t_0, \infty)$, donde $\bar{x}(t)$ y $y(\tau)$ son las soluciones de los problemas con límite o capa frontera reducidos (3.15) y (3.21). Más aún, dado cualquier $t_b > t_0$, existe $\epsilon^{**} \leq \epsilon^*$ tal que:

$$z(t, \epsilon) - h(t, \bar{x}(t)) = O(\epsilon)$$

se mantendrá uniforme para $t \in [t_b, \infty)$ cuando $\epsilon < \epsilon^{**}$.

Demostración. [40] ■

Comentario 3.8 *Si el sistema reducido (3.15) es autónomo, el conjunto Ω_x del teorema anterior puede ser cualquier subconjunto compacto de su región de atracción. Esta es una consecuencia (converso de Lyapunov) del teorema [40] que proporciona una función de Lyapunov $V(x)$ tal que cualquier subconjunto compacto de la región de atracción está en el interior de un conjunto compacto de la forma $\{V(x) \leq c\}$.*

3.2.4. Múltiples lentos y rápidos.

En esta sección, se dará un enfoque geométrico del comportamiento de dos escalas de tiempo de las soluciones de (3.1) y como trayectorias en \mathbb{R}^{n+m} . Para usar el concepto de múltiples invariantes, se restringirá la discusión a sistemas autónomos. Es más, para simplificar la notación, se considera σ y ϕ independientes de ϵ . Por lo tanto, se asume la siguiente forma simple del sistema perturbado singularmente sin entrada de control, es decir, $u = 0$:

$$\dot{x} = Ax + W_1\sigma_1(V_1z) \tag{3.32}$$

$$\epsilon\dot{z} = Bz + W_2\sigma_2(V_2x) \tag{3.33}$$

Sea $z = h(x)$ una raíz aislada de:

$$0 = Bz + W_2\sigma_2(V_2x)$$

y supóngase que las consideraciones del primer teorema se satisfacen para esta raíz. La ecuación $z = h(x)$ describe un múltiple $n - dimensional$ en el espacio de estado $(n + m) - dimensional$ de (x, z) . Este es un múltiple invariante para el sistema:

$$\dot{x} = Ax + W_1\sigma_1(V_1z) \quad (3.34)$$

$$0 = Bz + W_2\sigma_2(V_2x) \quad (3.35)$$

ya que la trayectoria de (3.34) y (3.35) empieza en el múltiple $z = h(x)$ que permanecerá en el múltiple para todo el tiempo futuro (para que la solución este definida). El movimiento en este múltiple esta descrito por el modelo reducido:

$$\dot{x} = Ax + W_1\sigma_1(V_1, h(x))$$

El primer teorema muestra que las trayectorias de (3.32) y (3.33), que empiezan en una $O(\epsilon)$ de la vecindad de $z = h(x)$, permanecerá con una $O(\epsilon)$ en la vecindad de $z = h(x)$. Esto motiva la siguiente cuestión: ¿Existe un análogo del múltiple invariante $z = h(x)$ para $\epsilon > 0$?. Se produce que, bajo las consideraciones del primer teorema, existe una cuasi múltiple invariante para (3.32) y (3.33) que con una $O(\epsilon)$ en una vecindad de $z = h(x)$. Entonces se observa que el múltiple invariante para (3.32) y (3.33) en la forma:

$$z = H(x, \epsilon) \quad (3.36)$$

donde H es una función suficientemente suave (es decir, suficientemente continuamente diferenciable) de x y ϵ . La expresión (3.36) define un múltiple $n - dimensional$, dependiente de ϵ , en el espacio de estado $(n + m) - dimensional$ de (x, z) . Para $z = H(x, \epsilon)$ será un múltiple invariante de (3.32) y (3.33), entonces necesita ser verdad que:

$$z(0, \epsilon) - H(x(0, \epsilon), \epsilon) = 0 \implies z(t, \epsilon) - H(x(t, \epsilon), \epsilon) \equiv 0, \quad \forall t \in J \subset [0, \infty)$$

donde J es cualquier intervalo en que la solución $[x(t, \epsilon), z(t, \epsilon)]$ existe. Diferenciando ambos lados de (3.36) con respecto a t , multiplicando después por ϵ , y sustituyendo para \dot{x} , $\epsilon \dot{z}$ y z de (3.32) y (3.33) y (3.36), respectivamente, se obtiene la condición múltiple:

$$0 = \phi(x, H(x, \epsilon)) - \epsilon \frac{\partial H}{\partial x} \sigma(x, H(x, \epsilon)) \quad (3.37)$$

donde la relación $H(x, \epsilon)$ necesita satisfacer para todas las x en la región de interés y toda $\epsilon \in [0, \epsilon_0]$. En $\epsilon = 0$, la ecuación diferencial parcial (3.37) degenera en :

$$0 = \phi(x, H(x, 0))$$

que muestra que $H(x, 0) = h(x)$. Ya que $0 = \phi(x, z)$ puede tener más que una raíz aislada $z = h(x)$, se puede buscar un múltiple invariante para (3.32) y (3.33) en la vecindad de cada raíz. Se puede mostrar que existe $\epsilon^* > 0$ y una función $H(x, \epsilon)$ que satisface la condición múltiple (3.37) para todo $\epsilon \in [0, \epsilon^*]$ y:

$$H(x, \epsilon) - h(x) = O(\epsilon)$$

para x acotado. El múltiple invariante $z = H(x, \epsilon)$ es llamado un múltiple lento para (3.32) y (3.33). Para cada múltiple lento, le corresponde un modelo lento:

$$\dot{x} = Ax + W_1 \sigma_1(V_1, H(x, \epsilon))$$

que describe exáctamente el movimiento del múltiple.

En la mayoría de los casos, no se posible resolver la condición múltiple (3.37) exáctamente, pero es posible aproximar $H(x, \epsilon)$ arbitrariamente cerca con series de Taylor en $\epsilon = 0$. El procedimiento de aproximación empieza sustituyendo en (3.37) en series de Taylor para $H(x, \epsilon)$, es decir:

$$H(x, \epsilon) = H_0(x) + \epsilon H_1(x) + \epsilon^2 H_2(x) + \dots$$

y para calcular $H_0(x)$, $H_1(x)$, y así por el estilo, para términos de la ecuación para potencias de ϵ . Esto requiere que las funciones σ y ϕ sean continuamente diferenciables en

sus argumentos en un número suficiente de veces. Es claro que $H_0(x) = H(x, 0) = h(x)$.

La ecuación para $H_1(x)$ es:

$$\frac{\partial \phi}{\partial z}(x, h(x)) H_1(x) = \frac{\partial h}{\partial x} \sigma(x, h(x))$$

y tendrá una única solución si el jacobiano $\left[\frac{\partial \phi}{\partial z}\right]$ en $z = h(x)$ es no singular.

La no singularidad del jacobiano está implicada por la condición del valor propio de la condición (3.24).

Similar a H_1 , las ecuaciones para términos de alto orden serán lineales y solubles si el jacobiano $\left[\frac{\partial \phi}{\partial z}\right]$ es no singular.

Para introducir la noción de un múltiple rápido, examinando (3.32) y (3.33) en la escala de tiempo $\tau = \frac{t}{\epsilon}$. En $\epsilon = 0$, $x(\tau) \equiv x(0)$, mientras $z(\tau)$ se desarrolla de acuerdo a:

$$\frac{dz}{d\tau} = \phi(x(0), z)$$

aproximándose al punto de equilibrio $z = h(x(0))$. Este movimiento describe trayectorias (x, z) en \mathbb{R}^{n+m} , que, para todo $x(0)$ dado, enlaza un múltiple rápido F_x definido por $x = x(0) = \text{constante}$ y rápidamente descendente para el múltiple $z = h(x)$. Para ϵ más grande que cero, pero pequeño, los múltiples rápidos son foliaciones de soluciones aproximándose rápidamente al múltiple lento.

Capítulo 4

Estabilidad de redes neuronales con diferentes escalas de tiempo

Las dinámicas de una clase grande de sistemas no lineales son descritos por ecuaciones diferenciales no lineales con restricciones algebraicas. Por lo tanto, el modelo es de la forma llamado perturbado singularmente. Si las restricciones algebraicas pueden ser resueltas para las variables de estado desconocidas, las dinámicas pueden ser convertidas en ecuaciones de formato normal.

Una diversa clase de sistemas físicos en ingeniería contienen ambos fenómenos dinámicos lentos y rápidos que ocurren en escalas de tiempo separadas. Las redes neuronales artificiales tienen muchas aplicaciones en control, procesamiento de señales, reconocimiento de patrones, procesamiento de imágenes, asociación, etc. Algunas de estas aplicaciones requieren que los puntos de equilibrio de la red neuronal diseñada sean estables [37]. De tal manera, que es importante estudiar la estabilidad de redes neuronales dinámicas artificiales. Las redes neuronales diferenciales [66] con diferentes escalas de tiempo pueden modelar las dinámicas de ambos niveles de actividad neuronal, la memoria de término corto (STM), y las dinámicas de modificaciones sinápticas no supervisadas, la memoria de término largo (LTM). Su capacidad de almacenamiento de patrones deseados como puntos de equilibrio estables requieren los criterios de estabilidad que incluyen la mutua interferencia entre las neuronas

y las dinámicas de aprendizaje.

Las dinámicas de redes neuronales diferenciales con diferentes escalas de tiempo pueden ser extremadamente complejas, exhibiendo convergencia para puntos atractores y atractores periódicos [2].

4.1. Pasividad de redes neuronales de escalas de dos tiempos

Una red neuronal dinámica general con dos escalas de tiempo puede ser expresada como:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(V_1[x, z]^T) + W_3\phi_1(V_3[x, z]^T)u \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(V_2[x, z]^T) + W_4\phi_2(V_4[x, z]^T)u\end{aligned}\tag{4.1}$$

donde $x \in R^n$ y $z \in R^n$ son los estados lento y rápido, $W_i \in R^{n \times 2n}$ ($i = 1 \dots 4$) son los pesos de las capas de salida, $V_i \in R^{2n \times 2n}$ ($i = 1 \dots 4$) son los pesos de las capas ocultas, $\sigma_k = [\sigma_k(x_1) \dots \sigma_k(x_n), \sigma_k(z_1) \dots \sigma_k(z_n)]^T \in R^{2n}$ ($k = 1, 2$), $\phi(\cdot) \in R^{2n \times 2n}$ es una matriz diagonal, $\phi_k(x, z) = \text{diag}[\phi_k(x_1) \dots \phi_k(x_n), \phi_k(z_1) \dots \phi_k(z_n)]$ ($k = 1, 2$), $u(k) = [u_1, u_2 \dots u_m, 0, \dots, 0]^T \in R^{2n}$. $A \in R^{n \times n}$, $B \in R^{n \times n}$ son matrices de estabilidad (Hurwitz). El parámetro ϵ es una pequeña constante positiva. La estructura de las redes neuronales (4.1) se muestra en la figura (3.4). Cuando $\epsilon = 0$, la red neuronal dinámica (4.1) que ya ha sido discutida por muchos autores, por ejemplo [61], [66], [72] y [102]. Se puede observar que el modelo de Hopfield [31], es un caso especial de este tipo de redes neuronales con: $A = \text{diag}\{a_i\}$, $a_i := -1/R_i C_i$, $R_i > 0$ y $C_i > 0$. R_i y C_i son la resistencia y capacitancia del i -ésimo nodo de la red respectivamente. La sub-estructura $W_1\sigma_1(V_1[x, z]^T) + W_3\phi_1(V_3[x, z]^T)u$ es una estructura de perceptrón multicapa.

Entonces se quiere estudiar si es posible utilizar la teoría de pasividad, para analizar la estabilidad de sistemas perturbados singularmente con redes neuronales recurrentes dinámicas de la siguiente forma:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(V_1z) + W_2\phi_1(V_2x)u \\ \epsilon\dot{z} &= Bz + W_3\sigma_2(V_3x) + W_4\phi_2(V_4z)u\end{aligned}\quad (4.2)$$

donde: $x \in \mathbb{R}^n, z \in \mathbb{R}^m$, la entrada $u \in \mathbb{R}^m$, A, B son matrices de estabilidad, generalmente Hurwitz; W_i y V_i son las matrices de pesos de dimensiones adecuadas, σ_i, ϕ_i son funciones no lineales suaves y continuas y ϵ es el parámetro positivo que indica la dinámica rápida del sistema. Por la teoría básica de teoría de sistemas perturbados singularmente se tiene que posicionar $\epsilon = 0$ para reducir el sistema completo al sistema reducido siguiente:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(V_1z) + W_2\phi_1(V_2x)u \\ 0 &= Bz + W_3\sigma_2(V_3x) + W_4\phi_2(V_4z)u \\ z &= -(W_3\sigma_2(V_3x) + W_4\phi_2(V_4z)u)B^{-1}\end{aligned}\quad (4.3)$$

entonces:

$$\dot{x} = Ax + W_1\sigma_1(V_1[-(W_3\sigma_2(V_3x) + W_4\phi_2(V_4z)u)B^{-1}]) + W_2\phi_1(V_2x)u \quad (4.4)$$

que es el sistema de orden reducido ó dinámica lenta del sistema completo, mientras que el sistema con nivel o capa frontera es:

$$\frac{dz}{dt} = Bz + W_3\sigma_2(V_3x) + W_4\phi_2(V_4z)u \quad (4.5)$$

Se pretende en esta sección explotar el concepto de pasividad en el análisis de estabilidad de sistemas interconectados o sistemas con diferentes escalas de tiempo, que son una consecuencia de sistemas perturbados singularmente. Para simplificar ciertas pruebas, se asume sin pérdida de generalidad que los sistemas están inicialmente relajados o en reposo. Y es importante estudiar las propiedades de combinación de sistemas pasivos.

Para simplificar el análisis teórico, se consideran las capas ocultas como: $V_i = I$. Analizando una red neuronal de capa sencilla perturbada singularmente de la manera siguiente:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(x, z) + W_3\phi_1(x, z)u \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(x, z) + W_4\phi_2(x, z)u\end{aligned}\quad (4.6)$$

Porque las matrices A y B son Hurwitz, existen matrices definidas positivas Q_1 y $Q_2 \in \mathfrak{R}^{n \times n}$ tal que las siguientes funciones de Lyapunov tienen soluciones definidas positivas P_1 y P_2

$$\begin{aligned} P_1 A + A^T P_1 &= -Q_1 \\ P_2 B + B^T P_2 &= -Q_2 \end{aligned} \quad (4.7)$$

Teorema 4.1 *Si los pesos $W_{1,t}$ y $W_{2,t}$ de las redes neuronales dinámicas definidas por (4.6) están actualizadas como:*

$$\begin{aligned} \dot{W}_1 &= -P_1 x \sigma_1^T + \frac{1}{2} x u^T \\ \dot{W}_2 &= -\frac{1}{\epsilon} P_2 z \sigma_2^T + \frac{1}{2} z u^T \\ \dot{W}_3 &= -P_1 x (\phi_1 u)^T + \frac{1}{2} x u^T \\ \dot{W}_4 &= -\frac{1}{\epsilon} P_2 z (\phi_2 u)^T + \frac{1}{2} z u^T \end{aligned} \quad (4.8)$$

donde P_1 y P_2 son la solución de (4.1), entonces la red neuronal dinámica (4.6) es estrictamente pasiva de la entrada u a la salida $y \in \mathbb{R}^{2n}$

$$y = x^T (W_1 + W_3) + z^T (W_2 + W_4) \quad (4.9)$$

Demostración. Seleccionando una función de Lyapunov S_t (o llamada función de almacenamiento) [70] como:

$$S_t(x, z) = x^T P_1 x + z^T P_2 z + tr \{ W_1^T W_1 \} + tr \{ W_2^T W_2 \} + tr \{ W_3^T W_3 \} + tr \{ W_4^T W_4 \} \quad (4.10)$$

donde $P \in \mathfrak{R}^{n \times n}$ es una matriz definida positiva, $tr \{ \cdot \}$ significa la traza de una matriz y esta definida como la suma de todos los elementos diagonales de la matriz.

De acuerdo a esto (4.10), su derivada es:

$$\begin{aligned} \dot{S}_t(x, z) &= x^T (P_1 A + A^T P_1) x + \frac{1}{\epsilon} z^T (P_2 B + B^T P_2) z \\ &+ 2x^T P_1 W_1 \sigma_1 + \frac{2}{\epsilon} z^T P_2 W_2 \sigma_2 + 2x^T P_1 W_3 \phi_1 u + \frac{2}{\epsilon} z^T P_2 W_4 \phi_2 u \\ &+ 2tr \left\{ \dot{W}_1^T W_1 \right\} + 2tr \left\{ \dot{W}_2^T W_2 \right\} + 2tr \left\{ \dot{W}_3^T W_3 \right\} + 2tr \left\{ \dot{W}_4^T W_4 \right\} \end{aligned}$$

adicionando y sustrayendo $x^T (W_1 + W_3) u + z^T (W_2 + W_4) u$ y usando (4.7), se obtiene:

$$\begin{aligned} \dot{S}_t(x, z) &= -x^T Q_1 x - \frac{1}{\epsilon} z^T Q_2 z + [x^T (W_1 + W_3) + z^T (W_2 + W_4)] u \\ &+ 2tr \left\{ \dot{W}_1^T W_1 \right\} + 2x^T P_1 W_1 \sigma_1 - x^T W_1 u \\ &+ 2tr \left\{ \dot{W}_2^T W_2 \right\} + \frac{2}{\epsilon} z^T P_2 W_2 \sigma_2 - z^T W_2 u \\ &+ 2tr \left\{ \dot{W}_3^T W_3 \right\} + 2x^T P_1 W_3 \phi_1 u - x^T W_3 u \\ &+ 2tr \left\{ \dot{W}_4^T W_4 \right\} + \frac{2}{\epsilon} z^T P_2 W_4 \phi_2 u - z^T W_3 u \end{aligned}$$

Utilizando la ley actualizada como (4.8), se obtiene:

$$\dot{S}_t(x, z) = -x^T Q_1 x - \frac{1}{\epsilon} z^T Q_2 z + [x^T (W_1 + W_3) + z^T (W_2 + W_4)] u \quad (4.11)$$

De la *definición 1* se observa que si la entrada esta definida como u y la salida como $[x^T (W_1 + W_3) + z^T (W_2 + W_4)]$, entonces la red neuronal dinámica dada por (4.6) es estrictamente pasiva con:

$$V_t = x^T Q_1 x + \frac{1}{\epsilon} z^T Q_2 z \geq 0$$

■

Comentario 4.1 Cuando existen capas ocultas V_i , de tal manera que $\sigma(\cdot)$ y $\phi(\cdot)$ estan acotadas, la propiedad de pasividad no está relacionada con V_i . Los pesos de la capa oculta pueden ser fijos. Lo que es más, se puede también concluir que las propiedades de estabilidad de las redes neuronales dinámicas (4.6) no están influenciadas por las capas ocultas.

Corolario 4.1 Si las redes neuronales dinámicas (4.6) no son forzadas ($u = 0$), la ley de actualización (4.8) hará que los puntos de equilibrio $x = z = 0$ sean estables.

Demostración. Puesto que el sistema dinámico (4.6) es pasivo, la función de almacenamiento $S(x_t)$ satisface (la propiedad 1)

$$\dot{S}_t(x, z) \leq u^T y = 0$$

Porque $S_t(x, z)$ es definida positiva, el equilibrio $x = z = 0$ es estable. ■

Corolario 4.2 *Si la entrada de control de las redes neuronales dinámicas (4.6) son seleccionadas como:*

$$u = -\mu y = -\mu [x^T (W_1 + W_3) + z^T (W_2 + W_4)], \quad \mu > 0 \quad (4.12)$$

la ley actualizada (4.8) hará el equilibrio $x_t = 0$ asintóticamente estable.

Demostración. Porque $y = h(x)$ es independiente de u , el lazo de retroalimentación con $u = -\mu y$ esta bien presentado. Para $u = -\mu y$, la derivada en el tiempo de S satisface:

$$\dot{S}_t(x, z) \leq -\mu y^T y \leq 0$$

Por lo tanto equilibrio $x = z = 0$ es estable. Basado en el principio de invariancia ver Sección 2.3 de [80], las soluciones acotadas de $\dot{x} = f(x, z, -y)$ y $\epsilon \dot{z} = g(x, z, -y)$ convergen al conjunto total invariante de $\dot{x} = f(x, z, 0)$ y $\dot{z} = g(x, z, 0)$. Ya que esta contenido en $E_1 = \{x \mid h_1(x) = 0\}$ y $E_2 = \{z \mid h_2(z) = 0\}$, este conjunto es $x = 0$ y $z = 0$, entonces el equilibrio $x = z = 0$ es asintóticamente estable. ■

Teorema 4.2 *Si las cotas superiores de los pesos W_i ($i = 1 \dots 4$) satisfacen*

$$\begin{aligned} \lambda_{\min}(Q_1) &\geq \lambda_{\max}(\overline{W}_1 + \overline{W}_3) \\ \frac{\lambda_{\min}(Q_2)}{\epsilon} &\geq \lambda_{\max}(\overline{W}_2 + \overline{W}_4) \end{aligned} \quad (4.13)$$

donde $\overline{W}_i = \sup \|W_i\|_{\Lambda_i^{-1}}^2$, la norma de la matriz esta definida como $\|A\|_B^2 := A^T B A$, Λ_i es una matriz definida positiva, $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ son el máximo y mínimo de los valores propios de la matriz. La ley actualizada (4.8) hará que las redes neuronales dinámicas (4.6) sea estable de entrada estado (ISS).

Demostración. En vista de que la desigualdad matricial:

$$X^T Y + (X^T Y)^T \leq X^T \Lambda^{-1} X + Y^T \Lambda Y \quad (4.14)$$

que es válida para cualquier $X, Y \in \mathfrak{R}^{n \times k}$ y para cualquier matriz definida positiva $0 < \Lambda = \Lambda^T \in \mathfrak{R}^{n \times n}$, (4.14) puede ser representada como:

$$\begin{aligned} \dot{S}_t(x, z) &= -x^T Q_1 x - \frac{1}{\epsilon} z^T Q_2 z + [x^T (W_1 + W_3) + z^T (W_2 + W_4)] u \\ &\leq -\lambda_{\min}(Q_1) \|x\|^2 - \frac{\lambda_{\min}(Q_2)}{\epsilon} \|z\|^2 + u^T (\Lambda_1 + \Lambda_2 + \Lambda_3 + \Lambda_4) u_t \\ &\quad + x^T (W_1 \Lambda_1^{-1} W_1^T + W_3 \Lambda_3^{-1} W_3^T) x + z^T (W_2 \Lambda_2^{-1} W_2^T + W_4 \Lambda_4^{-1} W_4^T) z \\ &\leq -\lambda_{\min}(Q_1) \|x\|^2 - \frac{\lambda_{\min}(Q_2)}{\epsilon} \|z\|^2 + u^T (\Lambda_1 + \Lambda_2 + \Lambda_3 + \Lambda_4) u_t \\ &\quad + x^T (\overline{W}_1 + \overline{W}_3) x + z^T (\overline{W}_2 + \overline{W}_4) z \\ &\leq -\alpha_1 \|x\| - \alpha_2 \|z\| + \beta \|u_t\| \end{aligned}$$

donde $\alpha_1 = [\lambda_{\min}(Q_1) - \lambda_{\max}(\overline{W}_1 + \overline{W}_3)] \|x\|$, $\alpha_2 = [\frac{\lambda_{\min}(Q_2)}{\epsilon} - \lambda_{\max}(\overline{W}_2 + \overline{W}_4)] \|z\|$, $\beta_{\|u_t\|} = \lambda_{\max}(\Lambda_1 + \Lambda_2 + \Lambda_3 + \Lambda_4) \|u_t\|$. Si la condición (4.13) se satisface, α_1 , α_2 y β son \mathcal{K}_∞ funciones, S_t es una función de ISS-Lyapunov. Usando el *teorema 1* de [85], las redes neuronales dinámicas (4.6) son estables de entrada estado (ISS). ■

Comentario 4.2 *Considerar la Propiedad 2; ISS significa que el comportamiento de las redes neuronales dinámicas permanecerá acotado cuando sus entradas son acotadas. Si perturbaciones o disturbios acotados son también considerados como entradas, las redes neuronales son también acotadas con respecto a las perturbaciones. Por lo tanto las redes neuronales dinámicas (4.6) son estables acotadas a la entrada-acotadas a la salida (BIBO)*

Comentario 4.3 *Si P_1 y P_2 son seleccionadas suficientemente grandes, entonces la condición (4.13) no es difícil a ser satisfecha. De (4.7) se tiene*

$$\begin{aligned} T_1 P_1 A + T A^T P_1 &= -T_1 Q \\ T_2 P_2 B + T B^T P_2 &= -T_2 Q \end{aligned}, \quad T_i = T_i^T > 0 \quad (4.15)$$

Si la matriz definida positiva T_i es suficientemente grande, (4.15) se satisface, i.e. (ϵ es una constante positiva pequeña)

$$\begin{aligned}\lambda_{\max}(\overline{W}_1 + \overline{W}_3) &\leq \lambda_{\min}(T_1 Q_1) \\ \lambda_{\max}(\overline{W}_2 + \overline{W}_4) &\leq \frac{\lambda_{\min}(T_2 Q_2)}{\epsilon}\end{aligned}\tag{4.16}$$

4.2. Estabilidad exponencial

La estabilidad exponencial es un t3pico bastante importante, porque establece las cualidades de robustez de un sistema, ante perturbaciones externas. Muy a menudo en el an3lisis de sistemas din3micos, se usan modelos de orden reducido, obtenidos por la supresi3n de peque1os par3metros par3sitos. Esta reducci3n en el orden del modelo puede ser representado como un problema de perturbaci3n singular, donde el modelo perturbado singularmente completo representa el sistema actual con los par3metros par3sitos y el modelo reducido es el modelo simplificado usado en el an3lisis. Es totalmente razonable asumir que el modelo con nivel o capa frontera es exponencialmente estable en el origen. En efecto, si las din3micas asociadas con los elementos par3sitos fueran inestables, no tendr3an que ser evitadas en primer lugar. Los tecnicismos de asumir estabilidad exponencial en lugar de s3lo estabilidad asint3tica, o asumir que la estabilidad exponencial se cumple uniformemente, es totalmente razonable en la mayor3a de los casos.

Consid3rese un caso simple de (4.6), con capa sencilla $V_1 = V_2 = I$

$$\begin{aligned}\dot{x} &= Ax + W_1 \sigma_1(z) + W_3 \phi_1(x) u \\ \epsilon \dot{z} &= Bz + W_2 \sigma_2(x) + W_4 \phi_2(z) u\end{aligned}\tag{4.17}$$

entonces se tiene el siguiente teorema:

Teorema 4.3 *Consid3rese la red neuronal recurrente modelada en forma de perturbaci3n singular estandar (4.17), se asume que las siguientes consideraciones se satisfacen para todo:*

$$(t, x, \epsilon) \in [0, \infty) \times B_r \times [0, \epsilon_0]\tag{4.18}$$

Donde $t \in \geq 0$, x esta acotada ($x \in B_r$, B_r es una esfera de radio r), ϵ es una constante positiva, las funciones no lineales

$\sigma_k(\cdot)$ y $\phi_k(\cdot)$ son funciones sigmoidales.

a) $\sigma(t, 0, 0, \epsilon) = 0$ y $\phi(t, 0, 0, \epsilon) = 0$

b) La ecuación o sistema reducido posicionando $\epsilon = 0$ es:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(z) + W_3\phi_1(x)u \\ 0 &= Bz + W_2\sigma_2(x) + W_4\phi_2(z)u\end{aligned}\tag{4.19}$$

tiene una raíz aislada única: $z = h(x)$ tal que $h(0) = 0$.

c) Las funciones $\sigma_k(\cdot)$, $\phi_k(\cdot)$ y $h(x)$, y sus derivadas parciales hasta segundo orden están acotadas por: $z - h(x) \in B_\rho$.

d) El origen del sistema reducido es exponencialmente estable:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(h(x)) + W_3\phi_1(x)u \\ \dot{x} &= Ax + W_1\sigma_1(-[W_2\sigma_2(x) + W_4\phi_2(z)u]B^{-1}) + W_3\phi_1(x)u\end{aligned}\tag{4.20}$$

e) Si se define $y = z - h(x)$, $\tau = \frac{t}{\epsilon}$ entonces el sistema con nivel o capa frontera:

$$\frac{dy}{d\tau} = B(h(x) + y) + W_2\sigma_2(x) + W_4\phi_2(h(x) + y)u\tag{4.21}$$

es exponencialmente estable, uniforme en (x) , entonces existe $\epsilon^* > 0$ tal que para todo $\epsilon < \epsilon^*$, el origen de (4.20) es exponencialmente estable.

Demostración. De los resultados de [40], si las condiciones de este teorema se satisfacen (a) – (d), entonces existe una función de Lyapunov $V(t, x)$ tal que

$$\begin{aligned}c_1 \|x\|^2 &\leq V(t, x) \leq c_2 \|x\|^2 \\ \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} \sigma(t, x, h(t, x), 0) &\leq -c_3 \|x\|^2 \\ \left\| \frac{\partial V}{\partial x} \right\| &\leq c_4 \|x\|\end{aligned}$$

donde $c_i, i = 1, \dots, 4$, son constantes positivas. Para $x \in B_{r_0}$, $r_0 \leq r$, existe otra función de

Lyapunov $Q(t, x, y)$ para el sistema con nivel en la frontera tal que

$$\begin{aligned} b_1 \|y\|^2 &\leq Q(t, x, y) \leq b_2 \|y\|^2 \\ \frac{\partial Q}{\partial y} \phi(t, x, y + h(t, x), 0) &\leq -b_3 \|y\|^2 \\ \left\| \frac{\partial Q}{\partial y} \right\| &\leq b_4 \|y\| \\ \left\| \frac{\partial Q}{\partial t} \right\| &\leq b_5 \|y\|^2; \quad \left\| \frac{\partial Q}{\partial x} \right\| \leq b_6 \|y\|^2 \end{aligned}$$

donde b_i , $i = 1, \dots, 6$ son constantes positivas. Para $y \in B_{\rho_0}$, $\rho_0 \leq \rho$, aplicando un cambio de variables:

$$y = z - h(t, x)$$

se obtiene

$$\begin{aligned} \dot{x} &= Ax + W_1 \sigma_1(-W_2 \sigma_2(V_2, x) B^{-1}) \\ \epsilon \dot{y} &= [B(y + h(t, x)) + W_2 \sigma_2(V_2, x)] \\ &\quad - \epsilon \frac{\partial h}{\partial t} - \epsilon \frac{\partial h}{\partial t} [Ax + W_1 \sigma_1(-W_2 \sigma_2(V_2, x) B^{-1})] \end{aligned}$$

Ahora definiendo:

$$v(t, x, y) = V(t, x) + Q(t, x, y)$$

Como una candidato a función de Lyapunov. Debido a que σ_1 y σ_2 se desvanecen en el origen para todo $\epsilon \in [0, \epsilon_0]$, entonces son Lipschitz en ϵ en el estado (x, y) , en particular:

$$\begin{aligned} &\|(Ax + W_1 \sigma_1(V_1, z), \epsilon) - (Ax + W_1 \sigma_1(V_1, z))\| \\ &\quad \leq \epsilon L_1 (\|x\| + \|y\|) \\ &\|(Bz^* + W_2 \sigma_2(V_2, x), \epsilon) - (Bz^* + W_2 \sigma_2(V_2, x))\| \\ &\quad \leq \epsilon L_2 (\|x\| + \|y\|) \end{aligned}$$

también,

$$\begin{aligned} &\left\| \begin{array}{l} (Ax + W_1 \sigma_1(V_1, z), 0) \\ -(Ax + W_1 \sigma_1(V_1, z), 0) \end{array} \right\| \leq L_3 \|y\| \\ &\|(Ax + W_1 \sigma_1(V_1, z), 0)\| \leq L_4 \|x\| \\ &\left\| \frac{\partial h}{\partial t} \right\| \leq k_1 \|x\|; \quad \left\| \frac{\partial h}{\partial t} \right\| \leq k_2 \end{aligned}$$

Aquí utilizando el hecho de que $(Ax + W_1\sigma_1(V_1z), 0)$ y $h(t, x)$ se desvanecen en el origen en $x = 0$ para todo t . Usando estas estimaciones y las propiedades de las funciones V y Q , es posible verificar que la derivada de v a lo largo de las trayectorias se satisface la desigualdad:

$$\begin{aligned} \dot{v} \leq & -a_1 \|x\|^2 + \epsilon a_2 \|x\|^2 - \frac{a_3}{\epsilon} \|y\|^2 + a_4 \|y\|^2 \\ & + a_5 \|x\| \|y\| + a_6 \|x\| \|y\|^2 + a_7 \|y\|^3 \end{aligned}$$

con valores positivos a_1 y a_3 y valores no negativos a_2 y a_4 para a_7 .

De tal manera que para todo $\|y\| \leq \rho_0$, esta desigualdad se simplifica para:

$$\begin{aligned} \dot{v} \leq & -a_1 \|x\|^2 + \epsilon a_2 \|x\|^2 - \frac{a_3}{\epsilon} \|y\|^2 + a_8 \|y\|^2 + 2a_9 \|x\| \\ = & - \begin{bmatrix} \|x\| \\ \|y\| \end{bmatrix}^T \begin{bmatrix} a_1 - \epsilon a_2 & -a_9 \\ -a_9 & \left(\frac{a_3}{\epsilon}\right) - a_8 \end{bmatrix} \begin{bmatrix} \|x\| \\ \|y\| \end{bmatrix} \end{aligned}$$

Por lo tanto, existe $\epsilon^* > 0$ tal que para todo $0 < \epsilon < \epsilon^*$, se obtiene:

$$\dot{v} \leq -2\gamma v$$

para algún $\gamma > 0$. Entonces se deduce:

$$v(t, x(t), y(t)) \leq \exp[-2\gamma(t - t_0)] v(t_0, x(t_0), y(t_0))$$

y, de las propiedades de V y Q , obtenemos:

$$\begin{aligned} \left\| \begin{array}{l} x(t) \\ y(t) \end{array} \right\| & \leq K_1 \exp[-\gamma(t - t_0)] \left\| \begin{array}{l} x(t_0) \\ y(t_0) \end{array} \right\| \\ \left\| \begin{array}{l} x(t) \\ z(t) \end{array} \right\| & \leq K_2 \exp[-\gamma(t - t_0)] \left\| \begin{array}{l} x(t_0) \\ z(t_0) \end{array} \right\| \end{aligned}$$

aquí $\|h(t, x)\| \leq k_2 \|x\|$, que completa la prueba del teorema. ■

Comentario 4.4 *El teorema anterior es conceptualmente importante porque establece la robustez de estabilidad exponencial para dinámicas rápidas no modeladas (alta frecuencia). Muy a menudo en el análisis de sistemas dinámicos, se usarán modelos de orden reducido*

obtenidos por la omisión de parámetros parásitos pequeños. Esta reducción en el orden del modelo puede ser representado como un problema de perturbación singular, donde el modelo perturbado singularmente completo representa el sistema actual con los parámetros parásitos y el modelo reducido es el modelo simplificado usado en el análisis. Es totalmente razonable asumir que el modelo con nivel en la frontera tendrá un origen exponencialmente estable. En efecto, si las dinámicas asociadas con los elementos parásitos son inestables, no se tienen que evitar en el primer lugar. El tecnicismo de asumir estabilidad exponencial en lugar de solo estabilidad asintótica, o asumir que la estabilidad exponencial se mantiene uniforme, es totalmente considerado en la mayoría de las aplicaciones. Es suficiente mencionar que todas estos tecnicismos se mantendrán automáticamente cuando las dinámicas rápidas sean lineales. Cuando el origen del modelo reducido es exponencialmente estable, el teorema anterior no asegura que el origen del sistema actual será exponencialmente estable, con tal de que las dinámicas rápidas evitables sean suficientemente rápidas.

Comentario 4.5 Las condiciones para estabilidad exponencial (a) – (d) de este teorema son bastante difíciles de probar, de tal modo que solo se presentan las condiciones de estabilidad robusta y asintótica para redes neuronales diferenciales con diferentes escalas de tiempo. Entonces ejemplos posteriores ilustrarán cómo estas propiedades de robustez surgen en diseño de control.

4.3. Estabilidad robusta

Esta parte está enfocada con el problema de estabilidad robusta de sistemas con escalas de dos tiempos con incertidumbres o perturbaciones singulares, donde solo la información disponible para sistemas con incertidumbres son sus propias cotas superiores. La cota de estabilidad ϵ^* es deducida vía una transformación de estado y el uso constructivo de una función de Lyapunov.

Considerar un caso simple de (3.9) y (3.10), con entrada $u = 0$, con capa sencilla $V_1 = V_2 = I$

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(x, z) \\ \epsilon\dot{z} &= Bz + W_2\sigma_2(x, z)\end{aligned}\tag{4.22}$$

Porque $t \geq 0$, x esta acotada ($x \in B_r$, B_r es una esfera con radio r), ϵ es una constante positiva acotada, de tal manera que:

$$(t, x, \epsilon) \in [0, \infty) \times B_r \times [0, \epsilon_0]$$

$\sigma_k(x, z)$ y ϕ_k son funciones sigmoidales, sus derivadas parciales hasta segundo orden están acotadas. Un sistema reducido esta definido posicionando $\epsilon = 0$ en (4.22) para obtener:

$$\begin{aligned}\dot{x} &= Ax + W_1\sigma_1(x, z) \\ 0 &= Bz + W_2\sigma_2(x, z)\end{aligned}\tag{4.23}$$

Se asume que:

$$0 = Bz + W_2\sigma_2(x, z)$$

tiene una raíz única $z = h(x)$.

Si se define $y = z - h(x)$, $\tau = \frac{t}{\epsilon}$, sustituyendo en (4.23) se obtiene un sistema nivel o capa frontera como:

$$\frac{dy}{d\tau} = B[h(x) + y] + W_2\sigma_2(x, h(x) + y)\tag{4.24}$$

Adicionando y sustrayendo Cz y Dx en (4.22) se obtiene:

$$\begin{aligned}\dot{x} &= Ax + Cz - Cz + W_1\sigma_1(x, z) \\ \epsilon\dot{z} &= Bz + Dx - Dx + W_2\sigma_2(x, z)\end{aligned}\tag{4.25}$$

donde $C \in R^{n \times n}$, $D \in R^{n \times n}$ son matrices de dimensiones apropiadas. Entonces reescribiendo (4.25) como

$$\begin{aligned}\dot{x} &= Ax + Cz + f_1 \\ \epsilon\dot{z} &= Bz + Dx + f_2\end{aligned}\tag{4.26}$$

donde: $f_1 = W_1\sigma_1(x, z) - Cz$, $f_2 = W_2\sigma_2(x, z) - Dx$. Por la razón de que $\sigma_1(x, z)$ y $\sigma_2(x, z)$ son funciones sigmoideas, además W_1 y W_2 son acotadas, se obtiene que:

$$\begin{aligned}\|f_1\| &\leq \alpha_1 \|x\| + \beta_1 \|z\| \\ \|f_2\| &\leq \alpha_2 \|x\| + \beta_2 \|z\|\end{aligned}$$

donde α_i y β_i son constantes positivas, $i = 1, 2$. Primero se trata y discute el caso lineal, es decir, $f_1 = f_2 = 0$, (4.26) para obtener:

$$\begin{aligned}\dot{x} &= Ax + Cz \\ \epsilon \dot{z} &= Bz + Dx\end{aligned}\tag{4.27}$$

Entonces posicionando $\epsilon = 0$, se obtiene el sistema reducido en orden ó dinámica lenta del sistema lineal como:

$$\dot{x} = (A - CB^{-1}D)x\tag{4.28}$$

El problema de estabilidad de un sistema perturbado singularmente es: si el sistema de orden reducido (4.25) es estable (si $A_0 = A - CB^{-1}D$ es Hurwitz), esto implica, que el sistema completo (4.26) es también estable?. El siguiente lemma indica las condiciones bajo las cuales el sistema completo (4.26) es también asintóticamente estable para ϵ suficientemente pequeño.

Lemma 4.1 *Si B y A_0 son Hurwitz, entonces existe un ϵ^* tal que para todo $\epsilon \in [0, \epsilon^*)$, el sistema (4.26) es asintóticamente estable.*

La prueba puede ser encontrada en [40]. Para la red neuronal diferencial (4.25), el objetivo es encontrar cotas superiores de α_i, β_i y ϵ^* , tal que el sistema (4.26) sea asintóticamente estable, probando que D y A_0 son Hurwitz. Pero es necesario definir la siguiente transformación de estado

$$\begin{bmatrix} x \\ z \end{bmatrix} = T \begin{bmatrix} \theta \\ \eta \end{bmatrix}\tag{4.29}$$

donde $T = \begin{bmatrix} I & \epsilon H \\ -L & I - \epsilon LH \end{bmatrix}$, $H = AB^{-1}$ y $L = B^{-1}D$. Observar que T es no singular si B^{-1} existe. Por (4.29), se obtiene

$$\begin{bmatrix} \theta \\ \eta \end{bmatrix} = T^{-1} \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} I - \epsilon LH & -\epsilon H \\ L & I \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}$$

y

$$\begin{aligned} \dot{\theta} &= [A_0 + \alpha_1] \theta + \alpha_2 \eta + \bar{f}_1 \\ \epsilon \dot{\eta} &= \alpha_3 \theta + [W_2 + \alpha_4] \eta + \bar{f}_2 \end{aligned} \quad (4.30)$$

donde:

$$\begin{aligned} \alpha_1 &= -\epsilon H L A_0 \\ \alpha_2 &= \epsilon (A_0 H - H L W_1) \\ \alpha_3 &= \epsilon L A_0 \\ \alpha_4 &= W_2 + \epsilon L W_1 + \epsilon^2 H L A_0 H \\ \bar{f}_1 &= (I - \epsilon L H) f_1 [\theta + \epsilon H \eta, -L\theta + (I - \epsilon L H) \eta, t] \\ &\quad - H f_2 [\theta + \epsilon H \eta, -L\theta + (I - \epsilon L H) \eta, t] \\ \bar{f}_2 &= \epsilon L f_1 [\theta + \epsilon H \eta, -L\theta + (I - \epsilon L H) \eta, t] \\ &\quad - f_2 [\theta + \epsilon H \eta, -L\theta + (I - \epsilon L H) \eta, t] \end{aligned} \quad (4.31)$$

Observando que:

$$\begin{aligned} &\|f_i [\theta + \epsilon H \eta, -L\theta + (I - \epsilon L H) \eta, t]\| \\ &\leq \alpha_i [\|\theta\| + \epsilon \|H\| \|\eta\|] + \beta_i [\|L\| \|\theta\| + (1 + \epsilon \|LH\|) \|\eta\|] \end{aligned}$$

donde $i = 1, 2$. Por lo tanto, se puede mostrar que:

$$\|f_i\| \leq \bar{\alpha}_i \|\theta\| + \bar{\beta}_i \|\eta\|$$

Como la estabilidad de (4.30) implica que dé (4.27), entonces regresando a estudiar (4.30) con norma de incertidumbres acotadas. Entonces se define:

$$V[\theta, \eta] = \theta^T P \theta + \epsilon \eta^T S \eta$$

como la candidato a función de Lyapunov, donde P y S son las soluciones de las siguientes ecuaciones de Lyapunov:

$$\begin{aligned} A_0^T P + P A_0 &= -2Q_P \\ B^T S + S B &= -2Q_S \end{aligned} \quad (4.32)$$

donde Q_P y Q_S son cualesquiera matrices tal $Q_P = Q_P^T > 0$ y $Q_S = Q_S^T > 0$. Como A_0 y D son Hurwitz, existen soluciones $P = P^T > 0$ y $S = S^T > 0$. Por lo tanto, la derivada de $V[\theta, \eta]$ a lo largo de (4.30) produce

$$\begin{aligned} \dot{V}(\theta, \eta) &= \dot{\theta}^T P \theta + \theta^T P \dot{\theta} + \epsilon \dot{\eta}^T S \eta + \epsilon \eta^T S \dot{\eta} \\ &= \theta^T (A_0^T P + P A_0) \theta + \theta^T [\alpha_1^T P + P \alpha_1] \theta \\ &\quad + 2 \left[\theta^T P \alpha_2 \eta + \theta^T P \bar{f}_1 \right] \\ &\quad + \eta^T (W_2^T S + S W_2) \eta + \eta^T (\alpha_3^T S + S \alpha_3) \eta \\ &\quad + 2 \left[\theta^T \alpha_3^T S \eta + \eta^T S \bar{f}_2 \right] \\ &\leq -2\lambda_{\min}(Q_P) \|\theta\|^2 + 2\lambda_{\max}(P) \|\alpha_1\| \|\theta\|^2 \\ &\quad + 2\lambda_{\max}(P) \|\alpha_2\| \|\theta\| \|\eta\| \\ &\quad + 2\lambda_{\max}(P) \|\theta\| \left(\bar{\alpha}_1 \|\theta\| + \bar{\beta}_1 \|\eta\| \right) \\ &\quad - 2\lambda_{\min}(Q_S) \|\eta\|^2 + 2\lambda_{\max}(S) \|\alpha_3\| \|\eta\|^2 \\ &\quad + 2\lambda_{\max}(S) \|\alpha_2\| \|\theta\| \|\eta\| \\ &\quad + 2\lambda_{\max}(S) \|\eta\| \left(\bar{\alpha}_2 \|\theta\| + \bar{\beta}_2 \|\eta\| \right) \end{aligned}$$

Observar que para cualquier $\lambda > 0$,

$$2 \|\theta\| \|\eta\| \leq \gamma \|\theta\|^2 + \frac{1}{\gamma} \|\eta\|^2$$

Sustituyendo (4.31) en (4.30), se obtiene:

$$\dot{V}(\theta, \eta) \leq -a \|\theta\|^2 - b \|\eta\|^2 \quad (4.33)$$

donde

$$\begin{aligned}
a &= 2 \left[\lambda_{\min}(Q_P) - \lambda_{\max}(P) \|\alpha_{11}\| - \bar{\alpha}_1 \lambda_{\max}(P) \right] \\
&\quad - \gamma \left[\begin{array}{c} \lambda_{\max}(P) \|\alpha_{12}\| + \bar{\beta}_1 \lambda_{\max}(P) + \lambda_{\max}(S) \|\alpha_{21}\| \\ + \bar{\alpha}_2 \lambda_{\max}(S) \end{array} \right] \\
b &= 2 \left[\lambda_{\min}(Q_S) - \lambda_{\max}(S) \|\alpha_{22}\| - \bar{\beta}_2 \lambda_{\max}(S) \right] \\
&\quad - \frac{1}{\gamma} \left[\begin{array}{c} \lambda_{\max}(P) \|\alpha_{12}\| + \bar{\beta}_1 \lambda_{\max}(P) + \lambda_{\max}(S) \|\alpha_{21}\| \\ + \bar{\alpha}_2 \lambda_{\max}(S) \end{array} \right]
\end{aligned} \tag{4.34}$$

Entonces para que $\dot{V}(\theta, \eta) \leq 0$, es suficiente que $a \geq 0$ y $b \geq 0$, o de manera equivalente

$$\begin{aligned}
&2 \left[\lambda_{\max}(P) \|\alpha_{22}\| + \bar{\beta}_2 \lambda_{\max}(S) \right] \\
&+ \frac{1}{\gamma} \left[\begin{array}{c} \lambda_{\max}(P) \|\alpha_{12}\| + \bar{\beta}_1 \lambda_{\max}(P) \\ + \lambda_{\max}(S) \|\alpha_{21}\| + \bar{\alpha}_2 \lambda_{\max}(S) \end{array} \right] \\
&\leq 2\lambda_{\min}(Q_P) \\
&2 \left[\lambda_{\max}(S) \|\alpha_{11}\| + \bar{\alpha}_1 \lambda_{\max}(P) \right] \\
&+ \gamma \left[\begin{array}{c} \lambda_{\max}(P) \|\alpha_{12}\| + \bar{\beta}_1 \lambda_{\max}(P) \\ + \lambda_{\max}(S) \|\alpha_{21}\| + \bar{\alpha}_2 \lambda_{\max}(S) \end{array} \right] \\
&\leq 2\lambda_{\min}(Q_S)
\end{aligned} \tag{4.35}$$

de (4.31), se tiene que

$$\begin{aligned}
\|\alpha_{11}(\epsilon)\| &= \lambda_1 \epsilon \\
\|\alpha_{12}(\epsilon)\| &= \lambda_2 \epsilon + \lambda_3 \epsilon^2 \\
\|\alpha_{21}(\epsilon)\| &= \lambda_4 \epsilon \\
\|\alpha_{22}(\epsilon)\| &= \lambda_5 \epsilon + \lambda_6 \epsilon^2
\end{aligned} \tag{4.36}$$

donde $\lambda_1 = \|HLA_0\|$, $\lambda_2 = \|A_0H - HLW_1\|$, $\lambda_3 = \|HLA_0H\|$, $\lambda_4 = \|LA_0\|$, $\lambda_5 = \|LW_1\|$, $\lambda_6 = \|LA_0H\|$. De tal manera que

$$\begin{aligned}
(a_1 \epsilon^2 + a_2 \epsilon + a_3) &\leq 2\lambda_{\min}(Q_P) \\
(b_1 \epsilon^2 + b_2 \epsilon + b_3) &\leq 2\lambda_{\min}(Q_S)
\end{aligned} \tag{4.37}$$

donde

$$\begin{aligned}
a_1 &= \gamma (\lambda_3 + \mu_3) \lambda_{\max}(P) \\
a_2 &= [2 (\lambda_1 + \mu_1) + \gamma (\lambda_2 + \mu_4)] \lambda_{\max}(P) \\
&\quad + \gamma (\lambda_4 + \mu_6) \lambda_{\max}(S) \\
a_3 &= (2\mu_2 + \gamma\mu_5) \lambda_{\max}(P) + \gamma\mu_7 \lambda_{\max}(S) \\
b_1 &= \frac{1}{\gamma} (\lambda_3 + \mu_3) \lambda_{\max}(P) + 2 (\lambda_6 + \mu_8) \lambda_{\max}(S) \\
b_2 &= \frac{1}{\gamma} (\lambda_3 + \mu_4) \lambda_{\max}(P) + \left[\begin{array}{c} 2 (\lambda_5 + \mu_9) \\ + \frac{1}{\gamma} (\lambda_4 + \mu_6) \end{array} \right] \lambda_{\max}(S) \\
b_3 &= \frac{1}{\gamma} \mu_5 \lambda_{\max}(P) + \left(2\mu_{10} + \frac{1}{\gamma} \mu_7 \right) \lambda_{\max}(S)
\end{aligned} \tag{4.38}$$

Observar que si (4.30) se cumple, (4.26) siempre se cumple. Como es necesario encontrar $\epsilon^* > 0$, $a_3 < 2\lambda_{\min}(Q_P)$ y $b_3 < 2\lambda_{\min}(Q_S)$ se debe satisfacer lo siguiente

$$\begin{aligned}
(2\mu_2 + \gamma\mu_5) \lambda_{\max}(P) + \gamma\mu_7 \lambda_{\max}(S) &< 2\lambda_{\min}(Q_P) \\
\frac{1}{\gamma} \mu_5 \lambda_{\max}(P) + \left(2\mu_{10} + \frac{1}{\gamma} \mu_7 \right) \lambda_{\max}(S) &< 2\lambda_{\min}(Q_S)
\end{aligned} \tag{4.39}$$

Sea $\bar{a}_3 = 2\lambda_{\min}(Q_P) - a_3$, $\bar{b}_3 = 2\lambda_{\min}(Q_S) - b_3$, entonces $\bar{a}_3 > 0$, $\bar{b}_3 > 0$. Sin pérdida de generalidad, se asume que $a_1 \neq 0$. Por lo tanto, el rango aceptable de ϵ para las desigualdades $a_1\epsilon^2 + a_2\epsilon - \bar{a}_3 \leq 0$ y $b_1\epsilon^2 + b_2\epsilon - \bar{b}_3 \leq 0$ son

$$\begin{aligned}
0 \leq \epsilon \leq \frac{-a_2 + \sqrt{a_2^2 + 4a_1\bar{a}_3}}{2a_1} &= \epsilon_1^* \\
0 \leq \epsilon \leq \frac{-b_2 + \sqrt{b_2^2 + 4b_1\bar{b}_3}}{2b_1} &= \epsilon_2^*
\end{aligned} \tag{4.40}$$

Entonces ahora se plantea el siguiente teorema:

Teorema 4.4 *Elegimos los pesos W_1 y W_2 que son acotados, B es una matriz Hurwitz no singular, A , D y C son elegidas tal que $A - CB^{-1}D$ es Hurwitz, seleccionando Q_P , y Q_S tal que P y S son las soluciones de (4.34), eligiendo una γ positiva tal que la desigualdad (4.39) se satisface, entonces el sistema (4.34) es asintóticamente estable para $0 \leq \epsilon \leq \epsilon^*$, donde $\epsilon^* = \min(\epsilon_1^*, \epsilon_2^*)$ y ϵ_1^* y ϵ_2^* están dadas en (4.40).*

Comentario 4.6 *En el caso de que $a_1 = 0$, $a_2 \neq 0$, se obtiene $\epsilon_1^* = \frac{\bar{a}_3}{a_2}$, por lo tanto ϵ^* puede todavía ser obtenida o derivada de (4.40). Si $a_1 = a_2 = 0 \rightarrow \lambda_2 = \lambda_4 = 0 \rightarrow L = 0$*

y $H = 0 \rightarrow B = 0$ y $A = 0 \rightarrow b_1 = b_2 = 0$ como para cualquier matriz A , $\|A\| = 0 \rightarrow \text{Tr}(A^T A) = 0 \rightarrow A = 0$. En este caso, el sistema (4.33) es asintóticamente estable para todo $\epsilon \geq 0$ si las siguientes dos desigualdades se satisfacen

$$\begin{aligned} 2\alpha_1 \lambda_{\max}(P) + \gamma [\beta_1 \lambda_{\max}(P) + \alpha_2 \lambda_{\max}(S)] &\leq 2\lambda_{\min}(Q_P) \\ \frac{1}{\gamma} [\beta_1 \lambda_{\max}(P) + \alpha_2 \lambda_{\max}(S)] + 2\beta_1 \lambda_{\max}(S) &\leq 2\lambda_{\min}(Q_S) \end{aligned}$$

4.4. Simulaciones

4.4.1. Pasividad de redes neuronales con diferentes escalas de tiempo

Ejemplo 4.1 Para ilustrar los resultados teóricos, se presentan las siguientes simulaciones. La red neuronal dinámica es:

$$\begin{aligned} \dot{x} &= Ax + W_1 \sigma_1(x, z) + W_3 \phi_1(x, z)u \\ \dot{z} &= Bz + W_2 \sigma_2(x, z) + W_4 \phi_2(x, z)u \end{aligned} \quad (4.41)$$

donde $x \in \mathbb{R}^2$, $z \in \mathbb{R}^2$, $A = \begin{bmatrix} -5 & 0 \\ 0 & -5 \end{bmatrix}$, $B = \begin{bmatrix} -5 & 0 \\ 0 & -5 \end{bmatrix}$, $\epsilon = 0,05$. Las condiciones iniciales son: $x_0 = [1, 2]^T$, $z_0 = [1, 1]^T$, $W_{i,0} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$, $i = 1 \dots 4$. Las funciones de activación σ_k y ϕ_k son seleccionadas como $\tanh(\cdot)$. Primero, checando la pasividad de las redes neuronales (4.41). Si se selecciona $Q_1 = Q_2 = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$, la solución de la ecuación de Lyapunov (4.33) es $P_1 = P_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. La ley actualizada para los pesos W que se usa como (4.8). La Fig.(4.1)-(4.3) muestra las respuestas de las entradas acotadas ($u = [3 \sin(0,5t), 0, 0, 0]$), respuestas de entrada cero y salida retroalimentada adicional (4.13) $u = -\mu y$ ($\mu = 1$). Claramente, se puede ver que la ley de aprendizaje (4.8) hace las redes neuronales dinámicas (4.41) pasivas con entradas acotadas, la estabilidad con entrada cero y estabilidad asintótica con salida de control retroalimentada. También se encuentra que

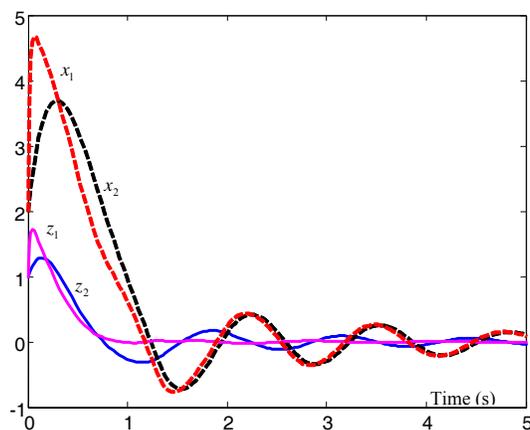


Figura 4.1: Respuesta a las entradas acotadas.

con el control retroalimentado, la propiedad de escalas multitiempo casi desaparece. Estos resultados aparecen ser de un pequeño uso para el almacenamiento de patrones, pero ellos son muy importantes para la identificación y control usando redes neuronales [33]. Segundo, checando la estabilidad de entrada estado (ISS). De (4.8) se conoce que la solución de la ecuación de Lyapunov (4.7) puede ser multiplicada por T . Sea T elegida como $T = 15$, así que $P_1 = P_2 = \begin{bmatrix} 15 & 0 \\ 0 & 15 \end{bmatrix}$. La ley actualizada es también (4.8). Si se usan las mismas entradas acotadas como $u = [3 \sin(0,5t), 0, 0, 0]$, es estable. Seleccionando $T = 0,1$, así que $P = \begin{bmatrix} 0,1 & 0 \\ 0 & 0,1 \end{bmatrix}$, la condición (4.8) no es satisfecha, la fig.4.4 muestra los resultados de inestabilidad. Se puede ver que la condición (4.8) es muy importante para la estabilidad de entrada estado.

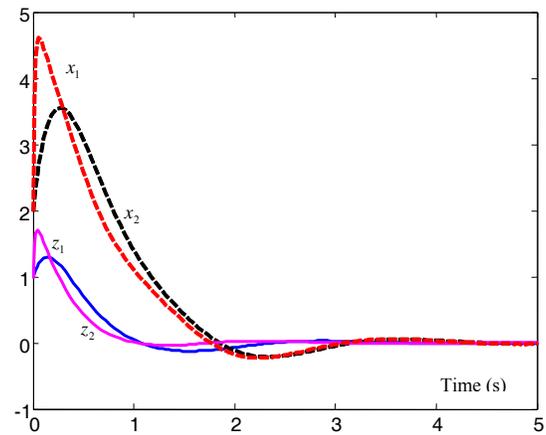


Figura 4.2: Respuestas a entrada cero.

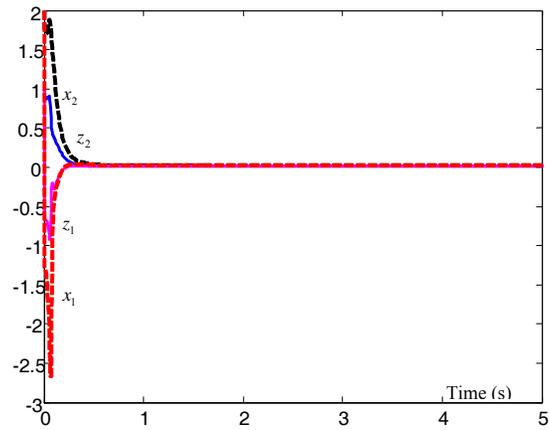


Figura 4.3: Respuesta a la salida retroalimentada.

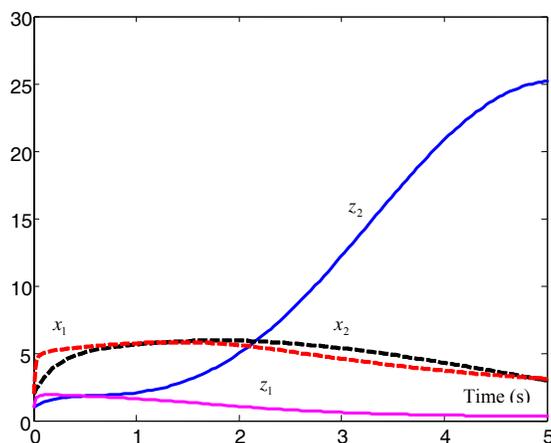


Figura 4.4: Para $T = 0,1$ es inestable.

4.4.2. Estabilidad exponencial de redes neuronales con diferentes escalas de tiempo

Ejemplo 4.2 Para las siguientes redes neuronales diferenciales con diferentes escalas de tiempo

$$\begin{aligned}\dot{x} &= x - z + W_1 \tanh(z) \\ \epsilon \dot{z} &= 2x - z + W_1 \tanh(x)\end{aligned}$$

donde $A = -3$, $C = -1$, $B = -1$, $D = -2$, $f_1 = W_1 \tanh(z)$, $f_2 = W_1 \tanh(x)$, $W_1 = 10$, $W_1 = -5$,. Checando las condiciones del teorema (2.1): 1) W_1 y W_2 son acotadas, 2) B es una matriz Hurwitz no singular, 3) $A - CB^{-1}D = -3 - (-1)(-1) \times (-2) = -1$ es Hurwitz, 4) seleccionando $Q_P = Q_S = 1$, las soluciones de (4.3) es $P = S = 1$, 5) ahora calculando

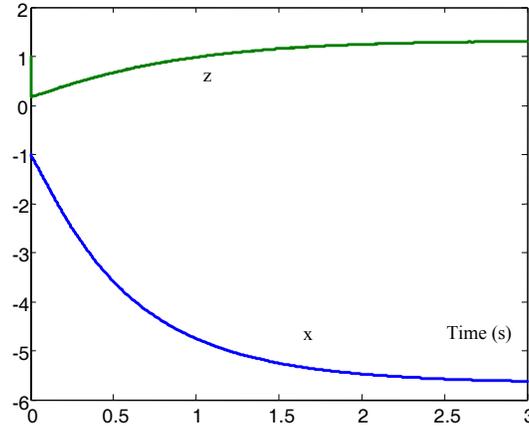


Figura 4.5: Estabilidad de la red neuronal diferencial (Ejemplo 1)

(4.38)

$$\begin{aligned}
 a_1 &= 2,5\gamma \\
 a_2 &= 5 + 6,25\gamma \\
 a_3 &= 1,5 + 0,75\gamma \\
 \bar{a}_3 &= 0,5 - 0,75\gamma \\
 b_1 &= 5 + \frac{2,5}{\gamma} \\
 b_2 &= 5 + \frac{6,25}{\gamma} \\
 b_3 &= 0,5 + \frac{0,75}{\gamma} \\
 \bar{b}_3 &= 1,5 - \frac{0,75}{\gamma}
 \end{aligned}$$

Porque se necesita $\bar{a}_3 > 0$ y $\bar{b}_3 > 0$, así que $\frac{1}{2} < \gamma < \frac{2}{3}$, seleccionando $\gamma = 0,558$. Ahora, variando el parámetro γ en (4.39) con un simple programa para maximizar ϵ^* . Se encuentra cuando ϵ^* es optimizado. En este caso, $\epsilon^* = 9,5 \times 10^{-3}$. De tal manera que el sistema (4.41) es asintóticamente estable cuando $\epsilon = 9 \times 10^{-3}$. La simulación resultante es mostrada en la Fig.(4.5)

Ejemplo 4.3 Se usará la red neuronal diferencial para aproximar el circuito mostrado en

la Fig. (4.6)

$$\begin{aligned}\dot{x} &= Ax + Cz - Cz + W_1\sigma_1(x, z) \\ \epsilon\dot{z} &= Bz + Dx - Dx + W_2\sigma_2(x, z) \\ \dot{x} &= -\frac{1}{R_2C_2}x + \frac{1}{R_2C_2}z - \frac{1}{R_2C_2}f(z) \\ \epsilon\dot{z} &= \frac{1}{R_2C_2}x - \left(\frac{1}{R_1C_2} + \frac{1}{R_2C_2}\right)z + \frac{1}{R_2C_2}f(z)\end{aligned}$$

donde $f(V_1)$ es una fuente de voltaje controlada por voltaje, y C_1 es un parámetro pequeño, $x = VC_2$, $z = V_1$, $y \epsilon = \frac{C_1}{C_2}$. Correspondiendo a redes neuronales diferenciales (4.26), $A = -\frac{1}{R_2C_2}$, $C = \frac{1}{R_2C_2}$, $B = -\left(\frac{1}{R_1C_2} + \frac{1}{R_2C_2}\right)$, $D = \frac{1}{R_2C_2}$, $W_1 = -\frac{1}{R_2C_2}$, $W_2 = \frac{1}{R_2C_2}$, $\sigma_1(x, z) = \sigma_2(x, z) = f(z)$. Checando las condiciones del teorema 2: 1) W_1 y W_2 son acotadas, 2) B es una matriz Hurwitz no singular, 3) $A - CB^{-1}D = -\frac{1}{R_2C_2} + \frac{1}{R_2C_2} \left(\frac{R_1R_2C_2}{R_1+R_2}\right) \times \frac{1}{R_2C_2} = -\frac{R_2}{(R_1+R_2)R_2C_2}$ es Hurwitz, 4) seleccionando $Q_P = Q_S = 1$, las soluciones de (4.32) es $P = 4 \times 10^{-6}$, $S = 1,25 \times 10^{-6}$, 5) ahora calculando (4.38)

$$\begin{aligned}a_1 &= 0,018 \\ a_2 &= 0,1436 + 0,44738\gamma \\ a_3 &= 0,0424 + 0,4968\gamma \\ b_1 &= 0,045 + \frac{1}{\gamma} \times 0,018 \\ b_2 &= 0,204 + \frac{1}{\gamma} \times 0,454 \\ b_3 &= 0,424 + \frac{1}{\gamma} \times 0,496 \\ \bar{a}_3 &= 1,576 - 0,4968\gamma \\ \bar{b}_3 &= 1,576 - \frac{1}{\gamma} \times 0,496\end{aligned}$$

seleccionando $f(z) = k_1z + k_2(z) \sin z$, donde k_1 es una constante y $k_2(z) \sin z$ es una perturbación no lineal, y $|k_2(z)| \leq 0,1$. Se asume que todos los resistores y capacitores tienen $\pm 5\%$ de tolerancia, y k_1 tolera $\pm 10\%$ de variación. Los valores nominales para R_1, R_2, C_2 y k_1 son $R_1^* = 10 \Omega$, $R_2^* = 50 \Omega$, $C_2^* = 0,1 \mu F$, $k_1^* = 2$. Después de algunas manipulaciones matemáticas y observando que $|\sin z| \leq |z|$, derivando la siguiente ecuación de estado de (4.26):

$$\begin{aligned}\dot{x} &= -2 \times 10^5 x - 2 \times 10^5 z + f_1(x, z, t) \\ \epsilon\dot{z} &= 2 \times 10^5 x - 8 \times 10^5 z + f_2(x, z, t)\end{aligned}$$

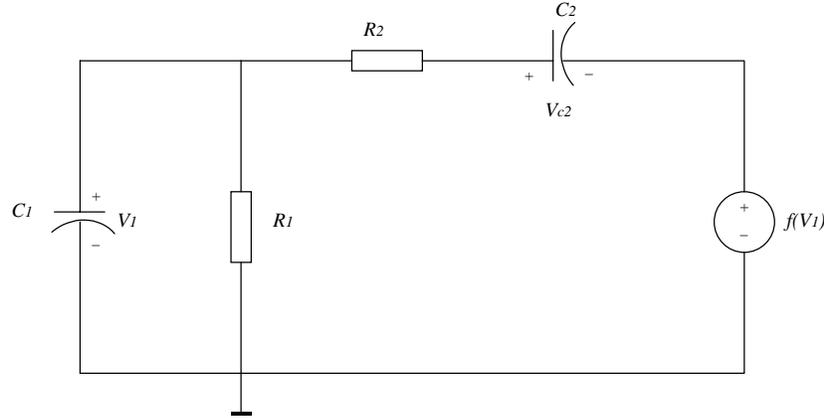


Figura 4.6: Circuito RC con fuente de voltaje dependiente.

porque se necesita que $\bar{a}_3 > 0$ y $\bar{b}_3 > 0$, tal que $0,315 < \gamma < 3,177$, seleccionando $\gamma = 0,98$. Por lo tanto, este sistema está garantizado que es estable para $\epsilon^* = 1,7$, $0 \leq C_1 \leq 0,17 \mu F$. La simulación resultante se muestra en la Fig.(4.7).

4.4.3. Estabilidad robusta de redes neuronales con diferentes escalas de tiempo

Ejemplo 4.4 En un caso vectorial la red neuronal diferencial (4.22) es $x = [x_1, x_2]^T$, $z = [z_1, z_2]^T$, $A = \begin{bmatrix} -0,2 & 0 \\ 0 & -0,2 \end{bmatrix}$, $B = \begin{bmatrix} -0,1 & 0 \\ 0 & -0,1 \end{bmatrix}$, $\sigma_1(x, z)$, $\sigma_2(x, z)$ son funciones \tanh , $W_1 = W_2 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$, eligiendo $C = D = 0,1I$, 1) W_1 y W_2 son acotadas, 2) B es una matriz Hurwitz no singular, 3) $A - CB^{-1}D = \begin{bmatrix} -0,1 & 0 \\ 0 & -0,1 \end{bmatrix}$ es Hurwitz, 4) seleccionando $Q_P = Q_S = 1$, $\epsilon^* = 0,3$. De tal manera que el sistema (4.22) es asintóticamente

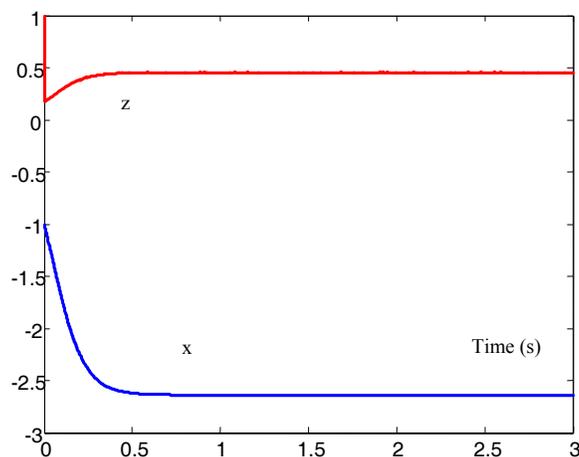


Figura 4.7: Estabilidad de la red neuronal diferencial (Ejemplo 2)

estable cuando $\epsilon = 0,2$. Por la Fig. (4.8) se puede observar que las trayectorias de los estados rápido y lento converge al punto de equilibrio requerido de la condición inicial $x(0) = [1, 2]^T$, $z(0) = [3, 4]^T$.

Si B es inestable, la red neuronal diferencial se tornará inestable. Por ejemplo, si seleccionamos $B = \begin{bmatrix} 0,1 & 0 \\ 0 & 0,1 \end{bmatrix}$.

4.5. Conclusión del capítulo

En esta parte de la tesis, las redes neuronales recurrentes dinámicas diferenciales con una actividad combinada y dinámica de pesos puede ser interpretada como sistemas no lineales perturbados singularmente. Se presentó un método de análisis de estabilidad global de un punto de equilibrio representando el patrón almacenado. El propósito de la función de Lyapunov tipo cuadrática supone un incremento monótono no lineal. Este método proporciona una cota superior en el parámetro de perturbación, y por lo tanto una estimación de una

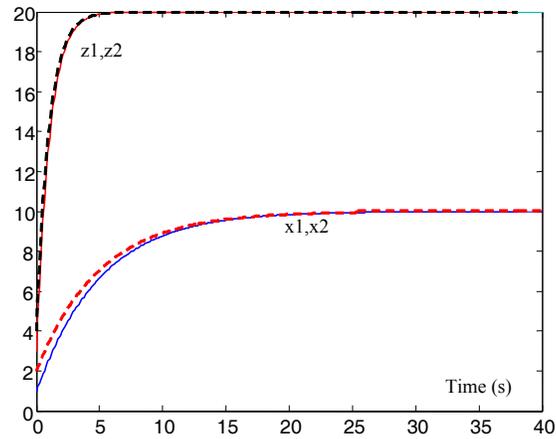


Figura 4.8: Estabilidad de lared neuronal diferencial (Ejemplo 3)

constante de tiempo neuronal positiva máxima, y una estimación del dominio de atracción del punto de equilibrio. También se plantea un método de análisis de estabilidad local para determinar la estabilidad alrededor de puntos de equilibrio individuales, y para analizar las partes rápidas y lentas del sistema neuronal.

Capítulo 5

Identificación vía redes neuronales con diferentes escalas de tiempo

Una diversa clase de sistemas físicos que contienen fenómenos dinámicos lentos y rápidos ocurren en escalas de tiempo separadas. Recientes resultados muestran que las técnicas de redes neuronales al parecer son muy efectivas para modelar e identificar una diversa clase de sistemas no lineales complejos con diferentes escalas de tiempo cuando no se tiene información completa del modelo, o cuando se asume la planta como una “*caja negra*”. Muchas aplicaciones muestran que la neuro-identificación ha emergido como una efectiva herramienta para sistemas no lineales desconocidos. Esta aproximación de modelo libre utiliza las agradables características de redes neuronales artificiales, pero la carencia de modelos hace difícil obtener resultados teóricos sobre estabilidad y realización de neuro-identificadores. Es muy importante para los ingenieros asegurar la estabilidad de neuro-identificadores en la teoría antes de aplicarlos a sistemas reales.

Las Redes Neuronales dinámicas con diferentes escalas de tiempo pueden modelar las dinámicas de la memoria de término corto (niveles de actividad neuronal) y la memoria de término largo (dinámicas de modificaciones sinápticas no supervisadas) [54].

Su capacidad de almacenamiento de patrones como puntos de equilibrio estables requiere los criterios de estabilidad que incluyen la mutua interferencia entre la neurona y dinámicas

de aprendizaje.

Las dinámicas de redes neuronales con diferentes escalas de tiempo son extremadamente complejas, exhibiendo convergencia de puntos atractores y atractores periódicos [2]. Las redes donde ambos términos corto y largo de memoria son variables dinámicas no pueden ser colocadas en forma de las ecuaciones de Cohen-Grossberg [22]. Dos tipos de estabilidad para neuro-identificadores han sido estudiados: 1) la estabilidad de redes neuronales puede ser encontrada en [89] y [100] la estabilidad de algoritmos de aprendizaje fué discutida por [88] y [61]. Se hará énfasis en esta sección sobre algoritmos de aprendizaje estables novedosamente derivados del neuro-identificador con diferentes escalas de tiempo.

Algunas de las aplicaciones de redes neuronales, tales como almacenamiento de patrones y solución de problemas de optimización, requieren que los puntos de equilibrio de la red diseñada sean estables [37]. De modo que, es importante estudiar la estabilidad de redes neuronales. La convergencia completa de redes neuronales con escalas de tiempo diferente esta probada en [99]. La estabilidad exponencial global de redes neuronales competitivas con retardo con diferentes escllas de tiempo es analizada en [50]. Por la técnica de perturbaciones singulares, [49] investiga los problemas de estabilidad para sistemas fuzzy en tiempo continuo y en tiempo discreto con diferentes escalas de tiempo. La estabilidad exponencial global de redes neuronales con diferentes escalas de tiempo es resuelta en [55]. Una clase grande de sistemas competitivos han sido identificados como al inicio "generalmente convergentes a sus puntos atractores aunque no se tengan funciones de Lyapunov para encontrar sus trayectorias.

La estabilidad de algoritmos de aprendizaje puede ser derivado por el análisis de identificación o errores de seguimiento de redes neuronales.

En [35], se estudiaron las condiciones de estabilidad de la ley de actualización cuando los perceptrones multicapa son usados para identificación y control de un sistema no-lineal. En [88] la retropropagación dinámica fué modificada con restricciones de estabilidad NLq . Debido a que las redes neuronales no pueden reproducir el sistema no lineal desconocido exactamente, algunas modificaciones robustas [34] pueden ser aplicadas en el algoritmo del gradiente normal o retropropagación [35], [72], [87], [102]. El enfoque de pasividad puede re-

solver sistemas no lineales mal definidos, generalmente por medio de cotas de sector, y ofrecer soluciones elegantes para la prueba de estabilidad absoluta. Puede llevar a la conclusión general sobre la estabilidad usando solo características de *entrada-salida*. Las propiedades de pasividad de redes neuronales multicapa fueron examinadas en [10]. Por medio del análisis de la interconexión de errores de modelado, ellos derivaron la relación entre pasividad y lazo cerrado estable. La técnica de pasividad puede ser también aplicada en redes neuronales dinámicas. Las propiedades de pasividad de redes neuronales dinámicas pueden ser encontradas en [100]. Esta técnica fue también extendida al neuro-identificador en el caso de capa sencilla [101]. Se concluye que los algoritmos de aprendizaje comúnmente usados con modificaciones robustas tal como zona-muerta [35] y *modification- σ* [72] no son necesarios.

En esta sección, se extienden los principales resultados obtenidos del neuro-identificador normal [100][101] al caso de escalas multi-tiempo. La teoría de pasividad es aplicada para analizar la estabilidad del neuro-identificador con diferentes escalas de tiempo. Pero lo mejor del conocimiento generado en este rubro, el análisis de estabilidad para neuro-identificación con diferentes escalas de tiempo no ha sido todavía establecido en la literatura especializada del tema. Se muestra que la ley de aprendizaje como el de retropropagación puede hacer el error de identificación estable, asintóticamente estable y estable de *entrada-estado*. Simulaciones de la identificación de la velocidad de arranque de un vehículo da la efectividad del algoritmo propuesto en esta sección.

5.1. Redes neuronales de capa simple

La principal preocupación de esta sección es entender algunos conceptos de pasividad y ISS. Considérese una clase de sistemas no lineales con dos escalas de tiempo

$$\begin{aligned}\dot{x}_t &= f(x_t, z_t, u_t, v) \\ \epsilon \dot{z}_t &= g(x_t, z_t, u_t, v) \\ y_t &= h(x_t, z_t, u_t)\end{aligned}\tag{5.1}$$

donde $x_t, z_t \in \mathfrak{R}^n$ es el estado, $u_t \in \mathfrak{R}^m$ es el vector de entrada, $y_t \in \mathfrak{R}^m$ es el vector de salida. $f, g : \mathfrak{R}^n \times \mathfrak{R}^m \rightarrow \mathfrak{R}^n$ son localmente Lipschitz, $h : \mathfrak{R}^n \times \mathfrak{R}^m \rightarrow \mathfrak{R}^m$ es continua. x_t es

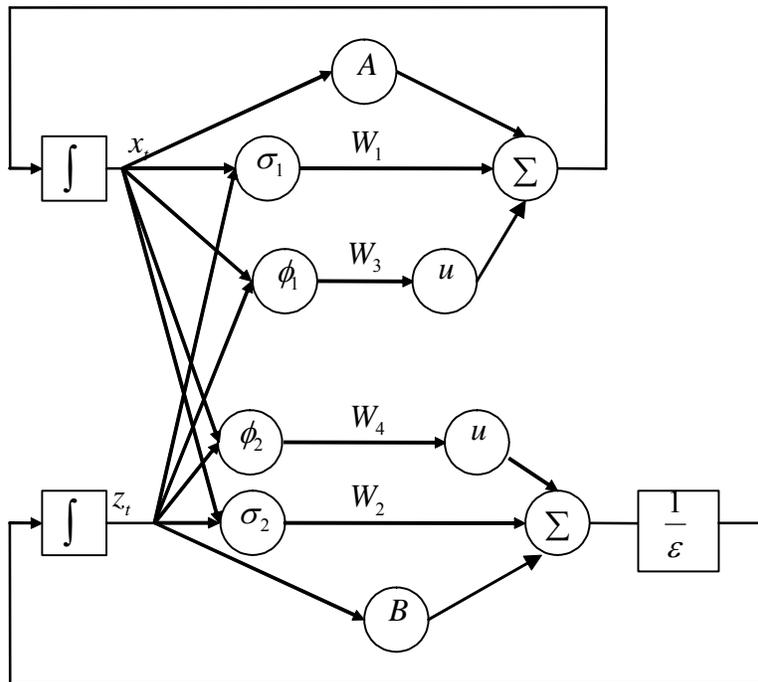


Figura 5.1: Red Neuronal Dinámica con dos escalas de tiempo.

el estado lento y z_t es el estado rápido . $\epsilon > 0$ es conocido.

Es también asumido que para cualquier $x^0 = x_0 \in \mathfrak{R}^n$, la salida $y_t = h(\Phi(t, x^0, u))$ del sistema (5.1) es tal que $\int_0^t |u_s^T y_s| ds < \infty$, para todo $t \geq 0$, *i.e.* la energía almacenada en el sistema (5.1) esta acotada. Recordando algunas propiedades de pasividad y estabilidad de sistemas pasivos ver [6] y [29]. Entonces definiendo $X_t = [x_t^T, z_t^T]^T$.

Construyendo la siguiente red neuronal dinámica con dos escalas de tiempo para identificar el sistema no lineal (5.1)

$$\begin{aligned}\dot{\hat{x}} &= A\hat{x} + W_1\sigma_1(V_1[\hat{x}, \hat{z}]^T) + W_3\phi_1(V_3[\hat{x}, \hat{z}]^T)u \\ \epsilon\dot{\hat{z}} &= B\hat{z} + W_2\sigma_2(V_2[\hat{x}, \hat{z}]^T) + W_4\phi_2(V_4[\hat{x}, \hat{z}]^T)u\end{aligned}\quad (5.2)$$

donde $\hat{x} \in R^n$ y $\hat{z} \in R^n$ son los estados lento y rápidos, $W_i \in R^{n \times 2n}$ ($i = 1 \dots 4$) son los pesos en las capas de salida, $V_i \in R^{2n \times 2n}$ ($i = 1 \dots 4$) son los pesos en las capas ocultas, $\sigma_k = [\sigma_k(\hat{x}_1) \dots \sigma_k(\hat{x}_n), \sigma_k(\hat{z}_1) \dots \sigma_k(\hat{z}_n)]^T \in R^{2n}$ ($k = 1, 2$), $\phi(\cdot) \in R^{2n \times 2n}$ es una matriz diagonal, $\phi_k(\hat{x}, \hat{z}) = \text{diag}[\phi_k(\hat{x}_1) \dots \phi_k(\hat{x}_n), \phi_k(\hat{z}_1) \dots \phi_k(\hat{z}_n)]$ ($k = 1, 2$), $u(k) = [u_1, u_2 \dots u_m, 0, \dots, 0]^T \in R^{2n}$. $A \in R^{n \times n}$, $B \in R^{n \times n}$ son matrices de estabilidad (Hurwitz). ϵ es una constante positiva pequeña. La estructura de las redes neuronales dinámicas (5.2) se muestran en la Fig.4.8. Cuando $\epsilon = 0$, las redes neuronales dinámicas (5.2) han sido discutidas por muchos autores, por ejemplo [61], [66], [72] y [102]. Se puede observar que el modelo de Hopfield [31] es un caso especial de este tipo de redes neuronales con $A = \text{diag}\{a_i\}$, $a_i := -1/R_i C_i$, $R_i > 0$ y $C_i > 0$. R_i y C_i son la resistencia y capacitancia del i -ésimo nodo de la red respectivamente. La sub-estructura $W_1\sigma_1(V_1[\hat{x}, \hat{z}]^T) + W_3\phi_1(V_3[\hat{x}, \hat{z}]^T)u$ es una estructura de perceptrón multicapa. Para simplificar el análisis teórico, se escogen las capas ocultas como $V_i = I$. Analizando una red neuronal de capa sencilla (Se observa que las capas ocultas no afectan la propiedad de pasividad en el comentario 1)

$$\begin{aligned}\dot{\hat{x}} &= A\hat{x} + W_1\sigma_1(\hat{x}, \hat{z}) + W_3\phi_1(\hat{x}, \hat{z})u \\ \epsilon\dot{\hat{z}} &= B\hat{z} + W_2\sigma_2(\hat{x}, \hat{z}) + W_4\phi_2(\hat{x}, \hat{z})u\end{aligned}\quad (5.3)$$

Generalmente, el sistema no lineal (5.1) puede ser escrito como

$$\begin{aligned}\dot{x}_t &= Ax_t + W_1^* \sigma_1(x, z) + W_3^* \phi_1(x, z)u - \tilde{f}_1 \\ \epsilon \dot{z} &= Bz + W_2^* \sigma_2(x, \hat{z}) + W_4^* \phi_2(x, z)u - \tilde{f}_2\end{aligned}\quad (5.4)$$

donde W_j^* ($j = 1 \cdots 4$) es una matriz acotada desconocida:

$$W_j^* \Lambda_j^{-1} W_j^{*T} \leq \overline{W}_j \quad (5.5)$$

\overline{W}_j es una matriz conocida previamente. La función vectorial \tilde{f}_i ($i = 1, 2$) puede ser considerada como error de modelado y disturbios o perturbaciones.

Definiendo el error de identificación como

$$\begin{aligned}\Delta_1 &= \hat{x}_t - x_t \\ \Delta_2 &= \hat{z}_t - z_t\end{aligned}$$

donde $\hat{X}_t = [\hat{x}_t^T, \hat{z}_t^T]^T$. Porque $\sigma_i(\cdot)$ y $\phi_i(\cdot)$ ($i = 1, 2$), son elegidas como funciones sigmoides, claramente ellas satisfacen la siguiente propiedad de *Lipschitz*:

$$\tilde{\sigma}_i^T \Lambda_i \tilde{\sigma}_i \leq \Delta_i^T D_{\sigma_i} \Delta_i, \quad \left(\tilde{\phi}_i \gamma(u_t) \right)^T \Lambda_i \left(\tilde{\phi}_i \gamma(u_t) \right) \leq \bar{u} \Delta_i^T D_{\phi_i} \Delta_i \quad (5.6)$$

donde $\tilde{\sigma}_i = \sigma_i(\hat{X}_t) - \sigma_i(X_t)$, $\tilde{\phi}_i = \phi_i(\hat{X}_t) - \phi_i(X_t)$, Λ_i , D_{σ_i} y D_{ϕ_i} son constantes positivas conocidas, $i = 1, 2$. Las dinámicas multicapa son obtenidas de (5.27) y (5.4)

$$\begin{aligned}\dot{\Delta}_1 &= A\Delta_1 + \tilde{W}_1 \sigma_1(\hat{X}) + \tilde{W}_3 \phi_1(\hat{X}_t) \gamma(u_t) + W_1^* \tilde{\sigma}_1 + W_3^* \tilde{\phi}_1 \gamma(u_t) + \tilde{f}_1 \\ \epsilon \dot{\Delta}_2 &= B\Delta_2 + \tilde{W}_2 \sigma_2(\hat{X}) + \tilde{W}_4 \phi_2(\hat{X}_t) \gamma(u_t) + W_2^* \tilde{\sigma}_2 + W_4^* \tilde{\phi}_2 \gamma(u_t) + \tilde{f}_2\end{aligned}\quad (5.7)$$

donde $\tilde{W}_j = W_j - W_j^*$ ($j = 1 \cdots 4$). Si se define

$$\begin{aligned}R_1 &= \overline{W}_1 + \overline{W}_2, & Q_1 &= D_{\sigma_1} + \bar{u} D_{\phi_1} + Q_{01} \\ R_2 &= \overline{W}_3 + \overline{W}_4, & Q_2 &= D_{\sigma_2} + \bar{u} D_{\phi_2} + Q_{02}\end{aligned}\quad (5.8)$$

y las matrices A , B , Q_{01} y Q_{02} son seleccionadas para cumplir las siguientes condiciones:

(1) los pares $(A, R_1^{1/2})$ y $(B, R_2^{1/2})$ son controlables, los pares $(Q_1^{1/2}, A)$ y $(Q_2^{1/2}, B)$ son observables,

(2) *las condiciones de frecuencia local* [95] se satisfacen, entonces la siguiente consideración puede ser establecida:

Existe una matriz de estabilidad A y una matriz definida estrictamente positiva Q_0 tal que la ecuación matricial de

$$\begin{aligned} P_1 A + A^T P_1 + P_1 R_1 P_1 &= -Q_1 \\ P_2 B + B^T P_2 + P_2 R_2 P_2 &= -Q_2 \end{aligned} \quad (5.9)$$

tendrá una solución positiva $P_i = P_i^T > 0$, ($i = 1, 2$).

Estas condiciones son fácilmente cumplidas seleccionando A como una matriz diagonal estable. El siguiente teorema plantea el procedimiento del neuroidentificador.

Teorema 5.1 *Dependiendo de la consideración (5.30) que sea satisfecha, si los pesos W_j ($j = 1 \cdots 4$) son actualizados como*

$$\begin{aligned} \dot{W}_1 &= -K_1 P_1 \Delta_1 \sigma_1^T(\hat{X}_t) \\ \dot{W}_2 &= -K_2 P_1 \phi_1(\hat{X}_t) \gamma(u_t) \Delta_1^T \\ \dot{W}_3 &= -K_3 P_2 \Delta_2 \sigma_2^T(\hat{X}_t) \\ \dot{W}_4 &= -K_4 P_2 \phi_2(\hat{X}_t) \gamma(u_t) \Delta_2^T \end{aligned} \quad (5.10)$$

donde P_1 y P_2 son la solución de la ecuación de Riccati (5.9), entonces las dinámicas de los errores de identificación (5.31) es estrictamente pasiva del error de modelado \tilde{f}_i para la identificación del error $2P\Delta_i$, $i = 1, 2$

Demostración. Seleccionando una función de Lyapunov (función de almacenamiento) como

$$S_t(x, z) = \Delta_1^T P_1 \Delta_1 + \Delta_2^T P_2 \Delta_2 + tr \left\{ \tilde{W}_1^T K_1^{-1} \tilde{W}_1 \right\} + tr \left\{ \tilde{W}_2^T K_2^{-1} \tilde{W}_2 \right\} + tr \left\{ \tilde{W}_3^T K_3^{-1} \tilde{W}_3 \right\} + tr \left\{ \tilde{W}_4^T K_4^{-1} \tilde{W}_4 \right\} \quad (5.11)$$

donde $P_i \in \Re^{n \times n}$ es una matriz definida positiva, $tr \{ \cdot \}$ significa la *traza* y es definida como la suma de todos los elementos diagonales de una matriz. De acuerdo a (5.31), la derivada

es

$$\begin{aligned} \dot{S}_t(\hat{x}, \hat{z}) = & \Delta_1^T (P_1 A + A^T P_1) \Delta_1 + \frac{1}{\epsilon} \Delta_2^T (P_2 B + B^T P_2) \Delta_1 + 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 \\ & + 2\Delta_1^T P_1 \tilde{W}_1 \sigma_1 + 2\Delta_1^T P_1 \tilde{W}_2 \phi_1 u + \frac{2}{\epsilon} \Delta_2^T P_2 \tilde{W}_3 \sigma_2 + \frac{2}{\epsilon} \Delta_2^T P_2 \tilde{W}_4 \phi_2 u \\ & + 2\Delta_1^T P_1 \left[W_1^* \tilde{\sigma}_1 + W_2^* \tilde{\phi}_1 \gamma(u_t) \right] + \frac{2}{\epsilon} \Delta_2^T P_2 \left[W_3^* \tilde{\sigma}_2 + W_4^* \tilde{\phi}_2 \gamma(u_t) \right] \\ & + 2 \left[tr \left\{ \tilde{W}_1^T K_1^{-1} \tilde{W}_1 \right\} + tr \left\{ \tilde{W}_2^T K_2^{-1} \tilde{W}_2 \right\} + tr \left\{ \tilde{W}_3^T K_3^{-1} \tilde{W}_3 \right\} + tr \left\{ \tilde{W}_4^T K_4^{-1} \tilde{W}_4 \right\} \right] \end{aligned}$$

Debido a que $\Delta_i^T P_i W_j^* \tilde{\sigma}_i$ es escalar, usando (5.30) y la desigualdad matricial

$$X^T Y + (X^T Y)^T \leq X^T \Lambda^{-1} X + Y^T \Lambda Y \quad (5.12)$$

donde $X, Y, \Lambda \in \mathfrak{R}^{n \times k}$ son cualesquiera matrices, Λ es cualquier matriz definida positiva, obteniendo

$$\begin{aligned} 2\Delta_1^T P_1 W_1^* \tilde{\sigma}_1 & \leq \Delta_1^T P_1 W_1^* \Lambda_1^{-1} W_1^{*T} P_1 \Delta_1 + \tilde{\sigma}_1^T \Lambda_1 \tilde{\sigma}_1 \leq \Delta_1^T (P_1 \bar{W}_1 P_1 + D_{\sigma_1}) \Delta_1 \\ 2\Delta_1^T P_1 W_2^* \tilde{\phi}_1 \gamma(u_t) & \leq \Delta_1^T (P_1 \bar{W}_2 P_1 + \bar{u} D_{\phi_1}) \Delta_1 \end{aligned} \quad (5.13)$$

De tal manera que se genera

$$\begin{aligned} \dot{S}_t & \leq \Delta_1^T [P_1 A + A^T P_1 + P_1 (\bar{W}_1 + \bar{W}_2) P_1 + (D_{\sigma_1} + \bar{u} D_{\phi_1} + Q_{01})] \Delta_1 \\ & + \frac{1}{\epsilon} \Delta_2^T [P_2 B + B^T P_2 + P_2 (\bar{W}_3 + \bar{W}_4) P_2 + (D_{\sigma_2} + \bar{u} D_{\phi_2} + Q_{02})] \Delta_2 \\ & + 2tr \left\{ \tilde{W}_1^T K_1^{-1} \tilde{W}_1 \right\} + 2\Delta_1^T P_1 \tilde{W}_1 \sigma_1 + 2tr \left\{ \tilde{W}_2^T K_2^{-1} \tilde{W}_2 \right\} + 2\Delta_1^T P_1 \tilde{W}_2 \phi_1 u \\ & + 2tr \left\{ \tilde{W}_3^T K_3^{-1} \tilde{W}_3 \right\} + \frac{2}{\epsilon} \Delta_2^T P_2 \tilde{W}_3 \sigma_2 + \frac{2}{\epsilon} tr \left\{ \tilde{W}_4^T K_4^{-1} \tilde{W}_4 \right\} + \frac{2}{\epsilon} \Delta_2^T P_2 \tilde{W}_4 \phi_2 u \\ & + 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 - \Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 \end{aligned}$$

Debido a que $\tilde{W}_j = \dot{W}_j$, si se usa la ley de actualización como en (5.10) y (5.30), se obtiene

$$\dot{S}_t \leq -\Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 + 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 \quad (5.14)$$

De la definición 1, si se define las entradas como \tilde{f}_i y las salidas como $2P_i \Delta_i$, entonces el sistema es estrictamente pasivo con $V_t = \Delta_1^T Q_{01} \Delta_1 + \Delta_2^T Q_{02} \Delta_2 \geq 0$. ■

Comentario 5.1 Debido a que la razón de actualización es $K_j P_i$ ($j = 1 \cdots 4, i = 1, 2$), y K_j puede ser seleccionada como cualquier matriz positiva, el proceso de aprendizaje de la red neuronal dinámica (5.16) es libre de la solución de la ecuación de Riccati (5.15).

Corolario 5.1 Si solo parámetros con incertidumbre están presentes ($\tilde{f}_i = 0, i = 1, 2$), entonces la ley de actualización como (5.16) puede hacer el error de identificación asintóticamente estable,

$$\lim_{t \rightarrow \infty} \Delta_i = 0, \quad W_j \in L_\infty, \quad \lim_{t \rightarrow \infty} \dot{W}_j = 0, \quad j = 1 \cdots 4, i = 1, 2 \quad (5.15)$$

Demostración. Debido a las dinámicas del error de identificación (5.31) es pasiva y $\tilde{f}_i = 0$, de la propiedad 1 la función de almacenamiento S satisface

$$\dot{S} \leq 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 = 0$$

La función definida positiva S implica que Δ_i y W_j son acotadas. De la ecuación del error (5.31) $\dot{\Delta}_i \in L_\infty$

$$\dot{S} \leq -\Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 \leq 0$$

Integrando (5.20) ambos lados

$$\int_0^\infty \|\Delta_1\|_{Q_{01}} + \|\Delta_2\|_{Q_{02}} \leq S_0 - S_\infty < \infty$$

Debido a que $\Delta_i \in L_2 \cap L_\infty$, usando el Lemma de Barbalat [34] obteniendo (5.21). Debido a que u_t, σ_i, ϕ_i y P_i están acotadas, $\lim_{t \rightarrow \infty} \dot{W}_j = 0$ y $\lim_{t \rightarrow \infty} \dot{W}_j = 0$. ■

Comentario 5.2 Para el caso del modelo que combina bien, el análisis tipo Lyapunov puede alcanzar el mismo resultado (ver Corollary 1 y Corollary 2) en [102]. Pero en presencia del error de modelado, se dará un resultado absolutamente diferente: el algoritmo del gradiente descendente (5.16) es también robusto respecto a la perturbación y dinámica no modelada.

Teorema 5.2 Si la consideración (5.30) se satisface y existe una matriz definida positiva Λ_f tal que

$$\begin{aligned} \lambda_{\max}(P_1 \Lambda_{f1} P_1) &\leq \lambda_{\min}(Q_{01}) \\ \lambda_{\max}(P_2 \Lambda_{f2} P_2) &\leq \lambda_{\min}(Q_{02}) \end{aligned} \quad (5.16)$$

entonces la ley actualizada (5.16) puede hacer las dinámicas del neuroidentificador(5.31) estable de entrada-estado (ISS).

Demostración. En vista de la desigualdad matricial (5.18),

$$\begin{aligned} 2\Delta_1^T P_1 \tilde{f}_1 &\leq \Delta_1^T P_1 \Lambda_{f_1} P_1 \Delta_1 + \tilde{f}_1^T \Lambda_{f_1}^{-1} \tilde{f}_1 \\ 2\Delta_2^T P_2 \tilde{f}_2 &\leq \Delta_2^T P_2 \Lambda_{f_2} P_2 \Delta_2 + \tilde{f}_2^T \Lambda_{f_2}^{-1} \tilde{f}_2 \end{aligned} \quad (5.17)$$

(5.20) puede ser representada como:

$$\begin{aligned} \dot{S}_t &\leq -\Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 + 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 \\ &\leq -\lambda_{\min}(Q_{01}) \|\Delta_1\|^2 - \lambda_{\min}(Q_{02}) \|\Delta_2\|^2 \\ &\quad + \Delta_1^T P_1 \Lambda_{f_1} P_1 \Delta_1 + \tilde{f}_1^T \Lambda_{f_1}^{-1} \tilde{f}_1 + \Delta_2^T P_2 \Lambda_{f_2} P_2 \Delta_2 + \tilde{f}_2^T \Lambda_{f_2}^{-1} \tilde{f}_2 \\ &\leq -\alpha_{\|\Delta_1\|} \|\Delta_1\| + \beta_{\|\tilde{f}_{t1}\|} \|\tilde{f}_{t1}\| - \alpha_{\|\Delta_2\|} \|\Delta_2\| + \beta_{\|\tilde{f}_{t2}\|} \|\tilde{f}_{t2}\| \end{aligned}$$

donde $\alpha_{\|\Delta_1\|} := [\lambda_{\min}(Q_{01}) - \lambda_{\max}(P_1 \Lambda_{f_1} P_1)] \|\Delta_1\|$, $\beta_{\|\tilde{f}_{t1}\|} := \lambda_{\max}(\Lambda_{f_1}^{-1}) \|\tilde{f}_{t1}\|$, $\alpha_{\|\Delta_2\|} := [\lambda_{\min}(Q_{02}) - \lambda_{\max}(P_2 \Lambda_{f_2} P_2)] \|\Delta_2\|$, $\beta_{\|\tilde{f}_{t2}\|} := \lambda_{\max}(\Lambda_{f_2}^{-1}) \|\tilde{f}_{t2}\|$. Es posible seleccionar matrices definidas positivas Λ_{f_1} y Λ_{f_2} tal que (5.22) es establecida. Debido a que α y β son \mathcal{K}_∞ funciones, S_t es una función ISS-Lyapunov. Usando el Teorema 1 de [85], la dinámica del error de identificación (5.31) es estable de entrada-estado. ■

Comentario 5.3 *Primero, se puede elegir A y Q_{01} tal que la ecuación de Riccati (5.15) tendrá solución positiva P_1 . Entonces Λ_{f_1} puede ser encontrada de acuerdo a la condición (5.22). Debido a que (5.23) es correcta para cualquier matriz definida positiva, (5.22) puede ser establecida si Λ_{f_1} es seleccionada como una matriz constante suficientemente pequeña. La condición (5.22) no tendrá efecto en las dinámicas de la red (5.8) y su entrenamiento (5.16).*

Corolario 5.2 *Si el error de modelado \tilde{f}_i es acotado, entonces la ley de actualización(5.16) puede hacer estable el procedimiento de identificación.*

$$\Delta_i \in L_\infty, \quad W_j \in L_\infty$$

Demostración. De la propiedad 2 se conoce que por medio de la estabilidad de entrada-estado el comportamiento de las redes neuronales dinámicas permanecerá acotada cuando sus entradas son acotadas. ■

Comentario 5.4 Debido a que el estado y las variables de salida son físicamente acotadas, el error de modelado \tilde{f}_i se puede considerar estar acotado también (ver, por ejemplo [35][66][72]). Como el tipo de leyes adaptivas robustas, tal como zona-muerta [66] y modificación- σ [61], la cota superior del error de modelado no es necesaria para el entrenamiento de la red.

Comentario 5.5 Es bien conocido que la estructura de incertidumbres causará la rigidez de parámetros para control adaptivo, de tal manera que se usa la modificación robusta para hacer estable el error de identificación [34]. Lo métodos adaptivos robustos pueden ser extendidos para neuro-identificación directamente [35][66] [72]. Pero la neuro-identificación es un tipo de método de caja-negra, la información de la estructura es inútil y todas las incertidumbres están dentro la caja-negra. Aunque los algoritmos adaptivos robustos son adecuados para neuro-identificación, no son tan sencillos. Por medio de técnicas de pasividad, exitosamente se prueba la siguiente conclusión: el algoritmo del gradiente descendente para los pesos es robusto con respecto a cualquier incertidumbre acotada.

5.2. Redes Neuronales multicapa

Construyendo la siguiente red neuronal dinámica con dos escalas de tiempo para identificar el sistema no lineal (5.18) como se muestra en la Fig.4.8

$$\begin{aligned}\dot{\hat{x}} &= A\hat{x} + W_1\sigma_1(V_1\hat{X}) + W_3\phi_1(V_3\hat{X})u \\ \epsilon\dot{\hat{z}} &= B\hat{z} + W_2\sigma_2(V_2\hat{X}) + W_4\phi_2(V_4\hat{X})u\end{aligned}\tag{5.18}$$

donde $\hat{X} = [\hat{x}, \hat{z}]^T$, $\hat{x} \in R^n$ y $\hat{z} \in R^n$ son los estados lento y rápido, $W_i \in R^{n \times 2n}$ ($i = 1 \cdots 4$) son los pesos en las capas de salida,

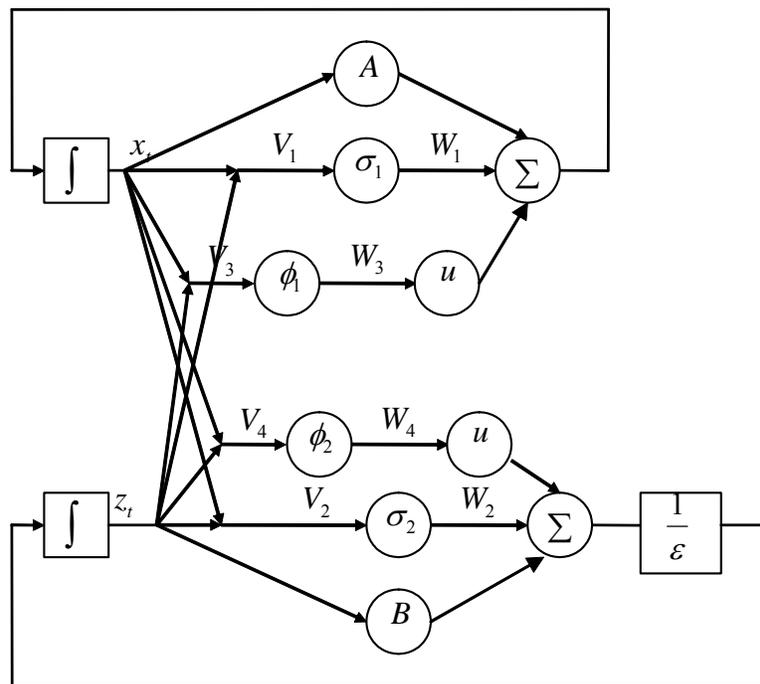


Figura 5.2: Red Neuronal Dinámica con dos escalas de tiempo.

$V_i \in R^{2n \times 2n}$ ($i = 1 \dots 4$) son los pesos en las capas ocultas, $\sigma_k = [\sigma_k(\hat{x}_1) \dots \sigma_k(\hat{x}_n), \sigma_k(\hat{z}_1) \dots \sigma_k(\hat{z}_n)]^T \in R^{2n}$ ($k = 1, 2$),

$\phi(\cdot) \in R^{2n \times 2n}$ es una matriz diagonal, $\phi_k(\hat{X}) = \text{diag}[\phi_k(\hat{x}_1) \dots \phi_k(\hat{x}_n), \phi_k(\hat{z}_1) \dots \phi_k(\hat{z}_n)]$ ($k = 1, 2$),

$u(k) = [u_1, u_2 \dots u_m, 0, \dots, 0]^T \in R^{2n}$. $A \in R^{n \times n}$, $B \in R^{n \times n}$ son matrices de estabilidad (Hurwitz). ϵ es una constante positiva pequeña. La estructura de las redes neuronales dinámicas (5.9) se muestran en Fig.4.8. Cuando $\epsilon = 0$, las redes neuronales dinámicas (5.8) han sido discutidas por muchos autores, por ejemplo [61], [66], [72] y [102]. Uno puede ver que el modelo de Hopfield [31] es un caso especial de este tipo de redes neuronales con $A = \text{diag}\{a_i\}$, $a_i := -1/R_i C_i$, $R_i > 0$ y $C_i > 0$. R_i y C_i son la resistencia y capacitancia en el i -ésimo nodo de la red respectivamente. La sub-estructura $W_1 \sigma_1(V_1 \hat{X}) + W_3 \phi_1(V_3 \hat{X})u$ es una estructura de perceptrón multicapa. La representación típica de los elementos $\sigma_i(\cdot)$ y $\phi_{ii}(\cdot)$ son funciones sigmoidales

$$\sigma_i(x_{i,t}) = a_i / (1 + e^{-b_i x_{i,t}}) - c_i$$

Comentario 5.6 *Las redes neuronales han sido discutidas por muchos autores, por ejemplo [72], [61], [66] y [102]. Se puede ver que el modelo de Hopfield [31] es el caso especial de estas redes con $A = \text{diag}\{a_i\}$, $a_i := -1/R_i C_i$, $R_i > 0$ y $C_i > 0$. R_i y C_i son la resistencia y capacitancia en el i -ésimo nodo de la red respectivamente.*

Generalmente, el sistema no lineal (5.18) puede ser representado como sigue:

$$\begin{aligned} \dot{x}_t &= Ax_t + W_1^* \sigma_1(V_1^0 X) + W_3^* \phi_1(V_3^0 X)u - \tilde{f}_1(V_1^0, V_3^0) \\ \epsilon \dot{z} &= Bz + W_2^* \sigma_2(V_2^0 X) + W_4^* \phi_2(V_4^0 X)u - \tilde{f}_2(V_2^0, V_4^0) \end{aligned} \quad (5.19)$$

donde W_j^* ($j = 1 \dots 4$) es una matriz acotada desconocida:

$$W_j^* \Lambda_j^{-1} W_j^{*T} \leq \overline{W}_j \quad (5.20)$$

\overline{W}_j es una matriz conocida previamente. La función vectorial \tilde{f}_i ($i = 1, 2$) puede ser considerada como perturbaciones y error de modelado, donde sus argumentos están definidos

como V_1^0 y V_2^0 que son matrices dadas a priori obtenidas del aprendizaje off-line. Definiendo el error de identificación como:

$$\begin{aligned}\Delta_1 &= \hat{x}_t - x_t \\ \Delta_2 &= \hat{z}_t - z_t \\ \Delta &= [\Delta_1, \Delta_2]^T\end{aligned}$$

$\tilde{\sigma}_i = \sigma_i(V_i^0 \hat{X}) - \sigma(V_i^0 X)$, $\tilde{\phi}_i = \phi_i(V_i^0 \hat{X})u - \phi_i(V_i^0 X)u$, $\tilde{\sigma}'_i = \sigma_i(V_i \hat{X}) - \sigma_i(V_i^0 \hat{X})$, $\tilde{\phi}'_i = \phi_i(V_i \hat{X})u - \phi_i(V_i^0 \hat{X})u$, $\tilde{V}_i = V_i - V_i^0$, $\tilde{W}_i = W_i - W_i^*$, $i = 1 \dots 4$. Porque $\sigma_i(\cdot)$ y $\phi_i(\cdot)$ son elegidas como funciones sigmoidales, claramente estas satisfacen la condición de Lipschitz,

$$\begin{aligned}\tilde{\sigma}_i^T \Lambda_i \tilde{\sigma}_i &\leq \Delta^T \Lambda_{\sigma_i} \Delta, & \tilde{\phi}_i^T \Lambda_i \tilde{\phi}_i &\leq \bar{u} \Delta^T \Lambda_{\phi_i} \Delta \\ \tilde{\sigma}'_i &= D_{\sigma_i} \tilde{V}_i \hat{X} + \nu_{\sigma_i}, & \tilde{\phi}'_i &= D_{\phi_i} \tilde{V}_{2,t} \hat{x}_t + \nu_{\phi_i}\end{aligned}\quad (5.21)$$

donde

$$\begin{aligned}D_{\sigma_i} &= \left. \frac{\partial \sigma_i^T(Z)}{\partial Z} \right|_{Z=V_i \hat{X}}, & \|\nu_{\sigma_i}\|_{\Lambda_i}^2 &\leq l_1 \left\| \tilde{V}_i \hat{X} \right\|_{\Lambda_i}^2 \\ D_{\phi_i} &= \left. \frac{\partial [\phi_i(Z)u]^T}{\partial Z} \right|_{Z=V_i \hat{X}}, & \|\nu_{\phi_i}\|_{\Lambda_i}^2 &\leq l_2 \left\| \tilde{V}_i \hat{X} \right\|_{\Lambda_i}^2\end{aligned}$$

$l_1 > 0$, $l_2 > 0$, Λ_i , Λ_{σ_i} y Λ_{ϕ_i} son matrices definidas positivas. La dinámica del error de identificación es obtenida de (5.27) y (5.10)

$$\begin{aligned}\dot{\Delta}_1 &= A\Delta_1 + \tilde{W}_1 \sigma_1(V_1 \hat{X}) + \tilde{W}_3 \phi_1(V_3 \hat{X})u + W_1^* \tilde{\sigma}_1 + W_1^* \tilde{\sigma}'_1 + W_3^* \tilde{\phi}_1 + W_3^* \tilde{\phi}'_1 + \tilde{f}_1(V_1^0, V_3^0) \\ \epsilon \dot{\Delta}_2 &= B\Delta_2 + \tilde{W}_2 \sigma_2(V_2 \hat{X}) + \tilde{W}_4 \phi_2(V_4 \hat{X})u + W_2^* \tilde{\sigma}_2 + W_2^* \tilde{\sigma}'_2 + W_4^* \tilde{\phi}_2 + W_4^* \tilde{\phi}'_2 + \tilde{f}_2(V_2^0, V_4^0)\end{aligned}\quad (5.22)$$

Entonces si se define:

$$\begin{aligned}R_1 &= \bar{W}_1 + \bar{W}_3, & Q_1 &= D_{\sigma_1} + \bar{u}D_{\phi_1} + Q_{01} \\ R_2 &= \bar{W}_2 + \bar{W}_4, & Q_2 &= D_{\sigma_2} + \bar{u}D_{\phi_2} + Q_{02}\end{aligned}\quad (5.23)$$

y las matrices A , B , Q_{01} y Q_{02} son seleccionadas para cumplir las siguientes condiciones:

1) los pares $(A, R_1^{1/2})$ y $(B, R_2^{1/2})$ son controlables, los pares $(Q_1^{1/2}, A)$ y $(Q_2^{1/2}, B)$ son observables,

2) si las condiciones de frecuencia local [95] satisfacen, entonces la siguiente consideración puede ser establecida:

Existen matrices de estabilidad A y B y estrictamente matrices definidas positivas Q_{01} y Q_{02} tal que la ecuación matricial de Riccati

$$\begin{aligned} P_1 A + A^T P_1 + P_1 R_1 P_1 &= -Q_1 \\ P_2 B + B^T P_2 + P_2 R_2 P_2 &= -Q_2 \end{aligned} \quad (5.24)$$

tendrá una solución positiva $P_i = P_i^T > 0$, ($i = 1, 2$).

Estas condiciones fácilmente se cumplen si se seleccionan A y B como matrices diagonales estables. El siguiente teorema plantea el procedimiento de aprendizaje del neuro-identificador.

Teorema 5.3 *Dependiendo de si la consideración (5.21) sea satisfecha, si los pesos W_i y V_i ($i = 1 \dots 4$) son actualizados como:*

$$\begin{aligned} \dot{W}_1 &= -K_1 P_1 \sigma_1(V_1 \hat{X}) \Delta_1^T + K_1 P_1 D_{\sigma_1}(V_1 - V_1^0) \hat{X} \Delta_1^T \\ \dot{W}_3 &= -K_3 P_1 \phi_1(V_3 \hat{X}) u \Delta_1^T + K_3 P_1 D_{\phi_1}(V_3 - V_3^0) \hat{X} u \Delta_1^T \\ \dot{W}_2 &= -K_2 P_2 \sigma_2(V_2 \hat{X}) \Delta_2^T + K_2 P_2 D_{\sigma_2}(V_2 - V_2^0) \hat{X} \Delta_2^T \\ \dot{W}_4 &= -K_4 P_2 \phi_2(V_4 \hat{X}) u \Delta_2^T + K_4 P_2 D_{\phi_2}(V_2 - V_2^0) \hat{X} u \Delta_2^T \\ \dot{V}_1 &= -L_1 P_1 W_1 D_{\sigma_1} \Delta_1 \hat{X}^T - \frac{l_1}{2} L_1 \Lambda_1 (V_1 - V_1^0) \hat{X} \hat{X}^T \\ \dot{V}_3 &= -L_3 P_1 W_3 D_{\phi_1} \hat{X} \Delta_1^T - \frac{l_3}{2} L_3 \Lambda_3 (V_3 - V_3^0) \hat{X} \hat{X}^T \\ \dot{V}_2 &= -L_2 P_2 W_2 D_{\sigma_2} \Delta_2 \hat{X}^T - \frac{l_2}{2} L_2 \Lambda_2 (V_2 - V_2^0) \hat{X} \hat{X}^T \\ \dot{V}_4 &= -L_4 P_2 W_4 D_{\phi_2} \hat{X} \Delta_2^T - \frac{l_4}{2} L_4 \Lambda_4 (V_4 - V_4^0) \hat{X} \hat{X}^T \end{aligned} \quad (5.25)$$

donde P_1 y P_2 son la solución de la ecuación de Riccati (5.15), entonces las dinámicas del error de identificación (5.31) es estrictamente pasivo del error de modelado \tilde{f}_i al error de identificación $2P\Delta_i$, $i = 1, 2$

Demostración. Seleccionando una función de Lyapunov (función de almacenamiento) como:

$$S_t = \Delta_1^T P_1 \Delta_1 + \Delta_2^T P_2 \Delta_2 + \sum_{i=1}^4 tr \left\{ \tilde{W}_i^T K_i^{-1} \tilde{W}_i \right\} + \sum_{i=1}^4 tr \left\{ \tilde{V}_i^T L_i^{-1} \tilde{V}_i \right\} \quad (5.26)$$

donde $P \in \mathfrak{R}^{n \times n}$ es una matriz definida positiva. De acuerdo a (5.31), la derivada de (5.17) es:

$$\begin{aligned} \dot{S}_t &= \Delta_1^T (P_1 A + A^T P_1) \Delta_1 + \frac{1}{\epsilon} \Delta_2^T (P_2 B + B^T P_2) \Delta_1 + 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 \\ &+ 2\Delta_1^T P_1 \tilde{W}_1 \sigma_1 + 2\Delta_1^T P_1 \tilde{W}_2 \phi_1 u + \frac{2}{\epsilon} \Delta_2^T P_2 \tilde{W}_3 \sigma_2 + \frac{2}{\epsilon} \Delta_2^T P_2 \tilde{W}_4 \phi_2 u \\ &+ 2\Delta_1^T P_1 \left[W_1^* \tilde{\sigma}_1 + W_3^* \tilde{\phi}_1 u + W_1^* \tilde{\sigma}'_1 + W_3^* \tilde{\phi}'_1 u \right] + \frac{2}{\epsilon} \Delta_2^T P_2 \left[W_2^* \tilde{\sigma}_2 + W_4^* \tilde{\phi}_2 u + W_2^* \tilde{\sigma}'_2 + W_4^* \tilde{\phi}'_2 u \right] \\ &+ \sum_{i=1}^4 \text{tr} \left\{ \tilde{W}_i^T K_i^{-1} \tilde{W}_i \right\} + \sum_{i=1}^4 \text{tr} \left\{ \tilde{V}_i^T L_i^{-1} \tilde{V}_i \right\} \end{aligned}$$

Debido a que $2\Delta_i^T P W_i^* \tilde{\sigma}_i$ es escalar, usando (5.30) y la desigualdad matricial:

$$X^T Y + (X^T Y)^T \leq X^T \Lambda^{-1} X + Y^T \Lambda Y \quad (5.27)$$

donde $X, Y, \Lambda \in \mathfrak{R}^{n \times k}$ son matrices, $\Lambda = \Lambda^T > 0$, obteniendo:

$$2\Delta_i^T P_i W_i^* \tilde{\sigma}_i \leq \Delta_i^T P_i W_i^* \Lambda_i^{-1} W_i^{*T} P_i \Delta_i + \tilde{\sigma}_i^T \Lambda_i \tilde{\sigma}_i \leq \Delta_i^T (P_i \bar{W}_i P_i + D_{\sigma_i}) \Delta_i \quad (5.28)$$

Similarmente se tiene $2\Delta_i^T P_i W_i^* \tilde{\phi}_i \leq \Delta_i^T (P_i \bar{W}_i P_i + \bar{u} D_{\phi_i}) \Delta_i$,

$$\begin{aligned} 2\Delta_i^T P_i W_i^* \tilde{\sigma}'_i &= 2\Delta_i^T P_i W_i^* D_{\sigma_i} \tilde{V}_i \hat{X} + 2\Delta_i^T P_i W_i^* \nu_{\sigma_i} \\ &\leq 2\Delta_i^T P_i (W_i - \tilde{W}_i) D_{\sigma_i} \tilde{V}_i \hat{X} + \Delta_i^T P_i \bar{W}_i P_i \Delta_i + l_i \hat{X}^T \tilde{V}_i^T \Lambda_i \tilde{V}_i \hat{X} \end{aligned}$$

También

$$2\Delta_i^T P_i W_i^* \tilde{\phi}'_i \leq 2\Delta_i^T P_i (W_i - \tilde{W}_i) D_{\phi_i} \tilde{V}_i \hat{X} + \Delta_i^T P_i \bar{W}_i P_i \Delta_i + l_i \hat{X}^T \tilde{V}_i^T \Lambda_i \tilde{V}_i \hat{X}$$

Así que se obtiene:

$$\begin{aligned} \dot{S}_t &\leq \Delta_1^T [P_1 A + A^T P_1 + P_1 (\bar{W}_1 + \bar{W}_3) P_1 + (D_{\sigma_1} + \bar{u} D_{\phi_1} + Q_{01})] \Delta_1 \\ &+ \frac{1}{\epsilon} \Delta_2^T [P_2 B + B^T P_2 + P_2 (\bar{W}_2 + \bar{W}_4) P_2 + (D_{\sigma_2} + \bar{u} D_{\phi_2} + Q_{02})] \Delta_2 \\ &+ 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 - \Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 \\ &+ \sum_{i=1}^4 \text{tr} \left\{ \tilde{W}_i^T K_i^{-1} \tilde{W}_i \right\} + \sum_{j=1}^2 \sum_{i=1}^4 2\Delta_j^T P_j \tilde{W}_i (\sigma_j(V_i \hat{X}) - D_{\sigma_j} \tilde{V}_i \hat{X}_t) \\ &+ \sum_{i=1}^4 \text{tr} \left\{ \tilde{V}_i^T L_i^{-1} \tilde{V}_i \right\} + \sum_{j=1}^2 \sum_{i=1}^4 2\Delta_j^T P W_i D_{\sigma_j} \tilde{V}_i \hat{X} + l_i \hat{X}^T \tilde{V}_i^T \Lambda_i \tilde{V}_i \hat{X} \end{aligned}$$

Entonces usamos la ley actualizada (5.16) y (5.30), se obtiene:

$$\dot{S}_t \leq -\Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 + 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 \quad (5.29)$$

De la definición 1, si se define la entrada como \tilde{f}_i y la salida como $2P\Delta_i$, entonces el sistema es estrictamente pasivo con $V_t = \Delta_1^T Q_{01} \Delta_1 + \Delta_2^T Q_{02} \Delta_2 \geq 0$. ■

Comentario 5.7 Debido a que las ganancias de actualización son $K_i P_j$ y $L_i P_j$ ($i = 1 \dots 4$) y K_i y L_i puede ser cualquier matriz positiva, los procesos de aprendizaje de la red neuronal dinámica (5.16) no dependen de la solución P de la ecuación de Riccati (5.15). De modo que la consideración (5.30) es para seleccionar A tal que (5.15) tendrá solución positiva. R esta relacionada a las cotas superiores de las matrices óptimas desconocidas W_i^* asumiendo que se conocen las cotas superiores. Q es libre par ser elegida porque de Q_0 . Para las ecuación matricial (5.15) se puede cambiar A , R y Q tal que P es positiva, de tal manera que casi siempre es posible para satisfacer (5.30).

Comentario 5.8 $W_1 D_{\sigma_1} \Delta_1$ es el error de retroalimentación para la capa oculta, \hat{X}^T es la entrada de la capa oculta; $\sigma_1(V_1 \hat{X})$ es la entrada para la capa de salida, así que las primeras partes $K_1 P_1 \sigma_1(V_1 \hat{X}) \Delta_1^T$ y $K_3 P_1 \phi_1(V_3 \hat{X}) u \Delta_1^T$ son las mismas que el esquema de retropropagación de los perceptrones multicapa. Las segundas partes son usadas para asegurar las propiedades de pasividad del error de identificación. Si solo parámetros con incertidumbre están presentes ($\tilde{f}_i = 0$, $i = 1, 2$), entonces la ley de actualización como (5.16) puede hacer el error de identificación asintóticamente estable:

$$\lim_{t \rightarrow \infty} \Delta_i = 0, i = 1, 2 \quad (5.30)$$

Demostración. Debido a que las dinámicas del error de identificación (5.31) es pasivo y $\tilde{f}_i = 0$, de la propiedad 1 la función de almacenamiento S satisface:

$$\dot{S} \leq 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 = 0$$

La matriz definida positiva S implica que Δ_i y W_j son acotadas. De la ecuación del error

$$(5.31) \quad \dot{\Delta}_i \in L_\infty$$

$$\dot{S} \leq -\Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 \leq 0$$

Integrando (5.20) ambos lados:

$$\int_0^\infty \|\Delta_1\|_{Q_{01}} + \|\Delta_2\|_{Q_{02}} \leq S_0 - S_\infty < \infty$$

Así que $\Delta_i \in L_2 \cap L_\infty$, usando el Lemma de Barbalat [34] se tiene (5.21). De modo que u_t , σ_i , ϕ_i y P_i están acotados, $\lim_{t \rightarrow \infty} \dot{W}_j = 0$ y $\lim_{t \rightarrow \infty} \dot{W}_j = 0$. ■

Teorema 5.4 *Si la consideración (5.30) se satisface y existe una matriz definida positiva Λ_f tal que*

$$\begin{aligned} \lambda_{\max}(P_1 \Lambda_{f1} P_1) &\leq \lambda_{\min}(Q_{01}) \\ \lambda_{\max}(P_2 \Lambda_{f2} P_2) &\leq \lambda_{\min}(Q_{02}) \end{aligned} \quad (5.31)$$

la ley de actualización (5.16) puede hacer las dinámicas del neuro-identificador(5.31) estable de entrada-estado (ISS).

Demostración. En vista de la desigualdad matricial (5.18),

$$\begin{aligned} 2\Delta_1^T P_1 \tilde{f}_1 &\leq \Delta_1^T P_1 \Lambda_{f1} P_1 \Delta_1 + \tilde{f}_1^T \Lambda_{f1}^{-1} \tilde{f}_1 \\ 2\Delta_2^T P_2 \tilde{f}_2 &\leq \Delta_2^T P_2 \Lambda_{f2} P_2 \Delta_2 + \tilde{f}_2^T \Lambda_{f2}^{-1} \tilde{f}_2 \end{aligned} \quad (5.32)$$

(5.20) puede ser representada como

$$\begin{aligned} \dot{S}_t &\leq -\Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 + 2\Delta_1^T P_1 \tilde{f}_1 + 2\Delta_2^T P_2 \tilde{f}_2 \\ &\leq -\lambda_{\min}(Q_{01}) \|\Delta_1\|^2 - \lambda_{\min}(Q_{02}) \|\Delta_2\|^2 \\ &\quad + \Delta_1^T P_1 \Lambda_{f1} P_1 \Delta_1 + \tilde{f}_1^T \Lambda_{f1}^{-1} \tilde{f}_1 + \Delta_2^T P_2 \Lambda_{f2} P_2 \Delta_2 + \tilde{f}_2^T \Lambda_{f2}^{-1} \tilde{f}_2 \\ &\leq -\alpha_{\|\Delta_1\|} \|\Delta_1\| + \beta_{\|\tilde{f}_{t1}\|} \|\tilde{f}_{t1}\| - \alpha_{\|\Delta_2\|} \|\Delta_2\| + \beta_{\|\tilde{f}_{t2}\|} \|\tilde{f}_{t2}\| \end{aligned}$$

donde $\alpha_{\|\Delta_1\|} := [\lambda_{\min}(Q_{01}) - \lambda_{\max}(P_1 \Lambda_{f1} P_1)] \|\Delta_1\|$, $\beta_{\|\tilde{f}_1\|} := \lambda_{\max}(\Lambda_{f1}^{-1}) \|\tilde{f}_1\|$, $\alpha_{\|\Delta_2\|} := [\lambda_{\min}(Q_{02}) - \lambda_{\max}(P_2 \Lambda_{f2} P_2)] \|\Delta_2\|$, $\beta_{\|\tilde{f}_2\|} := \lambda_{\max}(\Lambda_{f2}^{-1}) \|\tilde{f}_2\|$. Se puede seleccionar matrices definidas positivas Λ_{f1} y Λ_{f2} tal que (5.22) es establecida. Así que α y β son \mathcal{K}_∞ funciones, S_t es una función ISS-Lyapunov. Usando el teorema 1 de [85], la dinámica del error de identificación (5.31) es estable de entrada-estado. ■

Corolario 5.3 *Si el error de modelado \tilde{f}_i esta acotado, entonces la ley de actualización (5.16) puede hacer el proceso de identificación estable:*

$$\Delta_i \in L_\infty, \quad W_j \in L_\infty$$

Demostración. De la propiedad 2 se conoce la estabilidad de entrada-estado por medio del comportamiento de las redes neuronales dinámicas permanecerán acotadas cuando sus entradas están acotadas. ■

Comentario 5.9 *Es bien conocido que las incertidumbres estructurales causará parámetros rígidos para control adaptivo, así que se usará la modificación robusta para hacer estable la identificación [34]. Métodos adaptivos robustos pueden ser extendidos para neuro-identificación directamente [35][66] [72]. Pero la neuro-identificación es un tipo de método de “caja negra”, no se necesita información de la estructura y todas las incertidumbres están dentro la “caja-negra”. Aunque los algoritmos adaptivos robustos son adecuados para neuro-identificación, no son tan sencillos. Por medio de la técnica de pasividad, exitosamente se prueba la conclusión: el algoritmo de retropropagación o actualización de aprendizaje (5.16) es robusto con respecto a todos los tipos de incertidumbres acotadas para neuro-identificación multicapa.*

La condición de estabilidad de entrada-estado (ISS), puede ser establecida si Λ_f es seleccionada como una matriz pequeña. De tal manera que el estado y variables de salida son físicamente acotadas, el error de modelado \tilde{f}_t puede ser considerado estar acotado también (ver, por ejemplo) [35][66][72].

Para el caso del modelo que combina bien, el análisis tipo Lyapunov [102] puede alcanzar los mismos resultados como en el Corollary 1. Pero en el caso del error de modelado ($\tilde{f}_t \neq 0$), los términos de modificación robusta tienen que ser adicionados en la ley de actualización para asegurar la estabilidad [35][66] [72]. La modificación robusta generalmente depende de las cotas superiores del error de modelado \tilde{f}_t . Como en el caso de leyes adaptivas robustas, tal como zona-muerta [66] y modification- σ [61], la ley de actualización no necesita las cotas superiores de incertidumbres.

Teorema 5.5 Si el error de modelado $\tilde{f}_1(V_1^0, V_3^0)$ y $\tilde{f}_2(V_2^0, V_4^0)$ están acotadas como $\tilde{f}_1^T \Lambda_1 \tilde{f}_1 \leq \bar{\eta}_1$, $\tilde{f}_2^T \Lambda_2 \tilde{f}_2 \leq \bar{\eta}_2$, P_1 y P_2 son la solución de la ecuación de Riccati (5.15) con:

$$\begin{aligned} R_1 &= 2\bar{W}_1 + 2\bar{W}_3 + \Lambda_1, & Q_1 &= D_{\sigma_1} + \bar{u}D_{\phi_1} + Q_{01} \\ R_2 &= 2\bar{W}_2 + 2\bar{W}_4 + \Lambda_2, & Q_2 &= D_{\sigma_2} + \bar{u}D_{\phi_2} + Q_{02} \end{aligned} \quad (5.33)$$

entonces la ley de actualización (5.16) puede hacer que el error de identificación converja a:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\|\Delta_1\|_{Q_{01}}^2 + \|\Delta_2\|_{Q_{02}}^2 \right) dt \leq \bar{\eta}_1 + \bar{\eta}_2 \quad (5.34)$$

Demostración. Se define una función de Lyapunov como (5.17), en vista de la desigualdad matricial (5.18)

$$\begin{aligned} 2\Delta_1^T P_1 \tilde{f}_1 &\leq \Delta_1^T P_1 \Lambda_1 P_1 \Delta_1 + \tilde{f}_1^T \Lambda_1^{-1} \tilde{f}_1 \leq \Delta_1^T P_1 \Lambda_1 P_1 \Delta_1 + \bar{\eta}_1 \\ 2\Delta_2^T P_2 \tilde{f}_2 &\leq \Delta_2^T P_2 \Lambda_2 P_2 \Delta_2 + \tilde{f}_2^T \Lambda_2^{-1} \tilde{f}_2 \leq \Delta_2^T P_2 \Lambda_2 P_2 \Delta_2 + \bar{\eta}_2 \end{aligned}$$

Usando la ley actualizada (5.16), la derivada de la función de Lyapunov (5.17) es:

$$\begin{aligned} \dot{S}_t &\leq \Delta_1^T [P_1 A + A^T P_1 + P_1 (\bar{W}_1 + \bar{W}_3 + \Lambda_1) P_1 + (D_{\sigma_1} + \bar{u}D_{\phi_1} + Q_{01})] \Delta_1 - \Delta_1^T Q_{01} \Delta_1 + \bar{\eta}_1 \\ &\quad + \frac{1}{\epsilon} \Delta_2^T [P_2 B + B^T P_2 + P_2 (\bar{W}_2 + \bar{W}_4 + \Lambda_1) P_2 + (D_{\sigma_2} + \bar{u}D_{\phi_2} + Q_{02})] \Delta_2 - \Delta_2^T Q_{02} \Delta_2 + \bar{\eta}_2 \end{aligned}$$

De (6.6) se obtiene:

$$\dot{S}_t \leq -\Delta_1^T Q_{01} \Delta_1 - \Delta_2^T Q_{02} \Delta_2 + \bar{\eta}_1 + \bar{\eta}_2 \quad (5.35)$$

Integrando (6.8) de 0 hasta T produce:

$$S_T - S_0 \leq -\int_0^T [\Delta_1^T Q_{01} \Delta_1 + \Delta_2^T Q_{02} \Delta_2] dt + (\bar{\eta}_1 + \bar{\eta}_2) T$$

Así que:

$$\int_0^T [\Delta_1^T Q_{01} \Delta_1 + \Delta_2^T Q_{02} \Delta_2] dt \leq S_0 - S_T + (\bar{\eta}_1 + \bar{\eta}_2) T \leq S_0 + (\bar{\eta}_1 + \bar{\eta}_2) T$$

(6.7) es establecido. ■

Comentario 5.10 *El error de identificación tendrá convergencia al radio de la bola la cota superior de \tilde{f}_i , y es influenciada por la matriz conocida a priori V_i^0 . El teorema 2 muestra que V_i^0 no tiene influencia en la propiedad de estabilidad, pudiendo seleccionar cualesquiera valores para V_i^0 en primera. Del teorema 3 se conoce el algoritmo (5.16) que puede hacer el error de identificación convergente. V_i^0 puede ser seleccionado por los siguientes pasos off-line:*

1. Empezar de cualesquiera valores iniciales para V_i
2. Hacer identificación on-line con V_i^0
3. Sea V_i como una nueva condición inicial, i.e., $V_i^0 = V_i$
4. Si el error de identificación decrece, repetir el proceso de identificación, hasta 2. De lo contrario, detener la identificación off-line, ahora $V_{1,t}$ y $V_{2,t}$ son valores finales para V_i^0 .

Después de que se obtiene este final V_i^0 , se puede empezar la identificación on-line con ellos.

Comentario 5.11 *Debido a que la razón de actualización en el algoritmo de aprendizaje (5.16) es $K_i P$ ($i = 1 \dots 4$), y K_i puede ser seleccionado como cualquier matriz positiva, el proceso de aprendizaje de la red neuronal dinámica (5.16) esta libre de las soluciones de las dos ecuaciones de Riccati (5.15) y (6.6). Estas dos ecuaciones de Riccati son propuestas para probar la estabilidad resultante. Cuando usamos la ley de actualización para los pesos, solo necesitamos seleccionar buenas razones o contribuciones de actualización $k_i = K_i P$.*

5.3. Simulación

Para validar los resultados teóricos, se dan las siguientes simulaciones. La red neuronal dinámica es

$$\begin{aligned}\dot{\hat{x}} &= A\hat{x} + W_1\sigma_1(\hat{x}, \hat{z}) + W_3\phi_1(\hat{x}, \hat{z})u \\ \dot{\hat{z}} &= B\hat{z} + W_2\sigma_2(\hat{x}, \hat{z}) + W_4\phi_2(\hat{x}, \hat{z})u\end{aligned}\tag{5.36}$$

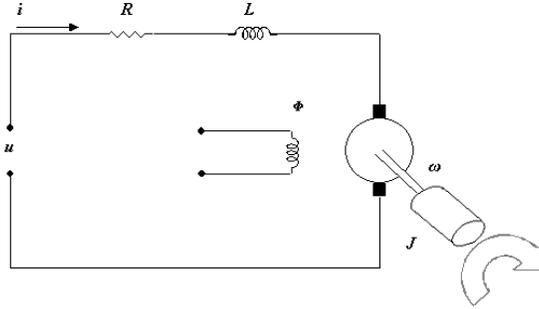


Figura 5.3: DC motor

Estas redes neuronales dinámicas con dos escalas de tiempo puede ilustrar la dinámica de un motor de D.C. como se muestra en Fig.5.3.

El modelado del servomotor de D.C. puede ser dividido en dos subsistemas eléctrico y mecánico. El subsistema eléctrico esta basado en la ley de voltaje de Kirchhoff:

$$L \frac{di}{dt} = -k\omega - Ri + u$$

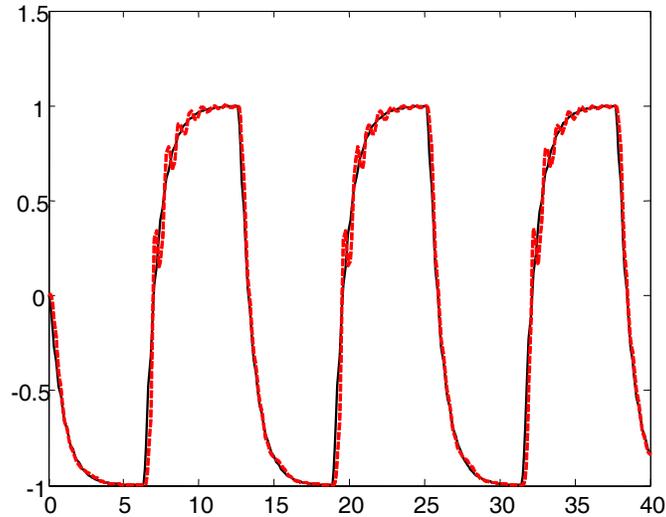
donde u es el voltaje de entrada, i es la corriente de armadura, R y L son la resistencia e inductancia de armadura, K es la fuerza electromotriz constante. El subsistema mecánico es:

$$J \frac{d\omega}{dt} = ki$$

donde J es el momento de inercia, donde k es el par constante del motor. Por una transformación de $\omega_r = \frac{\omega}{J}$, $i_r = \frac{iR}{Jk}$, $u_r = \frac{u}{Jk}$, (5.25) y (5.26) se convierte en

$$\begin{aligned} \dot{\omega}_r &= i_r \\ \varepsilon \dot{i}_r &= -\omega_r - i_r + u_r \end{aligned}$$

donde $\varepsilon = \frac{Lk^2}{JR^2} \ll 1$, i_r es el estado rápido.

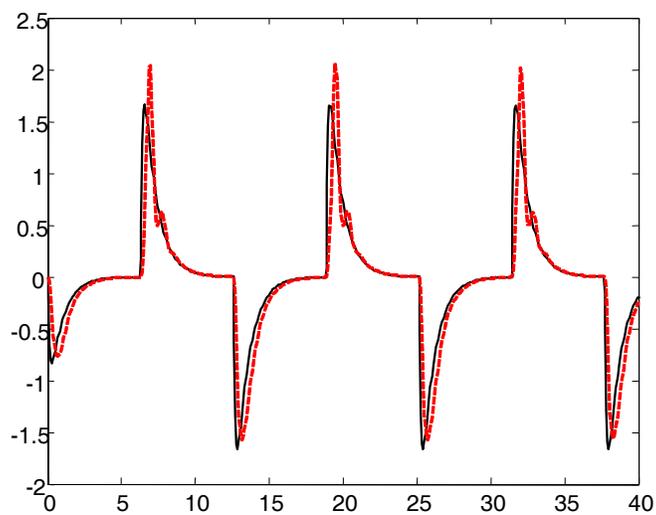
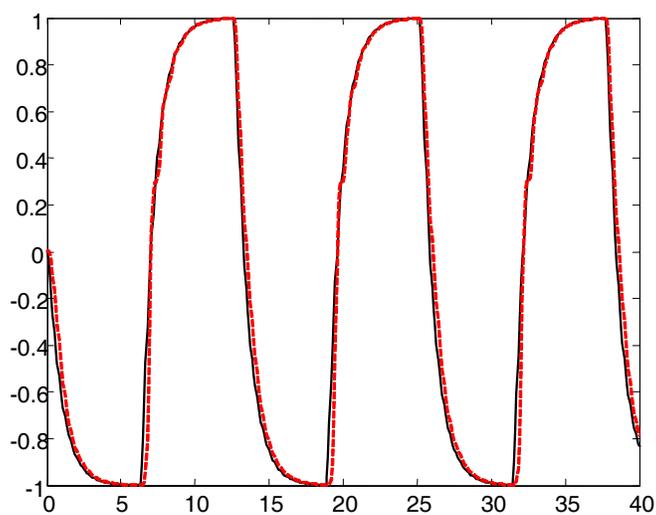
Figura 5.4: Identificación de ω_r

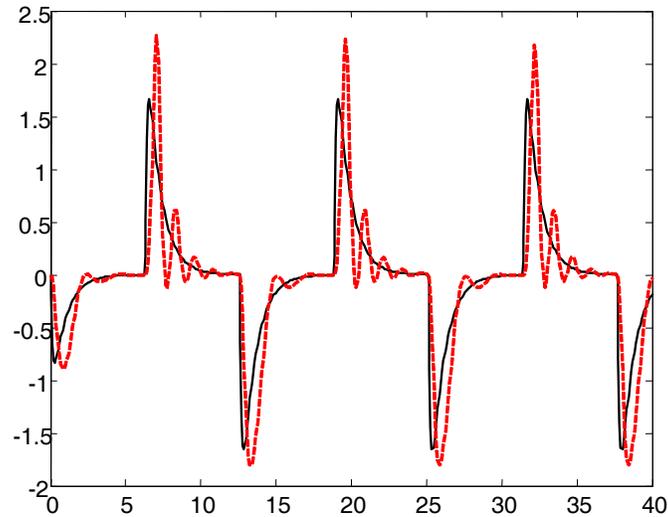
Para las redes neuronales dinámicas se elige \hat{x} y \hat{z} que son escalares, $A = -2$, $B = -2$, $\epsilon = 0,05$. Las condiciones iniciales son $\hat{x}_0 = 1$, $\hat{z}_0 = 1$. Los resultados de identificación son mostrados en Fig.5.4 y Fig.5.5.

Pero si se utilizan redes neuronales normales $\dot{\hat{x}} = A\hat{x} + W_1\sigma_1(\hat{x}, \hat{z}) + W_3\phi_1(\hat{x}, \hat{z})u$ para identificar i_r , los resultados son mostrados en las Fig.5.6 y Fig.5.7. Pdiendo ver que para el estado lento ω_r los resultados de identificación son casi los mismos, pero para el estado rápido i_r las redes neuronales con diferentes escalas de tiempo tienen mejor realización o comportamiento que las redes neuronales normales.

5.4. Conclusión del capítulo

En esta sección, el enfoque de pasividad es usado para probar que el algoritmo del gradiente descendente para el ajuste de pesos del neuro-identificador con escalas de tiempo diferentes es estable y robusto para cualquier incertidumbre acotada. Se obtienen algunos

Figura 5.5: Identificación de i_r .Figura 5.6: Identificación de ω_r .

Figura 5.7: Identificación de i_r .

nuevos resultados sobre la estabilidad de identificación neuronal con escalas de tiempo diferentes. Un algoritmo de aprendizaje del gradiente simple puede hacer el neuroidentificador pasivo, estable, asintóticamente estable, estable de entrada-estado y estable acotado a la entrada.

Además por medio de la técnica de pasividad, se dan algunos nuevos resultados sobre neuro-identificación con redes neuronales dinámicas. Comparadas con otros análisis de estabilidad de neuro-identificación, el algoritmo planteado es más simple porque modificaciones robustas no son aplicadas, así que el algoritmo propuesto en esta sección es más adecuado para aplicación ingenieril. Logrando probar que incluso el algoritmo de aprendizaje de retroalimentación puede garantizar la identificación del error estable robusto.

Capítulo 6

Aplicación a Robots de unión flexible

6.1. Introducción

La robótica es una disciplina tecnológica en constante crecimiento de la tecnología moderna que cruza los límites tradicionales de la ingeniería. Para entender la complejidad de los robots y sus aplicaciones se requiere conocimientos de ingeniería eléctrica, mecánica, industrial, ciencias de la computación, economía, y matemáticas. Nuevas disciplinas de la ingeniería, tales como ingeniería de manufactura, aplicaciones de ingeniería, y conocimiento ingenieril, están empezando a emerger para enfrentar la complejidad del campo de la robótica y su larga área de automatización industrial.

A pesar de tener relativamente pocos años de investigación, es posible que la ingeniería en robótica se planteará como una disciplina de ingeniería propia y distinta.

En esta investigación se explorarán la cinemática, dinámica y control de manipuladores robóticos. Haciendo esto se omiten muchas otras áreas tales como locomoción, visión de máquina, inteligencia artificial, arquitectura de computadoras, lenguajes de programación, diseño enfocado a la computación, tacto, agarre, y manipulación, que colectivamente componen la disciplina conocida como robótica. Mientras que los temas de las áreas que se han omitido son importantes para la ciencia de la robótica, para un firme entendimiento de la cinemática, dinámica, y control de manipuladores, es básico para el entendimiento y apli-

cación de estas otras áreas, para que un primer curso en robótica se pueda empezar con estos tres temas.

El término robot fué primero introducido en el vocabulario por el dramaturgo Karel Capek en su obra de 1920 *Rossum's Universal Robots*, la palabra robot indica la palabra checa "trabajo". Desde entonces el término ha sido aplicado para una gran variedad de dispositivos mecánicos, tales como tele-operadores, vehículos submarinos, excavadores de tierra autónomos, etc. Virtualmente cualquiera que opera con algún grado de autonomía, generalmente bajo control computarizado, tendrá algún punto que será llamado robot.

6.1.1. Componentes y estructura de robots

Los manipuladores robóticos están compuestos de eslabones conectados por juntas o uniones en una cadena cinemática abierta. Las juntas son típicamente rotatorias (revolución) o lineales (prismáticas). Una junta de revolución es como una bisagra y permite rotación relativa entre dos eslabones. Una junta prismática permite un movimiento relativo lineal entre dos eslabones. Se usará la convención (R) para representar juntas de revolución y (P) para juntas prismáticas. Cada junta representa la interconexión entre dos eslabones, se dice que, l_i y l_{i+1} .

Denotando el eje de rotación de una junta de revolución, o el eje a lo largo en que la junta prismática se desliza, por z_i si la unión es la interconexión de eslabones i y $i + 1$. Las juntas variables, denotadas por θ_i para una junta de revolución y d_i para una junta prismática, representan el desplazamiento relativo entre eslabones adyacentes. Las juntas de un manipulador pueden ser eléctricas, hidráulicas o actuadores neumáticos. El número de juntas o uniones determina los grados de libertad (DOF) del manipulador. Típicamente, una manipulador puede poseer al menos seis grados de libertad independientes: tres de posición y tres de orientación. Con pocos menos de seis (DOF) el brazo no puede alcanzar cualquier punto en su entorno de trabajo con orientación arbitraria. Ciertas aplicaciones tales como accesos circundantes o evitar obstáculos requiere más de seis grados de libertad (DOF). La dificultad de controlar un manipulador se incrementa rápidamente con el número de

eslabones. Un manipulador teniendo más grados de libertad que el mínimo necesario para posicionar y orientar un objeto es llamado Robot Redundante.

El espacio de trabajo de un manipulador es el espacio total amplio barrido por el efector final cuando el manipulador ejecuta todos los posibles movimientos. El espacio de trabajo esta restringido por la geometría del manipulador además de sus restricciones mecánicas en las juntas. Por ejemplo, una junta de revolución puede ser limitada para un valor menor que un movimiento de 360^0 . El espacio de trabajo a menudo es seccionado en un espacio de trabajo accesible y en un espacio de trabajo diestro. El espacio de trabajo accesible es el conjunto entero de puntos accesibles por el manipulador, mientras el espacio de trabajo diestro consiste de estos puntos que el manipulador puede alcanzar con una orientación arbitraria del efector final. Obviamente el espacio de trabajo es un subconjunto del espacio de trabajo accesible.

Un manipulador robótico puede ser visto como una serie de conexiones mecánicas. El brazo mecánico es justamente un componente para un sistema robótico completo, como se muestra en la 6.1, que consiste del brazo, fuente de poder externa, herramienta de brazo final, sensores externos e internos, servo, interface de cómputo, y computadora de control. Todo el software de programación puede ser considerado como parte integral del sistema completo, debido a que la manera en que el robot es programado y controlado puede tener un mejor impacto en su ejecución y subsecuente rango de aplicaciones.

Un robot es un manipulador reprogramable y multifuncional, diseñado para mover herramientas, materiales o dispositivos especializados, a través de movimientos programados variables, para el desempeño de una variedad de tareas y, básicamente están compuestos de juntas o uniones y eslabones. Como ya se mencionó antes, el número de juntas determina el número de grados de libertad (*DOF*). Los robots se pueden clasificar de varias formas:

1. De configuración articular (Antropomórfica)
2. De configuración esférica
3. SCARA (Selective Compliant Articulated Robot for Assembly)
4. Cartesiano
5. Cilíndrico

Los manipuladores robóticos son sistemas altamente no lineales, sistemas electro-mecánicos multi-ejes dinámicamente acoplados y que son comúnmente usados en tareas tales como pintado en aerosol, soldadura y manejo de material en manufactura, y en sistemas de posicionamiento de precisión, como por ejemplo, proyectiles anti-aereos, plataformas de misiles, máquinas de operación submarina, etc. En estas tareas, los efectores finales son requeridos para moverse de un lugar a otro, o para seguir alguna trayectoria deseada tan cerca como sea posible en una velocidad operacional rápida en un espacio de trabajo libre. El control de seguimiento de una trayectoria de un robot, es por lo tanto, de significado práctico, y es todavía simple la mayor tarea fundamental en control de robots [56]. Llevado por el deseo por un alto grado de automatización, la operación rápida de velocidad, y alto requerimiento de tolerancia de la industria, trabajo de investigación extensa ha sido llevado al diseño de controladores. Crecientes sofisticadas herramientas de teoría de control no lineal han sido desarrolladas para mejor ejecución de seguimiento. Avances actuales en tecnología de micro-procesadores han hecho la implementación de complicados algoritmos de control no lineal prácticamente factibles. Métodos de control diferentes han sido introducidos en la literatura tales como control PD de unión independienete [51], control de compensación hacía adelante [42], cómputo de control de par [24], modelo basado en control adaptivo [3], estructura de control variables [83], y control con redes neuronales [46].

A pesar del hecho de que en muchos sistemas, un controlador lineal opera satisfactoriamente, una consideración clave es hacer que el sistema opere sobre un pequeño rango estrechamente controlado, en que las dinámicas del sistema sean lineales. Sin embargo, el rango de operación normal de un robot es bastante grande, y su carga útil también varía. El modelo lineal obtenido en un punto de operación no es válido cuando el punto de operación cambia excepto para casos simples. El controlador lineal llevado a cabo es deficiente porque las no linealidades en el sistema no están apropiadamente compensadas. El control de unión independiente clásico de un robot esta diseñado basado en un modelo lineal teórico despreciando las fuerzas de acoplamiento no lineales asociadas con los movimientos mecánicos del robot. Actualmente, las mayoría de los controladores de robots industriales en uso están diseñados considerando cada unión independiente de los otras, y aplicando su control

proporcional más su derivada (*PD*) o control proporcional más su integral más su derivada (*PID*) para cada unión independiente de las otras. Los grados de ejecución de control como la operación de velocidad rápidamente se incrementan. Para alcanzar una precisión pre-especificada, la velocidad de movimiento y de acuerdo, al mejoramiento de productividad, esta ha sido restringida severamente. Como resultado, un manipulador robótico controlado en esta manera es solamente apropiado para movimiento relativamente lento y limitada tarea de precisión.

El control lineal no puede conocer las siempre crecientes demandas para realización de mejor control y rápida velocidad de operación de la industria. Diferentes técnicas de control no lineal están, sin embargo siendo investigadas. Avances actuales en tecnología de microprocesadores también han hecho factible la implementación práctica de complicados algoritmos de control no lineal. El cómputo de control del par es uno de los controladores más intuitivos, que ayuda en la cancelación exacta de todas las dinámicas no lineales del sistema. Las ventajas de tal controlador incluyen: a) el conocimiento exacto de las no linealidades es requerido, b) estabilidad de lazo cerrado robusta es difícil para garantizar si las aceleraciones no son medibles, que hace que el esquema sea poco atractivo para la industria.

Los robots pueden manipular cargas de diferentes pesos, tamaño y distribuciones de masa, en cuyo caso, las dinámicas de los robots también serán diferentes. El diseño de controladores para una carga útil nominal particular puede no ser capaz de controlar el sistema apropiadamente para todos los cambios en los parámetros. Sistemas de control adaptivo pueden ser diseñados para esta situación. Es bien conocido que las dinámicas de los robots son lineales en los parámetros. Algunos modelos basados en controladores adaptivos han sido propuestos basados linealmente en los modelos paramétricos. Los controladores son estables y robustos bajo ciertas condiciones y el llamado mecanismo de adaptación directa ha sido frecuentemente usado, que computa los parámetros del controlador directamente del error de seguimiento y los estados del sistema. El sistema puede alcanzar seguimiento asintótico sin usar retroalimentación de ganancia infinita a todo lo largo en la adaptación paramétrica cuando el sistema está sujeto solo a incertidumbres paramétricas [8] y [9]. Sin embargo, la

inestabilidad puede ocurrir debido a algunas perturbaciones inciertas acotadas o plantas no lineales [71].

Aunque algunas técnicas de modificación para la ley de adaptación tipo integral tal como modificación- σ [71] puede ser empleada para ensanchar la robustez del sistema, la precisión de seguimiento cero no puede ser grande para ser garantizado debido a que el error de seguimiento de estado continuo puede solo ser mostrado para permanecer en una esfera desconocida cuyo tamaño depende de las perturbaciones.

Para incertidumbres paramétricas de entrada y reducir la dependencia en los modelos, el control de estructura variable o control en modo deslizante [93] fue introducido, cuyo empleo de ciertas estructuras de control para atenuar los efectos de ambas incertidumbres paramétricas e incertidumbres no lineales. Sin modificación, este esquema de control inevitablemente introducirá el chattering indeseable en el sistema. En general, puede garantizarse la ejecución del transitorio y cierta precisión de seguimiento final. El controlador diseñado es generalmente muy conservativo porque no discrimina incertidumbres paramétricas de perturbaciones no lineales. Para reducir los errores de seguimiento, las ganancias de retroalimentación deben ser incrementadas, resultando en retroalimentación de alta ganancia e incrementando los requerimientos de ancho de banda del sistema de lazo cerrado. Sin embargo, el ancho de banda de los sistemas de lazo cerrado están limitados por el diseño mecánico de robots. Algunos esfuerzos se han hecho combinando las ventajas del modelado basado en control adaptivo y control robusto determinístico, el control adaptivo [18] han sido propuestos con énfasis en a) la selección de la estructura del controlador, y b) adaptación paramétrica para reducir las incertidumbres paramétricas. Por lo tanto, el problema de la realización del transitorio desconocido y la no robustez para no linealidades inciertas de control adaptivo pueden resultar inmediatamente. Por lo contrario para el control de modo deslizante puro, la aproximación discrimina la diferencia entre incertidumbres paramétricas e incertidumbres no lineales y el uso de adaptación paramétrica para reducir las incertidumbres paramétricas.

La ley de adaptación de compensación deseada ya propuesta, fue más adelante demostrada experimentalmente en [97], alcanzando una realización de seguimiento superior entre esquemas adaptivos existentes. Porque el regresor es calculado por referencia solamente de la

información de la trayectoria, la ley de adaptación resultante es menos sensible a señales de velocidad con ruido y tendrá mejor robustez en adición a cómputo on-line largamente reducido. La misma idea ha sido posteriormente extendida a control robusto adaptivo, y el diseño del controlador de robots requiere dinámicas del motor para evitar las medidas de velocidad y aceleración.

Sin embargo, el modelo basado en métodos de control depende cada uno en un modelo exacto o en una estructura de modelo exacto, tal como las funciones del regresor conocidas de la lineales en las dinámicas paramétricas de robots. A pesar de la existencia de programas computacionales simbólicos, es todavía muy tedioso y consume mucho tiempo para obtener las ecuaciones dinámicas de los robots, y por lo tanto no se mencionan las funciones de regresor explícito, excepto para casos muy simples. Es también extremadamente susceptible a errores hasta expresar las dinámicas como un producto del regresor y el vector de parámetros. Siempre para robots tienen muy pocas modificaciones mecánicas, el mismo proceso de modelado puede ser repetido para el diseño del controlador, no para mencionar que los robots tienen diferentes estructuras y diferentes números de grados de libertad (*DOF*).

Recientemente, algunas sugerencias han sido hechas para usar redes neuronales para mejorar la necesidad de modelos parametrizados linealmente. En muchos esquemas, las redes neuronales son primero entrenadas “off-line” para obtener suficiente aproximación de precisión en las funciones de entrada-salida de los sistemas, y las redes pueden ser apropiadamente usadas para construir controladores. Otros intentos fueron hechos para entrenar la red neuronal “on-line” durante las operaciones normales del sistema [46].

Es conocido que las redes neuronales solo tienen la capacidad de mapeo estático [73]. Sin embargo, cuando son usadas como controladores, estas deben ser capaces para realizar las dinámicas del sistema [46]. Apesar del hecho de que las redes neuronales son muy poderosas en el aprendizaje de complicadas dinámicas, el tamaño de las redes se conocen ser muy grandes, especialmente las dinámicas, que subsecuentemente conduce a la necesidad de poderosas instalaciones computacionales.

Mientras es verdad que las redes neuronales tienen capacidad de aprendizaje y pueden ser usadas para aproximar cualquier función dinámica con cualquier grado de precisión mientras

el tamaño de la red es suficientemente grande, no podemos abusar de sus capacidades y dejamos que aprendan todas las características de los sistemas sin ninguna discriminación. Se puede mostrar que, para explotar completamente las propiedades estructurales de un sistema particular, las dinámicas completas del sistema pueden ser aproximadas por redes neuronales estáticas de estructuras predeterminadas [20]. Como resultado, el tamaño del modelo de red neuronal de robots se tornará más pequeño que las dinámicas usuales. Esto es muy deseable para implementación actual en términos de menor poder computacional requerido, y como consecuencia, menor costo.

Porque de la naturaleza de aproximación subyacente de las redes neuronales, la estabilidad de lazo cerrado esta garantizada solo si los errores de aproximación de las redes neuronales correspondientes para diferentes funciones no lineales son tomados en cuenta. Un control de modo deslizante es a menudo usado para suprimir los errores y proporcionar la robustez de lazo cerrado. Otros problemas asociados con controladores con redes neuronales incluyen la dificultad en elegir el número de nodos, el número de capas, la elección de funciones básicas, la localización de cuadrícula y otros. Por más grande que sea el tamaño de la red neuronal, la generalización del error siempre existe.

En esta tesis de investigación, fundamentalmente, se usarán redes neuronales dinámicas, para modelar e identificar robots de unión flexible y aunque no se diseñarán controladores, si se mencionarán algunos tópicos relacionados con ellos.

6.2. Modelo de un robot de unión flexible

6.2.1. Modelado dinámico de robots

El modelado dinámico de manipuladores trata con la formulación matemática de las ecuaciones dinámicas del movimiento de un brazo de robot. La construcción del modelo es justificado:

1. Las simulaciones de la respuesta dinámica de robots proporciona un medio de diseño de prototipos de manipuladores y prueba estrategias de control sin incurrir en grandes gastos

de trabajo con robots actuales.

2. Todo buen regulador de un sistema necesita incluir, explícitamente o implícitamente, un modelo de tal sistema, es decir, éxito en la regulación del sistema implica que un modelo suficientemente similar puede primero ser construido.

3. El control basado en un modelo es superior a un control no basado en un modelo.

Historicamente, existen básicamente dos formulaciones diferentes, la forma cerrada o *formulación Euler-lagrange* y la forma recursiva hacia adelante-hacia atrás *formulación Euler-Newton*. El primero trata al manipulador como un todo y ejecuta el análisis usando la función Lagrangiana (*la diferencia entre la energía cinética y la energía potencial del sistema robótico*) basado en ambos el método de desplazamientos virtuales o el principio de Hamilton de acción mínima. El segundo, por lo contrario, trata cada eslabón del robot en turno e involucra ecuaciones que describen su movimiento lineal y su movimiento angular. Las dinámicas del sistema completo pueden ser descritas por las llamadas ecuaciones recursivas hacia adelante-hacia atrás. Inicialmente se consideró que la formulación de Euler-Newton era más adecuada para cómputo recursivo que la formulación Euler-Lagrange, cuando los grados de libertad, n , son grandes. Sin embargo, la visión actual es que ambas formulaciones son equivalentes [82]. La principal razón para tener dos diferentes tipos de formulaciones es que estas proporcionan significados diferentes.

El análisis dinámico del sistema significa encontrar la relación entre las coordenadas generalizadas q y las fuerzas generalizadas τ . Puede ser justo decir que la formulación Euler-Newton es superior para determinar las fuerzas generalizadas necesarias para realizar una evolución en el tiempo dada de las coordenadas generalizadas, de tal manera que solo existe interés en conocer la función dependiente en el tiempo $\tau(t)$ que produce una trayectoria particular $q(t)$. No obstante, para resolver la evolución en el tiempo de las coordenadas generalizadas y la facilidad del diseño del controlador, en esta tesis, solo la formulación de Euler-Lagrange será discutida.

El concepto de modelado estructural de red usando las mismas funciones del sistema han sido propuesto en [21] para eliminar los problemas asociados con la naturaleza de aproximación de redes neuronales. En la práctica, el tamaño máximo de la red estructural es fija.

En [46], el modelado estructural de redes de robots ha sido brevemente discutido usando el operador de *Kronecker*.

El modelado dinámico y el diseño de controladores para robots de unión flexible ha atraído la atención de muchos investigadores desde que fueron indicados para muchos robots industriales, particularmente los equipados con manipuladores armónicos usados para reducción de velocidad, exhibiendo significativa flexibilidad de unión [91]. También se ha mostrado que en el caso donde los robots son manejados por motores de ejecución directa, la ejecución falsa de los manipuladores directos también tiene flexibilidad torsional [4]. Otras fuentes de flexibilidad de unión surge de los dientes de los engranajes, correas de transmisión y transductores insertados en uniones para medir el par de la unión. Por lo tanto, si alta ejecución es requerida, la flexibilidad de la unión debe ser tomada en cuenta en modelado y control. Para la última década, considerable trabajo de investigación ha sido tomado en el modelado y control de robots de unión flexible [5].

Cuando la unión flexible esta incluida en el modelo dinámico usado para el diseño de un sistema de control, muchas de las propiedades estructurales matemáticas que facilitan el diseño de control para robots rígidos, tales como actuación completa, es decir un control independiente para cada grado de libertad, y la pasividad de los pares del motor para las velocidades de eslabón, no son muy aplicables. Lo que es más, el número de grados de libertad dobles y la propiedad de competición entre no linealidades y las entradas es desaprovechada. La flexibilidad de unión también resulta en resonancia estructural especialmente durante movimiento de alta velocidad, que reduce las limitaciones del ancho de banda en cualquier algoritmo de control que es diseñado en la base de rigidez absoluta. Por lo tanto, la realización de las técnicas de control de un robot de cuerpo rígido bien establecidas son limitadas o hasta inaceptables. Dos enfoques de modelado y control bien establecidas para robots de unión flexible son la técnica de *perturbación singular* y el método de *linealización con retroalimentación* [86].

El modelado y control adaptivo de robots de unión flexible han sido estudiados por muchos investigadores [11], en un intento para mejorar la robustez de sistemas de lazo cerrado para incertidumbres paramétricas y cambios de modelo tales como variaciones en la carga útil. Uno

de los controladores adaptivos simples para robots de unión flexible fue sugerido basado en la *Teoría de Perturbación Singular* usando solo posición y velocidad con retroalimentación. La idea está basada en la adición de un término de corrección para la ley de control para robots de cuerpos rígidos para amortiguar la oscilación elástica en las uniones, y modelando las fuerzas elásticas de la unión como las variables rápidas y las variables del eslabón como las variables lentas. Algunos trabajos posteriores [11] están basados en el resultante modelo de perturbación singular. Aunque la mayoría de los resultados son solo válidos para intervalos finitos, estos pueden ser fácilmente extendidos para mantenerse dentro de un intervalo infinito como el resultado de dinámicas de tiempo lentas que coinciden con un modelo de robot rígido, para que exponencialmente existan controladores estables. Estudios comparativos entre controladores en intervalos de tiempo finitos e infinitos fueron ya considerados.

Otro controlador adaptivo simple fue presentado en [18] basado en *Teoría de Perturbación Singular* usando solo posición y velocidad retroalimentada, también por el modelado de los errores de seguimiento del motor Ke_m como las variables rápidas en lugar de las fuerzas elásticas de la unión, donde K representa la rigidez de la unión y e_m son los errores de seguimiento del motor. Como una consecuencia, las dinámicas lentas son también diferentes de las que se dan en cierta literatura; existen las leyes de control resultante. Por lo tanto se muestra que el modelo de *perturbación singular* para el mismo sistema no es único. Elecciones diferentes para variables rápidas proporciona diferentes significados en el problema del diseño del controlador de robots de unión flexible.

6.2.2. Modelo de un robot de unión flexible

Representación en espacio de estado

Es deseable expresar las dinámicas de robots en variables de espacio cartesiano o espacio de tarea que en lugar de variables de espacio de unión q debido a que las tareas de manipuladores robóticos son a menudo expresados en espacios cartesianos. Por simplicidad, solamente robots no redundantes son considerados aquí. Sea $x = [\mathbf{r}^T \ \boldsymbol{\theta}^T]^T \in \mathbb{R}^n$ con \mathbf{r} y $\boldsymbol{\theta}$ que pertenecen a la posición y orientación en el marco base o de referencia. De acuerdo a la

cinemática directa, \mathbf{x} puede ser expresada como una función no lineal de \mathbf{q} como $\mathbf{x} = \mathbf{h}(\mathbf{q})$.

Por lo tanto, se tiene:

$$\dot{x} = J(q) \dot{q} \quad (6.1)$$

$$\ddot{x} = J(q) \ddot{q} + \dot{J}(q) \dot{q} \quad (6.2)$$

donde la matriz jacobiana $J(q)$ es definida como $J(q) \triangleq \frac{\partial h(q)}{\partial q}$ que es una matriz cuadrada. Asumiendo que el manipulador robótico esta a distancia del espacio de trabajo de singularidades, es decir, $|J(q)| \neq 0$ y $J^{-1}(q)$ existe.

Sea F_x la fuerza en el espacio cartesiano que causa el cambio en x . La relación entre F_x y τ esta dada por:

$$\tau = J^T(q) F_x \quad (6.3)$$

Sustituyendo las ecuaciones (6.1) y (6.2) en la ecuación matricial siguiente:

$$D(q) \ddot{q} + C(q, \dot{q}) \dot{q} + G(q) = \tau \quad (6.4)$$

produce las dinámicas cartesianas como:

$$D_x(q) \ddot{x} + C_x(q, \dot{x}) \dot{x} + G_x(q) = F_x \quad (6.5)$$

donde:

$$\begin{aligned} D_x(q) &= J^{-T}(q) D(q) J^{-1}(q) \\ C_x(q, \dot{q}) &= J^{-T}(q) \left(C(q, \dot{q}) - D(q) J^{-1}(q) \dot{J}(q) \right) J^{-1}(q) \\ G_x(q) &= J^{-T}(q) G_x(q) \end{aligned}$$

Puede verse que $D_x(q)$, $C_x(q, \dot{q})$, y $G_x(q)$ son funciones de q y \dot{q} . Por lo tanto hablando estrictamente, las dinámicas cartesianas no son completamente dadas en términos de x, \dot{x}, \ddot{x} .

La mayoría de las propiedades para dinámicas de espacio de unión de robots son aplicables para dinámicas de espacios cartesianos mientras $J(q)$ es no singular [84] y [46]. Por ejemplo

1. $D_x(q)$ es simétrica, definida positiva y acotada por arriba y abajo.
2. La matriz $N_x(q) \triangleq \dot{D}_x(q) - 2C_x(q, \dot{q})$ es simétrica-torcido si $C(q, \dot{q})$ es definido usando los símbolos de Christoffel.
3. La propiedad de linealidad en los parámetro se cumple, es decir,:

$$D_x(q) \ddot{x}_r + C_x(q, \dot{q}) \dot{x}_r + G_x(q) = Y_x(x, \dot{x}, \ddot{x}_r) P_x$$

donde P_x es el vector de los parámetros del robot, y $Y_x(x, \dot{x}, \ddot{x}_r)$ es la matriz regresiva cartesiana conocida.

4. El sistema es linealizable retroalimentado, es decir, existe una transformación no lineal tal que el sistema transformado es un sistema controlable lineal.

Considérese las dinámicas de robots de cuerpo rígido dados por un conjunto de ecuaciones diferenciales de segundo orden como:

$$D(q) \ddot{q} + C(q, \dot{q}) \dot{q} + G(q) = \tau$$

Primero, se escriben las ecuaciones dinámicas en forma de espacio de estado. Dejando los vectores $x_1 = q$, y $x_2 = \dot{q}$, y las ecuaciones dinámicas pueden ser expresadas en la forma de espacio de estado como:

$$\dot{x}_1 = x_2 \tag{6.6}$$

$$\dot{x}_2 = D^{-1}(x_1) (\tau - C(x_1, x_2) x_2 - G(x_1)) \tag{6.7}$$

Enseguida, se define:

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

las ecuaciones (6.5 y 6.6) pueden ser escritas convenientemente como:

$$\dot{x} = F(x) + B(x)\tau \quad (6.8)$$

donde:

$$F(x) = \begin{bmatrix} x_2 \\ -D^{-1}(x_1)(C(x_1, x_2)x_2 + G(x_1)) \end{bmatrix} \quad (6.9)$$

$$B(x) = \begin{bmatrix} 0 \\ D^{-1}(x_1) \end{bmatrix} \quad (6.10)$$

Enseguida, considerando las dinámicas de los robots de unión flexible y reescribiendo su forma de espacio de estado estandard. Las dinámicas de robots de unión flexible son:

$$\begin{aligned} D(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) + K(q - q_m) &= 0 \\ J\ddot{q}_m - K(q - q_m) &= u \end{aligned} \quad (6.11)$$

Se eligen los vectores $x_1 = q, x_2 = \dot{q}, x_3 = q_m, x_4 = \dot{q}_m$, las dinámicas pueden ser expresadas en forma de espacio de estado como:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= D^{-1}(x_1)(-K(x_1 - x_3) - C(x_1, x_2)x_2 - G(x_1)) \\ \dot{x}_3 &= x_4 \\ \dot{x}_4 &= J^{-1}(u + K(x_1 - x_3)) \end{aligned}$$

Además, se define:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

las ecuaciones pueden ser escritas concisamente como:

$$\dot{x} = F(x) + B(x)u$$

donde:

$$F(x) = \begin{bmatrix} x_2 \\ D^{-1}(x_1)(-K(x_1 - x_3) - C(x_1, x_2)x_2 - G(x_1)) \\ x_4 \\ J^{-1}K(x_1 - x_3) \end{bmatrix}$$

$$B(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ J^{-1} \end{bmatrix}$$

Por lo tanto, la representación de espacio de estado de ambos robots de cuerpo rígido y unión flexible puede ser capturada por la forma general de espacio de estado:

$$\dot{x}(t) = f(x(t), u(t)) \quad (6.12)$$

donde x y u son los estados y entradas de dimensiones apropiadas, respectivamente.

Robots de unión flexible vía Teorema de Tychonov

La técnica de perturbación singular ha sido aplicada para muchos problemas de modelado y control [43] donde las dinámicas son transformadas en modelos de escalas de dos tiempos, es decir, en dinámicas lentas y dinámicas rápidas. El diseño del controlador para las dinámicas desacopladas se tornará la suma de dos subcontroladores, uno diseñado para las dinámicas lentas y otro para las dinámicas rápidas usando el llamado estrategia de control compuesto. Para robots de unión flexible, la unión rígida es relativamente más grande que otros parámetros en situaciones más prácticas. Por lo tanto, es realista asumir flexibilidad de unión débil. Esta propiedad permite explotar la transformación del modelo de robot de unión flexible en un modelo de escala de dos tiempos usando teoría de perturbación singular.

En esta sección, se analizarán los dos modelos diferentes de perturbación singular para robots de unión flexible. El primero es obtenido por el *modelado de las fuerzas elásticas como las variables rápidas y las variables del eslabón como las variables lentas*, mientras la segunda es desarrollada por el modelado de los errores de seguimiento del motor como las variables rápidas en lugar de las fuerzas elásticas de unión.

Es importante mencionar la aplicabilidad del Teorema de Tychonov a un robot de unión flexible.

Teorema 6.1 *Teorema de Tychonov:*

Dadas las ecuaciones de estado en forma de perturbación singular estandard:

$$\dot{x} = f(x, z, \epsilon) \quad (6.13)$$

$$\epsilon \dot{z} = g(x, z, \epsilon) \quad (6.14)$$

entonces, probando algunas condiciones técnicas que sean cumplidas, el teorema de Tychonov plantea: si las dinámicas reducidas tienen una solución única en el intervalo $t \in [0, t_1]$ y si la capa frontera es exponencialmente estable, uniforme en (t, x) , entonces existe un ϵ^* tal que para todo $0 < \epsilon < \epsilon^*$

$$\begin{aligned} x(t) &= x_{slow}(t) + O(\epsilon) \\ z(t) &= z_{slow}(t) + z_{fast}(\epsilon) + O(\epsilon) \end{aligned}$$

Para un robot de unión flexible los estados son:

$$\begin{aligned} x &= \begin{bmatrix} \theta^T, \dot{\theta}^T \end{bmatrix}^T \\ z &= \begin{bmatrix} z^T, \dot{z}^T \end{bmatrix}^T \end{aligned}$$

donde θ es el ángulo del eslabón, z es el par de la unión, y $\epsilon = \sqrt{\mu}$.

6.2.3. Dinámicas de robots de eslabón flexible

En [81, 26, 46], se muestra que las dinámicas de cualquier robot multi eslabón de eslabón flexible puede ser representado por:

$$M(q)\ddot{q} + D(q, \dot{q})\dot{q} + Kq + F(q, \dot{q}) + G(q) = B(q)u \quad (6.15)$$

con

$$q = \begin{bmatrix} q_r \\ q_f \end{bmatrix} \quad (6.16)$$

donde q_r es el vector de modos rígidos (coordenadas de unión generalizadas) y q_f es el vector de modos flexibles (las amplitudes de los modos flexibles). $M(q)$ representa la matriz de inercia, $D(q, \dot{q})$ la matriz de coriolis y centrífuga, K es la matriz de rigidéz, $F(q, \dot{q})$ es la matriz de fricción, $G(q)$ es la matriz de gravedad, $B(q)$ es una matriz de entrada dependiente de las condiciones acotadas seleccionadas en la forma del método asumido, y u incluye los pares de control aplicados a cada unión.

Existen dos principales enfoques de modelado de robots de eslabón flexible, a saber, el método del elemento finito (*FEM*) y el método del modo supuesto (*AMM*). Aquí, el método de robot de eslabón flexible es usado. Para un robot con n_r eslabones flexibles y n_e elementos para cada eslabón, las dinámicas están dadas por:

$$D(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) + Kq = \tau = Bu \quad (6.17)$$

donde $q = [q_r^T \ q_f^T]^T \in \mathbb{R}^n$, $n = (2 \times n_e + 1)n_r$ denota el vector de variables rígidas y flexibles con $q_r \in \mathbb{R}^{n_r}$ como el vector de variables rígidas y $q_f \in \mathbb{R}^{n_f}$, $n_f = (2 \times n_e + 1)n_r$ como el vector de variables flexibles; $D(q) \in \mathbb{R}^{n \times n}$ es la matriz de inercia definida positiva simétrica; $C(q, \dot{q})\dot{q} \in \mathbb{R}^n$ representa la fuerza de coriolis y centrífuga; $G(q) \in \mathbb{R}^n$ es el vector de fuerza de gravedad; $K \in \mathbb{R}^{n \times n}$ es la matriz de rigidez de los eslabones flexibles; $Bu \in \mathbb{R}^{n_r}$ es el vector del par generalizado con $B = [I \ 0 \dots 0]^T \in \mathbb{R}^{n \times n}$, $I \in \mathbb{R}^{n_r \times n_r}$ es la matriz identidad; y $u \in \mathbb{R}^{n_r}$ es el vector del par del actuador.

Las dinámicas del robot de eslabón flexible está caracterizado por tener n coordenadas generalizadas pero solo n_r entradas de control.

La síntesis de la ley de control retroalimentada no lineal no es tan simple para robots rígidos donde existe una entrada de control independiente para cada grado de libertad. El uso de *teoría de perturbación singular* en explotar la natural separación en escalas de tiempo entre las dinámicas del modo flexible rápido y las dinámicas del modo rígido lento para obtener una reducción de orden del modelo conducente a identificación y control compuesto de sistemas de orden completo ofrece una posible solución. Entonces para aplicar la técnica de perturbación singular, se necesita reescribir la ecuación (6.17) en la forma estandard:

$$\begin{bmatrix} D_{rr} & D_{rf} \\ D_{fr} & D_{ff} \end{bmatrix} \begin{bmatrix} \ddot{q}_r \\ \ddot{q}_f \end{bmatrix} + \begin{bmatrix} C_{rr} & C_{rf} \\ C_{fr} & C_{ff} \end{bmatrix} \begin{bmatrix} \dot{q}_r \\ \dot{q}_f \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & K_{ff} \end{bmatrix} \begin{bmatrix} q_r \\ q_f \end{bmatrix} + \begin{bmatrix} G_r \\ 0 \end{bmatrix} = \begin{bmatrix} I \\ B_f \end{bmatrix} u \quad (6.18)$$

o de manera más simplificada como:

$$\begin{bmatrix} D_{rr} & D_{rf} \\ D_{fr} & D_{ff} \end{bmatrix} \begin{bmatrix} \ddot{q}_r \\ \ddot{q}_f \end{bmatrix} + \begin{bmatrix} H_r \\ H_f \end{bmatrix} + \begin{bmatrix} 0 \\ K_{ff}q_f \end{bmatrix} = \begin{bmatrix} u \\ 0 \end{bmatrix} \quad (6.19)$$

donde:

$$H_r = C_{rr}\dot{q}_r + C_{rf}\dot{q}_f + G_r \quad (6.20)$$

$$H_f = C_{fr}\dot{q}_r + C_{ff}\dot{q}_f \quad (6.21)$$

Debido a que la matriz de inercia D en (6.18) es definida positiva, su inversa existe y esta denotada por M como:

$$M = \begin{bmatrix} M_{rr} & M_{rf} \\ M_{fr} & M_{ff} \end{bmatrix} = \begin{bmatrix} D_{rr} & D_{rf} \\ D_{fr} & D_{ff} \end{bmatrix}^{-1} \quad (6.22)$$

donde:

$$M_{rr} = (D_{rr} - D_{rf}D_{ff}^{-1}D_{fr})^{-1} \quad (6.23)$$

$$M_{rf} = -D_{rr}^{-1} D_{rf} (D_{ff} - D_{fr}D_{rr}^{-1}D_{rf})^{-1} \quad (6.24)$$

$$M_{fr} = -D_{ff}^{-1} D_{fr} (D_{ff} - D_{rf}D_{ff}^{-1}D_{fr})^{-1} \quad (6.25)$$

$$M_{ff} = (D_{ff} - D_{fr}D_{rr}^{-1}D_{rf})^{-1} \quad (6.26)$$

Resolviendo (6.19) para \ddot{q}_r y \ddot{q}_f se produce:

$$\ddot{q}_r = -M_{rr}H_r - M_{rf}H_f - M_{rf}K_{ff}q_f + M_{rr}u \quad (6.27)$$

$$\ddot{q}_f = -M_{fr}H_r - M_{ff}H_f - M_{ff}K_{ff}q_f + M_{fr}u \quad (6.28)$$

Los subíndices r y f denotando la asociación de los modos de los componentes rígidos y flexibles respectivamente.

Entonces una aproximación al sistema original (6.18) es obtenido por la aplicación de la técnica de perturbaciones singulares como sigue, descomponiendo el sistema (6.19) en un subsistema lento y un subsistema rápido y permitiendo de manera mas sencilla el modelado e identificación y posteriormente, el exitoso diseño de un control compuesto para el sistema original (6.18).

Introduciendo un factor de escala ϵ y definiendo $q_f = \epsilon^2 z$ y $K_{ffs} = \epsilon^2 K_{ff}$ donde $\frac{1}{\epsilon^2}$ es la rigidez pequeña en K_{ff} . El parámetro ϵ es suficientemente pequeño que significa que el orden de rigidez es suficientemente largo. Por lo tanto, el modelo perturbado singularmente del robot de unión flexible es obtenido como

$$\begin{aligned} \ddot{q}_r &= -M_{rr}H_r - M_{rf}H_f - M_{rf}K_{ffs}z + M_{rr}u \\ \epsilon^2 \ddot{z} &= -M_{fr}H_r - M_{ff}H_f - M_{ff}K_{ffs}z + M_{fr}u \end{aligned} \quad (6.29)$$

Donde $H_r = C_{rr}\dot{q}_r + C_{rf}\dot{q}_f + G_r$, $H_f = C_{fr}\dot{q}_r + C_{ff}\dot{q}_f + G_r$. Se define el esfuerzo del control compuesto

$$u = \bar{u} + u_F$$

donde \bar{u} es el control lento y u_F es el control rápido. Para identificar el subsistema lento, el modelo perturbado singularmente es evaluado con $\epsilon = 0$ lo que produce:

$$\ddot{\bar{q}}_r = -\bar{M}_{rr}\bar{H}_r - \bar{M}_{rf}\bar{H}_f - \bar{M}_{rf}K_{ffs}\bar{z} + \bar{M}_{rr}\bar{u} \quad (6.30)$$

$$0 = -\bar{M}_{fr}\bar{H}_r - \bar{M}_{ff}\bar{H}_f - \bar{M}_{ff}K_{ffs}\bar{z} + \bar{M}_{fr}\bar{u} \quad (6.31)$$

donde la sobre barra denota la evaluación de los elementos con $\epsilon = 0$. Resolviendo para \bar{z} de la ecuación anterior se obtiene:

$$\bar{z} = K_{ffs}^{-1}\bar{M}_{ff}^{-1}(-\bar{M}_{fr}\bar{H}_r - \bar{M}_{ff}\bar{H}_f + \bar{M}_{fr}\bar{u}) \quad (6.32)$$

El subsistema lento está por lo tanto formado por la sustitución de (6.32) en (6.30) y evaluando, es decir,:

$$\ddot{\bar{q}}_r = \left(\bar{M}_{rr} - \bar{M}_{rf}\bar{M}_{ff}^{-1}\bar{M}_{fr}\right)(-\bar{H}_r + \bar{u}) \quad (6.33)$$

Usando las expresiones (6.23)-(6.26) se produce:

$$\ddot{\bar{q}}_r = \bar{D}_{rr}^{-1}(-\bar{H}_r + \bar{u})$$

Usando la definición de H_r de (6.31) para \bar{H}_r , el subsistema lento puede ser obtenido como:

$$\ddot{\bar{q}}_r = \bar{D}_{rr}^{-1}\left(-\bar{C}_{rr}\dot{\bar{q}}_r - \bar{G}_r + \bar{u}\right) \quad (6.34)$$

que corresponde al modelo dinámico del robot rígido.

Para derivar el subsistema rápido, se define una escala de tiempo rápida $\tau = \frac{t}{\epsilon}$ y términos de corrección capa frontera $\rho_1 = z - \bar{z}$ y $\rho_2 = \epsilon \dot{z}$.

El sistema capa frontera puede ser escrito como:

$$\frac{d\rho_1}{d\tau} = \rho_2 \quad (6.35)$$

$$\frac{d\rho_2}{d\tau} = -M_{fr}H_r - M_{ff}H_f - M_{ff}K_{ffs}(\rho_1 + \bar{z}) + M_{fr}(\bar{u} + u_F) \quad (6.36)$$

debido a que $\frac{d\bar{z}}{d\tau} = \epsilon \dot{\bar{z}} = 0$. Posicionando $\epsilon = 0$, y sustituyendo para \bar{z} de (6.32) resulta en:

$$\frac{d\rho_2}{d\tau} = -\bar{M}_{ff}K_{ffs}\rho_1 + \bar{M}_{fr}u_F \quad (6.37)$$

Por lo tanto, el subsistema rápido puede ser obtenido como un sistema lineal con variables lentas \bar{q}_r como parámetros, es decir,:

$$\frac{d\rho}{d\tau} = A_F\rho + B_Fu_F \quad (6.38)$$

donde $\rho = [\rho_1^T \ \rho_2^T]^T$ y:

$$A_F = \begin{bmatrix} 0 & I \\ -\bar{M}_{ff}K_{ffs} & 0 \end{bmatrix}, \quad B_F = \begin{bmatrix} 0 \\ \bar{M}_{fr} \end{bmatrix} \quad (6.39)$$

Las dinámicas internas relativas al seguimiento de salida $q_r(t)$ esta dada por el subsistema rápido (6.38). El sistema lento (6.34) son las dinámicas del brazo rígido. De acuerdo al teorema de Tikhonov, la red resultante es que el sistema perturbado singularmente original puede ser descrito para orden ϵ usando (6.30) y (6.31) con:

$$\begin{aligned} q_r &= \bar{q}_r + O(\epsilon) \\ q_f &= \epsilon^2 (\bar{\xi} + \zeta_1) + O(\epsilon) \end{aligned} \quad (6.40)$$

con $O(\epsilon)$ denotando términos de orden ϵ .

Comentario 6.1 *Como se mostró en la sección previa, usando la teoría de perturbación singular el sistema completo puede ser modelado con dos subsistemas: las dinámicas rápidas lineales y las dinámicas no lineales lentas que generalmente contiene términos de coriolis, fuerza centrípeta, fuerza gravitacional y fuerza de fricción.*

6.3. Modelado e identificación vía redes neuronales con diferentes escalas de tiempo

Considérese dinámicas de robots de eslabón flexible con dos escalas de tiempo

$$\begin{aligned}\ddot{q}_{slow} &= -M_{rr}(q_s, q_f) H_r - M_{rf}(q_s, q_f) H_f - M_{rf}(q_s, q_f) K_{ffs} q_f + M_{rr}(q_s, q_f) u \\ \epsilon^2 \ddot{q}_{fast} &= -M_{fr} H_r - M_{ff} H_f - M_{ff} K_{ffs} q_f + M_{fr} u\end{aligned}\quad (6.41)$$

Construyendo la siguiente red neuronal dinámica con dos escalas de tiempo para identificar el sistema no lineal (6.17)

$$\begin{aligned}\dot{\hat{q}}_{slow} &= A\hat{q}_s + W_1\sigma_1(V_1[\hat{q}_s, \hat{q}_f]^T) + W_3\phi_1(V_3[\hat{q}_s, \hat{q}_f]^T)u \\ \epsilon\dot{\hat{q}}_{fast} &= B\hat{q}_f + W_2\sigma_2(V_2[\hat{q}_s, \hat{q}_f]^T) + W_4\phi_2(V_4[\hat{q}_s, \hat{q}_f]^T)u\end{aligned}\quad (6.42)$$

Analizando una red neuronal de capa sencilla (Se observa que las capas ocultas no afectan la propiedad de pasividad en el comentario 1)

$$\begin{aligned}\dot{\hat{q}}_{slow} &= A\hat{q}_s + W_1\sigma_1(\hat{q}_s, \hat{q}_f) + W_3\phi_1(\hat{q}_s, \hat{q}_f)u \\ \epsilon\dot{\hat{q}}_{fast} &= B\hat{q}_f + W_2\sigma_2(\hat{q}_s, \hat{q}_f) + W_4\phi_2(\hat{q}_s, \hat{q}_f)u\end{aligned}$$

Generalmente, el sistema no lineal (6.17) puede ser escrito como:

$$\begin{aligned}\dot{q}_s &= Aq_s + W_1^*\sigma_1(q_s, q_f) + W_3^*\phi_1(q_s, q_f)u - \tilde{f}_1 \\ \epsilon\dot{q}_f &= Bq_f + W_2^*\sigma_2(q_s, q_f) + W_4^*\phi_2(q_s, q_f)u - \tilde{f}_2\end{aligned}\quad (6.43)$$

donde W_j^* ($j = 1 \dots 4$) es una matriz acotada desconocida:

$$W_j^* \Lambda_j^{-1} W_j^{*T} \leq \overline{W}_j$$

donde \overline{W}_j es una matriz conocida previamente. La función vectorial \tilde{f}_i ($i = 1, 2$) puede ser considerada como error de modelado y disturbios o perturbaciones.

Definiendo el error de identificación como:

$$\begin{aligned} \Delta_1 &= \hat{q}_s - q_s \\ \Delta_2 &= \hat{q}_f - q_f \end{aligned} \quad (6.44)$$

Entonces dependiendo de la consideración (5.30) que sea satisfecha, si los pesos W_j ($j = 1 \cdots 4$) son actualizados como previamente ya se había establecido:

$$\begin{aligned} \dot{W}_1 &= -K_1 P_1 \Delta_1 \sigma_1^T(\hat{X}_t) \\ \dot{W}_2 &= -K_2 P_1 \phi_1(\hat{X}_t) \gamma(u_t) \Delta_1^T \\ \dot{W}_3 &= -K_3 P_2 \Delta_2 \sigma_2^T(\hat{X}_t) \\ \dot{W}_4 &= -K_4 P_2 \phi_2(\hat{X}_t) \gamma(u_t) \Delta_2^T \end{aligned} \quad (6.45)$$

donde $\hat{X}_t = [\hat{q}_s^T, \hat{q}_f^T]^T$, P_1 y P_2 son la solución de la ecuación de Riccati (5.15), entonces las dinámicas de los errores de identificación (5.31) es estrictamente pasiva del error de modelado \tilde{f}_i para la identificación del error $2P\Delta_i$, $i = 1, 2$.

Comentario 6.2 Aunque la técnica de redes neuronales artificiales con diferentes escalas de tiempo, no es la única para el modelado e identificación y control de sistemas robotizados, si es al menos más flexible para su aplicabilidad que otras técnicas conocidas. Principalmente por las amigables propiedades que presentan tanto los conceptos de perturbaciones singulares como los de redes neuronales artificiales.

6.4. Simulación

A continuación se muestran una simulaciones para un robot de unión flexible en Fig.6.1, separados en diferentes escalas de tiempo, utilizando redes neuronales multicapa.

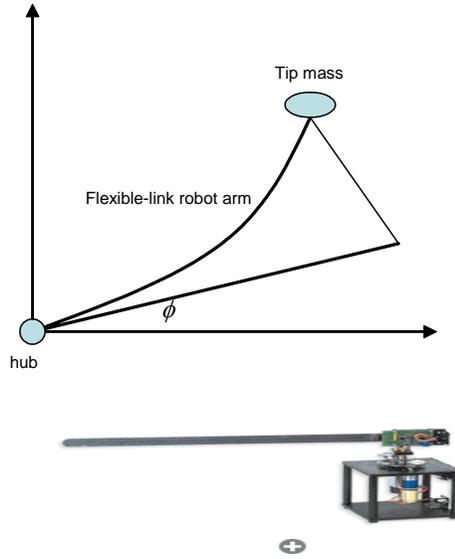


Figura 6.1: Un robot de unión flexible

Para validar los resultados teóricos, se dan las siguientes simulaciones. La red neuronal dinámica es:

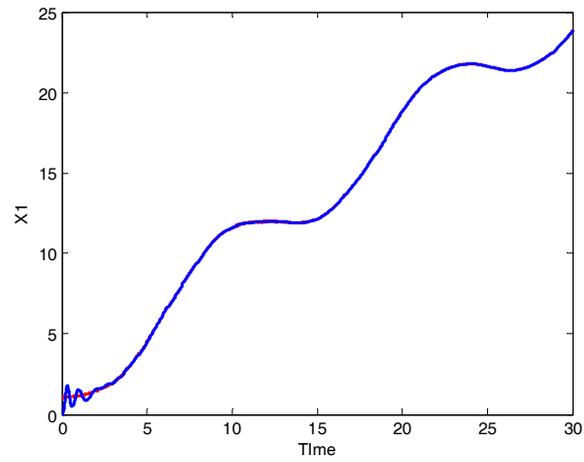
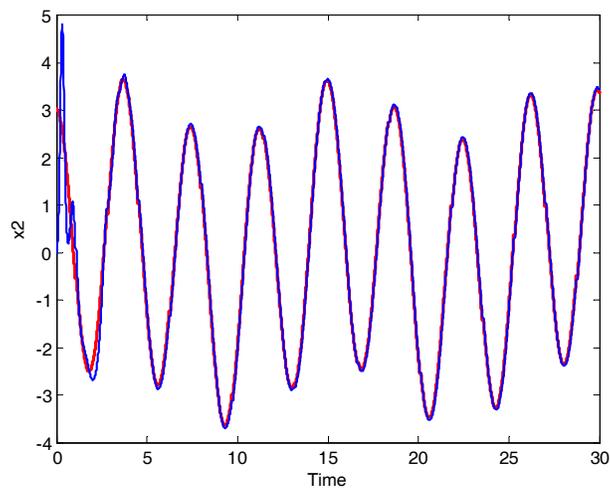
$$\begin{aligned}\dot{\hat{x}} &= A\hat{x} + W_1\sigma_1(\hat{x}, \hat{z}) + W_3\phi_1(\hat{x}, \hat{z})u \\ \epsilon\dot{\hat{z}} &= B\hat{z} + W_2\sigma_2(\hat{x}, \hat{z}) + W_4\phi_2(\hat{x}, \hat{z})u\end{aligned}\quad (6.46)$$

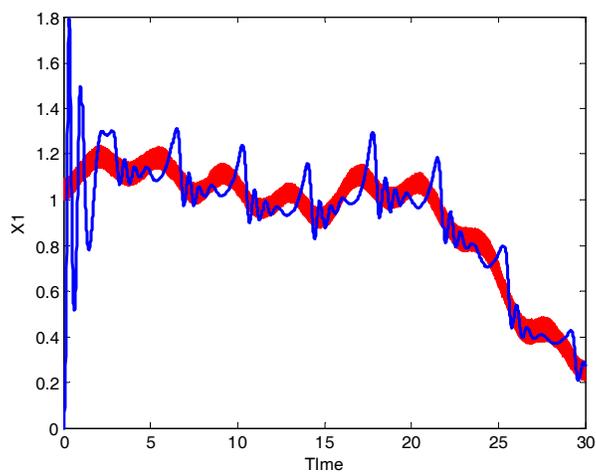
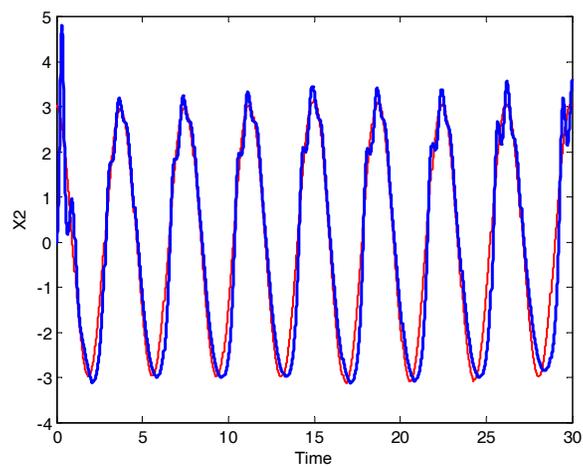
Estas redes neuronales dinámicas con dos escalas de tiempo puede ilustrar la dinámica de un robot de unión flexible:

$$\begin{aligned}\ddot{q}_r &= -M_{rr}H_r - M_{rf}H_f - M_{rf}K_{ffs}z + M_{rr}u \\ \epsilon^2\ddot{z} &= -M_{fr}H_r - M_{ff}H_f - M_{ff}K_{ffs}z + M_{fr}u\end{aligned}\quad (6.47)$$

Para las redes neuronales dinámicas se elige \hat{x} y \hat{z} que son escalares, $A = -3$, $B = -3$, $\epsilon = 0,05$. Las condiciones iniciales son $\hat{x}_0 = 0$, $\hat{z}_0 = 0$. Los resultados de identificación son mostrados en Fig.6.2 y Fig.6.3. Para las dos primeras figuras se utilizó una señal senoidal.

Mientras que para las dos siguientes figuras se utilizó una señal aleatoria cuyos resultados de identificación son mostrados en Fig.6.4 y Fig.6.5.

Figura 6.2: Identificación de x_1 Figura 6.3: Identificación de x_2

Figura 6.4: Identificación de x_1 Figura 6.5: Identificación de x_2

6.5. Conclusión del capítulo

La técnica de perturbaciones singulares con redes neuronales artificiales para identificar un sistema robótico de unión flexible, es posible aplicarlo en este caso, ya que un sistema robótico como el que se simuló, se puede separar en diferentes escalas de tiempo. Considerando la unión del eslabón como la dinámica rápida del sistema, mientras que el eslabón mismo se puede considerar como la dinámica lenta del sistema completo. Contribuyendo a que un sistema dinámico no lineal con cierto grado de complejidad, se puede simplificar su análisis, separando el sistema completo, en dos subsistemas dinámicos con menor grado de complejidad que todo el sistemas completo.

Capítulo 7

Conclusiones y trabajos futuros

Conclusiones

Con base en los resultados obtenidos en la presente tesis, se concluye que:

Para el modelado e identificación de sistemas no lineales separándolos en diferentes escalas de tiempo, con la técnica de perturbaciones singulares resulta relativamente más sencillo analizar tales sistemas como subsistemas rápidos y lentos.

Una vez que se logra separar un sistema en diferentes escalas de tiempo, la técnica de redes neuronales artificiales, resultan bastante eficientes debido a que permite obtener modelos confiables a partir de datos experimentales, y por la característica fundamental de las redes neuronales de aproximadores universales, es posible obtener modelos de identificación bastantes cercanos al modelo real a ser identificado.

Trabajos futuros

En base a los resultados obtenidos sobre la viabilidad de la metodología, se proponen como trabajos futuros las siguientes actividades:

Control de sistemas no lineales sometidos a perturbaciones singulares, con redes neuronales artificiales utilizando diferentes escalas de tiempo.

Adaptar redes neuronales artificiales con diferentes escalas de tiempo a sistemas no lineales multirango, es decir, sistemas con múltiples entradas y múltiples salidas.

Bibliografía

- [1] S. Amari, "Field theory of self-organizing neural nets," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, pp. 741–748, 1983.
- [2] S. Amari, "Competitive and cooperative aspects in dynamics of neural excitation and self-organization," *Competition Cooperation Neural Networks*, vol. 20, pp. 1–28, 1982.
- [3] C.H. An, C.G. Atkeson, and J.M. Hollerbach, *Model-Based Control of a Robot Manipulator*, MIT Press, Cambridge, M. A., 1988.
- [4] H. Asada and K. Youcef-Toumi, *Direct Drive Robots Theory and Practice*, MIT Press, Cambridge, MA, 1987.
- [5] B. Brogliato, R. Ortega, and R. Lozano, "Global tracking controllers for flexible-joint manipulators: A comparative study", *Automatica*, vol.31, No.7, pp.941-956, 1995.
- [6] Byrnes C.I., Isidori A., Willems J.C., "Passivity, Feedback Equivalence, and the Global Stabilization of Minimum Phase Nonlinear Systems," *IEEE Trans. Automat. Contr.*, 36, 1228-1240, (1991).
- [7] Burrows T. L. and Niranjani M. "The Use of Feed-forward and Recurrent Neural Networks for System Identification" CUED/F-INFENG/TR158. Cambridge University Engineering Department, 1993.
- [8] J. J. Craig, *Adaptive Control of Mechanical Manipulators*, Addison-Wesley, Reading, MA, 1989.

- [9] G. Cybenko, "*Approximation by superpositions of a sigmoidal functions*", Mathematics of Control, Signals and systems, vol.2, No.4, pp.303-314, 1989.
- [10] S.Commuri and F.L.Lewis, CMAC Neural Networks for Control of Nonlinear Dynamical Systems: Structure, Stability and Passivity, *Automatica*, vol.33, no.4, 635-641, 1996.
- [11] Y.Z. Chang and R. W. Daniel, "*On the adaptive control of flexible joint robots*", *Automatica*, vol. 28, No.5, pp.969-974, 1992.
- [12] S. Chen, S. A: Billings, and P. M. Grant, Recursive hybrid algorithm for non-linear system identification using radial basis function networks, "Int. J. Control, vol.55, No.5, pp. 1051-1070, 1992.
- [13] Y.C.Chu, J.Huang, A neural-network method for the nonlinear servomechanism problem, *IEEE Trans. Neural Networks*, Vol.10, No.6, 1412-1423, 1999.
- [14] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations*, 31:53-98, 1979.
- [15] M.Forti, S.Manetti and M.Marini, Necessary and Sufficient Condition for Absolute Stability of Neural Networks, *IEEE Trans. on Circuit and Systems-I*, Vol. 41, 491-494, 1994.
- [16] Fu, C.S. and Poch M. "*Application of a multiyered neural network to system identification*"Int. J. System SCI, vol.24 No.8, pp.1601-1609, 1993.
- [17] M. Galicki, L. Leistriz, and H.Witte, "Learning continuous trajectories in recurrent neural networks with time-dependent weights," *IEEE Trans. Neural Networks*, vol. 10, pp. 741-756, May 1999.
- [18] S. S. Ge, *Nonlinear Adaptive Control of Robots*, Ph. D. thesis, Mechanical Engineering Department, Imperial College of Science, Technology and Medicine, University of London, UK, 1992.

- [19] S. S. Ge, C. C. Hang, and L. C. Woon, "Adaptive neural network control of robot manipulators in task space," *IEEE Trans. Industrial Electronics*, vol. 44, No. 6, pp. 746-752, 1997.
- [20] S.S. Ge, Z. L. Wang, and Z. J. Chen, "Adaptive static neural network control of robots", in *Proc. IEEE Int. Conf. Industrial Technology*, Guangdong, Guangzhou, P. R. China, Dec. 1994, vol.1, pp. 240-244.
- [21] S.S. Ge, C. C. Hang, "Structural network modelling and adaptive control of dynamical systems", *May 1995, patent, Singapore.*
- [22] S. Grossberg, "Adaptive pattern classification and universal recording," *Biol. Cybern.*, vol. 23, pp. 121-134, 1976.
- [23] S. Grossberg, "Competition, decision and consensus," *J. Math. Anal. Applicat.*, vol. 66, pp. 470-493, 1978.
- [24] R. H. Middleton and G. C. Goodwin, "Adaptive computed torque control for rigid link manipulations," *Systems and Control Letters*, vol. 10, No.1, pp. 9-16, 1988.
- [25] Madan M. Gupta, Liang Jin, and Noriyasu Homma. *Static and Dynamic neural Networks. 2003 John Wiley & Sons, Inc.*
- [26] G.G. Hastings and W. J. Book, "Verification of a linear dynamic model for flexible robotic manipulators", in *Proc. IEEE Int. Conf. Robotics and Automation*, Apr. 1986, pp. 1024-1029.
- [27] Simon Haykin. *Neural Networks. Second Edition. Prentice Hall.*
- [28] D. Hebb, *The Organization of Behavior*. New York: J. Wiley, 1949.
- [29] Hill D., Moylan P., "Stability results for nonlinear feedback systems," *Automatica*, 13, 373-382, (1976).

- [30] F. C. Hoppensteadt. Singular perturbations on the infinite time interval. *Transactions of the American Mathematical Society*, 123:521-535, 1966.
- [31] J.J.Hopfield, "Neurons with grade response have collective computational properties like those of a two-state neurons", *Proc. of the National Academy of Science, USA*, vol. 81, 3088-3092, 1984.
- [32] Hrycej T. "Neurocontrol Towards an Industrial Control Methodology". Ed. Jhon Wiley & Sons, Inc, 1997.
- [33] K.J.Hunt, D.Sbarbaro, R.Zbikowski and P.J.Gawthrop, Neural Networks for Control Systems- A Survey, *Automatica*, Vol. 28, 1083-1112, 1992.
- [34] P.A.Ioannou and J.Sun, *Robust Adaptive Control*, Prentice-Hall, Inc, Upper Saddle River: NJ, 1996
- [35] S.Jagannathan and F.L. Lewis, "Identification of Nonlinear Dynamical Systems Using Multilayered Neural Networks", *Automatica*, vol.32, no.12, 1707-1712, 1996.
- [36] S. Jagannathan and F. L. Lewis, "Multilayer discrete-time neural-net controller with guarantees performance,"*IEEE Trans. Neural Networks*, vol. 7, No.1, pp. 107-130, 1996.
- [37] L. Jin and M. Gupta, "Stable dynamic backpropagation learning in recurrent neural networks," *IEEE Trans. Neural Networks*, vol. 10, pp. 1321-1334, Nov. 1999.
- [38] K.Kaszakurewics and A.Bhaya, On a Class of Globally Stable Neural Circuits, *IEEE Trans. on Circuit and Systems-I*, Vol. 41, 171-174, 1994.
- [39] J. Kevorkian & J.D.Cole. *Multiple Scale and Singular Perturbation Methods*. 1996 Springer-Verlag New York, Inc.
- [40] Hassan K. Khalil. *Nonlinear Systems. Third Edition. Prentice Hall*.

- [41] H.K.Khalil, *Nonlinear Systems*, 2nd Edition, Prentice-Hall, Inc., 1996M. Lemmon and V. Kumar, "Emulating the dynamics for a class of laterally inhibited neural networks", *Neural Networks*, vol. 2, pp. 193–214, 1989.
- [42] P. K. Khosla and T. Kanade, "Experimental evaluation of nonlinear feedback and feed-forward control schemes for manipulators," *Int. J. Robotics Research*, vol. 7, No. 1, pp. 18-28, 1988.
- [43] P. V. Kokotovic, H. Khalil, and J. O'reilly, *Singular Perturbation Methods in Control: Analysis and Design*, Academic Press, London, 1986.
- [44] M. Lemmon and V. Kumar, "Emulating the dynamics for a class of laterally inhibited neural networks", *Neural Networks*, vol. 2, pp. 193–214, 1989.
- [45] J.J. Levin and N. Levinson. Singular perturbations of nonlinear systems of ordinary differential equations and an associated boundary layer equation. *Journal of Rational Mechanics and Analysis*, 3: 267-270, 1959.
- [46] F.L.Lewis, K.Liu and A.Yesildirek , Multilayer Neural-Net Robot Controller with Guaranteed Tracking Performance, *IEEE Trans. on Neural Network*, Vol. 7, No 2, 388-398, 1996.
- [47] F. L. Lewis, K. Liu, and A. Yesildirek, "Neural net robot controller with guaranteed tracking performance," *IEEE Trans. Neural Networks*, vol. 6, No. 3, pp. 703-715, 1995.
- [48] J. Lin and F. L. Lewis, "*Dynamic equations of a manipulators with rigid and flexible links: Derivation and symbolic computation*" in Proc. American Control Conf. June 1993, pp. 2868-2872.
- [49] H.Liu, F.Sun and Z.Sun, Stability Analysis and Synthesis of Fuzzy Singularly Perturbed Systems, *IEEE Transactions on Fuzzy Systems*, Vol.13, No.2, 273-284, 2005.
- [50] H.Lu and Z.He, Global exponential stability of delayed competitive neural networks with different time scales, *Neural Networks*, Vol.18, No.3, 243-250, 2005.

- [51] J. Y. S. Luh, Conventional controller design for industrial robots-a tutorial,"IEEE Trans. Systems, Man and Cybernetics, vol. 13, No. 3, pp. 298-316, 1983.
- [52] Horacio J. Márquez. *Nonlinear Control -Systems analysis and design-*. Wiley-Interscience. A John Wiley & Sons,Inc.,Publication.
- [53] K.Matsouka, Stability Conditions for Nonlinear Continuous Neural Networks with Asymmetric Connection Weights, *Neural Networks*, Vol.5, 495-500, 1992
- [54] A. Meyer-Bäse, F. Ohl, and H. Scheich, "Singular perturbation analysis of competitive neural networks with different time-scales," *Neural Comput.*, vol. 8, pp. 545-563, 1996.
- [55] A. Meyer-Baese, S. S. Pilyugin, and Y. Chen, Global Exponential Stability of Competitive Neural Networks With Different Time Scales, *IEEE Trans. on Neural Networks*, VOL.14, NO.3, 716-719, 2003.
- [56] R. M. Murray, Z. Li, and S. S. sastry, A Mathematical Introduction to Robotic Manipulation, CRC Press, Boca Raton, FL, 1994.
- [57] Anke Meyer-Bäse, Frank Ohl, Henning Scheich. *Singular Perturbation Analysis of Competitive Neural Networks with Different Time-Scales*.
- [58] Mills P., Zomaya A. A neural network approach to on-line identification of nonlinear systemsCybernetics and Systems. An International Journal, 24:pp. 171-195, 1993.
- [59] E.F. Miscenko. Asymptotic calculation of periodic solutions of systems containing small parameters in the derivatives. *Translations of the American Mathematical Society, Series 2, 18: 199-230, 1961*.
- [60] Kevin L. Moore and D. Subbaram Naidu. *Singular Perturbations and Time Scales in Artificial Neural Networks*. Proceedings of the 30th Conference on Decision and Control. Brighton, England. December 1991.

- [61] E.B.Kosmatopoulos, M.M.Polycarpou, M.A.Christodoulou and P.A.Ioannpu, "High-Order Neural Network Structures for Identification of Dynamical Systems", *IEEE Trans. on Neural Networks*, Vol.6, No.2, 442-431, 1995
- [62] D.S. Naidu. *Singular perturbation methodology, in control systems. 1988 Peter Peregrinus Ltd.*
- [63] Narendra K.S. and Parthasarathy K. "Identification and Control of Dynamical Systems Using Neural Networks." *IEEE Transactions on Neural Networks*. Vol. I, No. 1, March 1990.
- [64] R. Ortega and M. W. Spong, *Adaptive motion control of rigid robots: a tutorial*", *Automatica*, vol.25, No.6, pp.877-888, 1989.
- [65] P. C. Parks and V. Hahn. *Stability Theory*. Prentice Hall, 1993.
- [66] Alexander S. Poznyak, Edgar N. Sánchez and Wen Yu Liu. *Differential Neural Networks for Robust Nonlinear Control. 2001 World Scientific Publishing Co. Pte. Ltd.*Singapore, 2001.
- [67] A.S.Poznyak, W.Yu, E.N.Sanchez and J.P.Perez, Nonlinear Adaptive Trajectory Tracking Using Dynamic Neural Networks, *IEEE Trans. on Neural Networks*, Vol.10, No.5, 1402-1411, 1999.
- [68] G.V.Puskorius and L.A.Feldkamp, Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks, *IEEE Trans. on Neural Networks*, Vol.5, No.2, 279-297, 1994.
- [69] J.R. Raol and H. Madhuranath, "Neural Network architectures for parameter estimation of dynamical systems,"*IEE Proc. Control Theory Applications*, vol. 143, No. 4, pp. 387-394, 1996.

- [70] Juan Reyes-Reyes, Wen Yu and Alexander S. Poznyak, Passivation and Control of Partially Known SISO Nonlinear Systems via Dynamic Neural Networks, *Mathematical Problems in Engineering*, *Mathematical Problems in Engineering*, Vol.6, No.1,61-83,2000.
- [71] J.S. Reed and P.A. Ioannou, "*Instability analysis and robust adaptive control of robotic manipulators*", *IEEE Trans. Robotics and Automation*, vol.5, No.3, pp. 381-386, 1989.
- [72] George A. Rovithakis, and Manolis A. Christodoulou, *Adaptive Control of Unknown Plants Using Dynamical Neural Networks*. *IEEE Transactions on Systems, Man, and Cybernetics*. Vol. 24, No. 3, March 1994.
- [73] D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. *MIT Press, Cambridge, MA, 1986*.
- [74] Ali Saberi and Hassan Khalil. *Quadratic-Type Lyapunov Functions for Singularly Perturbed Systems*. *IEEE Transactions on Automatic Control*, Seiten 542-550, June 1984.
- [75] E.N.Sanchez and J.P.Perez, Input-to-State Stability Analysis for Dynamic Neural Networks, to appear in *IEEE Trans. on Circuit and Systems-I*.
- [76] E.N.Sanchez, J.P. Perez, Input-to-State Stability (ISS) Analysis for Dynamic Neural Networks, *IEEE Trans. Circuits and Systems, Part I*, Vol.46 , Nov.11, 1395 -1398, 1999.
- [77] Shankar Sastry. *Nonlinear Systems. Analysis, Stability, and Control*. 1999 *Springer-Verlag New York, Inc.*
- [78] M.E. Sawan and Z.H. Shao. Robust stability of singularly perturbed systems. *Int. J Control*, 1993, vol.,58, No. 6 1469-1476.
- [79] Z.H. Shao. *Robust Stability of Two-Time-Scale Systems With Nonlinear Uncertainties*. *IEEE Transactions on Automatic Control*. Vol. 49, pp. 258-261. No. 2. February 2004.

- [80] R.Sepulchre, M.Jankovic and P.V.Kokotovic, *Constructive Nonlinear Control*, Springer-Verlag, London, 1997.
- [81] S. Centikunt, B. Siciliano and W. J. Book, "*Symbolic modeling and dynamic analysis of flexible manipulators*" in Proc. IEEE Int. Conf. Systems, Man and Cybernetics, Oct. 1986, pp. 798-803.
- [82] W.M. Silver, "*On the equivalence of lagrangian and Newton-Euler dynamics for manipulators*," Int. J. Robotics Research, vol.1, No.2, pp.60-70, 1982.
- [83] S. N. Singh, "*Adaptive model following control of nonlinear robotic systems*," IEEE Trans. Automatic Control, vol.30, No. 11, pp. 1099-1100. 1985. K.-K. D. Young, "*A variable structure model following control design for robotics applications*," IEEE J. Robotics and Automation, vol.4, No.5, pp.556-561, 1988.
- [84] J.J. E. Slotine and W. Li, "*On the adaptive control of robot manipulators*", Int. J. Robotics Research, vol. 6, no.3, pp.49-59, 1987.
- F.L.Lewis, C.T. Abdallah, and D. M. Dawson, *Control of Robot Manipulators*, MacMillan, New York, 1993.
- [85] E.D.Sontag and Y.Wang, On Characterization of the Input-to-State Stability Property, *System & Control Letters*, Vol.24, 351-359, 1995.
- [86] M. W. Spong, "*Modeling and control of elastic joint robots*", *Trans. ASME J. Dynamic Systems, Measurement and Control*, vol.109, No.4, pp.310-319, 1987.
- [87] J. Suykens, B. Moor, and J. Vandewalle, "Robust local stability of multilayer recurrent neural networks," *IEEE Trans. Neural Networks*, vol. 11, pp. 222–229, Jan. 2000.
- [88] J.A.K. Suykens, J. Vandewalle, B. De Moor, NLq Theory: Checking and Imposing Stability of Recurrent Neural Networks for Nonlinear Modelling, *IEEE Transactions on Signal Processing (special issue on neural networks for signal processing)*, Vol.45, No.11, 2682-2691, 1997.

- [89] J.A.K.Suykens, B.De Moor, J.Vandewalle, NLq Theory: A Neural Control Framework with Global Asymptotic Stability Criteria, *Neural Networks*, Vol.10, No.4, 615-637, 1997.
- [90] J.A.K.Suykens, J.Vandewalle and B.De Moor, Lur'e Systems with Multilayer Perceptron and Recurrent Neural Networks; Absolute Stability and Dissipativity, *IEEE Transactions on Automatic Control*, Vol.44, 770-774, 1999.
- [91] L. M. Sweet and M. C. Good, *Re-definition of the robot motion control problem: Effects of plant dynamics, drive system constraints, and user requirements*", in Proc. IEEE Conf. Decision and Control, Las Vegas, NV, Dec. 1984, vol.2, pp.724-732.
- [92] A.N.Tikhonov, A.B. Vasil'eva, A.G.Sveshnikov. *Differential Equations*. Springer-Verlag.
- [93] V. I. Utkin, *"Variable structure systems with sliding modes"*, IEEE Trans. Automatic Control, vol.22, No.2, pp. 212-222,1977.
- [94] M. Vidyasagar. *Nonlinear Systems Analysis, Second Edition.1978 Prentice Hall*.
- [95] J.C.Willems, Least Squares Optimal Control and Algebraic Riccati Equations, *IEEE Transactions Automat. Contr.*, vol.16, 621-634, 1971
- [96] D. Willshaw and C. von der Malsburg, "How patterned neural connections can be set up by self-organization," *Proc. Roy. Soc. London B*, vol. 194, pp. 431-445, 1976.
- [97] L. L. Whitcomb, A. A. Rizzi, and D. E. Koditschek, *Comparative experiments with a new adaptive controller for robot arm*", IEEE Trans. Robotics and Automation, vol.9, No.1, pp.59-69, 1993.
- [98] T. Yabuta and T. Yamada, "Force control using neural networks," *Advance Robotics*, vol 7, No. 4, pp. 395-408, 1993.

- [99] M.Ye, Y.Zhang, Complete Convergence of Competitive Neural Networks with Different Time Scales, *Neural Processing Letters*, Vol.21, No.1, 53-60, 2005.
- [100] Wen Yu and Xiaou Li, Some Stability Properties of Dynamic Neural Networks, *IEEE Trans. Circuits and Systems, Part I*, Vol.48, No.1, 256-259, 2001.
- [101] Wen Yu and Xiaou Li, New Results on System Identification with Dynamic Neural Networks, *IEEE Trans. Neural Networks*, Vol.12, No.2, 412-417, 2001
- [102] Wen Yu, Alexander S.Poznyak, Indirect Adaptive Control via Parallel Dynamic Neural Networks, *IEE Proceedings - Control Theory and Applications*, Vol.37, No.1, 1999.
- [103] Zamarreño J.M., Vega P. "Neural predictive control: application to a highly non-linear process"13th World Congress International Federation of Automatic Control, vol.C: Control Design I, 19-24, 1996, IFAC'96.

Capítulo 8

Apéndice: artículos publicados

1. Alejandro Cruz Sandoval and Wen Yu. Some new stability properties of dynamic neural networks with different time-scales, *International Journal of Neural Systems*, Vol.16, No.3, 191-200, 2006 .
2. Alejandro Cruz Sandoval and Wen Yu. Some stability properties of dynamic neural networks with different time-scales, *Internal Joint Conference on Neural Networks, IJCNN'06*, Vancouver, 8334-8340, 2006 .
3. Alejandro Cruz Sandoval and Wen Yu. Passivity Analysis of Dynamic Neural Networks with Different Time-Scales, *Advances in Neural Networks -ISNN 2006, Srpingerverlag, Lecture Notes in Computer Science*, LNCS 3971, 86-92, 2006

Artículos en revisión

1. Alejandro Cruz Sandoval and Wen Yu. Passivity analysis of dynamic neural networks with different time-scales, sometido para *Neural Processing Letters*.