

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN.

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

DIVISIÓN DE ESTUDIOS DE POSGRADO



**PROBLEMAS DE CONTROL ÓPTIMO Y JUEGOS
DIFERENCIALES EN SISTEMAS NO LINEALES DEL
TIPO POLINOMIAL**

TESIS

GRADO DE DOCTOR EN INGENIERÍA FÍSICA INDUSTRIAL

PRESENTA

VICTORIA CELESTE RODRIGUEZ CARREÓN

DIRECTORES DE TESIS:

MANUEL JIMÉNEZ LIZARRAGA

MIKHAIL V. BASIN

Problemas de Control Óptimo y Juegos Diferenciales en Sistemas no Lineales del tipo Polinomial.

Victoria Celeste Rodríguez Carreón
Universidad Autónoma de Nuevo León
Facultad de Ciencias Físico Matemáticas
División de Estudios de Posgrado

12 de noviembre de 2013

DEDICATORIA

Gracias a esas personas que siempre estuvieron listas para brindarme toda su ayuda. Con todo mi cariño esta tesis se las dedico a ustedes:

 Mi esposo Pablo y a nuestra pequeña hija Paola

AGRADECIMIENTOS

A mis padres por su apoyo incondicional, al Dr. Manuel Jiménez por su valiosa guía, al Dr. Basin por sus excelentes consejos, al CONACyT por su financiamiento económico.

RESUMEN

Esta tesis presenta los siguientes resultados:

Consideramos juegos diferenciales polinomiales determinísticos y estocásticos, estos últimos sobre procesos de observación lineales y afectados por ruidos blancos Gaussianos, ambos en horizonte de tiempo finito e infinito y exploramos la llamada ecuación tipo Riccati dependiente de estado para encontrar un conjunto de estrategias que garanticen un equilibrio ε -Nash en lazo abierto para esta clase particular de juegos no lineales. Cada jugador tiene un funcional de costo cuadrático como el índice de desempeño individual. Este trabajo presenta también la solución del regulador óptimo robusto para sistemas no lineales polinomiales afectados por incertidumbres tanto paramétricas como acopladas, que están basados solo en la información parcial del estado. Los parámetros que describen la dinámica de la planta no lineal polinomial dependen de un vector de parámetros desconocidos, que están contenidos en un conjunto finito de parámetros, y la aplicación de cierta acción de control esta asociada con el valor mas desfavorable de los parámetros desconocidos. Por otro lado se propone un algoritmo de control óptimo para sistemas polinomiales con un criterio cuadrático sobre un horizonte de tiempo infinito. El regulador diseñado da una solución cerrada al problema del controlador óptimo en horizonte de tiempo infinito con un criterio infinito. Finalmente, estudiamos el problema de diseñar estrategias de incentivos para un juego diferencial de Stackelberg para dos jugadores afectados por algún tipo de incertidumbre.

ABSTRACT

This thesis presents the following results:

We considered polynomial differential games and stochastic polynomial differential games, over linear observation processes, affected by Gaussian white noise, both in finite as well as infinite time horizon and we explore the so-called state-dependent Riccati equations to find a set of strategies that guarantee an open loop-Nash equilibrium for this particular class of nonlinear games. Each player has a quadratic cost as the individual performance index. This work also presents a solution to a robust optimal regulation problem for a nonlinear polynomial system affected by parametric and matched uncertainties, which is based only on partial state information. The parameters describing the dynamics of the nonlinear polynomial plant depend on a vector of unknown parameters, which belongs to a finite parametric set, and the application of a certain control input is associated with the worst or least favorable value of the unknown parameter. The other hand is proposes an optimal control algorithm for a polynomial system with a quadratic criterion over infinite horizon. The designed regulator gives a closed form solution to the infinite horizon optimal control problem for a polynomial system with a quadratic criterion. Finally, we study the problem of designing an Incentive Strategy for a Two Players Stackelberg Differential Game affected by some sort of uncertainties.

ÍNDICE GENERAL

1. Introducción	1
1.1. Motivación.	1
1.2. Antecedentes	2
1.3. Aportaciones	4
1.3.1. Control Óptimo para un Sistema Polinomial con un Criterio Cuadrático sobre un Horizonte de Tiempo Infinito.	4
1.3.2. Equilibrio ε -Nash en lazo abierto para juegos diferenciales polinomiales via ecuación tipo Riccati dependiente del estado.	4
1.3.3. Equilibrio ε -Nash en lazo abierto sobre juegos diferenciales polinomiales estocásticos con información incompleta.	5
1.3.4. Control Min-Max para Sistemas Polinomiales Basado en la Salida.	5
1.3.5. Estrategias de Stackelberg con Incentivos Mini-Max para Juegos Diferenciales Lineal-Cuadráticos con Incertidumbres.	5
1.4. Organización de la Tesis	6
2. Marco Teórico	7
2.1. Control Óptimo.	7
2.1.1. Técnicas de Optimización.	7
2.1.2. Regulador Lineal Cuadrático.	10
2.2. Juegos Diferenciales no cooperativos.	12
2.2.1. Horizonte de Tiempo Finito.	13
2.2.2. Horizonte de Tiempo Infinito.	14
2.3. Modos Deslizantes.	15
2.3.1. Observador por modos deslizantes de Segundo Orden.	15
2.3.2. Diferenciador por Modos Deslizantes de Alto Orden.	16
2.4. Procesos Estocásticos.	16
2.4.1. Ecuaciones Diferenciales Estocásticas.	18
2.4.2. Teoría de Filtrado.	19

3. Control Óptimo para un Sistema Polinomial con un Criterio Cuadrático sobre un Horizonte de tiempo Infinito.	22
3.1. Planteamiento del Problema.	22
3.2. Diseño del Control I. Principio del Máximo.	23
3.2.1. Solución del Problema del Controlador Óptimo I.	23
3.2.2. Ejemplo I.	23
3.3. Diseño del Control II. Ecuación Hamilton-Jacobi-Bellman.	25
3.3.1. Solución del problema del controlador óptimo II.	25
3.3.2. Ejemplo II.	27
3.3.3. Ejemplo III.	28
3.3.4. Apéndice.	33
4. Equilibrio ε-Nash en lazo abierto para juegos diferenciales polinomiales via ecuación tipo Riccati dependiente del estado.	36
4.1. Planteamiento del problema.	36
4.1.1. Suposiciones y definiciones básicas.	37
4.2. Equilibrio ε -Nash en lazo abierto para juegos polinomiales.	39
4.2.1. Caso de tiempo finito.	39
4.2.2. Equilibrio ε -Nash en lazo abierto via ecuación tipo Riccati dependiente del estado	40
4.3. Caso Horizonte de tiempo Infinito	43
4.3.1. Solución de las ecuaciones algebraicas acopladas tipo Riccati dependientes del estado por un enfoque Hamiltoniano.	44
4.4. Ejemplos Numéricos.	46
4.4.1. Ejemplo 1	46
4.4.2. Ejemplo 2	46
4.4.3. Example 3	48
4.4.4. Ejemplo 4	52
4.4.5. Ejemplo 5.	53
4.5. Apéndice	55
5. Equilibrio ε-Nash en lazo abierto sobre juegos diferenciales polinomiales estocásticos con información incompleta.	61
5.1. Planteamiento del problema.	61
5.2. Diseño del Filtro.	63
5.3. Equilibrio ε -Nash en lazo abierto: Diseño del controlador con horizonte de Tiempo finito	64
5.3.1. Diseño del controlador óptimo.	64
5.3.2. Diseño de la estrategias del equilibrio ε -Nash	66

5.3.3.	Ejemplo I.	67
5.4.	Equilibrio ε -Nash en Lazo Abierto: Diseño del Controlador con Horizonte de Tiempo Infinito. . .	69
5.4.1.	Diseño del controlador sub-Óptimo.	69
5.4.2.	Resolviendo la ecuación algebraica acoplada tipo Riccati dependiente del estimado del estado por la aproximación Hamiltoniana	73
5.4.3.	Ejemplo II.	74
6.	Control Min-Max para Sistemas Polinomiales Basado en la Salida.	77
6.1.	Planteamiento del Problema.	77
6.2.	Control Min-Max para Sistemas Polinomiales.	79
6.3.	Diseño del Observador por Modos Deslizantes de Alto Orden para Sistemas Polinomiales.	80
6.4.	Modos Deslizantes integrales para la salida de Sistemas polinomiales con Incertidumbres.	81
6.5.	Ejemplos.	83
6.5.1.	Circuito No Lineal.	83
6.5.2.	Péndulo	84
7.	Estrategias de Stackelberg con Incentivos Mini-Max para Juegos Diferenciales Lineal-Cuadraticos con Incertidumbres.	91
7.1.	Descripción del problema y suposiciones básicas	91
7.1.1.	Breve reseña del formato de incentivo.	93
7.2.	Solución en Equipo Óptima Mini-Max.	94
7.2.1.	Diseño de incentivos para hacer cumplir la solución de equipo óptimo mini-max del líder .	95
7.3.	Incentivo Mini-Max para un Juego Estacionario de Stackelberg con Horizonte de Tiempo Infinito. .	98
7.3.1.	Algoritmo 1 Solución del problema de equipo óptimo mini-max	100
7.3.2.	Algoritmo 2 Encontrar la estrategia de incentivos Mini-Max.	100
7.4.	Ejemplos Numéricos	100
7.4.1.	Ejemplo escalar.	100
7.4.2.	Ejemplo vectorial.	104
8.	Conclusion y trabajo futuro.	108
8.1.	Conclusiones.	108
8.2.	Trabajo Futuro.	108
8.3.	Artículos.	109

1

INTRODUCCIÓN

1.1. MOTIVACIÓN.

Esta tesis se trata de la teoría de juegos diferenciales dinámicos no cooperativos. La teoría de juegos consiste en la toma de decisiones de varios participantes; es no cooperativo si cada participante involucrado persigue sus propios intereses que están en conflicto parcial o total con los demás.

A pesar de que la noción de conflicto es tan antiguo como la humanidad, el enfoque científico ha comenzado hace poco, en los alrededores de los años 30's, con una corriente creciente de publicaciones científicas. También vemos que cada vez mas disciplinas científicas dedican tiempo y atención al análisis de las situaciones conflictivas. Estas disciplinas son las matemáticas, la economía, la ingeniería, la aeronáutica, la sociología, la política y las matemáticas financieras.

Los participantes involucrados, también llamados jugadores o los tomadores de decisiones, o simplemente personas, no siempre tienen el control completo sobre los resultados. A veces hay incertidumbres que influyen en el resultado de un modo no predecible.

Los juegos diferenciales hasta hoy estudiados tienen dinámicas del sistema lineal y con funciones de desempeño cuadráticas para cada jugador. Sin embargo, en muchas aplicaciones, el modelado original lineal no se puede adaptar a todas las situaciones en la practica, que son principalmente no lineales por naturaleza. Por lo tanto, se necesita extender el concepto de equilibrio a cierta clase de sistemas no lineales, esto es, sistemas polinomiales el cual es el tema principal de esta tesis.

Las dinámicas polinomiales representa una clase importante de dinámicas no lineales, ya que pueden aproximarse a una gran variedad de funciones no lineales, manteniendo la complejidad en un nivel manejable y específico. A comparación del caso lineal cuadrático, no existen muchos trabajos de juegos diferenciales no lineales, y particularmente, no se han obtenido resultado para juegos diferenciales polinomiales.

El problema del regulador óptimo para sistemas no lineales esta abierto, dado que solo aproximaciones numéri-

cas y computacionales han sido propuestas para tales problemas. Mas frecuentemente, el sistema no lineal es reformulado como

$$\dot{z}(t) = A(z(t))z(t) + G(t)v(t),$$

donde $A(z(t))$ es una matriz cuadrada dependiente del estado. El problema del controlador óptimo es entonces resuelto para sistemas lineales en pequeños intervalos de tiempo, donde la matriz $A(z(t))$ se asume constante en el valor del intervalo inicial de $z(t)$. La solución en cada intervalo se busca con los métodos convencionales para el problema del regulador lineal cuadrático. Esta aproximación es entonces repetida en cada intervalo pequeño de tiempo.

1.2. ANTECEDENTES

Entre los diferentes campos como la ingeniería, ecología, gestión y economía nos encontramos con situaciones que involucran varios tomadores de decisiones (o jugadores) con diferentes metas u objetivos interrelacionados con la misma decision del proceso. De tal forma que todos ellos están interactuando y son influenciados cada uno por la decision de los demás: la decision tomada por cualquiera de los participantes afecta a los otros, y viceversa. Particularmente, cuando el proceso subyacente evoluciona con el tiempo, este tipo de problemas son a menudo optimizados usando teoría de juegos. Esta teoría fue iniciada en el trabajo de Isaacs [1]: se enfoco principalmente en los juegos de suma cero. Mas tarde, los juegos diferenciales de suma no cero fueron introducidos en [2] y [3]. En estos juegos, cada jugador busca minimizar su propio criterio. El articulo [2] deriva las condiciones suficientes de existencia del equilibrio retroalimentado en horizonte de tiempo finito, pero solo en el caso de juegos lineal cuadráticos regidos por una dinámica lineal y un criterio cuadrático, (ver [4] para mas detalles). Usualmente, para resolver una situación conflictiva el equilibrio de Nash es aplicado (o, en el caso general, equilibrio ε -Nash) [5], [6], [7]. Es conocido que el equilibrio de Nash es una solución natural en el contexto no cooperativo. Sin embargo, si nos referimos a las dinámicas complejas no lineales, es mas apropiado aplicar el concepto de ε -equilibrio de Nash, dado que proporciona mayor flexibilidad en la selección de las estrategias del equilibrio (ver [6], [8]). Unos pocos trabajos relacionados con juegos diferenciales no lineales pueden ser mencionados. El articulo [9], presenta una solución en el caso particular de un juego no lineal representando un problema de la contaminación y gestión de recursos. El articulo [10] identifica el comportamiento potencialmente caótico en un equilibrio de Nash Markoviano en un modelo de duopolio de tiempo discreto de competencia publicitaria. El reciente articulo [11] propone un método interactivo de programación dinámica adoptivo para resolver un tipo particular de juegos llamado juegos de dos jugadores de suma cero, Todas estas publicaciones expresan el interés en buscar estrategias de equilibrios en sistemas no lineales complejos.

Por otro lado, también se han desarrollado la contraparte estocástica (ver [4], [8], [12]) para juegos diferenciales con funcional de costo cuadrática, pero hasta la fecha no hubo ningún intento de resolver el problema del equilibrio de Nash en lazo abierto para juegos diferenciales polinomiales estocásticos con respecto a un funcional de costo cuadrático tipo Bolza-Meyer como índice de desempeño para cada jugador.

Desde la publicación del artículo innovador escrito por Kalman en 1960 [13], sentó la base de lo que después se conoció como problema del regulador lineal cuadrático, este trabajo tubo un gran impacto en el area de control automático e influyo a un gran numero de profesionales e investigadores del control. Eventualmente, después de esta publicación, la necesidad de extender el enfoque desarrollado se hizo evidente en algunas situaciones mas complejas o generales, tal como sistemas no lineales, incertidumbres en los modelos, e información incompleta en las variables de estado, En efecto, en algunas aplicaciones, el modelo lineal original no puede servir para todas las situaciones en la practica, que son generalmente de naturaleza no lineal. Por ejemplo, dos maneras de aproximar la obtención del controlador óptimo para sistemas no lineales son el principio del máximo de Pontryagin [14] y programación dinámica [15]. Desafortunadamente, estas herramientas no proveen una forma explícita para el control óptimo en la mayoría de los casos, Sin embargo, existen una la larga tradición de diseñar el control óptimo para sistemas no lineales (ver. por ejemplo, [16]–[17]) y el diseño de filtrado óptimo en forma cerrada para plantas no lineales [18, 19, 20], [21], en particular, sistemas polinomiales ([22], [23]). El segundo punto critico es la existencia de incertidumbres en los modelos y la necesidad de optimizar el sistema (ver [24], [25], [26]). Tal problema puede ser visto desde diferentes perspectivas y enfoques: uno de ellos es considerar que los parámetros del sistema se encuentran en un conjunto de parámetros finitos, donde cada valor de los parámetros caracteriza una posible dinámica del modelo, y el diseño del control robusto deberá enfrentar con todas simultáneamente. Para este tipo de problemas de optimización, la version robusta del tradicional principio de máximo fue desarrollada y se conoce como el principio del máximo robusto (RMP) (ver [27], [28], [29], y [30]) permite diseñar un ley de control de tipo min-max para el problema de multi parámetro, donde cada conjunto de parámetros es visto como una posible realización de la dinámica del sistema. Otra clase importante de incertidumbres son las incertidumbres acopladas, que han sido extensamente estudiadas en la teoría de sistemas de estructura variable y control por modos deslizantes. Recientes avances en el control y observación por modos deslizantes de alto orden en algoritmos para diseñar una amplia clase de leyes de control basado en la salida medible del sistema. La capacidad de estas aproximaciones se revela en el desarrollo de practicas de diferenciadores exactos robustos de orden arbitrario en tiempo real [31], [32], [33]. El diferenciador por modos deslizantes de alto orden es usado como observadores para estimación del estado del sistema en la presencia de disturbios externos desconocidos [34], [35] y [36], que también es reconstruido en linea por el observador.

El equilibrio de Stackelberg (ver [37], [38]) ha tenido un impacto a largo plazo en diferentes areas, tales como la macroeconomía (como por ejemplo [39], [40] y [41]), ciencia de gestión ([42]), ciencias ambientales ([43]), etc. Desde la formulación dinámica del equilibrio estático de Líder-Seguidor original debido a Von Stackelberg ([44]), el equilibrio L-S ha encontrado su lugar como una forma natural para modelar la competencia estratégica donde existen asimetrías o algún tipo de jerarquía entre los tomadores de decisiones ([8]). Una de las principales críticas que se pueden tener en contra de este tipo de equilibrio es el hecho de que el líder se supone tiene el conocimiento exacto de los parámetros del juego pero, para muchas aplicaciones, esta suposición no es realista (ver [41], [4]). Por otra parte el líder también enfrentan el problema de las reacciones no óptimas del seguidor que afectarán el valor final de su función de costo. Para mitigar esta última situación, diferentes autores han propuesto los llamados

mecanismos de incentivos (ver [45], [46], [47], [48], [49]), tal mecanismo esta basado en la idea de declarar una penalizacion para el seguidor dependiente de su acción o decision particular elegida, esto se hace para inducir un comportamiento particular del seguidor que es favorable para el líder. En los juegos de regulacion de Stackelberg e incentivos de Stackelberg, el líder optimiza de tal forma que cuando el seguidor optimice el objetivo del seguidor, el objetivo del líder se optimice también. En caso de regulacion de Stackelberg, el espacio de posibilidades para el líder esta restringido, y en general, no es capaz de alcanzar su óptimo global. En el caso de juegos con incentivos de Stackelberg, debido al especial diseño del incentivo, el líder siempre es capaz de alcanzar su óptimo global.

1.3. APORTACIONES

1.3.1. CONTROL ÓPTIMO PARA UN SISTEMA POLINOMIAL CON UN CRITERIO CUADRÁTICO SOBRE UN HORIZONTE DE TIEMPO INFINITO.

En el capitulo 3 se presenta un algoritmo de control óptimo para sistemas polinomiales con un criterio cuadrático con horizonte de tiempo infinito. El regulador diseñado da una solución en forma cerrada para el problema del controlador óptimo con horizonte de tiempo infinito para un sistema polinomial y un criterio cuadrático. Una ecuación algebraica tipo Riccati dependiente del estado es introducida y resuelta considerando el estado $z(t)$ como un argumento fijo y utilizando métodos numéricos. La solución obtenida consiste en una ley de control retroalimentada obtenida por resolver la ecuación algebraica tipo Riccati dependiente del estado.

1.3.2. EQUILIBRIO ε -NASH EN LAZO ABIERTO PARA JUEGOS DIFERENCIALES POLINOMIALES VIA ECUACIÓN TIPO RICCATI DEPENDIENTE DEL ESTADO.

En el capitulo 4 se estudia el juego diferencial polinomial en horizonte tanto de tiempo finito como infinito, en ambos casos, se explora la llamada ecuación tipo Riccati dependiente del estado para encontrar un conjunto de estrategias que garanticen un equilibrio de Nash en lazo abierto, para este particular caso de juegos diferenciales, tal método presenta ventajas en simplicidad del diseño de las estrategias de Nash y proporciona un algoritmo de solución computacional efectivo. Se demuestra que esta solución lleva al juego a un ε -equilibrio y se provee una cota superior de ε , la solución propuesta es dada como un conjunto de N ecuaciones diferenciales acopladas tipo Riccati dependientes del estado, donde cada ecuación incluye una forma p -lineal para representar esta parte del tensor polinomial. Se da un algoritmo para buscar las soluciones de las ecuaciones algebraicas tipo Riccati dependientes del estado en el caso de horizonte de tiempo infinito basado en la aproximación Hamiltoniana. Un procedimiento numérico es detallado para buscar la solución de este conjunto de estrategias. Ejemplos numéricos son presentados para ilustrar la efectividad de la aproximación.

1.3.3. EQUILIBRIO ε -NASH EN LAZO ABIERTO SOBRE JUEGOS DIFERENCIALES POLINOMIALES ESTOCÁSTICOS CON INFORMACIÓN INCOMPLETA.

En el capítulo 5 se presenta una solución equilibrio ε -Nash para el problema de buscar estrategias de Nash en lazo abierto para juegos diferenciales polinomiales estocásticos con información incompleta sobre observaciones lineales, también para ambos casos en horizonte de tiempo finito e infinito. Cada jugador tiene un costo lineal cuadrático como el índice de desempeño individual. El diseño de estrategias propuesto está basado en el llamado ecuación tipo Riccati dependiente del estado. También se presentan ejemplos ilustrativos para verificar la efectividad del algoritmo.

1.3.4. CONTROL MIN-MAX PARA SISTEMAS POLINOMIALES BASADO EN LA SALIDA.

En el capítulo 6 se presenta la solución al problema al regulador óptimo robusto para un sistema no lineal polinomial afectado por incertidumbres paramétricas y acopladas, que está basado solo en información parcial del estado. Los parámetros que describen la dinámica de la planta no lineal polinomial dependen de un vector de parámetros desconocidos, que se encuentran en un conjunto de parámetros finito, y la aplicación de una cierta entrada control asociada con el peor o menos favorable valor de los parámetros desconocidos. Un diferenciador por modos deslizantes de alto orden es diseñado para la planta no lineal de tal forma que el previo control puede ser diseñado para un sistema con información incompleta. Además, las incertidumbres acopladas son también compensadas por medio de la mismo regulador basado en la salida. El algoritmo obtenido es aplicado para regular un circuito inductor no lineal incierto de tercer orden y a un péndulo mecánico aproximado con un polinomio de tercer orden, para verificar la eficiencia de la aproximación desarrollada.

1.3.5. ESTRATEGIAS DE STACKELBERG CON INCENTIVOS MINI-MAX PARA JUEGOS DIFERENCIALES LINEAL-CUADRÁTICOS CON INCERTIDUMBRES.

En el capítulo 7 se estudia el problema de diseñar una estrategia con incentivos para un juego diferencial de Stackelberg de dos jugadores por algún tipo de incertidumbres. Como es tradicionalmente entendido en la teoría estándar de incentivos, el líder tiene completo conocimiento de los parámetros del sistema del juego, tal que el puede calcular su estrategia que lleve al juego a un óptimo global que es favorable para el líder. La mayor parte del trabajo existente se dedica a esta situación, sin embargo esta hipótesis no es realista. Este capítulo propone un plan de incentivos para el juego de Stackelberg en que los parámetros que describen la dinámica del juego dependen de un vector desconocido contenido en un conjunto de parámetros finito y la solución de las estrategias de incentivos es dada en términos del escenario del peor caso. Basado en el principio del máximo robusto el nuevo esquema de incentivos es presentado en la forma de control retroalimentado Min-Max. Ejemplos numéricos son presentados para ilustrar la efectividad de la aproximación.

1.4. ORGANIZACIÓN DE LA TESIS

En el capítulo 2 se presenta una síntesis teórica de control, juegos diferenciales no cooperativos, teoría de modos deslizantes, procesos estocásticos y teoría de filtrado. En el capítulo 3 se considera el problema del regulador óptimo para sistemas polinomiales en horizonte de tiempo infinito con criterio cuadrático. En el capítulo 4 se plantea y resuelve el problema de diseño de estrategias del equilibrio de Nash en lazo abierto para juegos diferenciales polinomiales vía ecuación tipo Riccati dependiente de estado. En el capítulo 5 se resuelve el problema de diseño de estrategias del equilibrio de Nash en lazo abierto sobre juegos diferenciales polinomiales estocásticos con información incompleta. En el capítulo 6 se obtiene la solución del problema de diseño de control Min-Max para sistemas polinomiales con información incompleta. En el capítulo 7 se considera el problema de diseño de estrategias con incentivos Min-Max para juegos diferenciales lineales cuadráticos con incertidumbres. En el capítulo 8 se presentan las conclusiones y los trabajos futuros a desarrollar.

2

MARCO TEÓRICO

2.1. CONTROL ÓPTIMO.

En un problema de control óptimo un criterio de optimalidad es dado, que asigna un cierto valor a cada evolución del sistema dinámico subyacente. El problema de control óptimo es entonces buscar un control admisible que minimice al criterio de optimalidad.

En esta sección derivaremos técnicas de optimización para el problema básico de optimización para buscar una acción de control $u(t)$ que minimiza el funcional de costo

$$J(x_0, u) = \int_0^{\infty} g(t, x(t), u(t))dt + h(x(T)), \quad (2.1)$$

donde la variable de estado $x(t)$ satisface la ecuación diferencial:

$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(t_0) = x_0. \quad (2.2)$$

Aquí, $x(t) \in \mathbb{R}^n$ es llamado el estado del sistema, x_0 es la condición inicial del sistema, $u(t) \in \mathbb{R}^m$ es la acción de control.

2.1.1. TÉCNICAS DE OPTIMIZACIÓN.

Ecuación de Euler-Lagrange.

En este caso se considera las siguientes suposiciones:

- $f(t, x(t), u(t))$ y $g(t, x(t))$ son funciones continuas en \mathbb{R}^{1+n+m} . Además, para f y g todas las derivadas parciales con respecto a x y u existen y son continuas.
- $h(x) \in \mathcal{C}^1$
- El conjunto de controles admisibles, \mathcal{U} , consiste del conjunto de funciones que son continuas en $[0, T]$.

De la teoría de optimización estática se introduce para cada $t \in [0, T]$ la cantidad $\lambda(t)[f(t, x(t), u(t)) - \dot{x}(t)]$, donde el multiplicador de Lagrange $\lambda(t)$ (también conocido como co-estado) es un vector fila arbitrario. Dado que $f(t, x(t), u(t)) - \dot{x}(t) = 0$ para todo t , en particular

$$\int_0^T [f(t, x(t), u(t)) - \dot{x}(t)] dt = 0. \quad (2.3)$$

Podemos añadir la cantidad (2.3) en la función de costo (2.1) para obtener una función de costo \mathbf{J} que coincide con la función de costo original J si la restricción dinámica de la ecuación (2.2) es satisfecha. Esto es,

$$\mathbf{J} = \int_0^T (g(t, x(t), u(t)) + \lambda(t)f(t, x(t), u(t)) - \lambda(t)\dot{x}(t))dt + h(x(T)).$$

Introduciendo la función Hamiltoniana H como

$$H(t, x, u, \lambda) = g(t, x, u) + \lambda(t)f(t, x, u), \quad (2.4)$$

se puede reescribir \mathbf{J} como

$$\mathbf{J} = \int_0^T (H(t, x, u, \lambda) - \lambda(t)\dot{x}(t))dt + h(x(T)). \quad (2.5)$$

Integrando por partes se muestra que

$$-\int_0^T \lambda(t)\dot{x}(t)dt = -\lambda(T)x(T) + \lambda(0)x_0 + \int_0^T \dot{\lambda}(t)x(t)dt. \quad (2.6)$$

Por lo tanto, sustituyendo este resultado en la ecuación (2.5), \mathbf{J} puede ser re-escrita como

$$\mathbf{J} = \int_0^T (H(t, x, u, \lambda) + \dot{\lambda}(t)x(t))dt + h(x(T)) - \lambda(T)x(T) + \lambda(0)x_0. \quad (2.7)$$

Esta expresión tiene tres términos aditivos,

$$\mathbf{J}_1 = \int_0^T (H(t, x, u, \lambda) + \dot{\lambda}(t)x(t))dt,$$

$$\mathbf{J}_2 = h(x(T)) - \lambda(T)x(T),$$

$$\mathbf{J}_3 = \lambda(0)x_0,$$

el primer término, \mathbf{J}_1 , se refiere a todo el periodo $[0, T]$, el término \mathbf{J}_2 se refiere exclusivamente al tiempo final, y el tercer término, \mathbf{J}_3 , se refiere solo al tiempo inicial.

De nuevo, el hecho de seleccionar $\lambda(t)$ no afecta al valor de \mathbf{J} , siempre que el estado $x(t)$ satisfaga la ecuación (2.2), o dicho de otra forma

$$\dot{x}(t) = \frac{\partial H}{\partial \lambda}, \quad (2.8)$$

para todo $t \in [0, T]$.

Teorema 2.1 Considere el problema de optimización dado por las ecuaciones (2.1) y (2.2). Sea $H(t, x, u, \lambda) = g(t, x, u) + \lambda(t)f(t, x, u)$. Asumamos que las funciones f, g y h satisfacen las suposiciones anteriores. Si $u(t) \in \mathcal{U}$

es un control que da un mínimo local para el funcional de costo (2.1), y $x^*(t)$ y $\lambda^*(t)$ son el correspondiente estado y co-estado, entonces es necesario que

$$\dot{x}(t) = f(t, x^*, u^*) \left(= \frac{\partial H(t, x, u, \lambda)}{\partial \lambda} \right), \quad (2.9)$$

$$\dot{\lambda}(t) = \frac{\partial H(t, x, u, \lambda)}{\partial x}, \quad \lambda^*(T) = \frac{\partial h(x^*(T))}{\partial x}, \quad (2.10)$$

y, para todo $t \in [0, T]$,

$$\frac{\partial H(t, x, u, \lambda)}{\partial u} = 0. \quad (2.11)$$

Principio del Máximo de Pontryagin.

En la sub-sección anterior se asumió que el control era continuo. En aplicaciones, esta suposición es, sin embargo, a menudo demasiado restrictiva. Por lo tanto, un problema mas general seria asumir que existe un subconjunto $U \in \mathbb{R}^m$ tal que para todo $t \in [0, T]$, $u(t) \in U$. Así que, $u(t)$ es restringido al conjunto U . Adicional a las suposiciones de la sub-sección anterior, se asume que el conjunto de controles admisibles \mathcal{U} ahora consiste del conjunto de funciones medibles de $[0, T]$ en U para que las ecuaciones diferenciales (2.2) tiene una solución única en $[0, T]$ y la función de costo (2.1) existe. Esto es,

$$\mathcal{U} = \{u(t) \mid u(t) \in U, \forall t \in [0, T]; u(t) \text{ es medible en } [0, T]; x(t, u(t)) \text{ es una función absolutamente continua que satisface (2.2) en } [0, T] \text{ excepto en un conjunto } S \text{ que tiene una medida de Lebesgue cero; y la ecuación (2.1) existe.}\}$$

Observaciones.

- Dado que f y g son continuas en u , $f(t, x(t), u(t))$ y $g(t, x(t), u(t))$ son también funciones medibles. Como consecuencia la función de costo (2.1) existe si también se asume que U es acotado.
- Si f satisface la desigualdad

$$|f(t, x(t), u(t))| \leq L |x| + N, \quad t \in [0, T], x \in \mathbb{R}^n, u \in U,$$

Para algunas constantes positivas L y N , la ecuación diferencial (2.2) tiene una solución local única para todo $u(t)$ medible.

Teorema 2.2 Considere el problema de optimización dado por las ecuaciones (2.1) y (2.2). Asumamos que las funciones f , g y h satisfacen las suposiciones de la sub-sección anterior e introduzcamos $H = g + \lambda f$. $u^*(t) \in \mathcal{U}$ es un control que da un mínimo local para el funcional de costo (2.1), y sea $x^*(t)$ la correspondiente trayectoria del estado. Entonces existe una función de co-estado $[\lambda^*]^T : [0, T] \rightarrow \mathbb{R}^n$ que satisface

$$\dot{x}^*(t) = f(t, x^*, u^*) \left(= \frac{\partial H(t, x^*, u^*, \lambda)}{\partial \lambda} \right), \quad x^*(0) = x_0, \quad (2.12)$$

$$\dot{\lambda}^*(t) = \frac{\partial H(t, x^*, u^*, \lambda^*)}{\partial x}, \quad \lambda^*(T) = \frac{\partial h(x^*(T))}{\partial x}, \quad (2.13)$$

y, para todo $t \in [0, T]$ en la que u es continua,

$$H(t, x^*, u^*, \lambda^*) = \min_{u \in U} H(t, x^*, u, \lambda^*), \quad (2.14)$$

esto es

$$u^*(t) = \arg \min_{u \in U} H(t, x^*, u, \lambda^*). \quad (2.15)$$

Note que, aparte del hecho de que consideramos un conjunto más amplio de funciones de control admisibles, la principal diferencia entre el teorema de Pontryagin y el teorema de Euler-Lagrange es que el principio del máximo nos dice que el hamiltoniano se reduce al mínimo en la trayectoria óptima y que también es aplicable cuando el mínimo se alcanza en la frontera de U .

2.1.2. REGULADOR LINEAL CUADRÁTICO.

Horizonte de Tiempo Finito.

Considere la minimización de la función de costo definida como

$$J = \int_0^T \{x^T(t)Qx(t) + u^T(t)Ru(t)\}dt + x^T(T)Q_Tx(T), \quad (2.16)$$

sujeto a

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad (2.17)$$

donde R es una matriz definida positiva y Q , R y Q_T son matrices simétricas. El problema es conocido en la literatura como el problema del controlador lineal cuadrático. La solución de este problema esta relacionado con la existencia de la solución simétrica de la siguiente ecuación diferencial matricial tipo Riccati (RDE).

$$\dot{K}(t) = -A^TK(t) - K(t)A + K(t)SK(t) - Q, \quad K(T) = Q_T, \quad (2.18)$$

donde $S = BR^{-1}B^T$.

Teorema 2.3 El problema del controlador lineal cuadrático (2.16) y (2.17) tiene, para cualquier condición inicial del estado x_0 , una solución si y solo si la ecuación diferencial tipo Riccati (2.18) tiene una solución simétrica $K(t)$ en $[0, T]$. Si el problema del controlador lineal cuadrático tiene una solución, entonces es única y el control óptimo en lazo cerrado es

$$u^*(t) = -R^{-1}B^TK(t)x(t), \quad (2.19)$$

Mas aun, $J(u^*) = x_0^TK(0)x_0$.

También considere el problema del controlador lineal cuadrático no homogéneo. Esto es, la minimización de

$$J = \int_0^T \{x^T(t)Qx(t) + u^T(t)Ru(t)\}dt + x^T(T)Q_Tx(T), \quad (2.20)$$

sujeto a

$$\dot{x}(t) = Ax(t) + Bu(t) + c(t), \quad x(0) = x_0, \quad (2.21)$$

bajo las mismas suposiciones de las matrices R , Q y Q_T . El termino no homogéneo, $c(t)$ se asume que es una función conocida que es tal que la solución $x(t)$ de la ecuación diferencial es definida únicamente en el sentido extendido.

Teorema 2.4 Sea $c(t)$ una función arbitraria cuadrático integrable. Consideremos la minimización del funcional de costo lineal cuadrático (2.20) sujeto a la dinámica del estado

$$\dot{x}(t) = Ax(t) + Bu(t) + c(t), \quad x(0) = x_0, \quad (2.22)$$

Entonces, el problema lineal cuadrático (2.20) y (2.22) tiene una solución para toda $x_0 \in \mathbb{R}^n$ si y solo si la ecuación diferencial tipo Riccati (2.18) tiene una solución simétrica $K(t)$ en $[0, T]$. Mas aun, el control óptimo en lazo cerrado es

$$u^*(t) = -R^{-1}B^T(K(t)x(t) + m(t)), \quad (2.23)$$

donde $m(t)$ es la solución de la ecuación diferencial lineal

$$\dot{m}(t) = (K(t)S - A^T)m(t) - K(t)c(t), \quad m(T) = 0, \quad (2.24)$$

y $x(t)$ es la solución de la ecuación diferencial

$$\dot{x}(t) = (A - SK(t))x(t) - Sm(t) + c(t), \quad x(0) = x_0.$$

Horizonte de Tiempo Infinito.

En esta sub-sección se considera el problema de buscar una función de control $u(t) = Fx(t)$ (donde F es una matriz invariante en el tiempo) para cada $x_0 \in \mathbb{R}^n$ que minimice el funcional de costo

$$J(x_0, u) = \int_0^\infty \{x^T(t)Qx(t) + u^T(t)Ru(t)\}dt, \quad (2.25)$$

bajo la restricción adicional que $\lim_{t \rightarrow \infty} x(t) = 0$. Aquí $Q = Q^T$, $R > 0$, y la variable de estado x es la solución de

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad (2.26)$$

En esta sub-sección consideraremos que el par (A, B) es estabilizable. Por otra parte, se introduce el conjunto de matrices de retroalimentación invariante en el tiempo, estables y lineales, esto es,

$$\mathcal{F} = \{F \mid A + BF \text{ es estable}\},$$

Por conveniencia de notación se usara $S = BR^{-1}B^T$.

Teorema 2.5 Supongamos que (A, B) es estabilizante y $u = Fx$, con $F \in \mathcal{F}$. El problema del controlador lineal cuadrático (2.25) y (2.26) tiene un mínimo $\hat{F} \in \mathcal{F}$ para cada x_0 si y solo si la ecuación algebraica tipo Riccati definida como

$$Q + A^T X + XA - XSX = 0, \quad (2.27)$$

tiene una solución simétrica estable K . Si el problema del controlador lineal cuadrático tiene una solución, entonces la solución es única y esta dada por $\hat{F} = -R^{-1}B^TK$ y el control óptimo en lazo cerrado es

$$u^*(t) = -R^{-1}B^TK(t)x(t). \quad (2.28)$$

Además, $J(u^*) = x_0^TKx_0$.

2.2. JUEGOS DIFERENCIALES NO COOPERATIVOS.

En esta sección consideraremos la situación donde existen mas de un jugador. Cada uno de estos N jugadores tienen una función de costo cuadrática que desean minimizar.

$$J_i = \int_0^T \left\{ x^T(t)Q^i x(t) + \sum_{j=1}^N u^{jT}(t)R^{ij}u^j(t) \right\} dt + x^T(T)Q_f^i x(T), \quad i = 1, 2, \dots, N. \quad (2.29)$$

donde Q^i , R^{ij} y Q_f^i son matrices simétricas y R^{ij} es definida positiva. En la minimización una variable de estado $x(t)$ ocurre. Esta es una dinámica que puede ser influenciada por todos los jugadores. Esta es,

$$\dot{x}(t) = Ax(t) + B_1u_1(t) + \dots + B_Nu_N(t), \quad x(0) = x_0, \quad (2.30)$$

donde A y B_i , $i = 1, \dots, N$, son matrices constantes, y u_i es un vector que puede ser manipulado por el i -ésimo jugador. El objetivo de cada jugador es la minimización de su propia función de costo por la elección de los controles apropiados para el subyacente sistema dinámico lineal. El aspecto no cooperativo implica que los jugadores se supone que no colaboran con este objetivo. Dependiendo de la información disponible que tienen los jugadores en el juego, denotado por $\eta_i(t)$, $t \in [0, T]$, y al conjunto de estrategias que los jugadores pueden elegir (que depende obviamente de la información que tienen los jugadores en el juego), denotado por Γ_i , las acciones (o controles) de los jugadores son determinados por la relación

$$u_i = \gamma_i(\eta_i), \quad \text{donde } \gamma_i \in \Gamma_i.$$

Sustituyendo estos controles en (2.29) y (2.30) muestra que el funcional de costo J_i depende de la información que tienen los jugadores y los espacios de estrategias. Dependiendo de estos parámetros el valor del funcional de costo obviamente también depende para cada jugador i sobre la acción llevada a cabo de los otros jugadores.

Definición 2.6 Un conjunto de acciones admisibles (u_1^*, \dots, u_N^*) es un equilibrio de Nash para un juego de N jugadores, donde cada jugador tiene un funcional de costo $J_i(u_1, \dots, u_N)$, si para todo admisible (u_1, \dots, u_N) la desigualdad siguiente se mantiene:

$$J_i(u_1^*, \dots, u_{i-1}^*, u_i^*, u_{i+1}^*, \dots, u_N^*) \leq J_i(u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_N^*).$$

Así, el equilibrio de Nash es definido tal que tiene la propiedad que no existen incentivos para cualquier desviación unilateral de algún jugador. Tenga en cuenta, en general, no se puede esperar tener un único equilibrio de Nash. Además, es fácil verificar que siempre que un conjunto de acciones (u_1^*, \dots, u_N^*) es un equilibrio de Nash para el juego con funcionales de costo J_i , $i = 1, \dots, N$, estas acciones también constituyen un equilibrio de Nash para el juego con funcionales de costo $\alpha_i J_i$, $i = 1, \dots, N$, para cualquier selección de $\alpha_i > 0$.

2.2.1. HORIZONTE DE TIEMPO FINITO.

En esta sub-sección se restringirá el análisis al caso de dos jugadores, como un caso particular del caso mas general de N -jugadores. Considere el sistema descrito por:

$$\dot{x}(t) = Ax(t) + B_1u_1(t) + B_2u_2(t), \quad x(0) = x_0, \quad (2.31)$$

donde x es el estado n dimensional del estado del sistema, u_i es un vector (control) m_i dimensional del jugador i -ésimo, $i = 1, 2$, se puede manipular, x_0 es la condición inicial del estado del sistema, A , B_1 y B_2 son matrices constantes de dimensiones apropiadas. El funcional de costo del i -ésimo jugador, $i = 1, 2$, que tiene como objetivo minimizar es:

$$J_1(u_1, u_2) = \int_0^T \{x^T(t)Q^1x(t) + u_1^T(t)R^{11}u_1(t) + u_2^T(t)R^{12}u_2(t)\} + x^T(T)Q_f^1x(T), \quad (2.32)$$

y

$$J_2(u_1, u_2) = \int_0^T \{x^T(t)Q^2x(t) + u_1^T(t)R^{21}u_1(t) + u_2^T(t)R^{22}u_2(t)\} + x^T(T)Q_f^2x(T), \quad (2.33)$$

en que todas las matrices son simétricas y, además, R^{ii} son definidas positivas. Usando la notación compacta $S_i = B_i(R^{ii})^{-1}B_i^T$, el siguiente teorema puede ser planteado.

Teorema 2.7 Asumamos:

1. Que el conjunto de ecuaciones diferenciales acopladas tipo Riccati

$$\dot{P}_1 = -A^T P_1 - P_1 A - Q^1 + P_1 S_1 P_1 + P_1 S_2 P_2, \quad P_1(T) = Q_f^1,$$

$$\dot{P}_2 = -A^T P_2 - P_2 A - Q^2 + P_2 S_1 P_1 + P_2 S_2 P_2, \quad P_2(T) = Q_f^2,$$

tienen una solución P_i , $i = 1, 2$, en $[0, T]$; y

2. Que las dos ecuaciones diferenciales tipo Riccati,

$$\dot{K}_i = -A^T K_i - K_i A - Q^i + K_i S_i K_i, \quad K_i(T) = Q_f^i, \quad i = 1, 2,$$

tiene una solución simétrica $K(t)$ en $[0, T]$.

entonces el juego diferencial lineal cuadrático (2.31)-(2.33) tiene un único equilibrio de Nash para cualquier condición inicial del estado. Además, el conjunto de acciones del equilibrio esta dado por:

$$u_i^*(t) = -(R^{ii})^{-1}B_i P_i \phi(t, 0)x_0, \quad i = 1, 2.$$

Aquí $\phi(t, 0)$ satisface la ecuación de transición

$$\dot{\phi}(t, 0) = (A - S_1 P_1 - S_2 P_2)\phi(t, 0), \quad \phi(t, t) = I.$$

2.2.2. HORIZONTE DE TIEMPO INFINITO.

En esta sub-sección consideraremos el índice de desempeño de cada jugador $i = 1, 2$, que quieren minimizar es:

$$\lim_{T \rightarrow \infty} J_i(x_0, u_1, u_2, T), \quad (2.34)$$

donde

$$J_i = \int_0^{\infty} \{x^T(t)Q^i x(t) + \sum_{j=1}^2 u_j^T(t)R^{ij}u_j(t)\}dt,$$

sujeto la ecuación dinámica del estado

$$\dot{x}(t) = Ax(t) + B_1u_1(t) + B_2u_2(t), \quad x(0) = x_0. \quad (2.35)$$

Aquí las matrices R^{ii} , es definida positiva y las matrices Q^i , R^{ij} son simétricas, $i, j = 1, 2$. Se asume que el par de matrices (A, B_i) , $i = 1, 2$, son estables. Así, en principio, cada jugador es capaz de estabilizar el sistema por su cuenta. La información disponible para cada jugador al principio del juego es similar al caso de horizonte de tiempo finito. Cada jugador solo conoce el estado inicial del sistema. Las acciones de control admisible son ahora funciones de tiempo, donde el tiempo corre desde cero hasta infinito.

Teorema 2.8

Asumamos:

1. Que el conjunto de ecuaciones algebraicas acopladas tipo Riccati

$$0 = A^T P_1 + P_1 A + Q^1 - P_1 S_1 P_1 - P_1 S_2 P_2, \quad (2.36)$$

$$0 = A^T P_2 + P_2 A + Q^2 - P_2 S_2 P_2 - P_2 S_1 P_1, \quad (2.37)$$

tienen un conjunto de soluciones P_i , $i = 1, 2$, tal que $A - S_1 P_1 - S_2 P_2$ es estable; y

2. Que las dos ecuaciones algebraicas tipo Riccati,

$$0 = A^T K_i + K_i A - K_i S_i K_i + Q^i, \quad (2.38)$$

tienen una solución simétrica $K_i(t)$ tal que $A - S_i K_i$ es estable, $i = 1, 2$.

Entonces el juego diferencial lineal cuadrático (2.34) y (2.35) tiene un equilibrio de Nash para cualquier estado inicial. Por otra parte, el conjunto de acciones del equilibrio esta dada por:

$$u_i^*(t) = -(R^{ii})^{-1} B_i^T P_i \phi(t, 0) x_0, \quad i = 1, 2. \quad (2.39)$$

Aquí $\phi(t, 0)$ satisface la ecuación de transición

$$\dot{\phi}(t, 0) = (A - S_1 P_1 - S_2 P_2) \phi(t, 0), \quad \phi(t, t) = I.$$

2.3. MODOS DESLIZANTES.

En esta sección se propone un observador por modos deslizantes para una clase de sistemas no lineales.

2.3.1. OBSERVADOR POR MODOS DESLIZANTES DE SEGUNDO ORDEN.

La dinámica no lineal bajo estudio esta descrita por

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= f(t, x_1, x_2, u) + \xi(t, x, u), \\ y &= x_1,\end{aligned}\tag{2.40}$$

donde $x = [x_1, x_2]^T$ es el estado, y y es la salida del sistema. El sistema dinámico nominal es representado por la función $f(t, x_1, x_2, u)$, la incertidumbre interna es denotada por el termino $\xi(t, x, u)$. La solución para el sistema (2.40) esta dada en el sentido de Filippov. Asumamos que f y ξ son medibles según Lebesgue en cualquier region compacta de x . Aquí es solo considerado el caso escalar $x_1, x_2 \in \mathbb{R}$ porque el método diseñado puede ser extendido a un vector fácilmente. Mas aun, es fácil ver que el grado relativo de u para y es dos. Consideremos además las siguientes suposiciones:

Suposición 1. Existen dos constantes positivas k_1 y k_2 tal que

$$\begin{aligned}|f(t, x_1, x_2, u) - f(t, \hat{x}_1, \hat{x}_2, u)| &\leq k_1 |x_1 - \hat{x}_1|, \\ \left| \frac{df(t, x_1, x_2, u)}{dt} - \frac{df(t, \hat{x}_1, \hat{x}_2, u)}{dt} \right| &\leq k_2 |x_2 - \hat{x}_2|\end{aligned}$$

Suposición 2. La incertidumbre $\xi(t, x, u)$ satisface

$$\begin{aligned}|\xi(t, x, u)| &< \xi^+, \\ \left| \frac{d\xi(t, x, u)}{dt} \right| &< \delta\xi^+, \end{aligned}$$

donde ξ^+ y $\delta\xi^+$ son dos numeros positivos.

Basado en la dinámica del sistema (2.40), un observador por modos deslizantes de segundo orden es propuesto como sigue

$$\begin{aligned}\dot{\hat{x}}_1 &= \hat{x}_2 + z_1, & \hat{x}_1(0) &= x_1, \\ \dot{\hat{x}}_2 &= f(t, \hat{x}_1, \hat{x}_2, u) + z_2, & \hat{x}_2(0) &= 0, \\ \dot{\hat{y}} &= \hat{x}_1,\end{aligned}\tag{2.41}$$

donde \hat{x}_1 y \hat{x}_2 son las estimaciones del estado, z_1 y z_2 son las variables de corrección, que están definidas como

$$\begin{cases} z_1 = \lambda_1 |\tilde{x}_1|^{1/2} \text{sign}(\tilde{x}_1) + \nu_1, \\ \dot{\nu}_1 = \alpha_1 \text{sign}(\tilde{x}_1), \end{cases}\tag{2.42}$$

$$\begin{cases} z_2 = 0, & \text{si } \tilde{x}_1 \neq 0, \tilde{x}_1 \neq 0, \\ z_2 = \lambda_2 |z_1|^{1/2} \text{sign}(z_1) + \nu_2, & \text{si } \tilde{x}_1 = 0, \tilde{x}_1 = 0, \\ \dot{\nu}_2 = \alpha_2 \text{sign}(z_1), \end{cases} \quad (2.43)$$

donde $\tilde{x}_1 = x_1 - \hat{x}_1$, $\tilde{x}_2 = x_2 - \hat{x}_2$, y sign es la función signo discontinua.

Teorema 2.9 El primer par de variables (\hat{x}_1, \hat{x}_2) convergen a (x_1, x_2) en tiempo finito, si las condiciones

$$\begin{aligned} |f(t, x_1, x_2, u) - f(t, \hat{x}_1, \hat{x}_2, u) + \xi(t, x_1, x_2, u)| &< k_1 |x_1 - \hat{x}_1| + \xi^+, \\ \frac{|df(t, x_1, x_2, u) - df(t, \hat{x}_1, \hat{x}_2, u) + d\xi(t, x_1, x_2, u)|}{dt} &< k_2 |x_2 - \hat{x}_2| + \delta\xi^+, \end{aligned}$$

se satisfacen para el sistema (2.40), y los parámetros del observador (2.41) son seleccionados de acuerdo al siguiente criterio:

$$\begin{aligned} \alpha_1 &> k_1 \tilde{x}_{1M} + \xi^+ \text{ o } \alpha_1 > \sqrt{k_1} \tilde{x}_{1_0} + \xi^+, \\ \lambda_1 &> \frac{4\alpha_1}{\sqrt{\alpha_1 - \xi^+}}, \end{aligned}$$

2.3.2. DIFERENCIADOR POR MODOS DESLIZANTES DE ALTO ORDEN.

Sea $f(t) = f_0(t) + n(t)$ una función en $[0, \infty]$, donde $f_0(t)$ es una función desconocida con una constante L de Lipshitz de la n -ésima derivada, y $n(t)$ es una incertidumbre acotada medible con características desconocidas. El problema del diseño del diferenciador robusto por modos deslizantes de alto orden es buscar un estimador robusto en tiempo real de $\dot{f}_0(t)$, $\ddot{f}_0(t)$, \dots , $f_0^{(n)}(t)$ sea exacto cuando $n(t) = 0$. El diferenciador por modos deslizantes de alto orden toma la siguiente forma.

$$\begin{aligned} \dot{z}_0 &= \nu_0, \nu_0 = -\lambda_0 |z_0 - f(t)|^{n/(n+1)} \text{sign}(z_0 - f(t)) + z_1, \\ \dot{z}_1 &= \nu_1, \nu_1 = -\lambda_1 |z_1 - \nu_0|^{(n-1)/n} \text{sign}(z_1 - \nu_0) + z_2, \\ &\vdots \\ \dot{z}_{n-1} &= \nu_{n-1}, \nu_{n-1} = -\lambda_{n-1} |z_{n-1} - \nu_{n-2}|^{1/2} \text{sign}(z_{n-1} - \nu_{n-2}) + z_n, \\ \dot{z}_n &= -\lambda_n \text{sign}(z_n - \nu_{n-1}), \end{aligned}$$

donde $\lambda_0, \lambda_1, \dots, \lambda_n$ son parámetros positivos.

2.4. PROCESOS ESTOCÁSTICOS.

Un proceso estocástico $X(t, \omega)$ con $t \in T$ es una familia de variables aleatorias indexada por el conjunto de parámetros T . El parámetro t se refiere al tiempo. Si la variables aleatorias $X(t, \omega)$ tiene realizaciones discretas, diremos que el proceso estocástico tiene un espacio de estados discreto. Si tiene realizaciones continuas diremos que el proceso estocástico tiene un espacio de estados continuo. El conjunto de parámetro T también puede tener valores discretos o continuos. Un espacio de estados continuo, de un proceso con parámetros continuos sera llamado función aleatoria. Además, observe que un proceso estocástico $X(t, \omega)$ es de hecho una función de dos variables, el parámetro del tiempo t y el parámetro probabilístico ω . Para cada t , $X(t, *)$ es una variable aleatoria. Para cada ω , $X(*, \omega)$ es una realización del proceso, es decir, una función simple.

Proceso de Wiener.

Un proceso con parámetros continuos $\{x_t, t \in T\}$ tiene incrementos independientes si, para cualquier conjunto finito $\{t_i \mid t_i \leq t_{i+1}\} \in T$, las variables aleatorias

$$x_{t_2} - x_{t_1}, x_{t_3} - x_{t_2}, \dots, x_{t_n} - x_{t_{n-1}},$$

son independientes. El proceso x_t se dice que tiene incrementos independientes estacionarios si, además,

$$x_{t+h} - x_{\tau+h},$$

tiene la misma distribución como $x_t - x_\tau$ para todo $t > \tau \in T$ y cualquier $h > 0$. Un proceso estocástico que es de gran importancia en la teoría y aplicaciones, y que juega un rol importante en esta tesis, es el proceso de Wiener.

Un proceso con parámetros continuos $x_t, t \geq 0$ es un proceso de Wiener si

1. $x_t, t \geq 0$ tiene incrementos independientes estacionarios
2. para todo $t \geq 0$, x_t tiene una distribución normal
3. para todo $t \geq 0$, $E(x_t) = 0$
4. $P(x_0 = 0) = 1$

La función de covarianza de un proceso de Wiener esta dada por

$$E[(x_t - E(x_t))(x_\tau - E(x_\tau))^T] = Q \text{mín}(t, \tau),$$

donde Q es una matriz definida positiva.

Ruido Blanco.

Un ruido blanco Gaussiano es un proceso Gaussiano con

$$E[(x_t - E(x_t))(x_\tau - E(x_\tau))^T] = Q\delta(t - \tau),$$

también un ruido blanco Gaussiano puede ser considerado como la derivada en promedio cuadrático débil de un proceso de Wiener, definido como sigue. Sea $\{x_t, t \geq 0\}$ un proceso de Wiener, del cual sabemos que su función de covarianza es

$$P_x(t, \tau) = Q \text{mín}(t, \tau),$$

la función de covarianza del proceso $\{\frac{dx_t}{dt}, t \geq 0\}$ es

$$P_{\dot{x}}(t, \tau) = \frac{\partial^2 P_x(t, \tau)}{\partial t \partial \tau},$$

por lo tanto

$$P_{\dot{x}}(t, \tau) = Q \frac{\partial^2 \text{mín}(t, \tau)}{\partial t \partial \tau},$$

como sabemos

$$\min(t, \tau) = \begin{cases} \tau, & \tau < t, \\ t, & \tau > t, \end{cases}$$

entonces la primera derivada parcial de esta función es

$$\frac{\partial}{\partial t} \min(t, \tau) = \begin{cases} 0, & \tau < t, \\ 1, & \tau > t, \end{cases}$$

que es la función Heaviside o función de salto (en τ) y su derivada con respecto a τ es la delta de Dirac $\delta(t - \tau)$.

Obteniendo

$$P_{\dot{x}}(t, \tau) = Q\delta(t - \tau),$$

y \dot{x}_t es un proceso con una delta-covarianza. Entonces, formalmente, un ruido blanco Gaussiano es la derivada en promedio cuadrático débil del proceso de Wiener.

2.4.1. ECUACIONES DIFERENCIALES ESTOCÁSTICAS.

Sistemas dinámicos continuos con estado finito-dimensionales, que son sujetos a disturbios aleatorios, pueden ser representados por ecuaciones diferenciales ordinarias no lineales. Sean x_t y w_t , respectivamente, el vector de estado n -dimensional y el disturbio aleatorio m -dimensional en el tiempo t . Entonces una ecuación diferencial general del tipo descrito puede ser escrita como

$$\frac{dx_t}{dt} = f(x_t, w_t, t), \quad t \geq t_0, \quad (2,36)$$

donde f es una función no lineal real n -dimensional. La ecuación (2.36) es llamada ecuación diferencial estocástica, sistema dinámico estocástico continuo. La función del disturbio aleatorio w_t es llamado función driving. La condición inicial para (2.36) puede ser una constante o una variable aleatoria x_{t_0} con una distribución específica. La ley de probabilidad del proceso $\{w_t, t \geq t_0\}$ se asume especificada. Un caso especial importante de (2.36) es la ecuación diferencial estocástica con un aditivo ruido blanco Gaussiano

$$\frac{dx_t}{dt} = f(x_t, t) + G(x_t, t)\psi_t, \quad t \geq t_0, \quad (2,37)$$

donde G es una función matricial $n \times m$, y la condición inicial x_{t_0} que es independiente del ruido blanco Gaussiano $\{\psi_t, t \geq t_0\}$. Como el proceso $\{\psi_t\}$ es delta-correlacionado, y por lo tanto, el ψ_t no es integrable en promedio cuadrático, y dado el resultado de la sección anterior que un ruido blanco Gaussiano es la derivada en promedio cuadrático débil de un proceso de Wiener. Sea $\{W_t, t \geq t_0\}$ un proceso de Wiener independiente. Entonces

$$\psi_t \sim \frac{dW_t}{dt},$$

y podemos escribir una forma equivalente de (2.37) como

$$dx_t = f(x_t, t)dt + G(x_t, t)dW_t, \quad t \geq t_0. \quad (2,38)$$

Ahora (2.38) solo tiene sentido cuando la integral existe

$$x_t - x_{t_0} = \int_{t_0}^t f(x_\tau, \tau) d\tau + \int_{t_0}^t G(x_\tau, \tau) dW_\tau,$$

donde la primera integral puede ser definida como una integral en promedio cuadrático o como una integral ordinaria, la segunda integral sera definida como una integral estocástica de Itô, y por lo tanto (2.38) es una ecuación diferencial estocástica de Itô.

Calculo Estocástico de Itô.

Lema 2.10 Sea x_t es la solución única de la ecuación diferencial estocástica

$$dx_t = f(x_t, t)dt + G(x_t, t)dW_t, \quad t \geq t_0, \quad (2.44)$$

donde x y f son vectores n -dimensionales, G es $n \times m$, y $\{W_t, t \geq t_0\}$ es un proceso de Wiener m -dimensional con $E(dW_t dW_t^T) = Qdt$. Sea $\phi(x_t, t)$ una función real escalar, continuamente diferenciable en t y tenga segunda derivada parcial mixta continua con respecto a los elementos de x . Entonces la diferencial estocástica $d\phi$ de ϕ es

$$d\phi = \phi_t dt + \phi_x^T dx_t + \frac{1}{2} tr(GQG^T) \phi_{xx} dt, \quad (2.45)$$

donde

$$\phi_t = \frac{\partial \phi}{\partial t}; \quad \phi_x^T = \left[\frac{\partial \phi}{\partial x_1}, \dots, \frac{\partial \phi}{\partial x_n} \right],$$

$$\phi_{xx} = \begin{bmatrix} \frac{\partial^2 \phi}{\partial x_1^2} & \frac{\partial^2 \phi}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 \phi}{\partial x_1 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 \phi}{\partial x_n \partial x_1} & \frac{\partial^2 \phi}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 \phi}{\partial x_n^2} \end{bmatrix},$$

Observe que (2.45) puede ser escrito como

$$d\phi = (\phi_t + \phi_x^T f + \frac{1}{2} tr(GQG^T) \phi_{xx}) dt + \phi_x^T G dW_t. \quad (2.46)$$

2.4.2. TEORÍA DE FILTRADO.

Consideremos la ecuación dinámica estocástica continua descrita por la ecuación diferencial estocástica.

$$dx_t = f(x_t, t)dt + G(x_t, t)dW_t, \quad t \geq t_0, \quad (2.47)$$

donde x y f son vectores n -dimensionales, G es $n \times m$, y $\{W_t, t \geq t_0\}$ es un proceso de Wiener m -dimensional con $E(dW_t dW_t^T) = Qdt$. La ecuación (2.47) es equivalente a

$$\frac{dx_t}{dt} = f(x_t, t) + G(x_t, t)\psi_t,$$

donde $\{\psi_t, t \geq t_0\}$ es un ruido blanco Gaussiano, $\psi_t \sim N(0, Q(t))$. Supongamos observaciones continuas son tomadas en el sistema (2.47), de la forma

$$dY_t = h(x_t, t)dt + dW_t', \quad (2.48)$$

donde Y_t y h son vectores m -dimensionales y $\{W'_t, t \geq t_0\}$ es un proceso de Wiener m -dimensional con $E(dW'_t dW'^T_t) = R(t)dt$, $R(t) > 0$. Supongamos que $\{W_t\}$, $\{W'_t\}$ y x_{t_0} son independientes. La ecuación (2.48) es equivalente a

$$y_t = h(x_t, t) + \psi'_t, \quad (2.49)$$

con la identificación

$$y_t \sim \frac{dY_t}{dt}, \quad \psi'_t \sim \frac{dW'_t}{dt},$$

y con $\{\psi'_t, t \geq t_0\}$ un ruido blanco Gaussiano, $\psi'_t \sim N(0, R(t))$. Las ecuaciones (2.47) y (2.48), constituyen un vector de ecuaciones diferenciales estocásticas. El problema de estimación es buscar el estimado óptimo $\hat{x}(t)$ del sistema de estado $x(t)$ basado en el proceso de observación $Y(t) = \{y(s), 0 \leq s \leq t\}$, que minimiza la norma euclidiana

$$J = E[(x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t)) | F_t^Y], \quad (2.50)$$

para cualquier momento del tiempo t .

Las ecuaciones de filtrado óptimo se obtienen usando la formula para la ecuación diferencial de itô de la esperanza condicional $m(t) = E(x(t) | F_t^Y)$

$$dm(t) = E(f(x, t) | F_t^Y)dt + E(x[h(x_t, t) - E(h(x_t, t) | F_t^Y)]^T | F_t^Y) \times (dy(t) - E(h(x_t, t) | F_t^Y))dt. \quad (2.51)$$

Filtro de Kalman-Bucy.

En el caso particular, si la función $f(x_t, t) = a_0(t) + a_1(t)x(t)$, $G(x_t, t) = B(t)$ y $h(x_t, t) = A_0(t) + A(t)x(t)$ la ecuación de filtrado óptimo de la esperanza condicional es

$$dm(t) = E(a_0(t) + a_1(t)x(t) | F_t^Y)dt + E(x[A_0(t) + A(t)x(t) - E(A_0(t) + A(t)x(t) | F_t^Y)]^T | F_t^Y) \times (B(t)B(t)^T)^{-1}(dy(t) - E(A_0(t) + A(t)x(t) | F_t^Y))dt, \quad (2.52)$$

de aquí obtenemos que

$$\begin{aligned} E(a_0(t) + a_1(t)x(t) | F_t^Y) &= a_0(t) + a_1(t)m(t) \\ E(x[A_0(t) + A(t)x(t) - E(A_0(t) + A(t)x(t) | F_t^Y)]^T | F_t^Y) & \\ &= E(x[x(t) - E(x(t) | F_t^Y)]^T | F_t^Y)A(t)^T = P(t)A(t)^T \\ E(A_0(t) + A(t)x(t) | F_t^Y) &= A_0(t) + A(t)m(t), \end{aligned}$$

por lo tanto la ecuación diferencial de la esperanza condicional es

$$dm(t) = (a_0(t) + a_1(t)m(t))dt + P(t)A(t)^T \times \quad (2.53)$$

$$(B(t)B(t)^T)^{-1}(dy(t) - A_0(t) - A(t)m(t))dt.$$

La ecuación (2.53) deberá tener una condición inicial $m(t_0) = E(x_0 | F_{t_0}^Y)$. Tratando de componer una forma cerrada para el sistema de ecuaciones de filtrado, la ecuación (2.53) deberá ser complementada con la ecuación para la varianza del error $P(t)$. Para este propósito usamos la formula diferencial de Itô de la varianza $P(t) = E((x(t) - m(t))(x(t) - m(t))^T | F_t^Y)$ dada por

$$\begin{aligned} dP(t) &= (E((x(t) - m(t))(f(x, t))^T | F_t^Y) + E((f(x, t))(x(t) - m(t))^T | F_t^Y) \\ &\quad b(t)b^T(t) - E(x[h(x_t, t) - E(h(x_t, t) | F_t^Y)]^T | F_t^Y)A(t)^T \times \\ &\quad (B(t)B(t)^T)^{-1}A(t)^T E([h(x_t, t) - E(h(x_t, t) | F_t^Y)]x^T | F_t^Y))dt + \\ &\quad E((x(t) - m(t))(x(t) - m(t))[h(x_t, t) - E(h(x_t, t) | F_t^Y)]^T | F_t^Y) \times \\ &\quad (B(t)B(t)^T)^{-1}(dy(t) - E(h(x_t, t) | F_t^Y))dt, \end{aligned} \quad (2.54)$$

sustituyendo la expresión para $h(x_t, t)$ la fórmula anterior toma la forma

$$\begin{aligned} dP(t) &= (E((x(t) - m(t))(f(x, t))^T | F_t^Y) + E((f(x, t))(x(t) - m(t))^T | F_t^Y) \\ &\quad b(t)b^T(t) - P(t)A(t)^T(B(t)B(t)^T)^{-1}A(t)^T P(t))dt + \\ &\quad E((x(t) - m(t))(x(t) - m(t))(x(t) - m(t))^T | F_t^Y) \times \\ &\quad A^T(t)(B(t)B(t)^T)^{-1}(dy(t) - A_0(t) - A(t)m(t))dt, \end{aligned} \quad (2.55)$$

como $f(x_t, t) = a_0(t) + a_1(t)x(t)$ obtenemos

$$\begin{aligned} E((x(t) - m(t))(a_0(t) + a_1(t)x(t))^T | F_t^Y) &= P(t)a_1(t)^T \\ E((a_0(t) + a_1(t)x(t))(x(t) - m(t))^T | F_t^Y) &= a_1(t)P(t), \end{aligned}$$

además

$$E((x(t) - m(t))(x(t) - m(t))(x(t) - m(t))^T | F_t^Y) = 0,$$

por lo tanto la ecuación para la varianza del error sera

$$\begin{aligned} dP(t) &= (P(t)a_1(t)^T + a_1(t)P(t) \\ &\quad b(t)b^T(t) - P(t)A(t)^T(B(t)B(t)^T)^{-1}A(t)^T P(t))dt, \end{aligned} \quad (2.56)$$

con la condición inicial $P(t_0) = E((x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y)$. Las ecuaciones (2.53) y (2.56) forman el filtro de Kalman-Bucy.

3

CONTROL ÓPTIMO PARA UN SISTEMA POLINOMIAL CON UN CRITERIO CUADRÁTICO SOBRE UN HORIZONTE DE TIEMPO INFINITO.

3.1. PLANTEAMIENTO DEL PROBLEMA.

El estado del sistema $z(t)$ esta dado por la ecuación diferencial no lineal

$$\frac{dz(t)}{dt} = h(z, t) + G(t)v(t), \quad z(t_0) = z_0, \quad (3,1)$$

Aquí, $h(z, t)$ es una función polinomial no lineal, $v(t) \in R^m$ es la entrada de control, $z(t) \in R^n$ es el vector de estado. El coeficiente $G(t)$ es una función matricial de dimension $n \times m$. La función vectorial $h(z, t) \in R^n$ es un polinomio de n componentes del vector de estado z , con coeficientes dependientes del tiempo (ver [50] para definición). Se sigue de [50], un polinomio de grado r del vector de estado esta definido como una forma r -lineal de sus componentes

$$h(z, t) = \beta_1(t)z + \beta_2(t)zz^T + \dots + \beta_r(t)z \dots r \text{ times} \dots z,$$

donde $\beta_i, i = 1, \dots, r$, es un tensor de dimension $n \times \dots (r+1) \text{ times} \dots \times n$, y $z \times \dots r \text{ times} \dots \times z$ es un tensor de dimension $n \times \dots r \text{ times} \dots \times n$. Cada componente del polinomio puede ser representado como una sumatoria

$$h_l(z, t) = \sum_j \beta_{1 \ l_j}(t)z_j(t) + \sum_{jk} \beta_{2 \ l_jk}(t)z_j(t)z_k(t) + \dots$$

$$+ \sum_{i_1 \dots i_r} \beta_{r \ k i_1 \dots i_r}(t) z_{i_1}(t) \dots z_{i_r}(t), \quad l, j, k, i_1 \dots i_r = 1, \dots, n.$$

El criterio J esta dado por el siguiente funcional cuadrático de z y v

$$J(u) = \frac{1}{2} \int_{t_0}^{\infty} v^T(t) R(t) v(t) dt + \frac{1}{2} \int_{t_0}^{\infty} z^T(t) L(t) z(t) dt, \quad (3,2)$$

donde R y L son matrices simétricas, con $R > 0$ y $L \geq 0$, y a^T denota la transpuesta de un vector (matriz) a .

El problema del controlador óptimo es minimizar el criterio J a lo largo de la trayectoria $z^*(t)$, $t \in [t_0, \infty]$ a través de la selección del control $v^*(t)$, $t \in [t_0, \infty]$, que es generado al sustituir $v^*(t)$ dentro de la ecuación de estado (3.1).

3.2. DISEÑO DEL CONTROL I. PRINCIPIO DEL MÁXIMO.

3.2.1. SOLUCIÓN DEL PROBLEMA DEL CONTROLADOR ÓPTIMO I.

El siguiente teorema da la solución para el problema del controlador óptimo planteado anteriormente.

Teorema 3.1 La ley de control

$$v^*(t) = -R^{-1}(t)G^T(t)m(z) \quad (3,3)$$

presenta el regulador óptimo para el sistema polinomial (3.1) con respecto al criterio cuadrático (3.2), donde la función matricial $m(z)$ es la solución de la ecuación

$$\begin{aligned} \frac{dm(z)}{dz}(\beta_1(t)z + \beta_2(t)zz^T + \dots + \beta_r(t)z \dots r \text{ times} \dots z - G(t)R^{-1}(t)G^T(t)m(z)) = \\ = -Lz - (\beta_1(t) + 2\beta_2(t)z + \dots + r\beta_r(t)z \dots r-1 \text{ times} \dots z)^T m(z) \end{aligned} \quad (3,4)$$

con la condición inicial $m(0) = 0$.

Sustituyendo el control óptimo (3.3) en el sistema polinomial (3.1) produce el estado controlado óptimamente gobernado por la ecuación

$$\frac{dz(t)}{dt} = h(z, t) - G(t)R^{-1}(t)G^T(t)m(z), \quad z(t_0) = z_0. \quad (3,5)$$

Demostración Referirse al Apéndice.

3.2.2. EJEMPLO I.

Considere una ecuación de estado escalar cuadrática

$$\dot{z}(t) = z^2(t) + v(t), \quad z(0) = z_0. \quad (3,6)$$

El problema de control es minimizar

$$J = \frac{1}{2} \int_0^{\infty} v^2(t) dt + \frac{1}{2} \int_0^{\infty} z^2(t) dt. \quad (3,7)$$

a través de la selección del control $v(t)$, $t \in [0, \infty]$, i.e., usando la mínima energía total del control v tal que minimice la energía total del estado z . El ejemplo considerado corresponde a la ecuación de velocidad de un oscilador mecánico, donde la fuerza que actúa depende del cuadrado de velocidad.

La ley de control usada en este ejemplo es calculada como (3.3) donde la matriz $m(z)$ es la solución de la ecuación diferencial (3.4), esto es $v^*(t) = -R^{-1}(t)G^T(t)m(z)$. De las ecuaciones (3.6) y (3.7), se puede ver que $G(t) = 1$ y $R(t) = 1$, por lo tanto, el control toma la forma

$$v(t) = -m(z); \quad (3,8)$$

donde $m(z)$ satisface la ecuación

$$\frac{dm(z)}{dz}(z^2 - m(z)) = -z - 2zm(z) \quad (3,9)$$

con la condición inicial $q(0) = 0$. La solución de la ecuación diferencial $m(z)$ esta dado por

$$m(z) = z^2 + z\sqrt{z^2 + 1}, \quad (3,10)$$

entonces, la forma del control es

$$v(t) = -z^2 - z\sqrt{z^2 + 1}, \quad (3,11)$$

y la ecuación de estado (3.6) toma la forma

$$\dot{z}(t) = z^2(t) - z^2(t) - z(t)\sqrt{z^2(t) + 1}, \quad z(0) = z_0. \quad (3,12)$$

Considere la condición inicial $z(0) = 1$ para el sistema (3.6), controlado por (3.10)–(3.11), para llevar a cabo simulaciones numéricas.

La fig. 1 muestra el resultado de aplicar el regulador (3.10)–(3.11), que muestra la gráfica del estado (3.6) $z(t)$, controlado por (3.10), la función del control (3.10) $v(t)$, y la gráfica del criterio (3.7) $J(t)$ en el intervalo $[0, 20]$. En el momento final $T = 20$, el criterio (3.7) toma el valor $J(20) = 0,9428$.

EL regulador diseñado (3.10)–(3.11) es comparado con el regulador lineal para el sistema linealizado

$$\dot{z}(t) = 2z(t) + v(t), \quad z(0) = z_0. \quad (3,13)$$

La ley de control lineal esta dada por

$$v(t) = -M(t)z(t), \quad (3,14)$$

donde $M(t)$ satisface la ecuación algebraica de Riccati para el sistema linealizado

$$0 = 1 + 4M(t) - M(t)^2. \quad (3,15)$$

La solución de la ecuación algebraica de Riccati esta dada por

$$M = 2 + \sqrt{5}, \quad (3,16)$$

entonces, el control toma la forma

$$v(t) = -(2 + \sqrt{5})z, \quad (3,17)$$

y la ecuación de estado (3.6) toma la forma

$$\dot{z}(t) = z^2(t) - (2 + \sqrt{5})z(t), \quad z(0) = z_0. \quad (3,18)$$

Considerando de nuevo la condición inicial $z(0) = 1$ para el sistema (3.6), controlado por (3.14)–(3.17), para llevar a cabo simulaciones numéricas.

La fig. 1 muestra el resultado de aplicar el regulador (3.14)–(3.17), que muestra la gráfica del estado (3.6) $z(t)$, controlado por (3.17), la función de control (3.17) $v(t)$, y la gráfica del criterio (3.7) $J(t)$ en el intervalo $[0, 20]$. En el momento final $T = 20$, el criterio (3.7) toma el valor $J(20) = 1,3266$.

3.3. DISEÑO DEL CONTROL II. ECUACIÓN HAMILTON-JACOBI-BELLMAN.

3.3.1. SOLUCIÓN DEL PROBLEMA DEL CONTROLADOR ÓPTIMO II.

El siguiente teorema da otra solución para el problema del controlador óptimo planteado en la sección 3.1

Teorema 3.2. ley de control

$$v^*(t) = R^{-1}(t)G^T(t)M(t)z(t), \quad (3,19)$$

presenta el regulador óptimo para el sistema polinomial (3.1) con respecto al criterio cuadrático (3.2), donde la función matricial $M(t)$ es la solución de la ecuación algebraica tipo Riccati

$$0 = L(t) - [\beta_1(t) + \beta_2(t)z(t) + \beta_3(t)z(t)z^T(t) + \dots + \beta_r(t)z(t) \dots_{r-1 \text{ times}} \dots z(t)]^T M(t) - \quad (3,20)$$

$$M(t)[\beta_1(t) + \beta_2(t)z(t) + \beta_3(t)z(t)z^T(t) + \dots + \beta_r(t)z(t) \dots_{r-1 \text{ times}} \dots z(t)] - M(t)G(t)R^{-1}(t)G^T(t)M(t).$$

Sustituyendo el control óptimo (3.3) en el sistema polinomial (3.1) produce el estado controlado óptimamente gobernado por la ecuación

$$\frac{dz(t)}{dt} = h(z, t) + G(t)R^{-1}(t)G^T(t)M(t)z(t), \quad z(t_0) = z_0. \quad (3,21)$$

Demostración Referirse al Apéndice.

Observación 3.1. La ecuación algebraica (3.20) para $M(t)$ debería resolverse considerando al estado $z(t)$ como un argumento fijo. Observe que la solución analítica para (3.20) no puede ser obtenida siempre para un sistema polinomial arbitrario, especialmente de orden mayor que uno; por lo tanto, algunos métodos numéricos disponibles para resolver la ecuación matricial algebraica tipo Riccati deberían aplicarse (ver, por ejemplo, [51]). Sin embargo, esta solución puede ser obtenida analíticamente para algunos sistemas de orden uno y dos (ver las formulas (3.26), (3.39)–(3.42), y (3.53)–(3.56) en los ejemplos).

Observación 3.2. La estabilidad del sistema es derivado de la función de Lyapunov

$$V(t) = -\frac{1}{2}z(t)^T M(t)z(t).$$

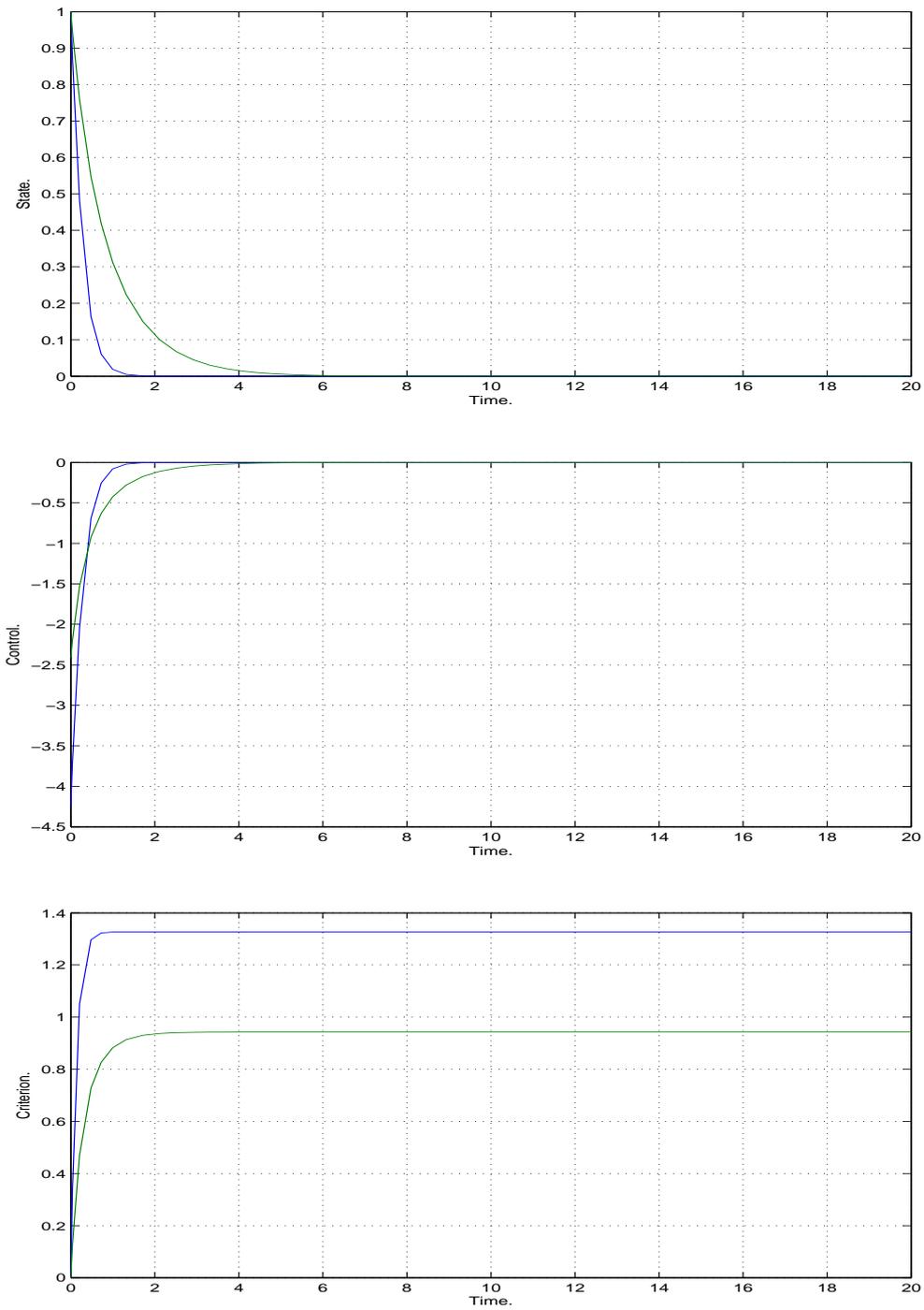


Figura 3.1: Regulador óptimo con respecto al criterio $J(t)$ vs. el regulador lineal para el sistema linealizado en $[0, 20]$. **1. Estado.** Gráfica del estado (3.6) $z(t)$ (Verde) controlado por (3.11), y el estado (3.6) $z(t)$ (Azul) controlado por (3.17), **2. Control.** la función de control (3.11) $v(t)$ (Verde) y el control lineal (3.17) $v(t)$ (Azul), y **3. Criterio.** el criterio (3.7) $J(t)$ producido por el control óptimo (Verde) y el control lineal (Azul).

Esto produce la derivada de $V(t)$ dado por

$$\dot{V}(t) = -\frac{1}{2}\dot{z}(t)^T M(t)z(t) - \frac{1}{2}z(t)^T M(t)\dot{z}(t),$$

teniendo en cuenta que $\dot{M}(t)$ esta fijo en cero (ver demostración del teorema 3.2). Sustituyendo (3.21) en la ecuación anterior conduce a

$$\begin{aligned}\dot{V}(t) &= -\frac{1}{2}(h(z,t)+G(t)R^{-1}(t)G^T(t)M(t)z(t))^T M(t)z(t) - \frac{1}{2}z(t)^T M(t)(h(z,t)+G(t)R^{-1}(t)G^T(t)M(t)z(t)) \\ &= -\frac{1}{2}z(t)^T M(t)h(z,t) - \frac{1}{2}h(z,t)^T M(t)z(t) - z(t)^T M(t)G(t)R^{-1}(t)G^T(t)M(t)z(t).\end{aligned}$$

Finalmente, sustituyendo el valor de $h(t, z)$ obtenemos

$$\begin{aligned}\dot{V}(t) &= -z(t)^T \{M(t)G(t)R^{-1}(t)G^T(t)M(t) \\ &+ \frac{1}{2}M(t)(\beta_1(t)+\beta_2(t)z+\dots+\beta_r(t)z \dots_{r-1 \text{ times}} \dots z) + \frac{1}{2}(\beta_1(t)+\beta_2(t)z+\dots+\beta_r(t)z \dots_{r-1 \text{ times}} \dots z)^T M(t)\} z(t), \\ &= -\frac{1}{2}z(t)^T \{M(t)G(t)R^{-1}(t)G^T(t)M(t) + L(t)\} z(t) < 0,\end{aligned}$$

que provee estabilidad asintótica del sistema controlado óptimamente (3.1).

3.3.2. EJEMPLO II.

Considere la ecuación del estado escalar cuadrática

$$\dot{z}(t) = z^2(t) + v(t), \quad z(0) = z_0. \quad (3,22)$$

El problema del controlador es minimizar

$$J = \frac{1}{2} \int_0^\infty v^2(t)dt + \frac{1}{2} \int_0^\infty z^2(t)dt. \quad (3,23)$$

a través de la selección del control $v(t)$, $t \in [0, \infty]$. Dado que las ecuaciones (3.22),(3.23) en este ejemplo coinciden con las ecuaciones (3.6),(3.7) del Ejemplo 1, ambos ejemplos tienen la misma interpretación mecánica.

La ley de control usado en este ejemplo es calculado como (3.19), donde la matriz $M(t)$ es la solución de la ecuación algebraica tipo Riccati (3.20), esto es, $v^*(t) = -R^{-1}(t)G^T(t)M(t)z(t)$. De las ecuaciones (3.22) y (3.23), es claro que $G(t) = 1$ y $R(t) = 1$, por lo tanto, la ley de control toma la forma

$$v(t) = -M(t)z(t), \quad (3,24)$$

donde $M(t)$ satisfaciendo la ecuación algebraica tipo Riccati

$$0 = 1 - z(t)^T M(t) - M(t)z(t) - M(t)M(t). \quad (3,25)$$

La solución de la ecuación algebraica tipo Riccati esta dada por

$$M(t) = z + \sqrt{z^2 + 1}, \quad (3,26)$$

entonces, la ley de control es

$$v(t) = -(z + \sqrt{z^2 + 1})z, \quad (3,27)$$

y la ecuación de estado (3.22) toma la forma

$$\dot{z}(t) = z^2(t) - z(t)\sqrt{z^2(t) + 1}, \quad z(0) = z_0. \quad (3,28)$$

Considerando la condición inicial $z(0) = 1$ para el sistema (3.22), controlado por (3.26)–(3.27), para llevar a cabo simulaciones numéricas.

La fig. 2 muestra el resultado de aplicar el regulador (3.26)–(3.27), que muestra la gráfica del estado (3.22) $z(t)$, controlado por (3.27), la función de control (3.27) $v(t)$, y la gráfica del criterio (3.23) $J(t)$ en el intervalo $[0, 20]$. En el momento final $T = 20$, el criterio (3.23) toma el valor $J(20) = 0,9428$.

Como en el Ejemplo 1, este ejemplo es comparado con el controlador lineal para el sistema linealizado dado por

$$v(t) = -(2 + \sqrt{5})z(t), \quad (3,29)$$

y la ecuación del estado (3.16) toma la forma

$$\dot{z}(t) = z^2(t) - (2 + \sqrt{5})z(t), \quad z(0) = z_0. \quad (3,30)$$

Se puede observar que el control diseñado por el algoritmo propuesto tiene un mejor rendimiento en comparación al del sistema linealizado. Note que la comparación es realizada por propósitos ilustrativos, dado que el regulador diseñado debería dar teóricamente un mejor resultado, como se sigue del teorema 3.1 y 3.2.

Observación 3.3. Como se mostró, el control obtenido por la ecuación algebraica tipo Riccati coincide con el control obtenido de la ecuación de co-estado.

3.3.3. EJEMPLO III.

Considere las ecuaciones de estado cuadráticas bidimensional

$$\dot{z}_1(t) = z_2(t), \quad z_1(0) = z_{10}, \quad (3,31)$$

$$\dot{z}_2(t) = z_1^2(t) + v(t), \quad z_2(0) = z_{20}. \quad (3,32)$$

El problema de control es minimizar

$$J = \frac{1}{2} \int_0^\infty v^2(t)dt + \frac{1}{2} \int_0^\infty (z_1^2(t) + z_2^2(t))dt. \quad (3,33)$$

a través de la selección del control $v(t)$, $t \in [0, \infty]$, i.e., usando la mínima energía total del control v tal que se minimice energía total del estado z . El ejemplo considerado corresponde a un oscilador mecánico incluyendo las ecuaciones de posición y velocidad, donde la fuerza que actúa depende del cuadrado de la velocidad.

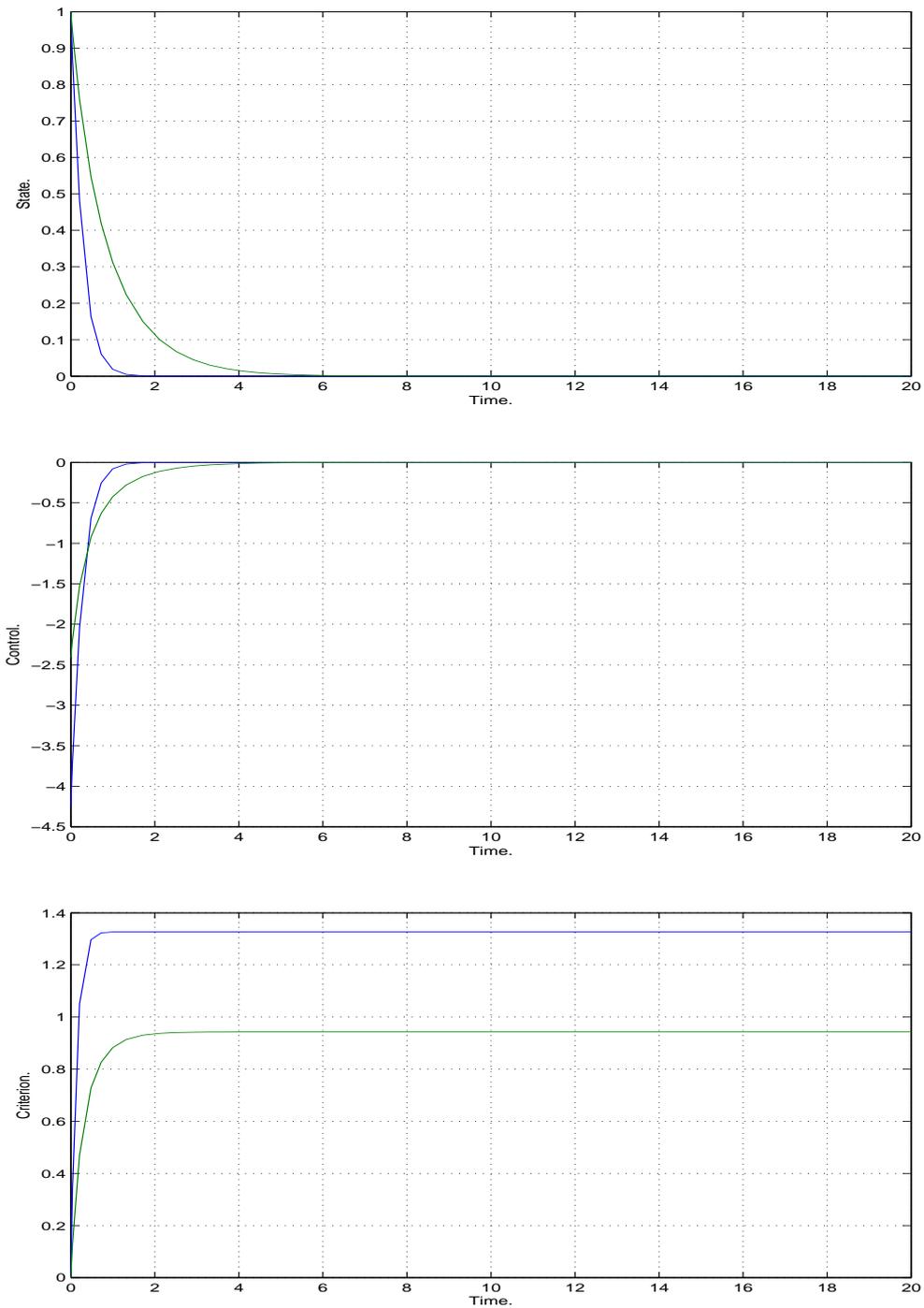


Figura 3.2: Regulador óptimo con respecto al criterio $J(t)$ vs. el regulador lineal para el sistema linealizado en $[0, 20]$. **1. Estado.** Gráfica del estado (3.16) $z(t)$ (Verde) controlado por (3.21), y el estado (3.24) $z(t)$ (Azul) controlado por (3.23), **2. Control.** la función de control (3.21) $v(t)$ (Verde) y el control lineal (3.23) $v(t)$ (Azul), y **3. Criterio.** el criterio (3.17) $J(t)$ producido por el control óptimo (Verde) y el control lineal (Azul).

La ley de control usada en este ejemplo es calculada como (3.19), donde la matriz $M(t)$ es la solución de la ecuación algebraica tipo Riccati (3.20), esto es, $v^*(t) = -R^{-1}(t)G^T(t)M(t)z(t)$. De las ecuaciones (3.31) y (3.32), se puede ver que $G(t) = [0 \ 1]^T$ y $R(t) = 1$, por lo tanto, el control toma la forma

$$v(t) = -M_{21}(t)z_1(t) - M_{22}(t)z_2(t), \quad (3,34)$$

donde $M(t)$ satisface la ecuación algebraica tipo Riccati

$$0 = 1 - z_1(t)^T M_{12}(t) - M_{21}(t)z_1(t) - M_{12}(t)M_{21}(t), \quad (3,35)$$

$$0 = -z_1(t)^T M_{22}(t) - M_{11}(t) - M_{12}(t)M_{22}(t), \quad (3,36)$$

$$0 = -M_{11}(t) - M_{22}(t)z_1(t) - M_{22}(t)M_{21}(t), \quad (3,37)$$

$$0 = 1 - M_{21}(t) - M_{12}(t) - M_{22}(t)M_{22}(t). \quad (3,38)$$

La solución a las ecuaciones algebraicas tipo Riccati esta dada por

$$M_{11}(t) = \sqrt{2z_1(t) + 1 + 2\sqrt{z_1^2(t) + 1}\sqrt{z_1^2(t) + 1}}, \quad (3,39)$$

$$M_{12}(t) = z_1(t) + \sqrt{z_1^2(t) + 1}, \quad (3,40)$$

$$M_{21}(t) = z_1(t) + \sqrt{z_1^2(t) + 1}, \quad (3,41)$$

$$M_{22}(t) = \sqrt{2z_1(t) + 1 + 2\sqrt{z_1^2(t) + 1}}, \quad (3,42)$$

la forma del control es

$$\begin{aligned} v(t) = & -(z_1(t) + \sqrt{z_1^2(t) + 1})z_1(t) \\ & - (\sqrt{2z_1(t) + 1 + 2\sqrt{z_1^2(t) + 1}})z_2(t), \end{aligned} \quad (3,43)$$

y las ecuaciones de estado (3.31) y (3.32) toman la forma

$$\dot{z}_1(t) = z_2(t), \quad z_1(0) = z_{10}, \quad (3,44)$$

$$\begin{aligned} \dot{z}_2(t) = & z_1^2(t) - (z_1(t) + \sqrt{z_1^2(t) + 1})z_1(t) \\ & - (\sqrt{2z_1(t) + 1 + 2\sqrt{z_1^2(t) + 1}})z_2(t), \quad z_2(0) = z_{20}. \end{aligned} \quad (3,45)$$

Considere las condiciones iniciales $z_1(0) = 1$ y $z_2(0) = 1$ para el sistema (3.31)–(3.32), controlado por (3.26)–(3.27), para llevar a cabo simulaciones numéricas.

La fig. 3 muestra los resultados de aplicar el regulador (3.39)–(3.43) al sistema (3.31),(3.32), que muestra las gráficas del estado (3.31) $z_1(t)$ y (3.32) $z_2(t)$, controlado por (3.43), la función de control (3.43) $v(t)$, y la gráfica del criterio (3.33) $J(t)$ en el intervalo $[0, 20]$. En el momento final $T = 20$, el criterio (3.33) toma la valor $J(20) = 5,5459$.

El regulador diseñado (3.39)–(3.43) es comparado con el regulador lineal del sistema linealizado

$$\dot{z}_1(t) = z_2(t), \quad z_1(0) = z_{10}, \quad (3,46)$$

$$\dot{z}_2(t) = 2z_1(t) + v(t), \quad z_2(0) = z_{20}. \quad (3,47)$$

La ley de control esta dada por

$$v(t) = -M(t)z(t), \quad (3,48)$$

donde $M(t)$ satisface las ecuaciones algebraicas de Riccati para el sistema linealizado

$$0 = 1 - 2M_{12}(t) - 2M_{21}(t) - M_{12}(t)M_{21}(t), \quad (3,49)$$

$$0 = -2M_{22}(t) - M_{11}(t) - M_{12}(t)M_{22}(t), \quad (3,50)$$

$$0 = -M_{11}(t) - 2M_{22}(t) - M_{22}(t)M_{21}(t), \quad (3,51)$$

$$0 = 1 - M_{21}(t) - M_{12}(t) - M_{22}(t)M_{22}(t). \quad (3,52)$$

La solución de las ecuaciones algebraicas de Riccati están dadas por

$$M_{11}(t) = \sqrt{5 + \sqrt{5}}\sqrt{5}, \quad (3,53)$$

$$M_{12}(t) = 2 + \sqrt{5}, \quad (3,54)$$

$$M_{21}(t) = 2 + \sqrt{5}, \quad (3,55)$$

$$M_{22}(t) = \sqrt{5 + \sqrt{5}}, \quad (3,56)$$

la forma del control es

$$v(t) = -(2 + \sqrt{5})z_1(t) - (\sqrt{5 + \sqrt{5}})z_2(t), \quad (3,57)$$

y las ecuaciones de estado (3.31) y (3.32) toma la forma

$$\dot{z}_1(t) = z_2(t), \quad z_1(0) = z_{10}, \quad (3,58)$$

$$\begin{aligned} \dot{z}_2(t) &= z_1^2(t) - (z_1(t) + \sqrt{z_1^2(t) + 1})z_1(t) \\ &\quad - (\sqrt{2z_1(t) + 1} + 2\sqrt{z_1^2(t) + 1})z_2(t), \quad z_2(0) = z_{20}. \end{aligned} \quad (3,59)$$

Considere las condiciones iniciales $z_1(0) = 1$ y $z_2(0) = 1$ para el sistema (3.58) y (3.59), controlado por (3.53)–(3.57), para llevar a cabo simulaciones numéricas.

La fig. 3 muestra el resultado de aplicar el regulador (3.48)–(3.57) al sistema (3.58),(3.59), que muestra las gráficas del estado (3.58) $z_1(t)$ y (3.59) $z_2(t)$, controlado por (3.57), la función de control (3.57) $v(t)$, y la gráfica del criterio (3.33) $J(t)$ en el intervalo $[0, 20]$. En el momento final $T = 20$, el criterio (3.33) toma el valor $J(20) = 6,9481$.

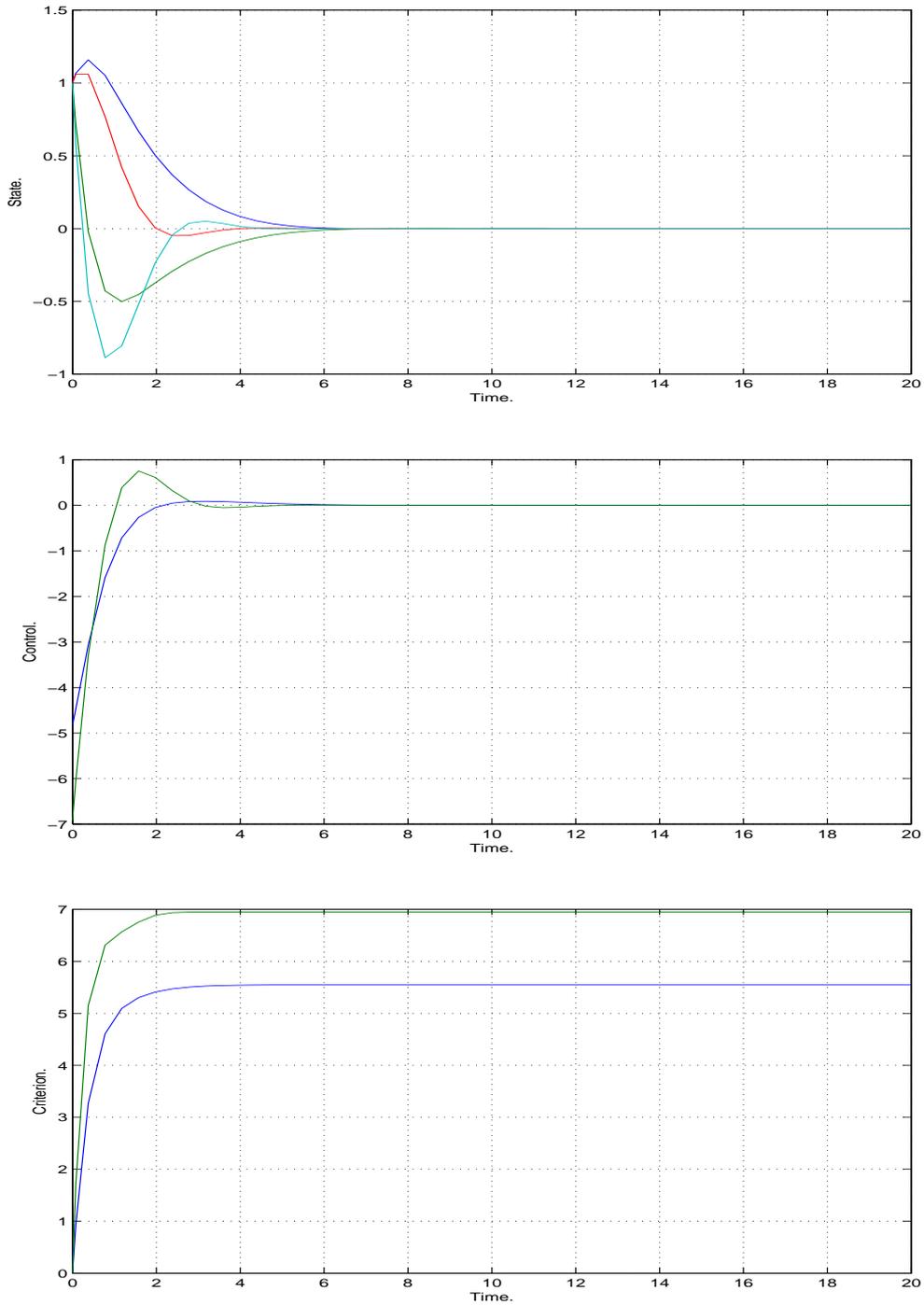


Figura 3.3: Regulador óptimo con respecto al criterio $J(t)$ vs. regulador lineal para el sistema linealizado en $[0, 20]$.

1. Estado. Gráfica del estado (3.31) $z_1(t)$ (Azul) y (3.32) $z_2(t)$ (Verde) controlado por (3.43), y el estado (3.58) $z_1(t)$ (Rojo) y (3.59) $z_2(t)$ (Celeste), controlado por (3.57), **2. Control.** la función del control (3.43) $v(t)$ (Azul) y el control lineal (3.57) $v(t)$ (Verde), y **3. Criterio.** el criterio (3.33) $J(t)$ producido por el control óptimo (Azul) y por el control lineal (Verde).

3.3.4. APÉNDICE.

Demostración del Teorema 3.1

Considere la ecuación Hamiltoniana para el problema del control óptimo (3.1),(3.2) dado por

$$H(z, m, v, t) = \frac{1}{2}(z^T(t)L(t)z(t) + v^T(t)R(t)v(t)) + m^T(h(z, t) + G(t)v(t)). \quad (3,60)$$

La función Hamiltoniana tiene un mínimo si

$$\frac{\partial H}{\partial v} = 0,$$

i.e.,

$$R(t)v(t) + G(t)^T m = 0.$$

La ley de control es obtenida como

$$v(t)^* = -R(t)^{-1}G(t)^T m.$$

El sistema de ecuaciones del co-estado esta dada por

$$\dot{m}(t) = -\frac{\partial H}{\partial z} = -Lz - (\beta_1(t) + \beta_2(t)z + \dots + \beta_r(t)z \dots_{r-1 \text{ times}} \dots z)^T m(t). \quad (3,61)$$

Supongamos que m depende del estado z como

$$m(t) = m(z(t)).$$

Entonces, la forma del control óptimo es

$$v(t)^* = -R(t)^{-1}G(t)^T m(z(t)). \quad (3,62)$$

La expresión para la derivada total $\dot{m}(t)$ toma la forma

$$\dot{m}(t) = \frac{dm(z)}{dz} \dot{z}(t). \quad (3,63)$$

Sustituyendo (3.63) y (3.1) en (3.61), la ecuación del co-estado es representado como

$$\begin{aligned} \frac{dm(z)}{dz}(\beta_1(t)z + \beta_2(t)zz^T + \dots + \beta_r(t)z \dots_{r \text{ times}} \dots z - G(t)R^{-1}(t)G^T(t)m(z)) = \\ = -Lz - (\beta_1(t) + 2\beta_2(t)z + \dots + r\beta_r(t)z \dots_{r-1 \text{ times}} \dots z)^T m(z) \end{aligned}$$

con la condición inicial $m(0) = 0$, que coincide con (3.4).

Demostración del Teorema 3.2

Considere el criterio dado por

$$J = \frac{1}{2} \int_{t_0}^T v^T(t)R(t)v(t)dt + \frac{1}{2} \int_{t_0}^T z^T(t)L(t)z(t)dt, \quad (3,64)$$

donde $T \in [t_0, \infty)$. Introduciendo la ecuación Hamilton-Jacobi-Bellman para el problema de control óptimo (3.1),(3.64) como

$$\begin{aligned} 0 &= \frac{\partial V}{\partial t} + \min_v \left\{ \frac{1}{2}(v^T R(t)v + z^T L(t)z) + \frac{\partial V^T}{\partial z} \dot{z}(t) \right\} \\ &= \frac{\partial V}{\partial t} + \min_v \left\{ \frac{1}{2}(v^T R(t)v + z^T L(t)z) + \frac{\partial V^T}{\partial z} (h(z, t) + G(t)v(t)) \right\}, \quad (3.65) \\ &V(z(T), T) = 0, \end{aligned}$$

donde

$$V(z, t) = \min_v \left\{ \frac{1}{2} \int_t^T v^T(t)R(t)v(t)dt + \frac{1}{2} \int_t^T z^T(t)L(t)z(t)dt \right\}$$

es la función de Bellman. Diferenciando con respecto a $v(t)$ para encontrar el mínimo en la ecuación HJB (3.65) obteniendo

$$0 = \frac{\partial}{\partial v} \left\{ \frac{\partial V}{\partial t} + \min_v \left\{ \frac{1}{2}(v^T R(t)v + z^T L(t)z) + \frac{\partial V^T}{\partial z} (h(z, t) + G(t)v(t)) \right\} \right\}$$

En consecuencia, se obtiene la ley de control óptimo como

$$0 = R(t)v + G(t)^T \frac{\partial V}{\partial z}.$$

$$v^*(t) = -R^{-1}(t)G^T(t) \frac{\partial V}{\partial z}.$$

Definamos $V(z, t)$ como una función cuadrático de $z(t)$ dada por

$$V(z, t) = -\frac{1}{2}z(t)^T M(t)z(t), \quad (3.66)$$

donde $M(t)$ es una matriz simétrica de dimension $n \times n$. Por otra parte, la derivada parcial de $V(t, z)$ con respecto a t y z estan dadas por

$$\frac{\partial V}{\partial z} \dot{z}(t) = -\frac{1}{2}\dot{z}(t)^T M(t)z(t) - \frac{1}{2}z(t)^T M(t)\dot{z}(t), \quad (3.67)$$

$$\frac{\partial V}{\partial t} = -\frac{1}{2}z(t)^T \dot{M}(t)z(t). \quad (3.68)$$

Esto proporciona la forma completa del control óptimo

$$v^*(t) = R^{-1}(t)G^T(t)M(t)z(t). \quad (3.69)$$

Usando la ecuación HJB y sustituyendo (3.67)-(3.69) en (3.65), se obtiene

$$\begin{aligned} \frac{1}{2}z(t)^T \dot{M}(t)z(t) &= \frac{1}{2}(z(t)^T M(t)G(t)R^{-1}(t)G^T(t)M(t)z(t) + z^T L(t)z) - \frac{1}{2}\dot{z}(t)^T M(t)z(t) - \frac{1}{2}z(t)^T M(t)\dot{z}(t) \\ &= \frac{1}{2}(z(t)^T M(t)G(t)R^{-1}(t)G^T(t)M(t)z(t) + z^T L(t)z) - \frac{1}{2}z(t)^T M(t)h(z, t) - \frac{1}{2}h(z, t)^T M(t)z(t) \\ &\quad - z(t)^T M(t)G(t)R^{-1}(t)G^T(t)M(t)z(t). \end{aligned}$$

Simplificando la ultima ecuación obtenemos

$$z(t)^T \dot{M}(t)z(t) = z^T L(t)z - z(t)^T M(t)G(t)R^{-1}(t)G^T(t)M(t)z(t) - z(t)^T M(t)h(z, t) - h(z, t)^T M(t)z(t).$$

Sustituyendo el valor de $h(z, t)$

$$h(z, t) = \beta_1(t)z(t) + \beta_2(t)z(t)z(t)^T + \dots + \beta_r(t)z \dots r \text{ times} \dots z,$$

se obtiene la siguiente ecuación incluyendo $M(t)$ obteniendo

$$\begin{aligned} z(t)^T \dot{M}(t)z(t) &= z(t)^T \{L(t) - M(t)G(t)R^{-1}(t)G^T(t)M(t) \\ &\quad - M(t)(\beta_1(t) + \beta_2(t)z + \dots + \beta_r(t)z \dots r-1 \text{ times} \dots z) \\ &\quad - (\beta_1(t) + \beta_2(t)z + \dots + \beta_r(t)z \dots r-1 \text{ times} \dots z)^T M(t)\}z(t), \end{aligned} \quad (3,70)$$

que se satisface si

$$\begin{aligned} \dot{M}(t) &= L(t) - M(t)G(t)R^{-1}(t)G^T(t)M(t) \\ &\quad - M(t)(\beta_1(t) + \beta_2(t)z + \dots + \beta_r(t)z \dots r-1 \text{ times} \dots z) \\ &\quad - (\beta_1(t) + \beta_2(t)z + \dots + \beta_r(t)z \dots r-1 \text{ times} \dots z)^T M(t). \end{aligned} \quad (3,71)$$

La condición frontera se sigue de la ecuación (3.66)

$$V(z(T), T) = -\frac{1}{2}z(T)^T M(T)z(T) = 0,$$

i.e.,

$$M(T) = 0. \quad (3,72)$$

Sustituyendo el control (3.69) en (3.1) obtenemos la solución del problema del control óptimo (3.1), (3.64), si $T < \infty$. Como T tiende a infinito, la solución de la ecuación (3.67) converge a un equilibrio, que implica que $\dot{M}(t) = 0$ cuando $T \rightarrow \infty$, i.e.,

$$\begin{aligned} 0 &= L(t) - M(t)G(t)R^{-1}(t)G^T(t)M(t) - M(t)(\beta_1(t) + \beta_2(t)z + \dots + \beta_r(t)z \dots r-1 \text{ times} \dots z) \\ &\quad - (\beta_1(t) + \beta_2(t)z + \dots + \beta_r(t)z \dots r-1 \text{ times} \dots z)^T M(t). \end{aligned} \quad (3,73)$$

La ecuacion (3.73) es una ecuación algebraica tipo Riccati dependiente del estado, que coincide con (3.20).

4

EQUILIBRIO ε -NASH EN LAZO ABIERTO PARA JUEGOS DIFERENCIALES POLINOMIALES VIA ECUACIÓN TIPO RICCATI DEPENDIENTE DEL ESTADO.

4.1. PLANTEAMIENTO DEL PROBLEMA.

Considere el siguiente juego diferencial polinomial, donde la dinámica de los jugadores esta regida por la ecuación diferencial:

$$\begin{aligned}\dot{x}(t) &= f(t, x) + \sum_{j=1}^N B^j(t)u^j(t) + d(t), \\ x(t_0) &= x_0,\end{aligned}\tag{4.1}$$

y un funcional de costo cuadrático para el índice de desempeño de cada jugador:

$$\begin{aligned}L_T^i(u^i, u^{\hat{i}}) &= \frac{1}{2}x^T(T)Q_f^i x(T) \\ &+ \frac{1}{2} \int_{t_0}^T \left(x^T(t)Q_1^i(t)x(t) + \sum_{j=1}^N u^{jT}(t)R^{ij}(t)u^j(t) \right) dt,\end{aligned}\tag{4.2}$$

donde $x(t) \in \mathbb{R}^n$ es el vector de estado de juego, u^j es el control (acción) del j -ésimo jugador, que varia dentro de una region dada $U^j \subset \mathbb{R}^{m_j}$, j denota el numero de jugador ($j = \overline{1, N}$), $B^j(t) \in \mathbb{R}^{n \times m_j}$ son los coeficientes matriciales de los controles y $d(t) \in \mathbb{R}^n$ es una señal de excitación continua conocida. El índice de desempeño $L_T^i(u^i, u^{\hat{i}})$ esta dado en la forma de Bolza, donde u^i es el control del i -ésimo jugador y $u^{\hat{i}}$ son los controles para el resto de los jugadores (\hat{i} es la colección de contra-coalición de jugadores contrarrestando al jugador con índice

i). Para cada jugador, el objetivo del juego es lograr la reducción al mínimo de su propio índice de desempeño por la selección de entradas apropiadas. También suponga que:

$$\begin{aligned} Q^i(t) = Q^{i\top}(t) \geq 0, \quad Q_f^i = Q_f^{i\top} \geq 0, \\ R^{ii}(t) = R^{ii\top}(t) > 0, \quad R^{ij}(t) = R^{ij\top}(t) \geq 0, \\ i \neq j. \end{aligned} \quad (4.3)$$

Considere la función no lineal $f(t, x)$ como un polinomio de n variables del vector de estado $x(t) \in \mathbb{R}^n$; esto requiere una especial definición del polinomio de orden $n > 1$. Siguiendo el trabajo anterior (ver [52]), un polinomio de p -ésimo grado del vector $x(t) \in \mathbb{R}^n$ es considerado como un forma p -lineal de n componentes de $x(t)$, es decir:

$$\begin{aligned} f(t, x) = a_0(t) + a_1(t)x + a_2(t)xx^\top + \dots \\ + a_s(t)x \cdots \textit{times} \cdots x. \end{aligned} \quad (4.4)$$

Aquí, los parámetros involucrados son: a_0 es un vector de dimensión n , a_1 es una matriz de dimensión $n \times n$, a_2 es un 3D tensor de dimensión $n \times n \times n$, y a_s es un $(s+1)$ D tensor de dimensión $n \times \cdots (s+1) \textit{times} \cdots \times n$, y $x \times \cdots \textit{times} \cdots \times x$ es un p D tensor de dimensión $n \times \cdots \textit{times} \cdots \times n$, obtenido por la multiplicación espacial del vector x p -veces por si mismo. Esto también se puede representar como un polinomio en forma de sumatoria:

$$\begin{aligned} f_k(t, x) = a_0 k(t) + \sum_i a_1 ki(t) x_i + \sum_{ij} a_2 kij(t) x_i x_j + \dots \\ + \sum_{i_1 \dots i_s} a_1 ki_1 \dots i_s(t) x_{i_1} \dots x_{i_s}, \\ k, i, j, i_1 \dots i_s = 1, \dots, n, \end{aligned} \quad (4.5)$$

Para cada conjunto de información disponible $\eta_i(t)$ y cada conjunto de estrategias $\gamma^i \in \Gamma^i (i \in \mathbb{N})$, la acción de control esta completamente determinada por la relación $u^i = \gamma^i(\eta_i)$. Sustituyendo el conjunto u^i dentro del funcional de consto (4.2) para un tiempo final fijo T , conduce al numero $L_T^i(u^i, u^{\hat{i}})$, ($i \in \mathbb{N}$), que es el costo que se origina para el jugador i definido en el espacio de acciones del control [8]. Para la condición inicial del estado fija x_0 , se obtiene el mapeo definido por

$$J_T^i : \Gamma^1 \times \Gamma^1 \times \cdots \times \Gamma^N \mapsto \mathbb{R}, \quad (\gamma^1, \gamma^2, \dots, \gamma^N) \mapsto L_T^i(u^i, u^{\hat{i}}),$$

el cual es conocido como el funcional de costo del i -ésimo jugador para el juego en forma estratégica.

4.1.1. SUPOSICIONES Y DEFINICIONES BÁSICAS.

A1 Los coeficientes matriciales del control $B^j(t)$ ($j = 1, \dots, N$), la función polinomial $f(t, x)$, y la señal de excitación $d(t)$ se consideran conocidas e integrables en $[0, T]$ para cualquier participante, i.e., $d(\cdot) \in \mathcal{L}^1(0, T; \mathfrak{R}^n)$.

A2 La clase U^i de acciones de control admisibles $u^i(t)$ ($i = 1, \dots, N$) contienen todos los controles no estacionarios satisfaciendo la condición uniformemente Lipschitz en t , esto es, para cualquier $t \geq 0$:

$$\|u^i(t) - u^i(t')\| \leq \bar{\delta}^i \|t - t'\|. \quad (4.6)$$

Además, todos los controles admisibles se consideran ser cuadráticamente integrables en t (así como las dinámicas correspondientes $x(t)$) sobre el intervalo de tiempo $[0, T]$ para todo x , esto es, para todo $x \in \mathbb{R}^n$

$$u^i(\cdot) \in \mathcal{L}^2(0, T; \mathbb{R}^{m_i}) \text{ y } x(t) \in \mathcal{L}^2(0, T; \mathbb{R}^n). \quad (4.7)$$

A3 Todos los jugadores tienen acceso solo a modelo de información en lazo abierto, esto es:

$$\eta_i(t) = \{x_0\}, \quad t \in [0, T]. \quad (4.8)$$

En el espacio de acciones de control, se define:

Definición 4.1[Equilibrio de ε -Nash] Las acciones de control $u^{i*}(\cdot)$ ($i = \overline{1, N}$) de los jugadores se dice que están en **Equilibrio ε -Nash** si para cualquier otra acción de control $u^i(\cdot)$ ($i = 1, \dots, N$) la siguiente desigualdad se satisface:

$$L_T^{i*} := L_T^i(u^{i*}, u^{i*}) \leq \inf_{u^i(\cdot) \in U^i} L_T^i(u^i, u^{i*}) + \varepsilon_T^i, \quad i = 1, \dots, N, \quad (4.9)$$

donde $\varepsilon_T^i \geq 0$ es un valor numérico que caracteriza hasta que punto el control $u^{i*}(\cdot)$ ($i = \overline{1, N}$) se desvía del equilibrio de Nash puro ($\varepsilon_T^i = 0$).

Es fácil observar de la definición, que en el caso del equilibrio de Nash puro ($\varepsilon^i = 0$), (si el funcional de costo es convexo con respecto a los primeros argumentos) el control u^{i*} que minimiza el funcional $L_T^i(u^i, u^{i*})$ para el i -ésimo jugador satisface exactamente $\varepsilon_T^i = 0$ (4.9). En el caso $\varepsilon^i > 0$, existe obligatoriamente un conjunto U^{i*} de controles $(u^{i*})'$, $(u^{i*})''$, etc., que satisface (4.9) y, por lo tanto, cada uno de ellos puede ser considerado como ε -equilibrio de Nash. El problema de encontrar los controles del equilibrio en una forma cerrada, como en el caso lineal, que admita una solución exacta de equilibrio ($\varepsilon_T^i = 0$) para el juego (4.1)-(4.2), puede no ser una tarea fácil. Sin embargo, proponemos una cierta solución en forma retroalimentada, que provee un *equilibrio no-cero*.

Evidentemente, al aplicar un conjunto de controles, el funcional de costo $\hat{L}_T^i(u^i, u^{i*})$, posee la misma estructura como el funcional (4.2) y coincide con $L_T^i(u^i, u^{i*})$ si son aplicados controles exactos, resultando ser dependiente de este conjunto de controles. Por lo tanto, ambos funcionales de costo en el lado derecho e izquierdo de la desigualdad resultan:

$$\hat{L}_T^{i*} := \hat{L}_T^i(u^{i*}, u^{i*}) \leq \inf_{u^i(\cdot) \in U^i} \hat{L}_T^i(u^i, u^{i*}) + \varepsilon_T^i, \quad (4.10)$$

también depende de los controles "no cero" así como ε_T^i . La relación entre L_T^{i*} y \hat{L}_T^{i*} es presentada en una nota dada en la siguiente sección.

4.2. EQUILIBRIO ε -NASH EN LAZO ABIERTO PARA JUEGOS POLINOMIALES.

4.2.1. CASO DE TIEMPO FINITO.

Para el problema de encontrar un conjunto de estrategias para el equilibrio de Nash en el juego (4.1)-(4.2), y dada la estructura de información en lazo abierto (4.8), el siguiente teorema contiene el caso en tiempo finito.

Teorema 4.1 Sean los controles u^{i*} , ($i = \overline{1, N}$) tales que existe la solución que consiste de las N funciones del co-estado ψ^i para el siguiente conjunto de ecuaciones diferenciales:

$$\begin{aligned} \dot{\psi}^i(t) &= -\frac{\partial H^i(t, \psi^i, x^*, u^{i*}, \hat{u}^{i*})}{\partial x} \\ &= -Q^i x^*(t) - \left(a_1 + \dots + a_n \underbrace{x^*(t) * \dots * x^*(t)}_{(n-1)\text{-times}} \right)^\top \psi^i, \end{aligned} \quad (4.11)$$

con la condición terminal $\psi^i(T) = Q_f^i x(T)$, donde $H^i(x, \psi^i, u^i, \hat{u}^i, t)$ es el Hamiltoniano para cada jugador definido como:

$$\begin{aligned} H^i(x, \psi^i, u^i, \hat{u}^i, t) &= \frac{1}{2} \left(x^T(t) Q^i x(t) + \sum_{j=1}^N u^{jT}(t) R^{ij} u^j(t) \right) + \\ \psi^{iT} &\left(a_1 x(t) + \dots + a_n \underbrace{x(t) * \dots * x(t)}_{n\text{-times}} + \sum_{j=1}^N B^j u^j(t) + d(t) \right), \\ & i = 1, \dots, N. \end{aligned} \quad (4.12)$$

el vector x^* satisface:

$$\dot{x}^* = f(t, x^*) + \sum_{j=1}^N B^j(t) u^{j*}(t) + d(t), \quad x^*(0) = x_0, \quad (4.13)$$

y i -ésimo control satisface la condición de minimalidad:

$$u^{i*}(t) = \arg \min_{u^i(\cdot) \in U^i} H^i(x^*, \psi^i, u^i, \hat{u}^{i*}, t). \quad (4.14)$$

Entonces, el conjunto de controles u^{i*} ($i = \overline{1, N}$) en la forma (4.15) provee la solución del equilibrio de Nash en lazo abierto.

$$u^{i*}(t) = - (R^{ii})^{-1} B^{i\top} \psi^i(t). \quad (4.15)$$

La demostración de este teorema emplea las mismas técnicas como el teorema 6.11 en [8], y se emite para evitar repeticiones.

Observación 4.1 La solución dada en el teorema 3.2 corresponde a la solución del equilibrio de Nash puro ($\varepsilon = 0$), esto es, una vez que la acción del control (4.15) son aplicadas al juego, la existencia de la solución para el problema de N puntos en la frontera provee la solución en lazo abierto del equilibrio de Nash.

En el caso lineal-cuadrático, la solución en lazo abierto del equilibrio de Nash, puede ser representada (ver [4]) en forma de retroalimentación del estado. En lo que sigue, se muestra que es realmente posible proporcionar

la representación en lazo cerrado para la solución dada por (4.2.1)-(4.15) para un juego diferencial polinomial, pero las acciones de control conducirán a una solución ε -equilibrio de Nash. Así, se consideraran los controles admisibles en la clase de funciones afines dependientes del estado (no estacionarios):

$$U_{fb}^i := \{u^i \in U^i \mid u^i(x, t) = K^i(t)x + k^i(t)\}. \quad (4.16)$$

Con el fin de estimar en que medida este conjunto de ε -controles se desvía del equilibrio cero se definen los siguientes procesos:

$x(t)$ es la dinámica del juego cuando todos los jugadores usan la solución exacta del equilibrio de Nash (4.15).

Esta trayectoria es además usada para evaluar el costo $L_T^i(u^{i*}, u^{\hat{i}*})$.

$\tilde{x}(t)$ es la dinámica del juego cuando los jugadores aplican el control en lazo cerrado en la forma (4.16); en este caso, esta trayectoria contribuye a $\hat{L}_T^i(u^{i*}, u^{\hat{i}*})$.

$\hat{x}(t)$ es la dinámica del juego cuando cada jugador j -ésimo usa el control (4.15), excepto el jugador i -ésimo, $j \neq i$, y u^i es un control admisible que satisface (4.6). El correspondiente costo es $L_T^i(u^i, u^{\hat{i}*})$.

$\bar{x}(t)$ es la dinámica del juego cuando cada jugador usa el control (4.16) excepto el jugador i -ésimo, u^i es un control admisible que satisface (4.6). El correspondiente costo es $\hat{L}_T^i(u^i, u^{\hat{i}*})$.

Dado que los sistemas en lazo cerrado están regulados con controles Nash (4.15) o con controles afín lineales (4.16), las trayectorias $x(t)$ y $\tilde{x}(t)$ deberían estar acotadas cuadráticamente, esto es, existe β_1, β_2 tales que

$$\int_0^T \|x\|^2 dt \leq \beta_2 < \infty, \quad \int_0^T \|\tilde{x}\|^2 dt \leq \beta_1 < \infty. \quad (4.17)$$

Los estados finales de estas trayectorias están también acotadas, esto es:

$$\|x(T)\|^2 \leq \gamma_2, \quad \|\tilde{x}(T)\|^2 \leq \gamma_1. \quad (4.18)$$

Debido a la suavidad de los controles admisibles (4.6)-(4.7), las trayectorias $\hat{x}(t)$ y $\bar{x}(t)$ están también acotadas, esto es, existen β_3, β_4 tales que

$$\int_0^T \|\hat{x}\|^2 dt \leq \beta_3 < \infty, \quad \int_0^T \|\bar{x}\|^2 dt \leq \beta_4 < \infty, \quad (4.19)$$

así como sus estados finales:

$$\|\hat{x}(T)\|^2 \leq \gamma_3, \quad \|\bar{x}(T)\|^2 \leq \gamma_4. \quad (4.20)$$

4.2.2. EQUILIBRIO ε -NASH EN LAZO ABIERTO VIA ECUACIÓN TIPO RICCATI DEPENDIENTE DEL ESTADO

El siguiente teorema usa las técnicas de la ecuación tipo Riccati dependiente del estado para representar la solución de lazo abierto en forma retroalimentada

Teorema 4.2 Para el juego diferencial polinomial de N -personas en tiempo finito (4.1)-(4.2) satisfaciendo las restricciones (4.3), si existe la solución de las ecuaciones acopladas tipo Riccati dependientes del estado

$$\begin{aligned}
& \dot{P}^i(t) + P^i(t) \left(a_1 + a_2 x(t) \dots + a_n \underbrace{x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right) \\
& + \left(a_1 + 2a_2 x(t) \dots + na_n \underbrace{x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right) P^i(t) \\
& + Q^i - P^i(t) \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} P^j(t) = 0; \\
& P^i(T) = Q_f^i,
\end{aligned} \tag{4.21}$$

y la solución del *shifting vector*

$$\begin{aligned}
& \dot{p}^i(t) + \left(a_1 + 2a_2 x(t) \dots + na_n \underbrace{x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right) p^i(t) \\
& - P^i(t) \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} p^j(t) + P^i(t) d(t) = 0, \\
& p^i(T) = 0.
\end{aligned} \tag{4.22}$$

entonces la acción de control estará dada por:

$$u^{i*}(t) = -(R^{ii})^{-1} B^{iT} [P^i(t)x(t) + p^i(t)]. \tag{4.23}$$

lleva al ε -equilibrio de Nash en lazo abierto para el juego polinomial (4.1)-(4.2). Esto es, para cualquier control admisible u^i satisfaciendo (4.6), la desigualdad

$$\hat{L}_T^{i*}(u^{i*}, u^{i*}) \leq \hat{L}_T^i(u^i, u^{i*}) + \varepsilon_T^i, \tag{4.24}$$

con

$$\begin{aligned}
\varepsilon_T^i & \leq \frac{1}{2} (\lambda_{\max}\{Q_f^i\} - \lambda_{\min}\{Q_f^i\}) \max\{\gamma_1, \gamma_2\} \\
& + \frac{1}{2} (\lambda_{\max}\{Q_f^i\} - \lambda_{\min}\{Q_f^i\}) \max\{\gamma_3, \gamma_4\} \\
& + \frac{1}{2} (\lambda_{\max}\{Q^i\} - \lambda_{\min}\{\bar{Q}^i\}) \max\{\beta_3, \beta_4\} \\
& + \frac{1}{2} \lambda_{\max}\{F^i\} + \lambda_{\max}\{\bar{M}^i\} \sqrt{\beta_3} + \xi_2 \\
& + \frac{1}{2} (\lambda_{\max}\{\bar{Q}^i\} - \lambda_{\min}\{Q^i\}) \max\{\beta_1, \beta_2\} \\
& + \frac{1}{2} \lambda_{\max}\{L^i\} + \lambda_{\max}\{\bar{M}^i\} \sqrt{\beta_1} + \xi_1,
\end{aligned} \tag{4.25}$$

se mantiene para un funcional \hat{L}_T^{i*} definido en (4.10). Las constantes $\beta_k = 1, \dots, 4$ y $\gamma_k = 1, \dots, 4$ son definidas en

(4.17) - (4.20), y las matrices involucradas se definen como

$$\begin{aligned}
\tilde{Q}^i &= Q^i + \sum_{j=1}^N P^j B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} P^j, \\
L^i &= \sum_{j=1}^N p^{jT} B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} p^j, \\
\tilde{M}^i &= \sum_{j=1}^N p^{jT} B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} P^j \\
\bar{Q}^i &= Q^i + \sum_{j=1, j \neq i}^N P^j B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} P^j, \\
F^i &= \sum_{j=1, j \neq i}^N p^{jT} B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} p^j, \\
\bar{M}^i &= \sum_{j=1, j \neq i}^N p^{jT} B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} P^j, \\
i &= 1, \dots, N.
\end{aligned} \tag{4.26}$$

La demostración del teorema esta dada en el apéndice.

Las ecuaciones (4.21) son también llamadas ecuaciones tipo Riccati dependientes del estado, que son comúnmente usadas en problemas de optimización de un jugador (ver [53], [54], [55], [52]). El punto principal en la búsqueda de la soluciones para (4.21) y (4.22) es que estas ecuaciones deben resolverse a lo largo de las trayectorias del estado; para satisfacer este objetivo se deberá usar una aproximación numérica, como el método de disparo [56].

Observación 4.2 En el problema de optimización de un jugador, las ecuaciones (4.21) y (4.22) corresponden a la solución conocida para el regulador polinomial dado en [55] y [52].

Observación 4.3 La condición para obtener el conjunto de controles (4.23) es que el vector del co-estado individual satisfaga la ecuación $\psi^i(t) = P^i(t)x + p^i(t)$. Sin embargo, en el caso matricial, como se a señalado en [53] para el problema de optimización de un jugador, esta aproximación provee solo una solución sub-óptima, porque la condición (4.2.1) puede ser satisfecha en una forma afín de $\psi^i(t)$ solo asintóticamente. Esta es la razón por la que el control en forma retroalimentada (4.23) puede dar solo un ε -equilibrio, esto es, no existe garantía que (4.2.1) sería satisfecha para todos los puntos de la trayectoria (4.13). No obstante, la trayectoria generada por (4.23) sería suficientemente cerca de la generada por (4.2.1)-(4.15); este efecto es también observado en la simulación numérica en la sección V.

El teorema da una cota superior para ε_T^i , que es útil para determinar que tan lejos esta el equilibrio ε -Nash del equilibrio de Nash puro.

La siguiente observación es valida en el caso presentado ([7]). Es fácil darse cuenta que la definición (4.10) puede ser re-escrita en la siguiente forma equivalente:

$$\begin{aligned}
\hat{L}_T^i(u^{i*}, u^{i*}) &\leq \inf_{u^i(\cdot) \in U^i} \hat{L}_T^i(u^i, u^{i*}) + \varepsilon_T^i = \\
L_T^i(u^{i*}, u^{i*}) &+ \left(\inf_{u^i(\cdot) \in U^i} \hat{L}_T^i(u^i, u^{i*}) - L_T^i(u^{i*}, u^{i*}) + \varepsilon_T^i \right) \\
&\leq L_T^i(u^{i*}, u^{i*}) + \tilde{\varepsilon}_T^i,
\end{aligned}$$

donde

$$\tilde{\varepsilon}_T^i := \varepsilon_T^i + \left| \inf_{u^i(\cdot) \in U^i} \hat{L}_T^i(u^i, u^{i*}) - L_T^i(u^{i*}, u^{i*}) \right|,$$

que da la comparación de los funcionales de costo correspondientes a las estrategias del equilibrio ε -Nash y a los controles del equilibrio de Nash puro. La modificación $\tilde{\varepsilon}_T^i$ puede ser estimada por cálculos similares a los de ε_T^i .

4.3. CASO HORIZONTE DE TIEMPO INFINITO

En esta sección, consideramos el problema de control con horizonte de tiempo infinito, esto es, el criterio que cada i -ésimo jugador desea minimizar esta dado por:

$$L_\infty^i(u_1, u_2) = \frac{1}{2} \int_0^\infty (x^T Q_i x + \sum_{j=1}^2 u_j^T R_{ij} u_j) dt, \quad (4.27)$$

sujeto a la dinámica polinomial (4.1), con $d(t) \equiv 0$. Así, el funcional de costo cuadrático con $Q_f^i = 0$ es estudiado. En este caso, el sistema de ecuaciones tipo Riccati dependientes del estado toman la forma

$$\begin{aligned} & P^i \left(a_1 + a_2 x(t) + \dots + \underbrace{a_n x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right) \\ & + \left(a_1 + 2a_2 x(t) + \dots + \underbrace{na_n x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right) P^i \\ & + Q^i - P^i \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} P^j = 0, \quad i = 1, \dots, N. \end{aligned} \quad (4.28)$$

Corolario 4.1 Para el juego no lineal polinomial con horizonte de tiempo infinito con las suposiciones (4.3), entradas no exogenas, y criterios cuadráticos (4.27), si las ecuaciones algebraicas tipo Riccati dependientes del estado (4.28) tienen solución definida positiva $P^i(x)$ ($i = 1, 2$), entonces los controles del equilibrio ε -Nash están dados por:

$$u^{i*}(t) = -(R^{ii})^{-1} B^{iT} P^i(t) x(t). \quad (4.29)$$

Observación 4.4 La principal diferencia entre la ecuación tipo Riccati del equilibrio estándar de Nash y de la ecuación tipo Riccati (4.28) es que la solución de (4.28) depende del vector $x(t)$. En el caso de optimización de un jugador, convencionalmente se considera buscar una solución en cada instante de tiempo, cuando los coeficientes dependientes del estado de (4.28) se consideran constantes, y entonces se aplican métodos estándar para resolver la ecuación algebraica tipo Riccati. Esto también es posible aplicar este enfoque para el caso de N -jugadores. En efecto, en vista de la estructura polinomial del juego, podemos imitar la aproximación Hamiltoniana para buscar los controles del equilibrio ε -Nash (4.29) en forma cerrada, esto es, como función que depende en términos polinomiales de $x(t)$. Desarrollaremos esta aproximación en la siguiente sección.

4.3.1. SOLUCIÓN DE LAS ECUACIONES ALGEBRAICAS ACOPLADAS TIPO RICCATI DEPENDIENTES DEL ESTADO POR UN ENFOQUE HAMILTONIANO.

La solución para el sistema de ecuaciones (4.28) está estrechamente relacionado (ver [57]) por la matriz Hamiltoniana dada por (por simplicidad, consideramos $N = 2$):

$$H = \begin{pmatrix} a_1 + \dots + a_n x^{n-1} & -S^1 & -S^2 \\ -Q^1 & -(a_1 + \dots + na_n x^{n-1})^\tau & 0 \\ -Q^2 & 0 & -(a_1 + \dots + na_n x^{n-1})^\tau \end{pmatrix}$$

donde $S^i = B^i (R^{ii})^{-1} B^{i\tau}$, $i = 1, 2$. Si el vector x es fijo, la matriz H coincide con el caso lineal cuadrático. En efecto, es conocido ([4] pag. 280) que para un vector fijo x si $V \subset \mathbb{R}^{3n \times n}$ es un subespacio invariante n -dimensional de H , $W_i \in \mathbb{R}^{n \times n}$, $i = 1, 2$, son tres matrices reales tales que:

$$V = \text{Im} \begin{bmatrix} W_0 \\ W_1 \\ W_2 \end{bmatrix},$$

y si W_0 es no singular, entonces $P_i := W_i W_0^{-1}$, $i = 1, 2$, es una solución para el sistema de ecuaciones acopladas tipo Riccati (4.28) (para $N = 2$). Por otra parte, la solución (P_1, P_2) , son independientes de la base V elegidas, que da un método para encontrar los controles del equilibrio Nash. Es también conocido que en el caso estándar (no dependiente de x), las ecuaciones acopladas tipo Riccati tienen una solución estabilizante (P_1, P_2) si y solo si la matriz H tiene un subespacio estable n -dimensional y H tiene $2n$ eigenvalores. Nosotros extendimos este enfoque para el caso polinomial como sigue.

El siguiente teorema establece las condiciones de existencia de un subespacio estable de la matriz H para un sistemas escalar ($n = 1$), que, a su vez, garantiza la existencia de una solución estable. Este resultado es formulado en términos de los coeficientes del juego polinomial de tercer orden escalar de dos jugadores, esto es, el costo (4.27) es minimizado con pesos escalares: $Q^i = q_i > 0 \in \mathbb{R}$, $R^{ij} = r_{ij} > 0 \in \mathbb{R}$, $i, j = 1, 2$, sujeto a la dinámica polinomial:

$$\begin{aligned} \dot{x}(t) &= a_1 x(t) + a_2 x(t)^2 + a_3 x(t)^3 + b_1 u_1(t) + b_2 u_2(t), \\ a_1, a_2, a_3, b_1, b_2 &\in \mathbb{R}. \end{aligned} \tag{4.30}$$

Teorema 4.3 Si para todo $x(t)$ los coeficientes $a_1, a_2, a_3, b_1, b_2, q_1, q_2, r_{11}$, y r_{22} para el juego polinomial de tercer orden para dos jugadores (4.27)-(4.30) satisface la condición

$$\frac{-\sqrt{r_{22}(r_{11}(4a_3x^2 + 3a_2x + 2a_1)^2 + 4b_1^2q_1) + 4b_2^2q_2r_{11}}}{\sqrt{r_{11}r_{22}}} < 2a_3x^2 + a_2x, \tag{4.31}$$

entonces existe un subespacio invariante H estable. Los controles ε -Nash (4.29) están dados por

$$\begin{aligned} u_1 &= \frac{q_1 r_{11} r_{22} \left(\frac{\sqrt{r_{22}(r_{11}(4a_3x^2 + 3a_2x + 2a_1)^2 + 4b_1^2q_1) + 4b_2^2q_2r_{11}}}{\sqrt{r_{11}\sqrt{r_{22}}} + 4a_3x^2 + 3a_2x + 2a_1} \right)}{2(b_1^2q_1r_{22} + b_2^2q_2r_{11})}, \\ u_2 &= \frac{q_2 r_{11} r_{22} \left(\frac{\sqrt{r_{22}(r_{11}(4a_3x^2 + 3a_2x + 2a_1)^2 + 4b_1^2q_1) + 4b_2^2q_2r_{11}}}{\sqrt{r_{11}\sqrt{r_{22}}} + 4a_3x^2 + 3a_2x + 2a_1} \right)}{2(b_1^2q_1r_{22} + b_2^2q_2r_{11})}, \end{aligned}$$

La demostración esta dada en el apéndice.

El teorema 4.3 plantea que para todo $x(t) \in \mathbb{R}$ que la desigualdad (4.31) se satisfaga para el conjunto de coeficientes en (4.27)–(4.30). En efecto, si para todo $x(t)$ la desigualdad es satisfecha, entonces la matriz H tiene un eigenvalor estable para toda $x(t)$.

Para el particular caso de $b_1 = 1, b_2 = 1, q_1 = 1, q_2 = 1, r_{11} = 1, y r_{22} = 1$, la desigualdad (4.31) puede ser reformulada como sigue.

Corolario 4.2 Si para toda $x(t)$ los coeficientes a_1, a_2 , y a_3 el juego polinomial de tercer orden escalar (4.27)–(4.30), satisface la condición

$$-\sqrt{(4a_3x^2 + 3a_2x + 2a_1)^2 + 8} < 2a_3x^2 + a_2x,$$

entonces existe un subespacio invariante H estable. Por otra parte, los controles ε -Nash

$$\begin{aligned} u^i(t) = & -\frac{1}{4} (2a_1 + 3a_2x(t) + 4a_3x(t)^2 \\ & + \sqrt{(4a_3x(t)^2 + 3a_2x(t) + 2a_1)^2 + 8}) x(t), \quad i = 1, 2, \end{aligned} \quad (4.32)$$

estabilizan globalmente la dinámica del juego (4.30) al origen.

La demostración se sigue del resultado obtenido en la demostración del teorema 3 (ver apéndice) que el eigenvalor dependiente del estado es negativo

$$\lambda(x) = -\frac{\sqrt{(4a_3x^2 + 3a_2x + 2a_1)^2 + 8}}{2} - \frac{1}{2} (2a_3x^2 + a_2x) < 0,$$

para toda $x(t)$. Por otra parte, sustituyendo los controles (4.32) en (4.30) obtenemos

$$\dot{x}(t) = x \left(\frac{1}{2} \left[- (2a_3x^2 + a_2x) - \sqrt{(4a_3x^2 + 3a_2x + 2a_1)^2 + 8} \right] \right). \quad (4.33)$$

El lado derecho de (4.33) es exactamente igual a $\lambda(x)x$. Dado que $\lambda(x) < 0$, la única raíz de la parte derecha de (4.33) es $x = 0$. Por lo tanto, $x = 0$ es el único equilibrio de (4.30), que es estable en vista del teorema 3.

Como se a señalado, la estructura polinomial del juego ayuda a encontrar los controles ε -Nash dependientes del estado x o a calcular los eigenvalores y eigenvectores de H . Para el caso general, Se propone el siguiente algoritmo heurístico para encontrar las soluciones estabilizantes para el juego de dos jugadores.

Algoritmo 1.- Se Asume que x es fijo como un valor constante.

2.- Se calculan los eigenvalores de H , que resultan ser dependientes de x .

3.- Para la elección particular de eigenvalores, calculamos los eigenvectores (también dependientes del estado x).

4.- Se calculan los eigenespacios de dimension $(3 * n) \times n$, generados por los eigenvectores linealmente independientes, que están dados por:

$$W = \begin{bmatrix} W_0 \\ W_1 \\ W_2 \end{bmatrix}.$$

5.- Las soluciones de la ecuación algebraica tipo Riccati dependiente del estado (4.28) están dadas por:

$$P_i = W_i W_0^{-1}. \quad (4.34)$$

Observación 4.5 El algoritmo desarrollado puede ser extendido al caso de N jugadores similarmente al caso lineal, como se describe en [4].

4.4. EJEMPLOS NUMÉRICOS.

4.4.1. EJEMPLO 1

Considere el juego escalar no lineal polinomial con 2 jugadores:

$$\dot{x} = x^3 + u_1 + u_2,$$

con la siguiente función de costo con horizonte de tiempo finito para cada jugador:

$$L_{10}^1(u_1, u_2) = \frac{1}{2} \int_0^{10} (x^2 + u_1^2 + u_2^2) dt + \frac{1}{2} x^2(10),$$

$$L_{10}^2(u_1, u_2) = \frac{1}{2} \int_0^{10} (0,1x^2 + u_1^2 + u_2^2) dt.$$

Las ecuaciones acopladas tipo Riccati toman la forma:

$$\dot{P}^1 = -(3xP^1 - \sum_{j=1}^2 P^1 P^j + 1),$$

$$\dot{P}^2 = -(3xP^2 - \sum_{j=1}^2 P^1 P^j + 0,1),$$

por lo tanto, los controles del ε -equilibrio de Nash para ambos jugadores esta dado por:

$$u^i = -P^i x; \quad i = 1, 2.$$

4.4.2. EJEMPLO 2

Considere el siguiente juego vectorial no lineal polinomial:

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = x_1^2 + u_1 + u_2,$$

con la función de costo cuadrática con horizonte de tiempo finito:

$$L_{10}^1(u_1, u_2) = \frac{1}{2} \int_0^{10} (0,1x_1^2 + x_2^2 + u_1^2 + u_2^2) dt$$

$$+ \frac{1}{2} (0,1x_1^2(T) + x_2^2(T)),$$

$$L_{10}^2(u_1, u_2) = \frac{1}{2} \int_0^{10} (x_1^2 + 0,1x_2^2 + u_1^2 + u_2^2) dt$$

$$+ \frac{1}{2} (x_1^2(T) + 0,1x_2^2(T)).$$

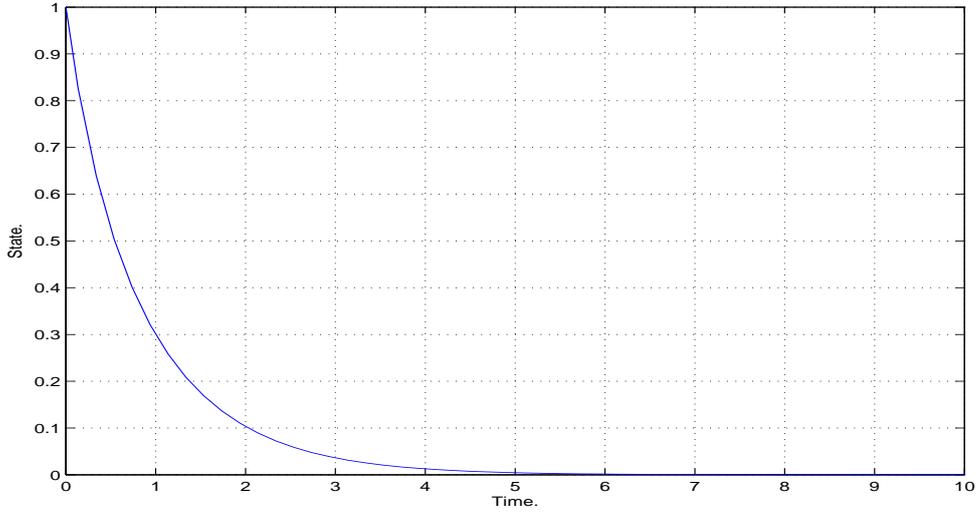


Figura 4.1: Trayectoria del estado del juego polinomial del ejemplo 1.

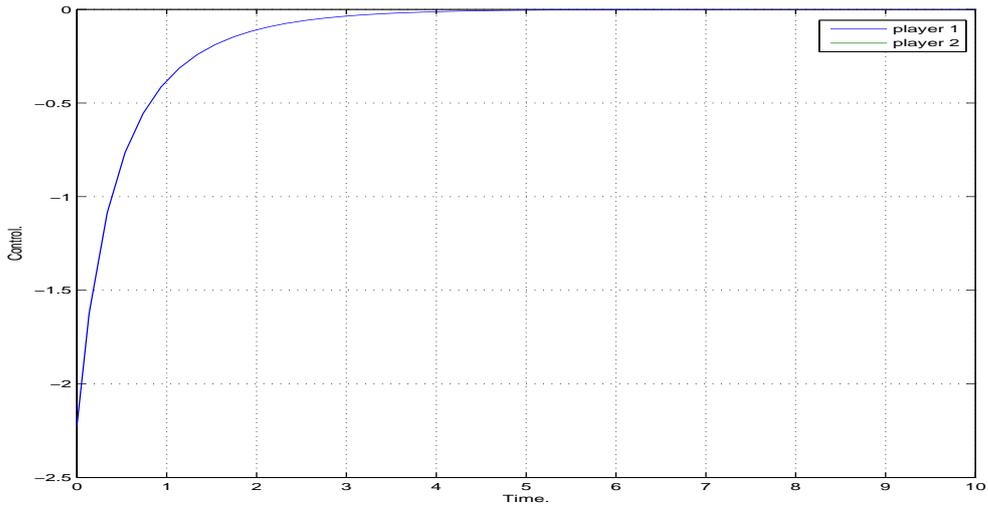


Figura 4.2: Acción de control del jugador 1 del ε -equilibrio de Nash.

Las ecuaciones acopladas tipo Riccati para el jugador 1 son:

$$\dot{P}_{11}^1 = -(x_1 P_{12}^1 + 2x_1 P_{21}^1 + 0,1 - \sum_{j=1}^2 P_{12}^1 P_{21}^j),$$

$$\dot{P}_{12}^1 = -(P_{11}^1 + 2x_1 P_{22}^1 - \sum_{j=1}^2 P_{12}^1 P_{22}^j),$$

$$\dot{P}_{21}^1 = -(P_{11}^1 + x_1 P_{22}^1 - \sum_{j=1}^2 P_{22}^1 P_{21}^j),$$

$$\dot{P}_{22}^1 = -(P_{12}^1 + P_{21}^1 + \frac{1}{47} - \sum_{j=1}^2 P_{22}^1 P_{22}^j),$$

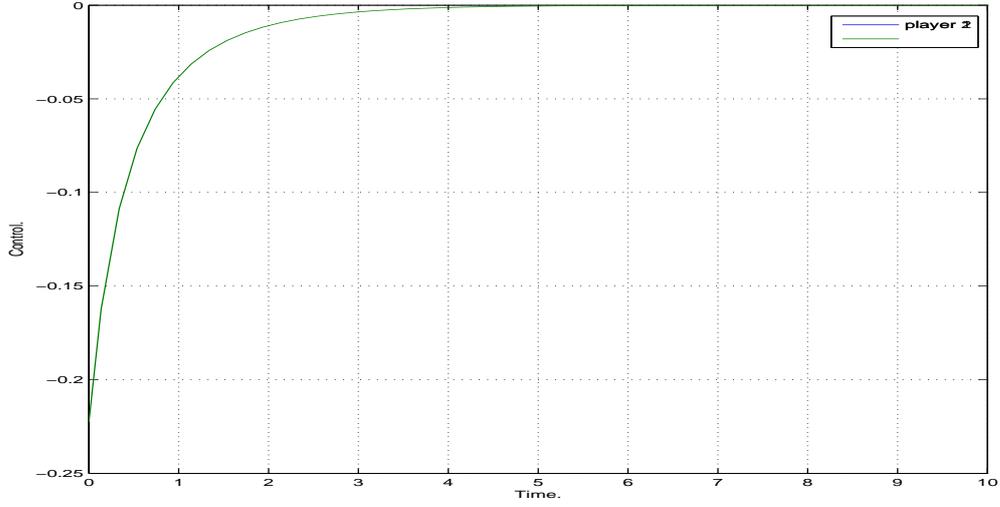


Figura 4.3: Acción de control del jugador 2 del ε -equilibrio de Nash.

con la condición final $P_{11}^1(10) = 0,1$, $P_{12}^1(10) = 0$, $P_{21}^1(10) = 0$, $P_{22}^1(10) = 1$, y las ecuaciones acopladas tipo Riccati para el jugador 2 son:

$$\begin{aligned}\dot{P}_{11}^2 &= -(x_1 P_{12}^2 + 2x_1 P_{21}^2 + 1 - \sum_{j=1}^2 P_{12}^2 P_{21}^j), \\ \dot{P}_{12}^2 &= -(P_{11}^2 + 2x_1 P_{22}^2 - \sum_{j=1}^2 P_{12}^2 P_{22}^j), \\ \dot{P}_{21}^2 &= -(P_{11}^2 + x_1 P_{22}^2 - \sum_{j=1}^2 P_{22}^2 P_{21}^j), \\ \dot{P}_{22}^2 &= -(P_{12}^2 + P_{21}^2 + 0,1 - \sum_{j=1}^2 P_{22}^2 P_{22}^j),\end{aligned}$$

con la condición final $P_{11}^2(10) = 1$, $P_{12}^2(10) = 0$, $P_{21}^2(10) = 0$, $P_{22}^2(10) = 0,1$. La acción de control para cada jugador es

$$u_i = -P_{21}^i x_1 - P_{22}^i x_2.$$

4.4.3. EXAMPLE 3

Para el siguiente juego no lineal escalar

$$\dot{x} = x^2 + u_1 + u_2,$$

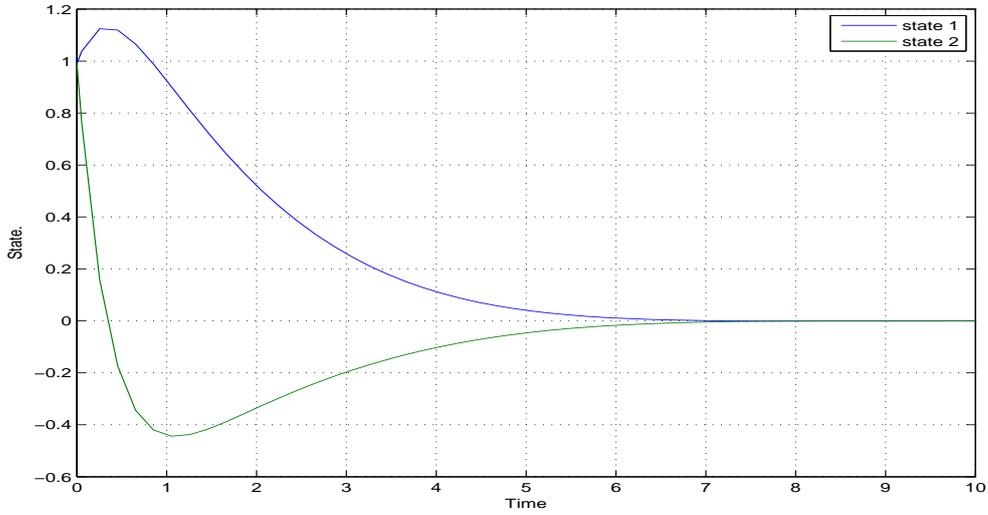


Figura 4.4: Trayectoria del estado del juego polinomial del ejemplo 2.

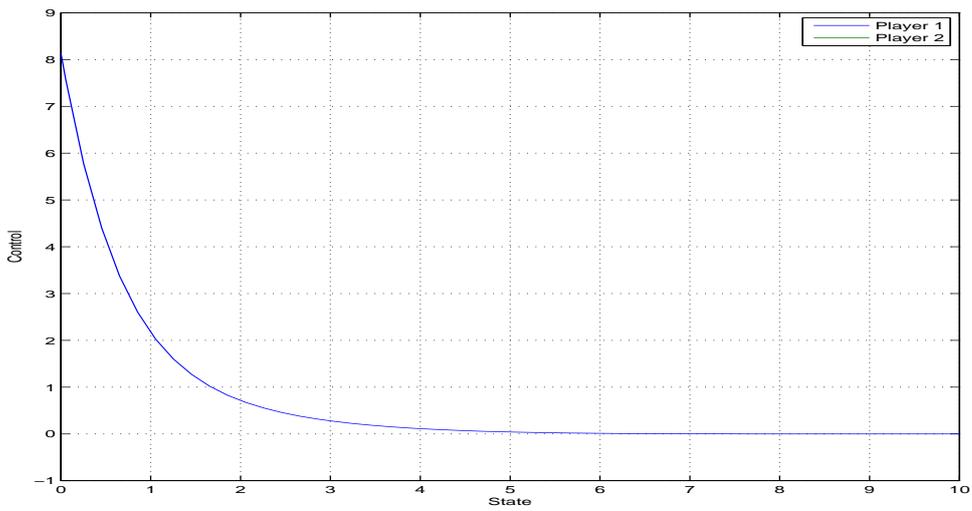


Figura 4.5: Acción de control del jugador 1 del ε -equilibrio de Nash.

con el funcional de costo cuadrático en horizonte de tiempo infinito

$$L_{\infty}^1(u_1, u_2) = \frac{1}{2} \int_0^{\infty} (x^2 + u_1^2 + u_2^2) dt,$$

$$L_{\infty}^2(u_1, u_2) = \frac{1}{2} \int_0^{\infty} (2x^2 + u_1^2 + u_2^2) dt,$$

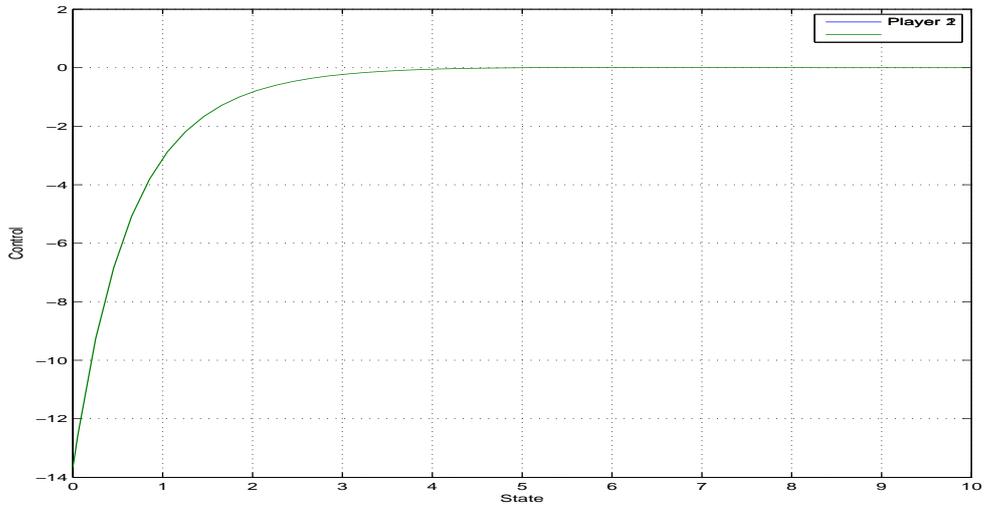


Figura 4.6: Acción de control del jugador 2 del ε -equilibrio de Nash.

Observe que los coeficientes polinomiales para este problema son iguales a $a_1 = 0$, $a_2 = 1$, $a_3 = 0$. Sustituyendo estos valores en la desigualdad (4.31) se obtiene

$$-\sqrt{9x^2 + 12} < x,$$

esta desigualdad es satisfecha para todos los valores de $x(t)$; entonces existe un eigenvalor con parte real negativa, como se muestra en la figura 5 que para todo valor de $x(t)$ el eigenvalor siempre es negativo.

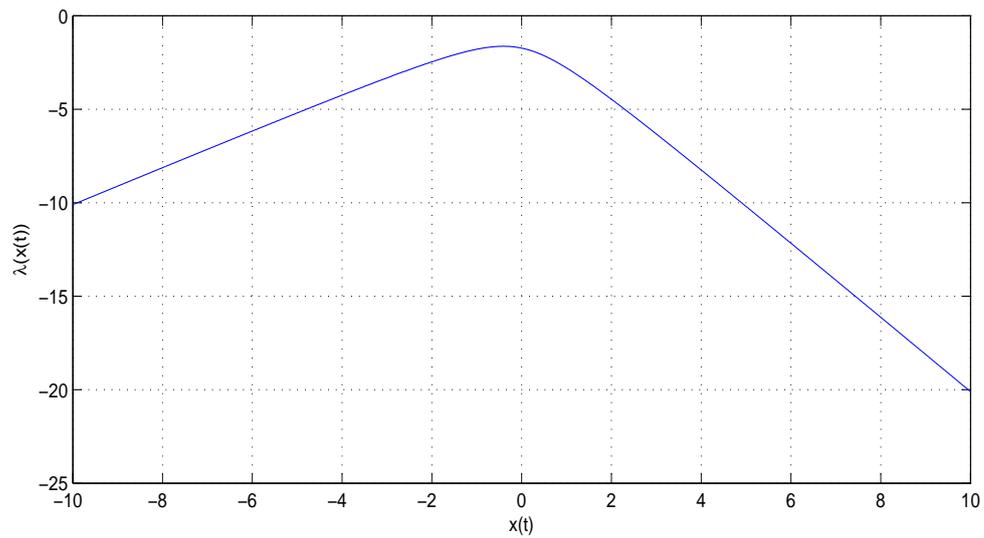


Figura 4.7: Eigenvalor dependiente del estado.

La matriz Hamiltoniana esta dada por

$$H = \begin{pmatrix} x & -1 & -1 \\ -1 & -2x & 0 \\ -2 & 0 & -2x \end{pmatrix}.$$

el polinomio característico

$$\begin{aligned} \det(H - \lambda I) &= \begin{vmatrix} x - \lambda & -1 & -1 \\ -1 & -2x - \lambda & 0 \\ -2 & 0 & -2x - \lambda \end{vmatrix} \\ &= (x - \lambda)(-2x - \lambda)^2 + 2x + \lambda + 4x + 2\lambda = 0, \end{aligned}$$

obteniendo

$$[(x - \lambda)(2x + \lambda) + 3](2x + \lambda) = 0,$$

que tiene las raíces $\lambda_1 = -2x$, and $\lambda^2 + x\lambda - (2x^2 + 3) = 0$ que dan:

$$\lambda_3 = \frac{-x + \sqrt{9x^2 + 12}}{2}, \quad \lambda_2 = \frac{-x - \sqrt{9x^2 + 12}}{2},$$

Sustituyendo el eigenvalor λ_2 en la matriz $H - \lambda I$

$$H - \lambda_2 = \begin{pmatrix} \frac{3x + \sqrt{9x^2 + 12}}{2} & -1 & -1 \\ -1 & \frac{-3x + \sqrt{9x^2 + 12}}{2} & 0 \\ -2 & 0 & \frac{-3x + \sqrt{9x^2 + 12}}{2} \end{pmatrix}$$

se obtiene el eigenvector asociado con el eigenvalor $\lambda_2 = \frac{-x - \sqrt{9x^2 + 12}}{2}$:

$$\begin{pmatrix} \frac{-3x + \sqrt{9x^2 + 12}}{2} \\ 1 \\ 2 \end{pmatrix}$$

Multiplicado por $\frac{-3x + \sqrt{9x^2 + 12}}{2}$:

$$\begin{pmatrix} 3 \\ \frac{3x + \sqrt{9x^2 + 12}}{2} \\ 3x + \sqrt{9x^2 + 12} \end{pmatrix}$$

se obtiene la solución dependiente del estado para la ecuación de Riccati:

$$P_1 = \frac{3x + \sqrt{9x^2 + 12}}{6}, \quad P_2 = \frac{3x + \sqrt{9x^2 + 12}}{3},$$

y los controles ε -Nash están dados por

$$u_1 = -\frac{3x + \sqrt{9x^2 + 12}}{6}x, \quad u_2 = -\frac{3x + \sqrt{9x^2 + 12}}{3}x.$$

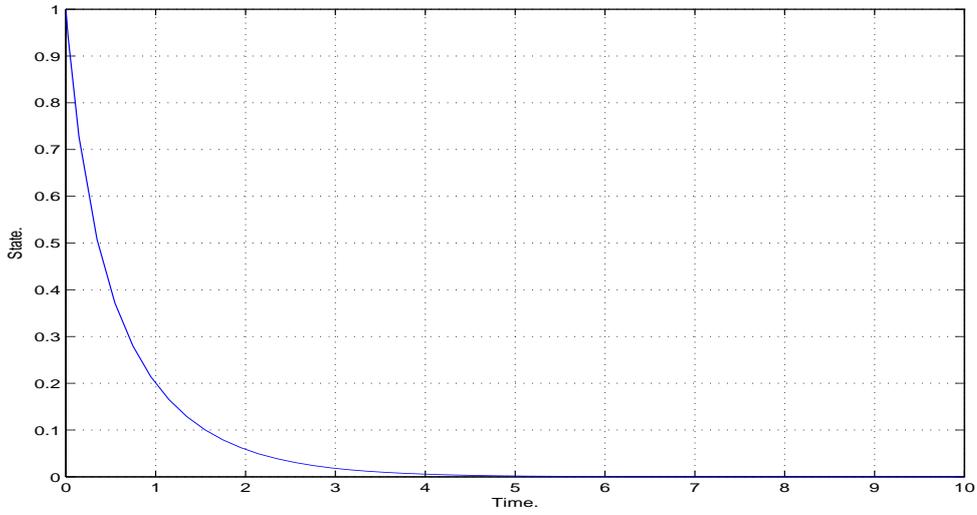


Figura 4.8: Trayectoria del estado del juego con horizonte de tiempo infinito del ejemplo 3.

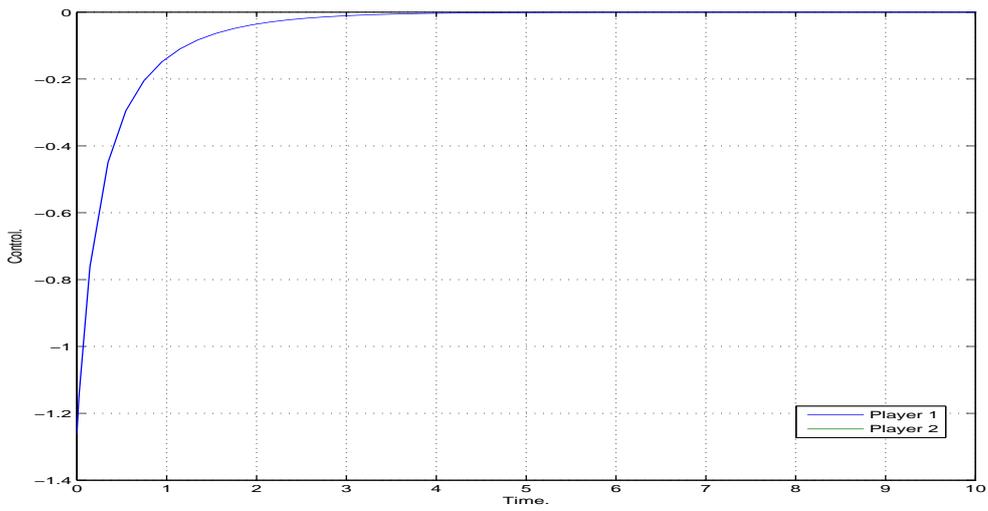


Figura 4.9: Acción del control del ε -equilibrio de Nash del jugador 1.

4.4.4. EJEMPLO 4

Ahora considere la siguiente dinámica polinomial:

$$\dot{x} = x + x^2 + x^3 + u_1 + u_2,$$

con el costo cuadrático con horizonte de tiempo infinito

$$J_i(u_1, u_2) = \int_0^{\infty} (x^2 + u_1^2 + u_2^2) dt, \quad i = 1, 2,$$

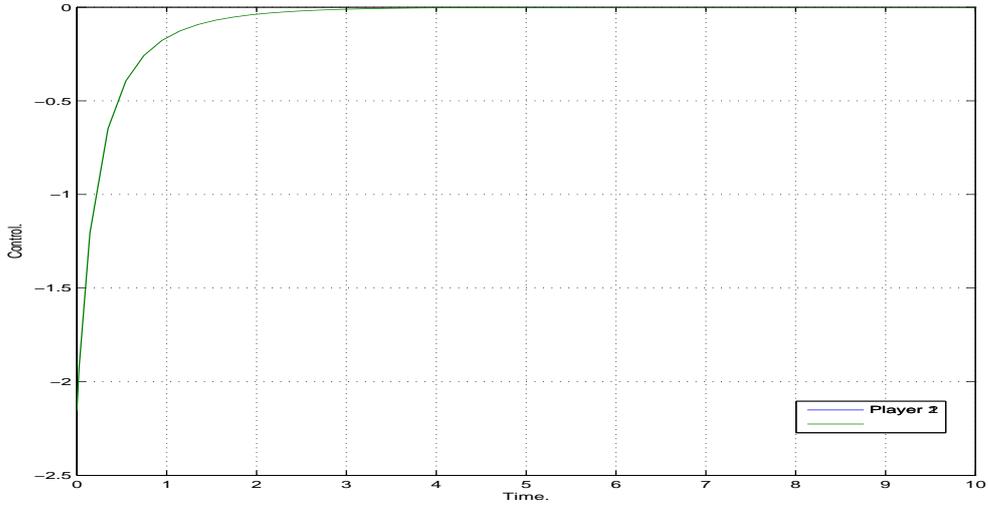


Figura 4.10: Acción del control del ε -equilibrio de Nash del jugador 2.

donde $u_1, u_2 \in R$ son las acciones de control para cada jugador.

Observe que los coeficientes de este juego satisfacen la condición del corolario 4.2, y la desigualdad toma la forma

$$-\sqrt{(2+x)^2 - 24} < 4 + 5x + 6x^2,$$

que es satisfecha para toda $x \in R$. Por lo tanto el eigenvalor con parte real negativa es igual a

$$\lambda_2(t) = -\frac{1}{2}\sqrt{(4x^2 + 3x + 2)^2 + 8} - \frac{1}{2}(x + 2x^2),$$

y las acciones de control para cada jugador son

$$u_i(t) = -\frac{1}{4}(2 + 3x + 4x^2 + \sqrt{(4x^2 + 3x + 2)^2 + 8}), \quad i = 1, 2.$$

Aplicando los controles al juego polinomial

$$\dot{x} = -\frac{1}{2}x(x + 2x^2 + \sqrt{(4x^2 + 3x + 2)^2 + 8}),$$

asegura la convergencia al único punto de equilibrios $x = 0$.

4.4.5. EJEMPLO 5.

Para el siguiente juego no lineal multi estado con dos jugadores

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_1^2 + u_1 + u_2, \end{aligned}$$

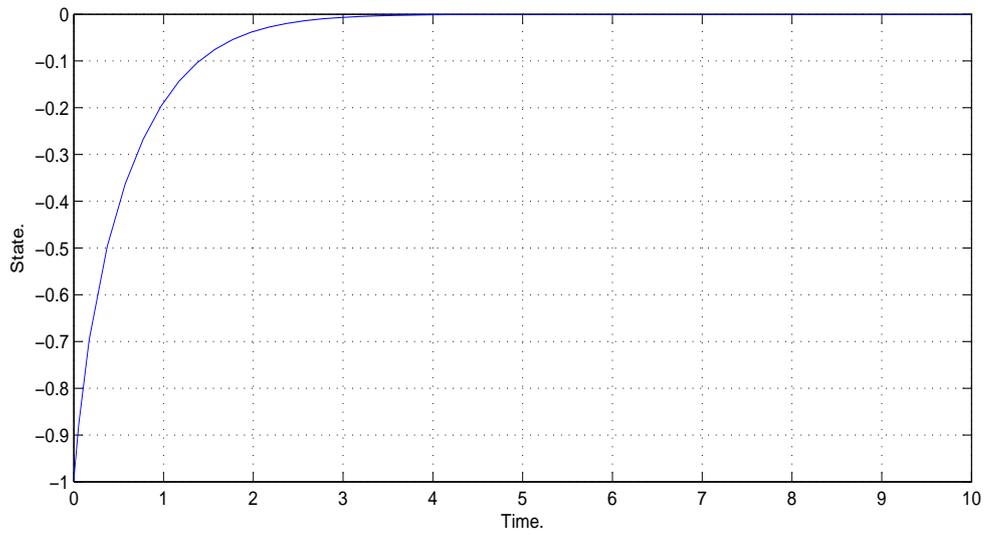


Figura 4.11: Estado del juego en horizonte de tiempo infinito. Ejemplo 4.

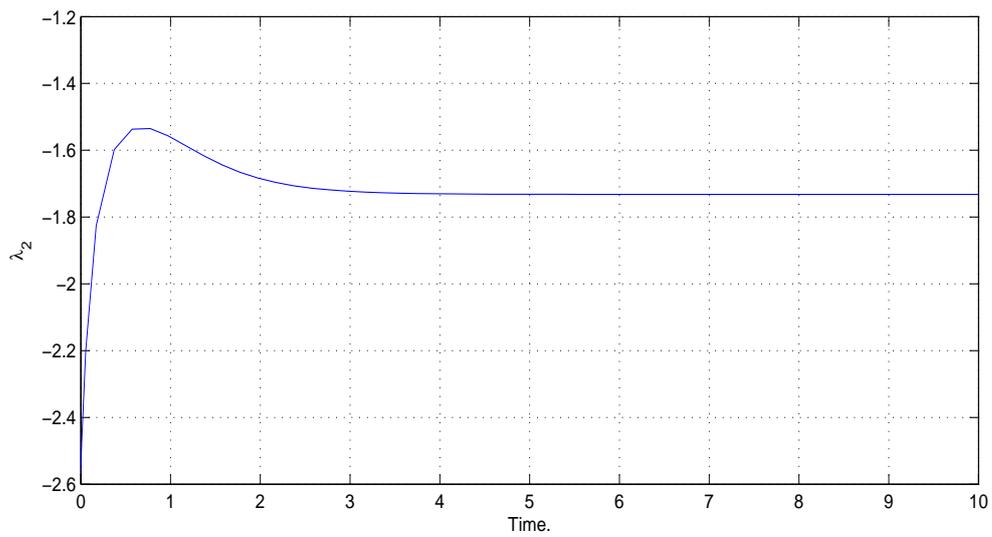


Figura 4.12: Eigenvalor dependiente del estado. Ejemplo 4.

cada jugador tiene la misma función de costo

$$J_i(u_1, u_2) = \frac{1}{2} \int_0^{\infty} (x_1^2 + x_2^2 + u_1^2 + u_2^2) dt.$$

La matriz Hamiltoniana para este juego es

$$H = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ x_1 & 0 & 0 & -1 & 0 & -1 \\ -1 & 0 & 0 & -2x_1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & -2x_1 \\ 0 & -1 & 0 & 0 & -1 & 0 \end{pmatrix}$$

Las soluciones para la ecuación tipo Riccati dependiente del estado

$$P_{21}^i = \frac{1}{4} \left(-2x_1 - \sqrt{3x_1 - \sqrt{x_1(x_1 + 12) - 4} + 2} + 2\sqrt{3x_1 + \sqrt{x_1(x_1 + 12) - 4} + 2} \right),$$

$$P_{22}^i = -\frac{\sqrt{3x_1 - \sqrt{x_1(x_1 + 12) - 4} + 2} + \sqrt{3x_1 + \sqrt{x_1(x_1 + 12) - 4} + 2}}{2\sqrt{2}},$$

y los controles del ε -equilibrio de Nash son

$$u_i = -P_{21}^i x_1 - P_{22}^i x_2.$$

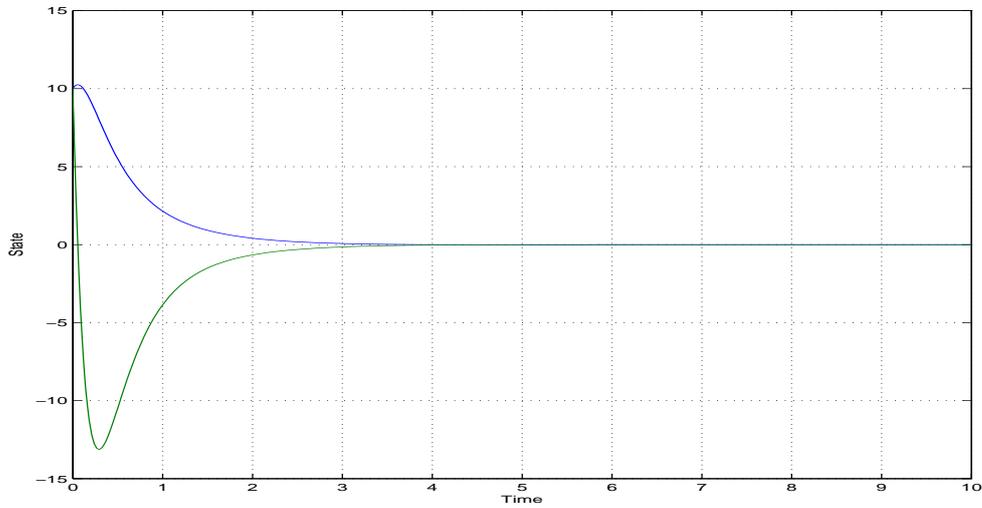


Figura 4.13: Trayectoria del estado del juego del ejemplo 4.

4.5. APÉNDICE

Demostración del teorema 4.1. Considerando la desigualdad (4.9) con $\varepsilon = 0$, esto es un equilibrio de Nash puro, donde la acción de control u^{i*} realiza el ínfimo del costo $L_T^i(u^{i*}, u^{\hat{i}*})$ con el control admisible en la region

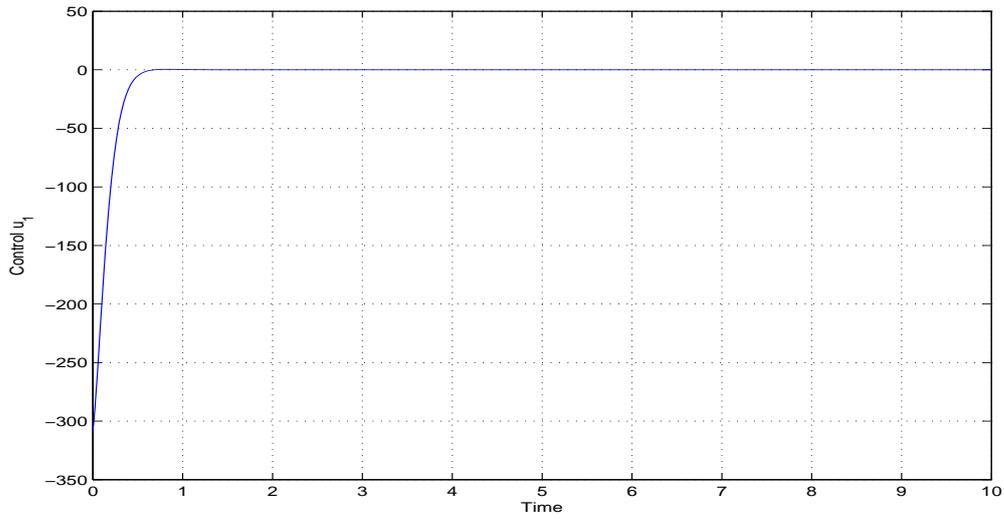


Figura 4.14: Acción del control del ε -equilibrio de Nash del jugador 1.

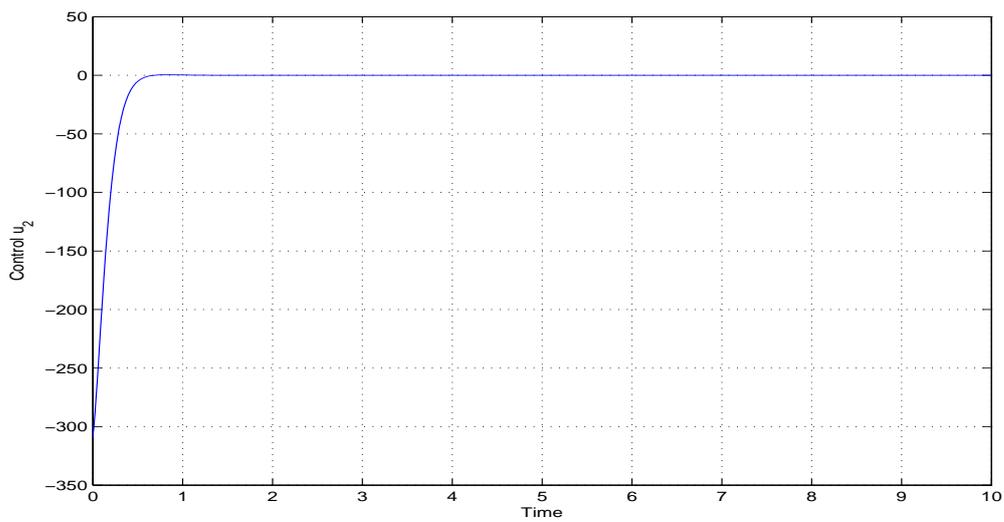


Figura 4.15: Acción del control del ε -equilibrio de Nash del jugador 2.

U^i , sujeto a

$$\dot{x}(t) = f(t, x) + \sum_{\substack{j=1 \\ j \neq i}}^N B^j(t)u^j(t) + B^i(t)u^i(t) + d(t),$$

$$x(t_0) = x_0,$$

donde se considera que todas las acciones de control ($j = \overline{1, N}$) ($j \neq i$), están fijas en sus controles del equilibrio de Nash, por lo tanto esto se convierte en un problema de control de un jugador (control óptimo) y las condiciones se siguen directamente del teorema del principio de Pontryagin.

Demostración del teorema 4.2. Dada la función Hamiltoniana individual (5.13) para cada i -jugador, la variable individual del co-estado $\psi^i(t)$ satisface la ecuación (4.2.1) con la condición terminal $\psi^i(T) = Q_f^i x(T)$. La condición (4.14) da los controles en lazo abierto:

$$u^{i*}(t) = -(R^{ii})^{-1} B^{iT} \psi^i(t).$$

Suponiendo que cada co-estado toma la forma lineal retroalimentada: $\psi^i(t) = P^i(t)x(t) + p^i(t)$, la ley de control se convierte:

$$u^{i*}(t) = -(R^{ii})^{-1} B^{iT} [P^i(t)x(t) + p^i(t)].$$

Tomando la derivada temporal de $\psi^i(t)$ conduce a:

$$\begin{aligned} & \dot{P}^i(t)x(t) + P^i(t)\dot{x}(t) + \dot{p}^i(t) = \dot{P}^i(t)x(t) + \dot{p}^i(t) \\ & + P^i(t) \left(a_1 x(t) + \dots + a_n \underbrace{x(t) * \dots * x(t)}_{n\text{-times}} - \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} [P^j(t)x(t) + p^j(t)] + d(t) \right) \\ & = - \left(a_1 + 2a_2 x(t) \dots + n a_n \underbrace{x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right) [P^i(t)x(t) + p^i(t)] - Q^i x(t), \end{aligned}$$

y se obtiene:

$$\begin{aligned} & \left[\dot{P}^i(t) + P^i(t) \left(a_1 + a_2 x(t) \dots + a_n \underbrace{x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right) + \right. \\ & \left. + \left(a_1 + 2a_2 x(t) \dots + n a_n \underbrace{x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right) P^i(t) + \right. \\ & \left. + Q^i - P^i(t) \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} P^j(t) \right] x(t) + \dot{p}^i(t) + P^i(t) d(t) + \\ & \left(a_1 + 2a_2 x(t) \dots + n a_n \underbrace{x(t) * \dots * x(t)}_{(n-1)\text{-times}} \right)^T p^i(t) - P^i(t) \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} p^j(t) = 0, \end{aligned}$$

que se cumple si (4.21) y (4.22) son satisfechas. La condición terminal es:

$$\begin{aligned} P^i(t) &= Q_f^i, \\ p^i(t) &= 0. \end{aligned}$$

Para obtener la cota superior para (4.25), se usa la definición del equilibrio de Nash:

$$L^i(u^{i*}, u^{\hat{i}*}) \leq L^i(u^i, u^{\hat{i}*}).$$

Recuerde que este costo es evaluado considerando los controles exactos en lazo abierto $u^{i*}(t) = -(R^{ii})^{-1} B^{iT} \psi^i(t)$, y u^i es un control admisible. Sumando y restando $\hat{L}^i(u^i, u^{\hat{i}*})$ y $\hat{L}^i(u^{i*}, u^{\hat{i}*})$ (definido en (4.10)), al lado derecho obteniendo:

$$L^i(u^{i*}, u^{\hat{i}*}) \leq L^i(u^i, u^{\hat{i}*}) + \hat{L}^i(u^{i*}, u^{\hat{i}*}) - \hat{L}^i(u^{i*}, u^{\hat{i}*}) + \hat{L}^i(u^i, u^{\hat{i}*}) - \hat{L}^i(u^i, u^{\hat{i}*}),$$

que puede ser representado como

$$\hat{L}^i(u^{i*}, u^{\hat{i}*}) \leq \hat{L}^i(u^i, u^{\hat{i}*}) + \varepsilon,$$

donde

$$\varepsilon = \hat{L}^i(u^{i*}, u^{\hat{i}*}) - L^i(u^{i*}, u^{\hat{i}*}) + L^i(u^i, u^{\hat{i}*}) - \hat{L}^i(u^i, u^{\hat{i}*}).$$

Ahora es necesario encontrar cotas para $\hat{L}^i(u^{i*}, u^{\hat{i}*}) - L^i(u^{i*}, u^{\hat{i}*})$ y $L^i(u^i, u^{\hat{i}*}) - \hat{L}^i(u^i, u^{\hat{i}*})$ para obtener una cota superior para ε . Para este propósito, procedemos como sigue:

$$\begin{aligned} \hat{L}^i(u^{i*}, u^{\hat{i}*}) - L^i(u^{i*}, u^{\hat{i}*}) &= \frac{1}{2} \tilde{x}^T(T) Q_f^i \tilde{x}(T) - \frac{1}{2} x^T(T) Q_f^i x(T) \\ &\quad + \frac{1}{2} \int_0^T (\tilde{x}^T Q^i \tilde{x} + \\ &\quad \sum_{j=1}^N [P^j \tilde{x} + p^j]^T B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} [P^j \tilde{x} + p^j]) dt \\ &\quad - \frac{1}{2} \int_0^T \left(x^T Q^i x + \sum_{j=1}^N \psi^{jT} B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} \psi^j \right) dt \\ &= \frac{1}{2} \tilde{x}^T(T) Q_f^i \tilde{x}(T) - \frac{1}{2} x^T(T) Q_f^i x(T) \\ &\quad + \frac{1}{2} \int_0^T (\tilde{x}^T \tilde{Q}^i \tilde{x} + L^i + 2\tilde{M}^i \tilde{x}) dt \\ &\quad - \frac{1}{2} \int_0^T \left(x^T Q^i x + \sum_{j=1}^N \psi^{jT} B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} \psi^j \right) dt, \end{aligned}$$

donde las matrices \tilde{Q}^i , L^i y \tilde{M}^i son definidas en (4.26). Teniendo en cuenta (4.17) y (4.18), obtenemos

$$\begin{aligned} \hat{L}^i(u^{i*}, u^{\hat{i}*}) - L^i(u^{i*}, u^{\hat{i}*}) &\leq \frac{1}{2} (\lambda_{\max} \{Q_f^i\} - \lambda_{\min} \{Q_f^i\}) \max \{\gamma_1, \gamma_2\} \\ &\quad + \frac{1}{2} (\lambda_{\max} \{\tilde{Q}^i\} - \lambda_{\min} \{Q^i\}) \max \{\beta_1, \beta_2\} + \frac{1}{2} \lambda_{\max} \{L^i\} \\ &\quad + \lambda_{\max} \{\tilde{M}^i\} \sqrt{\beta_1} + \xi_1. \end{aligned} \tag{4.35}$$

Por otro lado,

$$\begin{aligned} L^i(u^i, u^{\hat{i}*}) - \hat{L}^i(u^i, u^{\hat{i}*}) &= \frac{1}{2} \hat{x}^T(T) Q_f^i \hat{x}(T) - \frac{1}{2} \bar{x}^T(T) Q_f^i \bar{x}(T) \\ &\quad - \frac{1}{2} \int_0^T (\bar{x}^T Q^i \bar{x} + u^{iT} R^{ii} u^i + \\ &\quad \sum_{j=1, j \neq i}^N [P^j \bar{x} + p^j]^T B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} [P^j \bar{x} + p^j]) dt \\ &\quad + \frac{1}{2} \int_0^T \left(\hat{x}^T Q^i \hat{x} + u^{iT} R^{ii} u^i + \sum_{j=1, j \neq i}^N \psi^{jT} B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} \psi^j \right) dt \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \hat{x}^T(T) Q_f^i \hat{x}(T) - \frac{1}{2} \bar{x}^T(T) Q_f^i \bar{x}(T) \\
&\quad + \frac{1}{2} \int_0^T (\bar{x}^T \bar{Q}^i \bar{x} + F^i + 2\bar{M}^i \bar{x}) dt \\
&\quad - \frac{1}{2} \int_0^T \left(\hat{x}^T Q^i \hat{x} + \sum_{j=1, j \neq i}^N \psi^{jT} B^j (R^{jj})^{-1} R^{ij} (R^{jj})^{-1} B^{jT} \psi^j \right) dt,
\end{aligned}$$

donde \bar{Q} , \bar{Q}^i y F^i es definida en (4.26). Teniendo en cuenta (4.20) y (4.19), obtenemos

$$\begin{aligned}
L^i(u^i, u^{\hat{i}^*}) - \hat{L}^i(u^i, u^{\hat{i}^*}) &\leq \frac{1}{2} (\lambda_{\max}\{Q_f^i\} - \lambda_{\min}\{Q_f^i\}) \max\{\gamma_3, \gamma_4\} \\
&\quad + \frac{1}{2} (\lambda_{\max}\{Q^i\} - \lambda_{\min}\{\bar{Q}^i\}) \max\{\beta_3, \beta_4\} + \frac{1}{2} \lambda_{\max}\{F^i\} \\
&\quad + \lambda_{\max}\{\bar{M}^i\} \sqrt{\beta_3} + \xi_2.
\end{aligned} \tag{4.36}$$

Finalmente, (4.25) se sigue de (4.35) y (4.36).

Demostración del teorema 4.3 EL Hamiltoniano para el sistema es:

$$H = \begin{pmatrix} a_3 x^2 + a_2 x + a_1 & -S_1 & -S_2 \\ -q_1 & -3a_3 x^2 - 2a_2 x - a_1 & 0 \\ -q_2 & 0 & -3a_3 x^2 - 2a_2 x - a_1 \end{pmatrix} \tag{4.37}$$

donde $S_1 = b_1^2 r_{11}^{-1}$ y $S_2 = b_2^2 r_{22}^{-1}$, tiene un subespacio invariante V como

$$V = \begin{pmatrix} X_0 \\ X_1 \\ X_2 \end{pmatrix} \tag{4.38}$$

que satisface $Re(\lambda) \leq 0$ para $\forall \lambda \in \sigma(H|_V)$ y X_0 es invertible. Los eigenvalores de H son

$$\lambda_1(x) = -x(3a_3 x + 2a_2) - a_1, \tag{4.39}$$

$$\lambda_2(x) = -\frac{\sqrt{r_{22}(r_{11}(4a_3 x^2 + 3a_2 x + 2a_1)^2 + 4b_1^2 q_1) + 4b_2^2 q_2 r_{11}}}{2\sqrt{r_{11}}\sqrt{r_{22}}} - \frac{1}{2}x(2a_3 x + a_2), \tag{4.40}$$

$$\lambda_3(x) = \frac{\sqrt{r_{22}(r_{11}(4a_3 x^2 + 3a_2 x + 2a_1)^2 + 4b_1^2 q_1) + 4b_2^2 q_2 r_{11}}}{2\sqrt{r_{11}}\sqrt{r_{22}}} - \frac{1}{2}x(2a_3 x + a_2). \tag{4.41}$$

Todos los eigenvalores dependen del estado, y los correspondientes eigenvectores son:

$$v_1(x) = \left(0, -\frac{b_2^2 r_{11}}{b_1^2 r_{22}}, 1 \right), \tag{4.42}$$

$$v_2(x) = \left(\frac{\sqrt{r_{22}(r_{11}(4a_3 x^2 + 3a_2 x + 2a_1)^2 + 4b_1^2 q_1) + 4b_2^2 q_2 r_{11}}}{\sqrt{r_{11}}\sqrt{r_{22}}} - 4a_3 x^2 - 3a_2 x - 2a_1, \frac{q_1}{q_2}, 1 \right), \tag{4.43}$$

$$v_3(x) = \left(-\frac{\sqrt{r_{22}(r_{11}(4a_3 x^2 + 3a_2 x + 2a_1)^2 + 4b_1^2 q_1) + 4b_2^2 q_2 r_{11}}}{\sqrt{r_{11}}\sqrt{r_{22}}} + 4a_3 x^2 + 3a_2 x + 2a_1, \frac{q_1}{q_2}, 1 \right). \tag{4.44}$$

Se puede observar que los eigenvalores $\lambda_1(x)$ y $\lambda_3(x)$, no satisfacen la condición de construcción del subespacio V (4.38) solo el eigenvalor $\lambda_2(x)$ puede ser candidato. Para este propósito los parámetros $a_1, a_2, a_3, b_1, b_2, q_1, q_2, r_{11}$ y $r_{22} > 0$ deberán satisfacer:

$$\lambda_2(x) = -\frac{\sqrt{r_{22}(r_{11}(4a_3x^2 + 3a_2x + 2a_1)^2 + 4b_1^2q_1) + 4b_2^2q_2r_{11}}}{2\sqrt{r_{11}}\sqrt{r_{22}}} - \frac{1}{2}x(2a_3x + a_2) < 0,$$

lo que conduce a

$$-\frac{\sqrt{r_{22}(r_{11}(4a_3x^2 + 3a_2x + 2a_1)^2 + 4b_1^2q_1) + 4b_2^2q_2r_{11}}}{\sqrt{r_{11}r_{22}}} < 2a_3x^2 + a_2x.$$

5

EQUILIBRIO ε -NASH EN LAZO ABIERTO SOBRE JUEGOS DIFERENCIALES POLINOMIALES ESTOCÁSTICOS CON INFORMACIÓN INCOMPLETA.

5.1. PLANTEAMIENTO DEL PROBLEMA.

Sea (Ω, F, P) un espacio probabilístico completo con una familia de σ -álgebras con parte derecha creciente $F_t, t \geq t_0$, y sean $(W_1(t), F_t, t \geq t_0)$ y $(W_2(t), F_t, t \geq t_0)$ procesos estándar de Wiener independientes. El proceso aleatorio $(x(t), y(t))$ F_t -Medible representa el juego diferencial polinomial, donde la dinámica de los jugadores esta determinada por la ecuación diferencial

$$dx(t) = f(x, t)dt + \sum_{j=1}^N B^j(t)u^j(t)dt + b(t)dW_1(t), \quad x(t_0) = x_0, \quad (5.1)$$

y la ecuación diferencial lineal para el proceso de observación

$$dY(t) = A(t)x(t)dt + B(t)dW_2(t). \quad (5.2)$$

Aquí, $x(t) \in R^n$ es el vector del estado del juego, $u^j \in R^{m_j}$ es el control (acción) de cada j -ésimo jugador, j denota el numero del jugador ($j = \overline{1, N}$), y $y(t) \in R^m$ es el vector lineal de observación, $m \leq n$. La condición inicial $x_0 \in R^n$ es un vector Gaussiano tal que $x_0, W_1(t) \in R^p$, y $W_2(t) \in R^q$ son independientes. La matriz de observación $A(t) \in R^{m \times n}$ no se supone que sea no singular o cuadrada. Se asume que $b(t)b^T(t)$ es una matriz

definida positiva, por consiguiente, $m \leq q$. Todos los coeficientes en (5.1)–(5.2) son funciones determinísticas de dimensiones apropiadas.

La función no lineal $f(x, t)$ es considerada polinomial de n variables, componentes del vector del estado $x(t) \in R^n$, con coeficientes dependientes del tiempo. Ya que $x(t) \in R^n$ es un vector, se requiere una definición especial del polinomio para $n > 1$. Un polinomio de grado p del vector $x(t) \in R^n$ es considerado como una forma p -lineal de n componentes de $x(t)$

$$f(x, t) = a_0(t) + a_1(t)x(t) + a_2(t)xx^T + \dots + a_p(t)x \dots p \text{ times} \dots x,$$

donde $a_0(t)$ es un vector de dimension n , a_1 es una matriz de dimension $n \times n$, a_2 es un 3D tensor de dimension $n \times n \times n$, a_p es un $(p+1)$ D tensor de dimension $n \times \dots (p+1) \times \dots \times n$, y $x \times \dots p \text{ times} \dots \times x$ es un p D tensor de dimension $n \times \dots p \text{ times} \dots \times n$ obtenido por la multiplicación espacial de p veces del vector $x(t)$ por si mismo. Tal polinomio puede ser también expresado en la forma de sumatoria como

$$\begin{aligned} f_k(x, t) &= a_0 k(t) + \sum_i a_1 ki(t)x_i(t) + \sum_{ij} a_2 kij(t)x_i(t)x_j(t) \\ &+ \dots + \sum_{i_1 \dots i_p} a_p ki_1 \dots i_p(t)x_{i_1}(t) \dots x_{i_p}(t), \\ &k, i, j, i_1 \dots i_p = 1, \dots, n. \end{aligned}$$

Las ecuaciones de la dinámica del juego y del proceso de observación también se pueden reescribir en una forma alternativa

$$\dot{x}(t) = f(x, t) + \sum_{j=1}^N B^j(t)w^j(t) + b(t)\psi_1(t), \quad x(t_0) = x_0, \quad (5.3)$$

$$y(t) = A(t)x(t) + B(t)\psi_2(t), \quad (5.4)$$

donde $y(t) = \dot{Y}(t)$, y $\psi_1(t)$ y $\psi_2(t)$ son ruidos blancos Gaussianos, que son las derivadas en promedio cuadrático débil de los procesos estándar de Wiener $W_1(t)$ y $W_2(t)$ (ver [58]). Las representaciones (5.1),(5.2) y (5.3), (5.4) son equivalentes ([59]). Las ecuaciones (5.3), (5.4) presentan la forma convencional para las ecuaciones (5.1),(5.2), que son usadas actualmente en la practica.

En este capítulo, los criterios a minimizar son funcionales cuadráticas que representan los índices de desempeño individuales de cada jugador:

$$L_T^i(u^i, u^{\hat{i}}) = E \left\{ \frac{1}{2}x^T(T)Q_f^i x(T) + \frac{1}{2} \int_{t_0}^T \left(x^T(t)Q_1^i(t)x(t) + \sum_{j=1}^N u^{jT}(t)R^{ij}(t)u^j(t) \right) dt \right\}, \quad (5.5)$$

donde u^i es el control del i -ésimo jugador y $u^{\hat{i}}$ son los controles para el resto de los jugadores y (\hat{i} es la colección de contra-coalición de jugadores contrarrestando al jugador con índice i). Para cada jugador, el propósito del juego es obtener la minimización de su índice de desempeño por la selección de apropiadas entradas. También suponemos que:

$$\begin{aligned} Q^i(t) &= Q^{iT}(t) \geq 0, \quad Q_f^i = Q_f^{iT} \geq 0, \\ R^{ii}(t) &= R^{iiT}(t) > 0, \quad R^{ij}(t) = R^{ijT}(t) \geq 0, \quad i \neq j. \end{aligned} \quad (5.6)$$

5.2. DISEÑO DEL FILTRO.

Debido a la falta de información en la medición completa del estado, el estado polinomial inmensurable del juego $x(t)$, que satisface (5.3), es reemplazado con su estimado en promedio cuadrático $m(t)$ sobre observaciones lineales $y(t)$ (5.4), que es obtenido usando el filtro en promedio cuadrático para sistemas polinomiales (ver [60] para el planteamiento y solución del problema de filtrado correspondiente)

$$\begin{aligned} \dot{m}(t) &= E(f(x, t) | F_t^Y) + \sum_{j=1}^N B^j(t)w^j(t) \\ &+ P(t)A^T(t)(B(t)B^T(t))^{-1}(y(t) - A(t)m(t)), \end{aligned} \quad (5.7)$$

con la condición inicial $m(t_0) = E(x(t_0) | F_{t_0}^Y)$. La función matricial $P(t)$ satisface la ecuación matricial con coeficientes variantes en el tiempo

$$\begin{aligned} \dot{P}(t) &= E(f(x, t)(x(t) - m(t))^T | F_t^Y) \\ &+ E((x(t) - m(t))f(x, t)^T | F_t^Y) \\ &+ b(t)b^T(t) - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t), \end{aligned} \quad (5.8)$$

con la condición inicial $P(t_0) = E[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y]$

Recordar que $m(t)$ es el estimado en promedio cuadrático para el vector de estado del juego $x(t)$, basado en el proceso de observación $Y(t) = \{y(s), t_0 \leq s \leq t\}$, que minimiza la norma en promedio cuadrático

$$H = E[(x(t) - m(t))^T(x(t) - m(t)) | F_t^Y],$$

en cualquier momento del tiempo t . Aquí, $E[\xi(t) | F_t^Y]$ significa la esperanza condicional de un proceso estocástico $\xi(t) = (x(t) - m(t))^T(x(t) - m(t))$ con respecto a la σ -álgebra F_t^Y generada por el proceso de observación $Y(t)$ en el intervalo $[t_0, t]$. Como se sabe de [58], el estimado óptimo está dado por la esperanza condicional

$$m(t) = E(x(t) | F_t^Y),$$

del estado del juego $x(t)$ con respecto a la σ -álgebra F_t^Y generado por el proceso de observación $Y(t)$ en el intervalo $[t_0, t]$. Como siempre, la función matricial

$$P(t) = E[(x(t) - m(t))(x(t) - m(t))^T | F_t^Y],$$

es el estimado de la varianza del error.

Observación 5.1. Note que la ecuación (5.7) y (5.8) no tiene una forma cerrada el sistema de ecuaciones debido a la presencia del termino polinomial dependiente de x , $E(f(x, t) | F_t^Y)$, y $E(f(x, t)(x(t) - m(t))^T | F_t^Y)$, que no están expresadas aun como funciones de las variables del filtro, $m(t)$ y $P(t)$. Sin embargo, como se mostró en [60], la forma cerrada del sistema de ecuaciones de filtrado en promedio cuadrático puede ser obtenido para cualquier estado polinomial del juego (5.1) sobre observaciones lineales(5.2).

Se comprueba fácilmente que el problema del equilibrio de Nash en lazo abierto para el estado del sistema (5.1) y el índice de desempeño individual para cada jugador (5.5) es equivalente al problema del equilibrio de Nash para el estimado (5.7) y el índice de desempeño para cada jugador $L_T^i(u^i, u^{\hat{i}})$ representado como

$$L_T^i(u^i, u^{\hat{i}}) = E \left\{ \frac{1}{2} m^T(T) Q_f^i m(T) + \frac{1}{2} \int_{t_0}^T \left(m^T(t) Q_1^i(t) m(t) + \sum_{j=1}^N u^{jT}(t) R^{ij}(t) u^j(t) \right) dt + \frac{1}{2} \int_{t_0}^T \text{tr}[P(t) Q_1^i] dt + \frac{1}{2} \text{tr}[P(T)(T) Q_f^i] \right\}, \quad (5.9)$$

donde $\text{tr}[A]$ denota la traza de la matriz A . Dado que la última parte de $L_T^i(u^i, u^{\hat{i}})$ no depende del control (acción) de cada j -ésimo jugador $u^j \in R^{m_j}$ o del estado del juego $x(t)$, la función de costo efectiva reducida $M_T^i(u^i, u^{\hat{i}})$ para ser minimizada toma la forma

$$M_T^i(u^i, u^{\hat{i}}) = E \left\{ \frac{1}{2} m^T(T) Q_f^i m(T) + \frac{1}{2} \int_{t_0}^T \left(m^T(t) Q_1^i(t) m(t) + \sum_{j=1}^N u^{jT}(t) R^{ij}(t) u^j(t) \right) dt \right\}, \quad (5.10)$$

Así, la solución del problema del equilibrio de Nash en lazo abierto dado por (5.1),(5.5) puede ser encontrado resolviendo el problema del equilibrio de Nash en lazo abierto dado por (5.7),(5.10). Finalmente, el valor mínimo del criterio J_1 debería ser representado usando (5.9).

5.3. EQUILIBRIO ε -NASH EN LAZO ABIERTO: DISEÑO DEL CONTROLADOR CON HORIZONTE DE TIEMPO FINITO

5.3.1. DISEÑO DEL CONTROLADOR ÓPTIMO.

La solución óptima para el problema de buscar un conjunto de estrategias del equilibrio de Nash en el juego definido por (5.7),(5.10) está dado por en el siguiente teorema.

Teorema 5.1 El conjunto de estrategias del equilibrio de Nash para un juego diferencial polinomial (5.1) sobre observaciones lineales (5.2) con respecto a un criterio cuadrático (5.5) está dado por la ley de control para cada jugador

$$u^{i*}(t) = -(R^{ii})^{-1} B^{i\top} p^i(t), \quad (5.11)$$

donde la función $p^i(t)$ es la solución de la ecuación de co-estado

$$\begin{aligned} \dot{p}^i(t) &= -\frac{\partial H^i(t, m, u^{i*}, u^{\hat{i}*}, p^i)}{\partial x} \\ &= -Q^i m(t) - (a_1 + \dots + n a_n m(t) \times \dots (n-1) \text{times} \dots \times m(t))^T p^i \end{aligned} \quad (5.12)$$

con la condición terminal $p^i(T) = Q_f^i m(T)$, donde $H^i(t, m, u^{i*}, u^{\hat{i}*}, p^i)$ es el Hamiltoniano individual definido

como:

$$\begin{aligned}
H^i(t, m, u^{i*}, \hat{u}^{i*}, p^i) &= \\
&= \frac{1}{2} \left(m^T(t) Q^i m(t) + \sum_{j=1}^N u^{jT}(t) R^{ij} u^j(t) \right) + \\
&+ p^{iT} (a_1 m(t) + \dots + a_n m(t) \times \dots (n) \text{ times } \dots \times m(t) + \\
&\quad + \sum_{j=1}^N B^j u^j(t)), \quad i = 1, \dots, N.
\end{aligned} \tag{5.13}$$

Sustituyendo el conjunto de estrategias del Equilibrio de Nash anteriores (5.11) en la ecuación (5.7), se obtiene la ecuación del estimado del estado controlado óptimamente

$$\begin{aligned}
\dot{m}(t) &= E(f(x, t) | F_t^Y) + \sum_{j=1}^N B^j(t) (R^{jj})^{-1} B^{jT} p^j(t) \\
&+ P(t) A^T(t) (B(t) B^T(t))^{-1} (y(t) - A(t) m(t)),
\end{aligned} \tag{5.14}$$

con la condición inicial $m(t_0) = E(x(t_0) | F_{t_0}^Y)$.

Como se menciona en la Observación 4.1, las ecuaciones obtenidas (5.7), (5.8), (5.11)–(5.12), no presentaran una forma cerrada explícita del sistema de ecuaciones, mientras una forma específica del polinomio $f(x, t)$ en (5.1) no sea dada. Por lo tanto, en la siguiente sección, una forma cerrada de las ecuaciones de filtrado es obtenida en (5.7) y (5.8) para una función de tercer orden $f(x, t)$ en la ecuación (5.1), como sigue. Hay que señalar, sin embargo, que la aplicación del mismo procedimiento resultaría en el diseño de un sistema cerrado de las ecuaciones de filtrado para cualquier función polinomial $f(x, t)$ en (5.1).

Diseño del controlador para un estado polinomial de Tercer orden. I

Sea la función

$$f(x, t) = a_0 + a_1(t)x + a_2(t)xx^T + a_3(t)xxx^T, \tag{5.15}$$

un polinomio de tercer orden, donde x es un vector n -dimensional, $a_0(t)$ es un vector de dimension n , a_1 es una matriz de dimension $n \times n$, a_2 es un 3D tensor de dimension $n \times n \times n$, a_3 es un 4D tensor de dimension $n \times n \times n \times n$. En este caso, las ecuaciones de filtrado para el estimado óptimo $m(t)$, la matriz de ganancia $P(t)$, y la función de co-estado $p(t)$ para cada jugador son explícitamente obtenidas

$$\begin{aligned}
\dot{m}(t) &= a_0 + a_1 m(t) + a_2 m(t) m(t)^T + a_2 P(t) \\
&\quad + 3a_3 P(t) m(t) + a_3 m(t) m(t) m(t)^T \\
&\quad + \sum_{j=1}^N B^j(t) (R^{jj})^{-1} B^{jT} p^j(t) \\
&\quad + P(t) A^T(t) (B(t) B^T(t))^{-1} (y(t) - A(t) m(t)),
\end{aligned} \tag{5.16}$$

con la condición inicial $m(t_0) = E(x(t_0) | F_{t_0}^Y)$,

$$\begin{aligned}
\dot{P}(t) &= a(t)P(t) + P(t)a^T(t) + 2a_2(t)m(t)P(t) \\
&+ 2(a_2(t)m(t)P(t))^T + 3(a_3(P(t)P(t) + m(t)m^T(t)P(t))) \\
&\quad + 3(a_3(P(t)P(t) + m(t)m^T(t)P(t)))^T + b(t)b^T(t) \\
&\quad - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t),
\end{aligned} \tag{5.17}$$

con la condición inicial $P(t_0) = E[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y]$, y

$$\dot{p}^i(t) = -Q^i m(t) - (a_1 + 2a_2 m(t) + 3a_3 m(t)m(t))^T p^i, \quad (5.18)$$

con la condición terminal $p^i(T) = Q_f^i x(T)$. Consecuentemente, el resultado obtenido es formulado en el siguiente teorema.

Teorema 5.2 El equilibrio de Nash en lazo abierto para un juego diferencial estocástico polinomial de tercer orden (5.1) sobre observaciones lineales (5.2) con respecto a un criterio cuadrático (5.5) esta dado por la ley de control para cada jugador (5.11), la ecuación (5.16) para el estimado $m(t) = E(x(t) | F_t^Y)$, la ecuación (5.17) para la matriz de la varianza del error $P(t)$, y la ecuación (5.18) para la matriz de ganancia del control $p(t)$.

5.3.2. DISEÑO DE LA ESTRATEGIAS DEL EQUILIBRIO ε -NASH

El siguiente teorema usa las técnicas para la ecuación tipo Riccati dependiente del estado para presentar la solución en lazo abierto en una forma retroalimentada.

Teorema 5.3 La solución sub-óptima del equilibrio de Nash en forma retroalimentada para el juego diferencial polinomial (5.1) sobre observaciones lineales (5.2) con respecto al criterio cuadrático (5.5) esta dado por la ley de control para cada jugador

$$u^{i*}(t) = -(R^{ii})^{-1} B^{iT} [M^i(t)m(t)], \quad (5.19)$$

donde la función $M^i(t)$ es la solución de la ecuación acoplada tipo Riccati dependiente del estado

$$\begin{aligned} \dot{M}^i = & - \left(M^i (a_1 + \dots + a_n m(t) \times \dots_{(n-1)\text{times}} \dots \times m(t)) \right. \\ & + (a_1 + \dots + n a_n m(t) \times \dots_{(n-1)\text{times}} \dots \times m(t))^T M^i \\ & \left. - M^i \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} M^j + Q^i \right), \end{aligned} \quad (5.20)$$

con la condición terminal $M^i(T) = Q_f^i$.

Sustituyendo las estrategias sub-óptimas del equilibrio de Nash (5.19) en la ecuación (5.7), la siguiente ecuación del estimado del estado controlado del juego es obtenida

$$\begin{aligned} \dot{m}(t) = & E(f(x, t) | F_t^Y) + \sum_{j=1}^N B^j(t) (R^{jj})^{-1} B^{jT} [M^j(t)m(t)] \\ & + P(t) A^T(t) (B(t) B^T(t))^{-1} (y(t) - A(t)m(t)), \end{aligned} \quad (5.21)$$

con la condición inicial $m(t_0) = E(x(t_0) | F_{t_0}^Y)$.

Como se menciona en la Observación 4.1, las ecuaciones obtenida (5.7), (5.8), (5.19)–(5.20), no presentan una forma cerrada explícita del sistema de ecuaciones, mientras una forma específica del polinomio $f(x, t)$ en (5.1) no sea dada. Por lo tanto, en la siguiente sub-sección, una forma cerrada del sistema de filtrado es obtenida en (5.7) y (5.8) para una función de tercer orden $f(x, t)$ en la ecuación (5.1).

Diseño del controlador para un estado polinomial de Tercer orden. II

Sea la función

$$f(x, t) = a_0 + a_1(t)x + a_2(t)xx^T + a_3(t)xxx^T, \quad (5.22)$$

un polinomio de tercer orden. En este caso, las siguientes ecuaciones de filtrado para el estimado sub-óptimo $m(t)$, la matriz de ganancia del filtro $P(t)$ y la ecuación acoplada tipo Riccati dependiente del estado para cada jugador son obtenidas explícitamente

$$\begin{aligned}\dot{m}(t) = & a_0 + a_1 m(t) + a_2 m(t)m(t)^T + a_2 P(t) \\ & + 3a_3 P(t)m(t) + a_3 m(t)m(t)m(t)^T \\ & + \sum_{j=1}^N B^j(t)(R^{jj})^{-1} B^{jT} [M^i(t)m(t)] \\ & + P(t)A^T(t)(B(t)B^T(t))^{-1}(y(t) - A(t)m(t)),\end{aligned}\tag{5.23}$$

con la condición inicial $m(t_0) = E(x(t_0) | F_{t_0}^Y)$,

$$\begin{aligned}\dot{P}(t) = & a(t)P(t) + P(t)a^T(t) + 2a_2(t)m(t)P(t) \\ & + 2(a_2(t)m(t)P(t))^T + 3(a_3(P(t)P(t) + m(t)m^T(t)P(t))) \\ & + 3(a_3(P(t)P(t) + m(t)m^T(t)P(t)))^T + b(t)b^T(t) \\ & - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t),\end{aligned}\tag{5.24}$$

con la condición inicial $P(t_0) = E[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y]$.

$$\begin{aligned}\dot{M}^i = & - (M^i (a_1 + a_2 m(t) + a_3 m(t)m(t)) \\ & + (a_1 + 2a_2 m(t) + 3a_3 m(t)m(t))^T M^i \\ & - M^i \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} M^j + Q^i),\end{aligned}\tag{5.25}$$

con la condición terminal $M^i(T) = Q_f^i$. Consecuentemente, el resultado obtenido es formulado en el siguiente teorema.

Teorema 5.4 La solución sub-óptima del equilibrio de Nash para un juego diferencial estocástico polinomial de tercer orden (5.1) sobre observaciones lineales (5.2) con respecto al criterio cuadrático (5.5) esta dado por la ley de control para cada jugador (5.19), la ecuación (5.23) para el estimado $m(t) = E(x(t) | F_t^Y)$, la ecuación (5.24) para la matriz de varianza del error $P(t)$, y la ecuación (5.25) para la matriz de ganancia del control $M(t)$.

5.3.3. EJEMPLO I.

Esta sub-sección presenta un ejemplo del diseño controlador óptimo y sub-óptimo para un juego diferencial estocástico polinomial, donde la dinámica de los jugadores es gobernada por la ecuación diferencial (5.1) sobre observaciones lineales (5.2) con un criterio cuadrático (5.5) para cada jugador, usando el controlador óptimo (5.11),(5.16)–(5.18), y comparándolo con el controlador sub-óptimo (5.19),(5.23)–(5.25).

Considere la ecuación del estado escalar de orden 2 del juego para dos jugadores

$$\dot{x}(t) = 0,1x^2(t) + u^1(t) + u^2(t) + \psi_1, \quad x(0) = 1,\tag{5.26}$$

y el proceso de observación lineal

$$y(t) = x(t) + \psi_2(t),\tag{5.27}$$

donde $\psi_1(t)$ y $\psi_2(t)$ son ruidos blancos Gaussianos, que son la derivada en promedio cuadrático débil de procesos de Wiener (ver [58]). Las ecuaciones (5.26),(5.27) corresponden a la forma alternativa convencional de (5.3),(5.4) para las ecuaciones (5.1),(5.2).

El problema del equilibrio de Nash es buscar un conjunto de estrategias $u^1(t), u^2(t), t \in [0, T], T = 1$, que minimiza el criterio para cada jugador

$$\begin{aligned} L_T^1(u^1, u^2) &= E \left\{ \frac{1}{2} \int_0^T (x^2 + (u^1)^2 + (u^2)^2) dt + 0,1 \frac{1}{2} x^2(T) \right\}, \\ L_T^2(u^1, u^2) &= E \left\{ \frac{1}{2} \int_0^T (0,1 x^2 + (u^1)^2 + (u^2)^2) dt + \frac{1}{2} x^2(T) \right\}. \end{aligned} \quad (5.28)$$

aplicando el controlador óptimo diseñado (5.11),(5.16)–(5.18), la ley de control (5.11) para cada jugador esta dado por

$$\begin{aligned} u^1(t) &= -p^1(t), \\ u^2(t) &= -p^2(t), \end{aligned} \quad (5.29)$$

donde $m(t)$ satisface la ecuacion

$$\begin{aligned} \dot{m}(t) &= 0,1m^2(t) + 0,1P(t) + u^1(t) + u^2(t) \\ &\quad + P(t)[y(t) - m(t)], \end{aligned} \quad (5.30)$$

con la condición inicial $m(0) = E(x(0) | y(0)) = m_0$,

$$\dot{P}(t) = 1 + 0,4m(t)P(t) - P^2(t), \quad (5.31)$$

con la condición inicial $P(0) = E((x(0) - m(0))(x(0) - m(0))^T | y(0))$, y

$$\begin{aligned} \dot{p}^1(t) &= -0,1m(t) - 0,2m(t)p^1(t), \\ \dot{p}^2(t) &= -1m(t) - 0,2m(t)p^2(t), \end{aligned} \quad (5.32)$$

con la condición terminal $p^1(T) = 1m(T)$ y $p^2(T) = 0,1m(T)$.

El control óptimo diseñado (5.11),(5.16)–(5.18), es comparado con el controlador sub-óptimo (5.19),(5.23)–(5.25), la ley de control(5.19) para cada jugador esta dado por

$$\begin{aligned} u^1(t) &= M^1(t)m(t), \\ u^2(t) &= M^2(t)m(t), \end{aligned} \quad (5.33)$$

donde $m(t)$ satisface la ecuación

$$\begin{aligned} \dot{m}(t) &= 0,1m^2(t) + 0,1P(t) + u^1(t) + u^2(t) \\ &\quad + P(t)[y(t) - m(t)], \end{aligned} \quad (5.34)$$

con la condición inicial $m(0) = E(x(0) | y(0)) = m_0$,

$$\dot{P}(t) = 1 + 0,4m(t)P(t) - P^2(t), \quad (5.35)$$

con la condición inicial $P(0) = E((x(0) - m(0))(x(0) - m(0))^T | y(0))$, y

$$\begin{aligned}\dot{M}^1(t) &= 0,1 - 0,3m(t)M^1(t) - M^1(t)M^2(t) - (M^1(t))^2, \\ \dot{M}^2(t) &= 1 - 0,3m(t)M^2(t) - M^1(t)M^2(t) - (M^2(t))^2,\end{aligned}\tag{5.36}$$

con la condición terminal $Q^1(T) = -1$ y $Q^2(T) = -0,1$.

El sistema obtenido (5.29)–(5.32) y (5.33)–(5.36) puede ser resuelto usando métodos numéricos simples, tales como el *método de disparo*. Este método consiste en variar la condición inicial de (5.32) y (5.36) hasta que la condición terminal dada sea satisfecha.

Para la simulación numérica del sistema (5.26),(5.27), el controlador óptimo (5.29)–(5.32) y el controlador sub-óptimo (5.33)–(5.36), los valores iniciales $x(0) = 1$, $m(0) = 10$, y $P(0) = 100$ son asignados. El momento final del tiempo es fijado en $T = 1$. Los disturbios $\psi_1(t)$ en (5.26) y $\psi_2(t)$ en (5.27) son realizados usando la función incorporada de ruidos blancos de Matlab.

5.4. EQUILIBRIO ε -NASH EN LAZO ABIERTO: DISEÑO DEL CONTROLADOR CON HORIZONTE DE TIEMPO INFINITO.

5.4.1. DISEÑO DEL CONTROLADOR SUB-ÓPTIMO.

En esta sección, se considera el problema de control con horizonte de tiempo infinito, esto es, el criterio que cada i -ésimo jugador desea minimizar esta dado por:

$$L_\infty^i(u_1, u_2) = E \left\{ \frac{1}{2} \int_0^\infty (x^T Q_i x + \sum_{j=1}^2 u_j^T R_{ij} u_j) dt \right\},\tag{5.37}$$

sujeto a la dinámica estocástica polinomial (5.1) sobre observaciones lineales (5.2). Así, funcionales de costo cuadráticos con $Q_f^i = 0$ es estudiado. En este caso, el sistema de ecuaciones algebraicas tipo Riccati dependientes del estimado del estado toma la forma

$$\begin{aligned}M^i & (a_1 + \dots + a_n m(t) \times \dots_{(n-1) \text{ times}} \dots \times m(t)) + \\ & (a_1 + \dots + a_n m(t) \times \dots_{(n-1) \text{ times}} \dots \times m(t))^T M^i \\ & + Q^i - M^i \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} M^j = 0, \quad i = 1, \dots, N.\end{aligned}\tag{5.38}$$

Teorema 5.5 La solución sub-óptima del equilibrio de Nash en lazo abierto en forma retroalimentada para un juego diferencial estocástico polinomial con horizonte de tiempo infinito (5.1) sobre observaciones lineales (5.2) con respecto a un criterio cuadrático (5.37) esta dado por la ley de control para cada jugador

$$u^{i*}(t) = -(R^{ii})^{-1} B^{iT} [M^i(t)m(t)],\tag{5.39}$$

donde la función $M^i(t)$ es la solución de la ecuación algebraica tipo Riccati dependiente del estimado del estado (5.38).

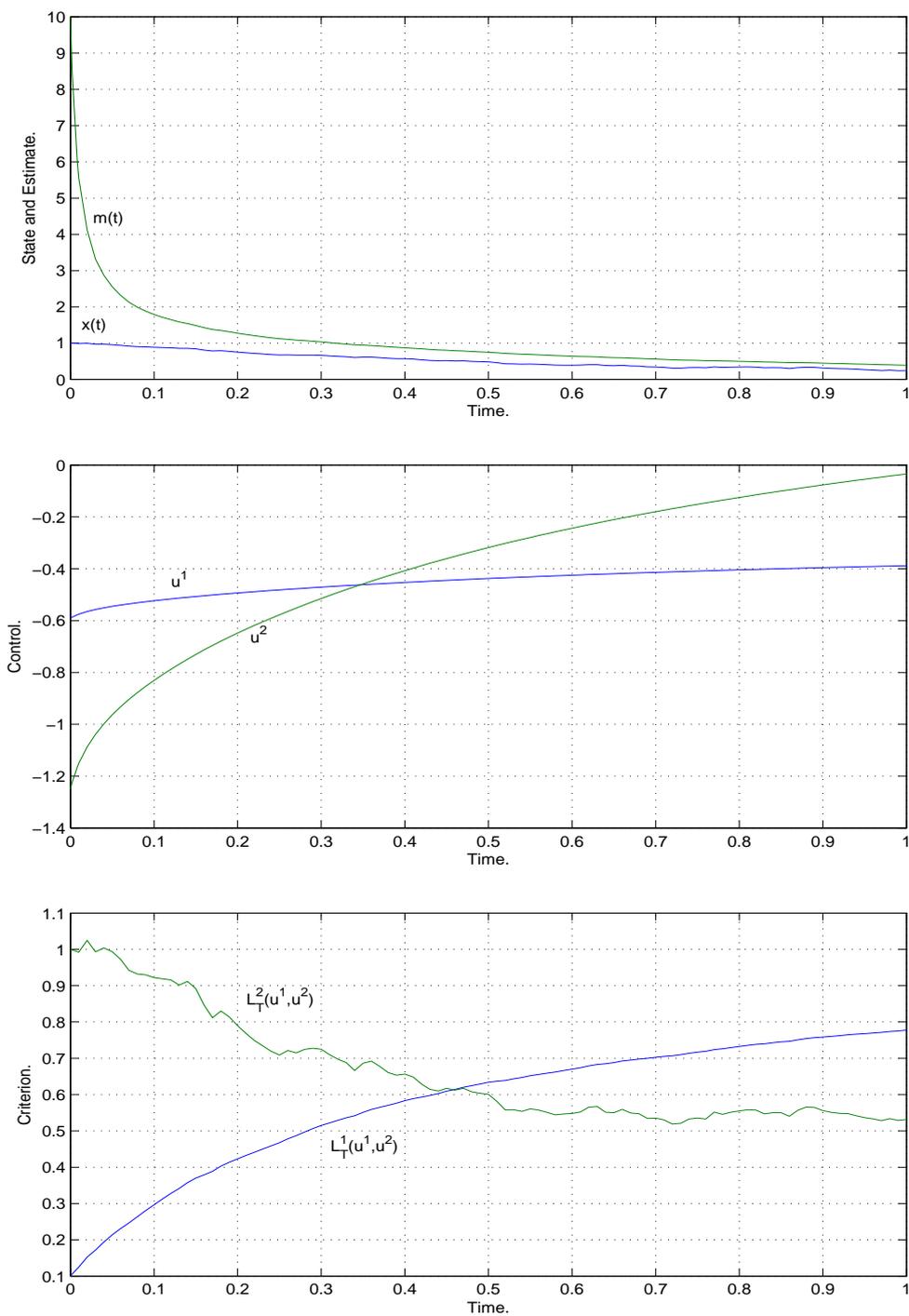


Figura 5.1: El controlador óptimo con respecto al criterio $L_T^1(u^1, u^2)$ y $L_T^2(u^1, u^2)$ en el intervalo entero de simulación $[0, 1]$. **1.** Gráfica del estado controlado (5.26) $x(t)$ (Azul) y el estimado (5.30) $m(t)$ (verde); **2.** Gráfica del control óptimo (5.29) $u^1(t)$ (Azul) y $u^2(t)$ (verde); **3.** Gráfica del criterio (5.28) $L_T^1(u^1, u^2)$ (azul) y $L_T^2(u^1, u^2)$ (verde) producidos por el controlador óptimo.

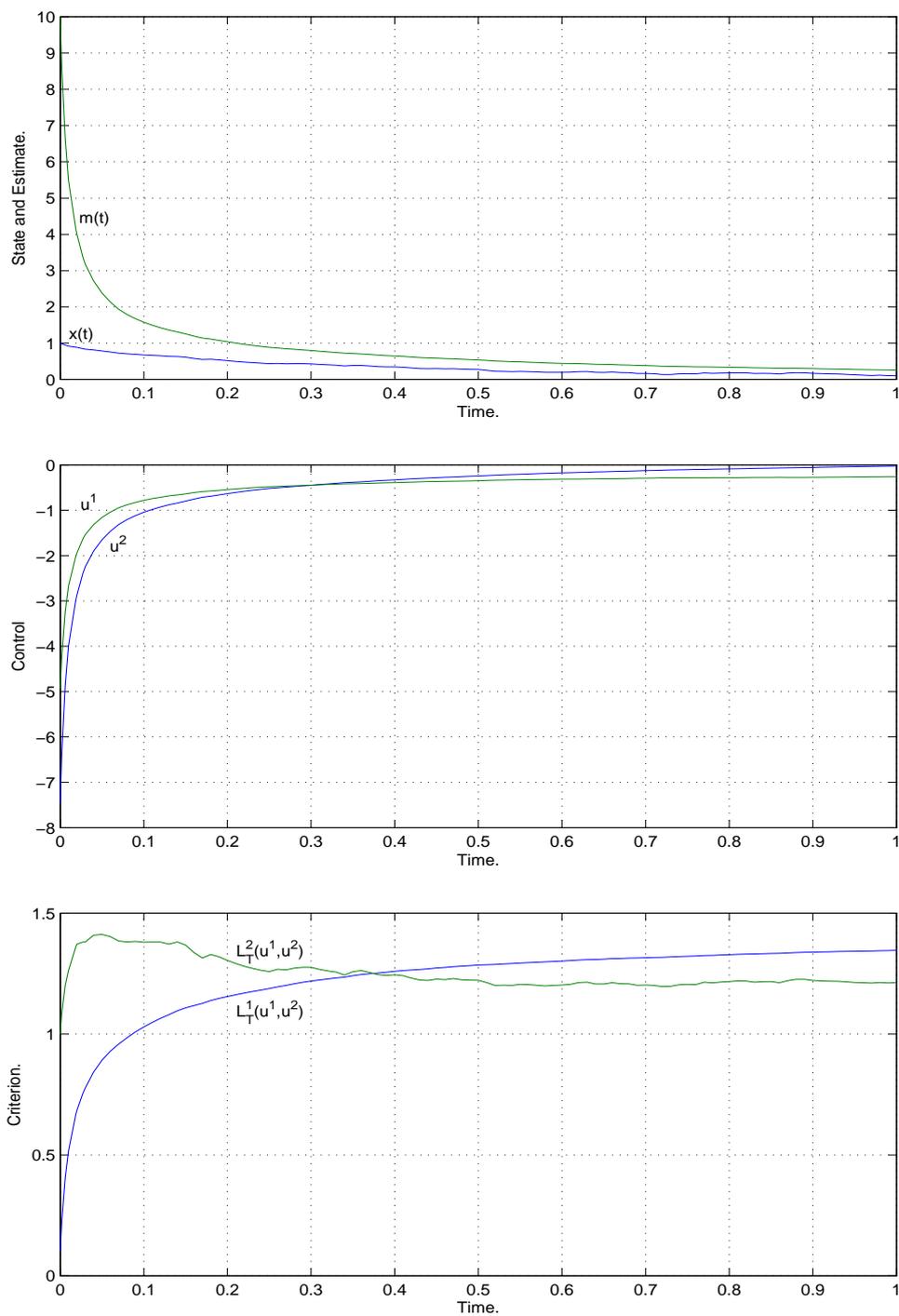


Figura 5.2: El controlador sub-óptimo con respecto al criterio $L_T^1(u^1, u^2)$ y $L_T^2(u^1, u^2)$ en el intervalo entero de simulación $[0, 1]$. **1.** Gráfica del estado controlado (5.26) $x(t)$ (Azul) y el estimado (5.34) $m(t)$ (verde); **2.** Gráfica del control óptimo (5.33) $u^1(t)$ (verde) y $u^2(t)$ (azul); **3.** Gráfica del criterio (5.28) $L_T^1(u^1, u^2)$ (azul) y $L_T^2(u^1, u^2)$ (verde) producidos por el controlador sub-óptimo.

Diseño del controlador para un estado polinomial de Tercer orden. III

Sea la función

$$f(x, t) = a_0 + a_1(t)x + a_2(t)xx^T + a_3(t)xxx^T, \quad (5.40)$$

un polinomio de tercer orden. En este caso, las siguientes ecuaciones de filtrado para el estimado sub-óptimo $m(t)$, la matriz de ganancia del filtro $P(t)$ y la ecuación algebraica acopladas tipo Riccati dependientes del estado para cada jugador son obtenidas explícitamente

$$\begin{aligned} \dot{m}(t) = & a_0 + a_1m(t) + a_2m(t)m(t)^T + a_2P(t) \\ & + 3a_3P(t)m(t) + a_3m(t)m(t)m(t)^T \\ & + \sum_{j=1}^N B^j(t)(R^{jj})^{-1}B^{jT} [M^i(t)m(t)] \\ & + P(t)A^T(t)(B(t)B^T(t))^{-1}(y(t) - A(t)m(t)), \end{aligned} \quad (5.41)$$

con la condición inicial $m(t_0) = E(x(t_0) | F_{t_0}^Y)$,

$$\begin{aligned} \dot{P}(t) = & a(t)P(t) + P(t)a^T(t) + 2a_2(t)m(t)P(t) \\ & + 2(a_2(t)m(t)P(t))^T + 3(a_3(P(t)P(t) + m(t)m^T(t)P(t))) \\ & + 3(a_3(P(t)P(t) + m(t)m^T(t)P(t)))^T + b(t)b^T(t) \\ & - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t), \end{aligned} \quad (5.42)$$

con la condición inicial $P(t_0) = E[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y]$.

$$\begin{aligned} 0 = & M^i (a_1 + a_2m(t) + a_3m(t)m(t)) \\ & + (a_1 + 2a_2m(t) + 3a_3m(t)m(t))^T M^i \\ & - M^i \sum_{j=1}^N B^j (R^{jj})^{-1} B^{jT} M^j + Q^i. \end{aligned} \quad (5.43)$$

Consecuentemente, el resultado obtenido es formulado en el siguiente teorema.

Teorema 5.6 La solución del equilibrio de Nash en lazo abierto para un juego diferencial estocástico de tercer orden con horizonte de tiempo infinito (5.1) sobre observaciones lineales (5.2) con respecto a un criterio cuadrático (5.37) esta dado por la ley de control para cada jugador (5.39), la ecuación (5.41) para el estimado $m(t) = E(x(t) | F_t^Y)$, la ecuación (5.42) para la matriz de varianza del error $P(t)$, y la ecuación (5.43) para la matriz de ganancia del control $M(t)$.

Observación 5.2 La principal diferencia de la ecuación tipo Riccati para el equilibrio de Nash estándar de (5.38) es que la solución de (5.38) depende del vector $m(t)$. En el caso de optimización de un jugador, lo que convencionalmente se considerada que se puede encontrar una solución para cada instante del tiempo, cuando los coeficientes dependientes del estimado de (5.38) se consideran constantes, y entonces se aplican los métodos estándar para resolver la ecuación algebraica tipo Riccati. Es posible aplicar tal aproximación para el caso de N -jugadores. En efecto, en vista de la estructura polinomial del juego, se pueden recrear la aproximación Hamiltoniana para buscar los controles sub-óptimos (5.39) en una forma cerrada, esto es, como función que dependen de los términos $m(t)$ del polinomio. Se desarrollara esta aproximación en la siguiente sub-sección.

5.4.2. RESOLVIENDO LA ECUACIÓN ALGEBRAICA ACOPLADA TIPO RICCATI DEPENDIENTE DEL ESTIMADO DEL ESTADO POR LA APROXIMACIÓN HAMILTONIANA

La solución para el sistema de ecuaciones (5.38) está muy relacionado (ver [57]) con la matriz Hamiltoniana dada por (para simplificar, se considera $N = 2$):

$$H = \begin{pmatrix} a_1 + \dots + a_n m^{n-1} & -S^1 & -S^2 \\ -Q^1 & -(a_1 + \dots + na_n m^{n-1})^\top & 0 \\ -Q^2 & 0 & -(a_1 + \dots + na_n m^{n-1})^\top \end{pmatrix}$$

donde $S^i = B^i (R^{ii})^{-1} B^{i\top}$, $i = 1, 2$. Si el vector m es fijo, la matriz H coincide con el caso lineal cuadrático. En efecto, Es conocido ([4] pág. 280) que para un vector fijo m si $V \subset \mathbb{R}^{3n \times n}$ es un espacio invariante n -dimensional de H , $W_i \in \mathbb{R}^{n \times n}$, $i = 1, 2$, son tres matrices reales tales que

$$V = Im \begin{bmatrix} W_0 \\ W_1 \\ W_2 \end{bmatrix},$$

y si W_0 es no singular, entonces $M_i := W_i W_0^{-1}$, $i = 1, 2$, es una solución para el sistema de ecuaciones acopladas tipo Riccati (5.38) (para $N = 2$). Por otro parte, las soluciones (M_1, M_2) son independientes de la base seleccionada V , que da un método para encontrar los controles para cada jugador. Como también se conoce en el caso estándar (no depende de m), las ecuaciones algebraicas acopladas tipo Riccati tiene solución estabilizante (M_1, M_2) si y solo si la matriz H tiene un subespacio n -dimensional estable y H tiene $2n$ eigenvalores. Se extenderá esta aproximación para el caso polinomial como sigue.

Como se a señalado, en vista de la estructura polinomial del juego, es posible encontrar controles sub-óptimos dependientes del estimado m o, equivalentemente, calculando los eigenvalores y eigenvectores de H , que dependen de m . Por este propósito, se define el siguiente algoritmo:

Algoritmo 1.- Suponga que m es fijo como un valor constante.

- 2.- Calcula los eigenvalores de H , que resultan ser dependientes de m .
- 3.- Para la selección particular de eigenvalores, calculamos los eigenvectores (también dependientes de m).
- 4.- Calcula el eigenespacio de dimension $(3 * n) \times n$, generado por los eigenvectore linealmente independientes, que esta dado por:

$$W = \begin{bmatrix} W_0 \\ W_1 \\ W_2 \end{bmatrix}.$$

- 5.- Las soluciones de las ecuaciones algebraicas tipo Riccati dependientes del estado (5.38) esta dadas por:

$$M_i = W_i W_0^{-1}. \quad (5.44)$$

Observación 5.3 El algoritmo desarrollado puede ser extendido al caso de N jugadores similarmente al caso lineal, como se describe en [4].

5.4.3. EJEMPLO II.

Esta sub-sección presenta un ejemplo del diseñado controlador sub-óptimo para un juego diferencial estocástico con horizonte de tiempo infinito, donde la dinámica de los jugadores es gobernada por la ecuación diferencial (5.1) sobre observaciones lineales (5.2) con un criterio cuadrático (5.37) para cada jugador, usando el controlador sub-óptimo (5.39),(5.41)–(5.43).

Considere el estado escalar de segundo orden del juego para dos jugadores:

$$\begin{aligned} \dot{x}_1(t) &= x_2(t), & x_1(0) &= 1, \\ \dot{x}_2(t) &= 0,1x_1^2(t) + u^1(t) + u^2(t) + \psi_1, & x_2(0) &= 1, \end{aligned} \quad (5.45)$$

y el proceso de observación lineal

$$y(t) = x_1(t) + \psi_2(t), \quad (5.46)$$

donde $\psi_1(t)$ y $\psi_2(t)$ son ruidos blancos Gaussianos.

El problema del equilibrio de Nash en lazo abierto es buscar un conjunto de estrategias $u^1(t)$, $u^2(t)$, $t \in [0, T]$, $T = 1$, que minimice el criterio para cada jugador:

$$\begin{aligned} L_\infty^1(u^1, u^2) &= E \left\{ \frac{1}{2} \int_0^\infty (x_1^2 + (u^1)^2) dt \right\}, \\ L_\infty^2(u^1, u^2) &= E \left\{ \frac{1}{2} \int_0^\infty (x_2^2 + (u^2)^2) dt \right\}. \end{aligned} \quad (5.47)$$

Aplicando el controlador sub-óptimo diseñado (5.39),(5.41)–(5.43), la ley de control (5.39) para cada jugador esta dado por

$$\begin{aligned} u^1(t) &= M_{21}^1(t)m_1(t) + M_{22}^1(t)m_2(t), \\ u^2(t) &= M_{21}^2(t)m_1(t) + M_{22}^2(t)m_2(t), \end{aligned} \quad (5.48)$$

donde $m(t)$ satisface la ecuación

$$\begin{aligned} \dot{m}_1(t) &= m_2(t) + P_{11}(t)[y(t) - m_1(t)], \\ \dot{m}_2(t) &= 0,1m_1^2(t) + 0,1P_{11}(t) + u^1(t) + u^2(t) \\ &\quad + P_{12}(t)[y(t) - m_1(t)], \end{aligned} \quad (5.49)$$

con la condición inicial $m(0) = E(x(0) | y(0)) = m_0$,

$$\begin{aligned} \dot{P}_{11}(t) &= 2P_{12}(t) - P_{11}^2(t), \\ \dot{P}_{12}(t) &= 0,2m_1(t)P_{11}(t) + P_{22}(t) - P_{11}P_{12}, \\ \dot{P}_{22}(t) &= 1 + 0,4m_1(t)P_{12}(t) - P_{12}^2(t), \end{aligned} \quad (5.50)$$

con la condición inicial $P(0) = E((x(0) - m(0))(x(0) - m(0))^T | y(0))$, y la matriz Hamiltoniana esta dada por

$$H = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0,1m_1(t) & 0 & 0 & -10 & 0 & -10 \\ 0 & 0 & 0 & -0,2m_1(t) & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & -0,2m_1(t) \\ 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \quad (5.51)$$

La solución de la ecuación algebraica tipo Riccati dependiente del estado (5.51) esta dada por:

$$M_{21}^1 = \frac{\left(\sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right) m_1^2}{200(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}$$

$$- \frac{\left(\sqrt{m_1^2 + 600m_1 + 6000} \sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} + 100 \sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}{200(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}$$

$$+ \frac{\left(\sqrt{m_1^2 + 600m_1 + 6000} \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - 100 \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right) m_1}{200(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}$$

$$+ \frac{1000 \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}{200(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)},$$

$$M_{22}^1 = - \frac{m_1 \sqrt{m_1^2 + 600m_1 + 6000}}{10\sqrt{5}(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)},$$

$$M_{21}^2 = \frac{m_1 \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}{40(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}$$

$$+ \frac{\sqrt{m_1^2 + 600m_1 + 6000} \sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - 100 \sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}}{40(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}$$

$$+ \frac{\sqrt{m_1^2 + 600m_1 + 6000} \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} + 100 \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}}{40(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)},$$

$$M_{22}^2 = \frac{\sqrt{m_1^2 + 600m_1 + 6000}}{2\sqrt{5}(m_1 - 5) \left(\sqrt{-\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100} - \sqrt{\sqrt{m_1^2 + 600m_1 + 6000} + 3m_1 + 100}\right)}.$$

Para la simulación numérica del sistema (5.45),(5.46), el controlador óptimo (5.48)–(5.51), la condición inicial $x_1(0) = 1$, $x_2(0) = 1$, $m_1(0) = 10$, $m_2(0) = 10$, $P_{11}(0) = 100$, $P_{12}(0) = 10$ y $P_{22}(0) = 100$ son asignados. Los disturbios $\psi_1(t)$ en (5.45) y $\psi_2(t)$ en (5.46) son realizados usando la funciones incorporadas de ruidos blancos de Matlab.

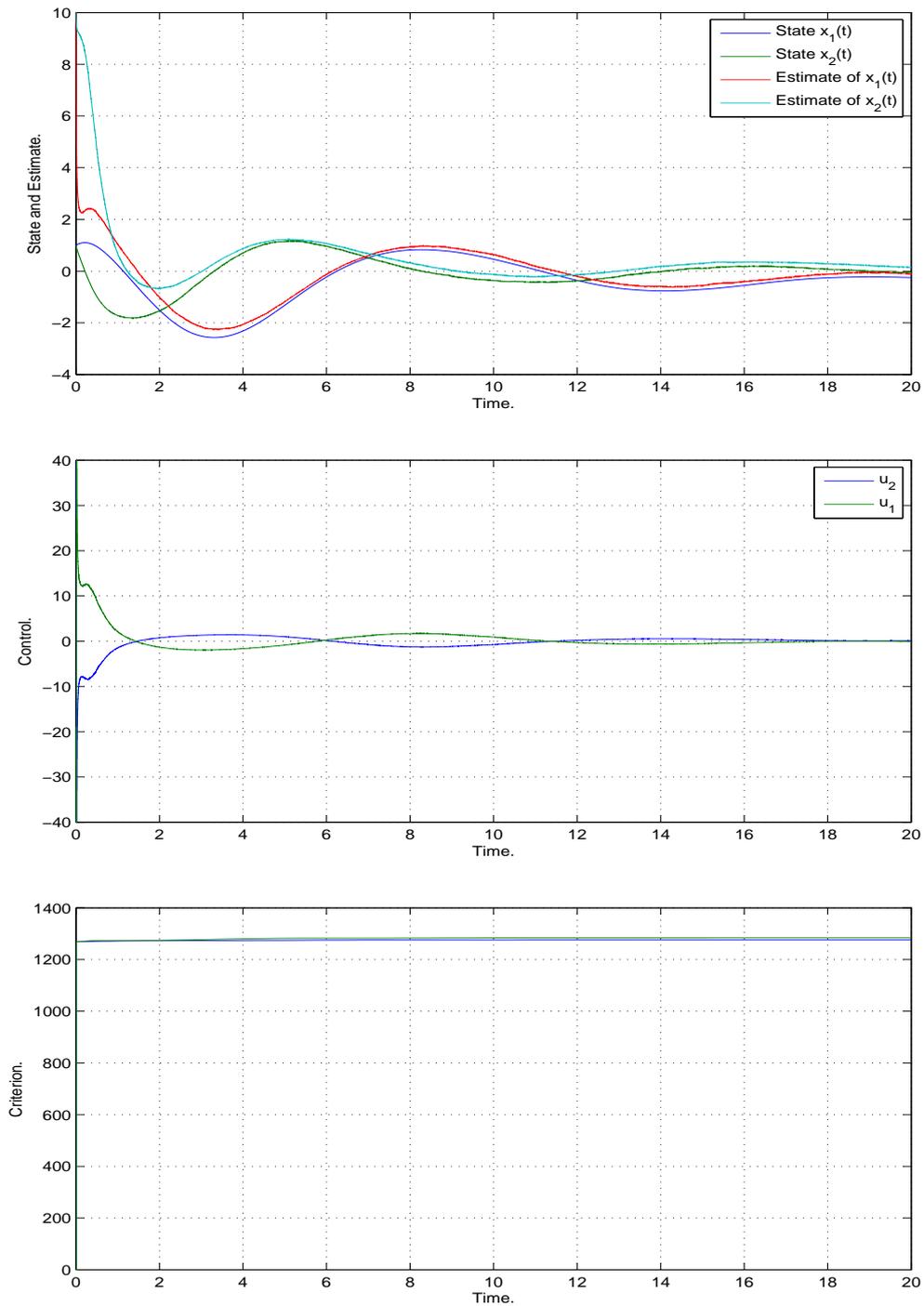


Figura 5.3: El controlador sub-óptimo con respecto al criterio $L_\infty^1(u^1, u^2)$ y $L_\infty^2(u^1, u^2)$ en el entero intervalo de simulación $[0, 20]$. **1.** Gráfica del estado controlado (5.45) $x_1(t)$ (azul) y $x_2(t)$ (verde) el estimado (5.49) $m_1(t)$ (rojo) y $m_2(t)$ (celeste); **2.** Gráfica del control sub-óptimo (5.33) $u^1(t)$ (verde) y $u^2(t)$ (azul); **3.** Gráfica del criterio (5.28) $L_\infty^1(u^1, u^2)$ (azul) y $L_\infty^2(u^1, u^2)$ (verde) producido por el controlador sub-óptimo

6

CONTROL MIN-MAX PARA SISTEMAS POLINOMIALES BASADO EN LA SALIDA.

6.1. PLANTEAMIENTO DEL PROBLEMA.

Considere la siguiente ecuación diferencial no lineal con un término polinomial que incluye un vector desconocido α , que introduce incertidumbre en el modelo, y una ecuación de salida con la información de estado disponible:

$$\begin{aligned}\dot{x}(t; \alpha) &= f(t, x, \alpha) + B(t; \alpha)u(t) + d(t; \alpha) + B(t; \alpha)\zeta(t), \\ y(t; \alpha) &= C(t; \alpha)x(t; \alpha) \quad x(t_0; \alpha) = x_0,\end{aligned}\tag{6.1}$$

donde $x(t; \alpha) \in \mathbb{R}^n$ es el vector de estado, $u(t)$ es el control de entrada que varía dentro de una región de controles dada por $U \subset \mathbb{R}^m$, $d(t; \alpha) \in \mathbb{R}^n$ es una función conocida continua, $B(t; \alpha) \in \mathbb{R}^{m \times n}$ es la matriz del control, y α es un parámetro contenido en un conjunto paramétrico \mathcal{A} , $\zeta(t) \in \mathbb{R}^n$ es una incertidumbre del tipo acoplada contenido en el rango de la matriz del control, $C(t; \alpha) \in \mathbb{R}^{p \times n}$ es la matriz de salida, y $y(t; \alpha) \in \mathbb{R}^p$ la señal de salida. En este trabajo, se considera \mathcal{A} como un conjunto finito, esto es, $\mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$, cada uno representa una posible realización o posible modelo del sistema. La variable de tiempo corre en un intervalo $t \in [t_0, T]$.

Observe que en este trabajo se enfrenta a dos tipos de incertidumbres, por lo que el desafío es diseñar una ley de control que trate efectivamente ambas y este basado solo en la información de salida. Para la incertidumbre no acoplada la solución o optimalidad puede ser dada en términos del peor caso del índice de desempeño, y para la incertidumbre acoplada se propone el modo deslizante integral basado en la salida del sistema. El control es diseñado en dos pasos, existen dos términos en la ley de control: $u = u_1 + u_2$. La primera parte del control intenta tratar con la incertidumbre paramétrica y la segunda parte trata con la incertidumbre acoplada.

Para cada parámetro fijo, el sistema (6.1) se supone que es uniformemente controlable; la definición de uniformemente controlable para sistemas no lineales puede ser encontrada en [61]. Para cada parámetro fijo α , la función

no lineal $f(t, x, \alpha)$ es considerada como un polinomio de n variables, componentes del vector de estado $x(t; \alpha) \in \mathbb{R}^n$; esto requiere una especial definición del polinomio de grado $n > 1$. Siguiendo el trabajo previo (see [52]), un polinomio de grado p del vector $x(t; \alpha) \in \mathbb{R}^n$ es considerado como una forma p -lineal de n componentes de $x(t)$, esto es:

$$f(t, x, \alpha) = a_0(t; \alpha) + a_1(t; \alpha)x + a_2(t; \alpha)xx^T + \cdots + a_s(t; \alpha)x \cdots s \text{ times} \cdots x. \quad (6.2)$$

Aquí, los parámetros considerados son: a_0 es un vector de dimension n , a_1 es una matriz de dimension $n \times n$, a_2 es un 3D tensor de dimension $n \times n \times n$, y a_s es un $(s+1)$ D tensor de dimension $n \times \cdots (s+1) \text{ times} \cdots \times n$, y $x \times \cdots s \text{ times} \cdots \times x$ es un p D tensor de dimension $n \times \cdots s \text{ times} \cdots \times n$, obtenido por la multiplicación espacial de p veces el vector x por si mismo. Esto también es posible representarlo como un polinomio en forma de sumatoria:

$$f_k(t, x, \alpha) = a_{0k}(t; \alpha) + \sum_i a_{1ki}(t; \alpha)x_i + \sum_{ij} a_{2kij}(t; \alpha)x_i x_j + \cdots + \sum_{i_1 \dots i_s} a_{1ki_1 \dots i_s}(t; \alpha)x_{i_1} \dots x_{i_s},$$

$$k, i, j, i_1 \dots i_s = 1, \dots, n, \quad (6.3)$$

donde la dependencia de $a_0(t; \alpha)$, $a_1(t; \alpha)$, ..., $a_s(t; \alpha)$ con respecto a α significa que un tensor de parámetros también pertenece al conjunto de parámetros.

Observación 6.1 Claramente, la incertidumbre en el parámetro realizado es representado por un valor de α . El parámetro α pertenece a un conjunto finito \mathcal{A} que contienen todos los posibles escenarios o realizaciones paramétricas de la plantas no lineales, que se fija durante el proceso actual sin posibilidad de cambiar una vez que el proceso a inicia. Así, cada trayectoria esta únicamente determinada por un conjunto de parámetros dados. Sin embargo, no existe información de cual trayectoria se realiza. De este modo, el control propuesto debería tratar con todos los parámetros y proporcionar un comportamiento aceptable para tal clase de sistemas.

El siguiente funcional de costo cuadrático es definido:

$$g(x(t; \alpha), u_1(t), \alpha) = \frac{1}{2}x^T(T; \alpha)Qx(T; \alpha) + \frac{1}{2} \int_{t_1}^T (x^T(t; \alpha)Lx(t; \alpha) + u_1^T(t)Ru_1(t)) dt. \quad (6.4)$$

Este índice de desempeño es dado en la forma estándar de Bolza, donde R es una matriz simétrica definida positiva, L y Q son dos matrices simétricas definidas no negativas. El diseño del control es realizado en dos pasos, primero, para el problema de multi-parámetro donde la solución es formulada como un extremo del siguiente problema Mini-max:

$$\min_{u(t)} \max_{\alpha \in \mathcal{A}} g(x(t; \alpha), u_1, \alpha). \quad (6.5)$$

Entonces, en el segundo paso, un modo deslizante integral es diseñado para eliminar la incertidumbre acoplada desde casi el inicio del proceso. Ambos controles son basados en un observador por modos deslizante de alto orden.

6.2. CONTROL MIN-MAX PARA SISTEMAS POLINOMIALES.

Primero se diseñara el control mini-max. Considere el siguiente sistema extendido incluyendo todos los escenarios sin la incertidumbre acoplada:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t)) + \mathbf{B}(t) u_1(t) + \mathbf{d}(t), \quad (6.6)$$

$$\mathbf{f} := \begin{bmatrix} f(t, x, \alpha_1) & \cdots & 0 \\ \vdots & . & \vdots \\ 0 & \cdots & f(t, x, \alpha_N) \end{bmatrix}, \quad \mathbf{x} := \begin{pmatrix} x(t, \alpha_1) \\ \vdots \\ x(t, \alpha_N) \end{pmatrix},$$

$$\mathbf{a}_i := \begin{bmatrix} a_i & .. & 0 \\ \vdots & \ddots & \vdots \\ 0 & .. & a_i \end{bmatrix}, \quad \mathbf{B}^\top := \begin{bmatrix} B^\top(t, \alpha_1) \\ \vdots \\ B^\top(t, \alpha_N) \end{bmatrix}, \quad (6.7)$$

$$i = 0, \dots, s$$

donde los términos tensoriales también aparecen en una forma extendida para los coeficientes del polinomio. El vector extendido \mathbf{d} es definido como

$$\mathbf{d} := (d^\top(t; \alpha_1), \dots, d^\top(t; \alpha_N)). \quad (6.8)$$

Se puede observar que la dependencia en el parámetro α a desaparecido; la nueva dinámica incluye familia completa de plantas pero el control sigue siendo el mismo para todas las plantas. El regulador mini-max que realiza (6.5) con respecto al criterio cuadrático (6.4) para el sistema polinomial (6.1) según [62] toma la forma:

$$u_1 = -R^{-1} \mathbf{B}^\top [\mathbf{M}_\lambda \mathbf{x} + \mathbf{p}_\lambda], \quad (6.9)$$

donde la función matricial \mathbf{M}_λ es la solución de la siguiente ecuación tipo Riccati:

$$\begin{aligned} \dot{\mathbf{M}}_\lambda + \Lambda \mathbf{L} + [\mathbf{a}_1 + 2\mathbf{a}_2 \mathbf{x} + 3\mathbf{a}_3 \mathbf{x} \mathbf{x}^\top + \cdots + s \mathbf{a}_s \mathbf{x} \cdots (s-1) \text{ times } \cdots \mathbf{x}]^\top \mathbf{M}_\lambda + \\ \mathbf{M}_\lambda [\mathbf{a}_1 + \mathbf{a}_2 \mathbf{x} + \mathbf{a}_3 \mathbf{x} \mathbf{x}^\top + \dots + \mathbf{a}_s \mathbf{x} \cdots (s-1) \text{ times } \cdots \mathbf{x}] - \mathbf{M}_\lambda \mathbf{B} R^{-1} \mathbf{B}^\top \mathbf{M}_\lambda = 0, \end{aligned} \quad (6.10)$$

con la condición terminal $\mathbf{M}_\lambda(T) = \Lambda \mathbf{Q}$, y la función vectorial parametrizada \mathbf{p}_λ es la solución de la ecuación lineal:

$$\begin{aligned} \dot{\mathbf{p}}_\lambda + \mathbf{M}_\lambda \mathbf{a}_0 + [\mathbf{a}_1 + 2\mathbf{a}_2 \mathbf{x} + 3\mathbf{a}_3 \mathbf{x} \mathbf{x}^\top + \cdots + s \mathbf{a}_s \mathbf{x} \cdots (s-1) \text{ times } \cdots \mathbf{x}]^\top \mathbf{p}_\lambda - \\ \mathbf{M}_\lambda \mathbf{B} R^{-1} \mathbf{B}^\top \mathbf{p}_\lambda + \mathbf{M}_\lambda \mathbf{d} = \mathbf{0}, \end{aligned} \quad (6.11)$$

con la condición terminal $\mathbf{p}_\lambda(T) = \mathbf{0}$. La matriz Λ contiene el peso óptimo que pertenece al siguiente simplex:

$$S^N := \left\{ \lambda \in \mathbb{R}^N : \lambda_i \geq 0; \sum_{i=1}^N \lambda_i = 1 \right\}. \quad (6.12)$$

Las matrices \mathbf{L} , \mathbf{Q} , Λ esta definido como

$$\begin{aligned} \mathbf{L} &:= \begin{bmatrix} L & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & L \end{bmatrix}, \quad \mathbf{Q} := \begin{bmatrix} Q & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Q \end{bmatrix}, \\ \Lambda &:= \begin{bmatrix} \lambda_1 I_{n \times n} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_2 I_{n \times n} \end{bmatrix}. \end{aligned} \tag{6.13}$$

La matriz Λ contiene los pesos óptimo del parámetro $\lambda = \lambda^*$ que resuelve el siguiente problema de optimización:

$$\begin{aligned} \lambda^* &= \min_{\lambda \in S^N} J(\lambda), \\ J(\lambda) &:= \max_{\alpha \in \mathcal{A}} g(x(t; \alpha), u_1, \alpha), \end{aligned} \tag{6.14}$$

con $u_1(t)$ dado en (6.9), que esta parametrizado por el vector $\lambda = (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*)^\top$ ($\lambda_i^* \in S^N$) a través de (6.10) y (6.11). Note que para obtener el vector de pesos óptimo resolviendo el problema anterior no necesariamente es una tarea simple. Se sigue de [62] para proveer un procedimiento numérico factible para buscar el peso mini-max en el caso de dos escenarios. Suponiendo que existe un estimado del proceso para los estados del sistema extendido $\hat{\mathbf{x}}$, la ley de control optima es implementada como:

$$u_1 = -R^{-1} \mathbf{B}^\top [\mathbf{M}_\lambda \hat{\mathbf{x}} + \mathbf{p}_\lambda]. \tag{6.15}$$

6.3. DISEÑO DEL OBSERVADOR POR MODOS DESLIZANTES DE ALTO ORDEN PARA SISTEMAS POLINOMIALES.

Considere otra vez el sistema extendido con medición parcial del estado y sin incertidumbres acopladas:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(t, \mathbf{x}(t)) + \mathbf{B}(t) u_1(t) + \mathbf{d}(t), \\ \mathbf{y}(t) &= \mathbf{C} \mathbf{x}(t), \end{aligned} \tag{6.16}$$

donde

$$\mathbf{C} := \begin{pmatrix} C(t, \alpha_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C(t, \alpha_N) \end{pmatrix} \tag{6.17}$$

Como se menciona, el observador por modos deslizantes de alto orden esta basado en el diferenciado exacto de Levant ([31] y [32]), siguiendo [34]. El observado toma la forma:

$$\begin{aligned}
\dot{\hat{x}}_{i,1} &= w_{i,1} = -\alpha_{r_i+1} M_i^{1/(r_i+1)} |\hat{y}_i - y_i|^{(r_i)/(r_i+1)} \text{sign}(\hat{y}_i - y_i) + \hat{x}_{i,2}, \\
\dot{\hat{x}}_{i,2} &= w_{i,2} = -\alpha_{r_i} M_i^{1/(r_i)} |\hat{x}_{i,2} - w_{i,1}|^{(r_i-1)/(r_i)} \text{sign}(\hat{x}_{i,2} - w_{i,1}) + \hat{x}_{i,3}, \\
&\vdots \\
\dot{\hat{x}}_{i,r_i} &= w_{i,r_i} = -\alpha_2 M_i^{1/(2)} |\hat{x}_{i,r_i} - w_{i,r_i-1}|^{(1)/(2)} \text{sign}(\hat{x}_{i,r_i} - w_{i,r_i-1}) + \hat{x}_{i,r_i+1}, \\
\dot{\hat{x}}_{i,r_i+1} &= -\alpha_1 M_i \text{sign}(\hat{x}_{i,r_i+1} - w_{i,r_i}), \\
\hat{y}_i &= \hat{x}_{i,1},
\end{aligned} \tag{6.18}$$

donde r_i es el grado relativo del i -ésimo componente de la salida $\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$; el procedimiento se repite para cada elemento de la salida. La solución de esta ecuación produce el estimado del estado

$$\hat{\mathbf{x}}_i^T = [\hat{x}_{i,1}^T \ \hat{x}_{i,2}^T \ \dots \ \hat{x}_{i,r_i}^T].$$

Ahora, se define como $\hat{\mathbf{x}}$ el conjunto de soluciones para cada componente de $\mathbf{y}(t)$.

$$\hat{\mathbf{x}}^T = [\hat{\mathbf{x}}_1^T \ \hat{\mathbf{x}}_2^T \ \dots \ \hat{\mathbf{x}}_m^T].$$

Los parámetros de ajuste M, α , son seleccionados acorde a [31].

6.4. MODOS DESLIZANTES INTEGRALES PARA LA SALIDA DE SISTEMAS POLINOMIALES CON INCERTIDUMBRES.

En esta sección, se presenta el diseño del control u_2 para hacer frente a la parte con incertidumbre acoplada, que esta basada en el modo deslizante integral. Considere el sistema extendido:

$$\begin{aligned}
\dot{\mathbf{x}}(t) &= \mathbf{f}(t, \mathbf{x}(t)) + \mathbf{B}(t)(u_1(t) + u_2 + \zeta(t)) + \mathbf{d}(t), \\
\mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t).
\end{aligned} \tag{6.19}$$

Se puede observar que la parte de la incertidumbre acoplada es la misma para todos los modelos. El propósito es eliminar la incertidumbre tan rápido como sea posible, de modo que el control mini-max hace el resto del trabajo. Para diseñar el modo deslizante integral de la salida, se define la siguiente superficie deslizante para la salida:

$$z(\mathbf{y}(t)) := \mathbf{G}\mathbf{y}(t) - \sigma(t) \tag{6.20}$$

donde: $\mathbf{G} = (\mathbf{C}\mathbf{B})^+ = [(\mathbf{C}\mathbf{B})^T(\mathbf{C}\mathbf{B})]^{-1}(\mathbf{C}\mathbf{B})^T$. Evaluando la derivada temporal de $z(\mathbf{y}(t))$ se obtiene:

$$\begin{aligned}
\dot{z}(\mathbf{y}(t)) &= \mathbf{G}\mathbf{C}(\mathbf{f}(x, t) + \mathbf{B}(u_1(t) + u_2(t) + \zeta(t)) \\
&\quad + \mathbf{d}(t)) - \dot{\sigma}(t)
\end{aligned} \tag{6.21}$$

Debido a que no se tiene la información completa del estado, se define $\dot{\sigma}(t)$ basado solo en el estimado $\hat{\mathbf{x}}$:

$$\begin{aligned}
\dot{\sigma}(t) &= \mathbf{G}\mathbf{C}(\mathbf{f}(\hat{\mathbf{x}}, t) + \mathbf{B}u_1(t) + \mathbf{d}(t)) \\
\sigma(0) &= \mathbf{G}\mathbf{y}(0),
\end{aligned} \tag{6.22}$$

lo que conduce a

$$\begin{aligned} \dot{z}(y(t)) &= \mathbf{GC}(\mathbf{f}(x, t) - \mathbf{f}(\hat{x}, t)) + u_2(t) + \zeta(t) \\ z(0) &= 0. \end{aligned} \quad (6.23)$$

La segunda parte del control u_2 , que es discontinua, es propuesto como:

$$u_2 = -\gamma(t) \frac{z(t)}{\|z(t)\|} \quad (6.24)$$

donde el escalar $\gamma(t)$ satisface:

$$\gamma(t) - q_a(t) - \|\mathbf{GC}\| \|\mathbf{f}(x, t) - \mathbf{f}(\hat{x}, t)\| \geq \lambda \geq 0,$$

λ es una constante. Para elegir correctamente los parámetros de la ley de control discontinuo y asegurar el movimiento del sistema a lo largo de la superficie deslizante, se propone la función de Lyapunov como $V = (1/2)\|z\|^2$.

La derivada temporal es calculada como:

$$\begin{aligned} \dot{V} &= z^T (\mathbf{GC}(\mathbf{f}(x, t) - \mathbf{f}(\hat{x}, t)) - \gamma(t) \frac{z(t)}{\|z(t)\|} + \zeta(t)) \\ &\leq -\|s\| (\gamma(t) - q_a(t) - \|\mathbf{GC}\| \|\mathbf{f}(x, t) - \mathbf{f}(\hat{x}, t)\|) \\ &\leq -\|s\| \lambda \leq 0 \end{aligned}$$

Debido a la selección de la condición inicial para z (6.23), $z(0) = 0$, se obtiene

$$V(s(t)) \leq V(s(0)) = \frac{1}{2} \|s(0)\|^2 = 0$$

que implica que $s(t) = \dot{s}(t) = 0$. La conclusion es que el modo deslizante integral no tiene la fase de alcance. Por lo tanto, el control equivalente se obtiene como:

$$u_{eq} = -\mathbf{GC}(\mathbf{f}(x, t) - \mathbf{f}(\hat{x}, t)) - \gamma(t) \quad (6.25)$$

Sustituyendo el control equivalente (6.25) en (6.19) se obtiene:

$$\begin{aligned} \dot{x} &= \mathbf{f}(x, t) - \mathbf{BGC}(\mathbf{f}(x, t) - \mathbf{f}(\hat{x}, t)) \\ &\quad + \mathbf{B}u_1(t) + \mathbf{d}(t) \\ y(t) &= \mathbf{C}x(t) \end{aligned} \quad (6.26)$$

que es el sistema extendido, donde la parte de la incertidumbre acoplada es eliminada y solo un termino dependiendo de la observación aparece; dicho termino se desvanece conforme el observado converge al estado real.

Observación 6.2 Dado que el modo deslizante integral elimina la parte de la incertidumbre acoplada, la combinación de las dos técnicas, el control mini-max y el modo deslizante integral junto con la rápida reconstrucción del modo deslizante de alto orden, presenta una poderosa herramienta para el controlador de sistemas polinomiales con información parcial del estado, como se ilustra en los siguientes ejemplos numéricos.

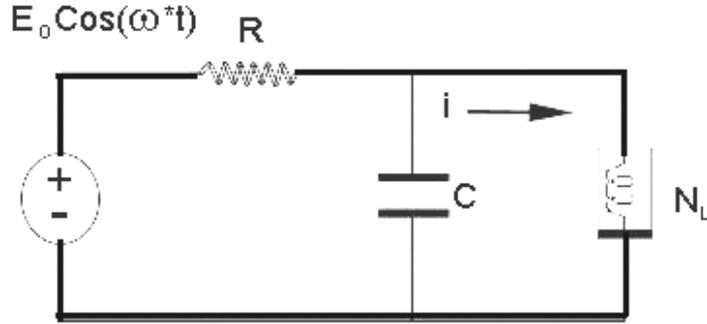


Figura 6.1: Circuito no Lineal.

6.5. EJEMPLOS.

6.5.1. CIRCUITO NO LINEAL.

El bien conocido oscilador electrico no lineal incluyendo un inductor no lineal (ver Figura 1) es modelado por un sistema polinomial bidimensional de tercer orden (ver [63], [64]). Considere el caso de dos posibles escenarios ($N = 2$) del siguiente circuito:

$$\dot{x}_1(t, 1) = x_2, \quad (6.27)$$

$$\dot{x}_2(t, 1) = -1,1x_1 - 0,4x_2 - x_1^3 + 2,05 \cos(1,8t) + \sin(1000t) - u, \quad (6.28)$$

$$\dot{x}_3(t, 2) = x_4,$$

$$\dot{x}_4(t, 2) = -1,15x_1 - 0,45x_2 - 1,05x_1^3 + 2 \cos(1,9t) + \sin(1000t) - u, \quad (6.29)$$

donde la parte de la incertidumbre conocida es $d(t, 1) = 2,05 \cos(1,8t)$, y la parte de la incertidumbre acoplada desconocida es $\zeta(t, 1) = \sin(1000t)$, con la siguiente salida medible para los dos modelos:

$$y(t, 1) = x_1(t, 1),$$

$$y(t, 2) = x_1(t, 2).$$

La ecuación para el estado reconstruido esta dada por:

$$\dot{\hat{x}}_1 = w_1 = \hat{x}_2 - \alpha_3 M^{1/3} |\hat{y}_1 - y_1|^{2/3} \text{sign}(\hat{y}_1 - y_1),$$

$$\dot{\hat{x}}_2 = w_2 = \hat{x}_3 - \alpha_2 M^{1/2} |\hat{x}_2 - w_1|^{1/2} \text{sign}(\hat{x}_2 - w_1),$$

$$\dot{\hat{x}}_3 = \alpha_1 M \text{sign}(\hat{x}_3 - w_2),$$

$$\hat{y}_1 = \hat{x}_1,$$

$$\begin{aligned}
\dot{\hat{x}}_4 &= w_4 = \hat{x}_5 - \alpha_6 M^{1/3} |\hat{y}_2 - y_2|^{2/3} \text{sign}(\hat{y}_2 - y_2), \\
\dot{\hat{x}}_5 &= w_5 = \hat{x}_6 - \alpha_5 M^{1/2} |\hat{x}_5 - w_4|^{1/2} \text{sign}(\hat{x}_5 - w_4), \\
\dot{\hat{x}}_6 &= \alpha_4 M \text{sign}(\hat{x}_6 - w_5), \\
\hat{y}_2 &= \hat{x}_4,
\end{aligned}$$

donde $\hat{x}_1, \hat{x}_2, \hat{x}_4, \hat{x}_5$ son los estimados de las variables de estado $x_1(t, 1), x_2(t, 1), x_1(t, 2)$, y $x_2(t, 2)$, respectivamente. La condición inicial para las variables del observador son fijadas como cero. Las siguientes constantes son usadas para afinar el observador:

$$M = 1, \alpha_1 = \alpha_2 = \alpha_4 = \alpha_5 = 20, \alpha_3 = \alpha_6 = 10.$$

Finalmente, la ley de control toma la forma:

$$u_1 = (P_{21} + P_{41})\hat{x}_1 + (P_{22} + P_{42})\hat{x}_2 + (P_{23} + P_{43})\hat{x}_3 + (P_{24} + P_{44})\hat{x}_4 + p_2 + p_4.$$

Las figuras están organizadas como sigue. Las figuras 2 y 3 muestran las variables de estado y el estimado de los modelos 1 y 2. Note que el estado de ambas plantas muestran un comportamiento idéntico, cuando el control robusto es aplicado, demostrando que el diseño robusto funciona bien para cualquiera de las dos plantas. Los estimados también se ven casi iguales todo el tiempo debido a la rápida razón de convergencia del observador. La figura 4 presenta los valores de los costos dependientes de λ , donde se puede observar que el índice de desempeño tiene el mínimo (del peor de los casos del costo) alrededor de $\lambda = 0,28$. La figura 5 corresponde al control min-max basado en la salida. Finalmente, la figura 6 muestra el control por modos deslizantes basado en la salida. Otra vez, se puede observar que con la aplicación del modo deslizante integral basado en la salida la perturbación no afecta la dinámica de cualquiera de las plantas, resultando en un buen desempeño de ambos controles.

6.5.2. PÉNDULO

Considere la siguiente ecuación para el péndulo simple (ver Fig. 7):

$$\frac{d^2x}{dt^2} = -\frac{g}{l} \sin(x) + u, \quad (6.30)$$

El termino no lineal puede ser aproximado por un polinomio usando la expansion de Taylor de la función $\sin(x)$:

$$\sin(x) = x - \frac{1}{6}x^3 + \dots + \frac{(-1)^{n+1}x^{2n-1}}{(2n-1)!} + \dots$$

Considere la aproximación por el polinomio de tercer orden, esto es, la ecuación (6.30) es aproximada como:

$$\frac{d^2x}{dt^2} = -\frac{g}{l} \left(x - \frac{1}{6}x^3 \right) + u.$$

Procediendo con el cambio de variables $x_1 = x$ y $x_2 = \dot{x}$, se obtiene:

$$\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= -\frac{g}{l} \left(x_1 - \frac{1}{6}x_1^3 \right) + u.
\end{aligned}$$

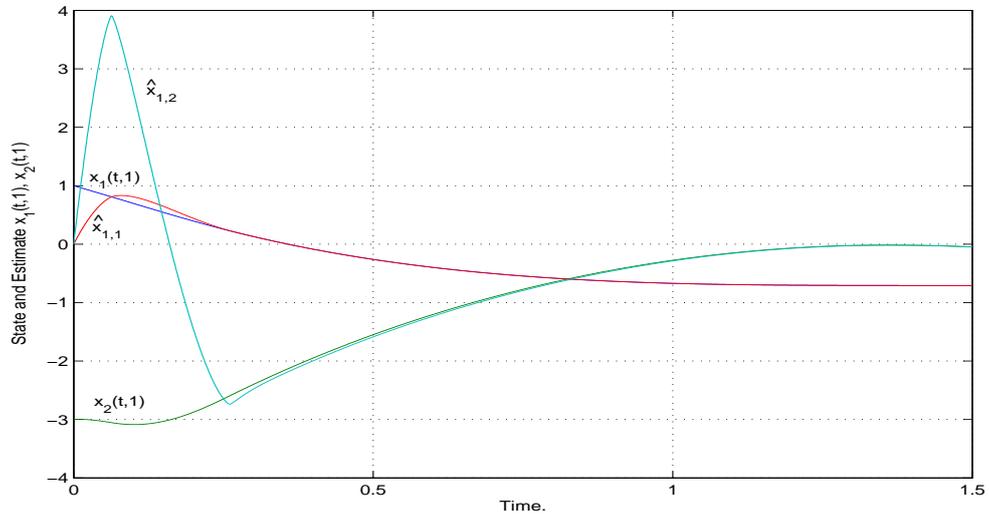


Figura 6.2: Estado y estimado. Planta 1.

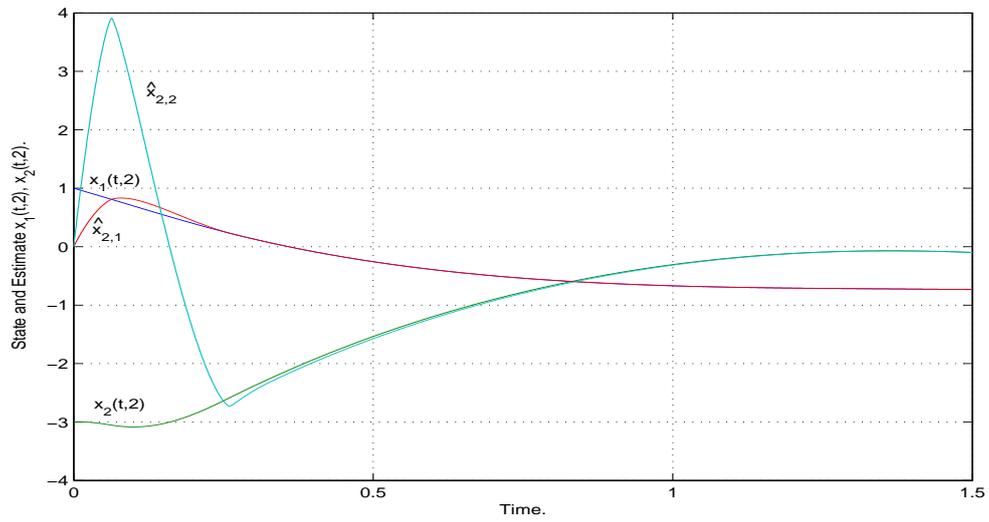


Figura 6.3: Estado y estimado. Planta 2.

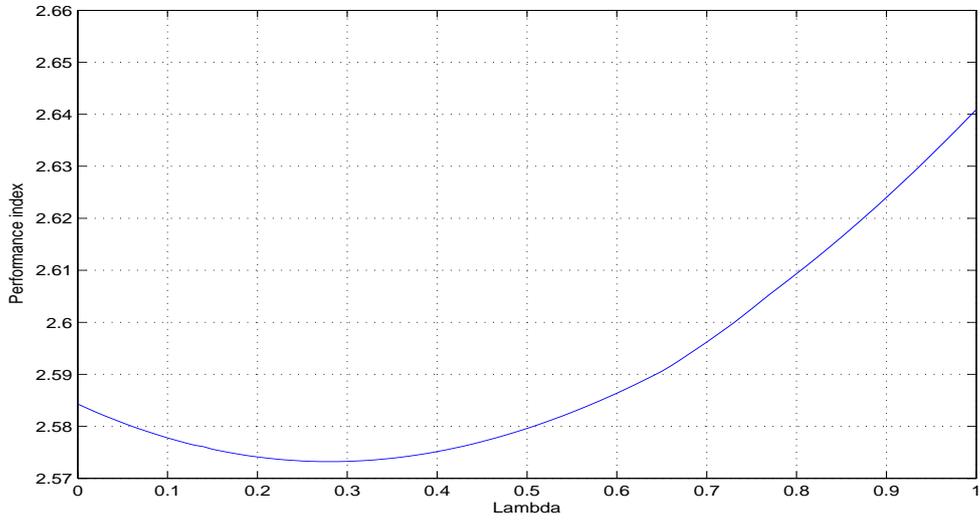


Figura 6.4: Índice de Desempeño $J(\lambda)$.

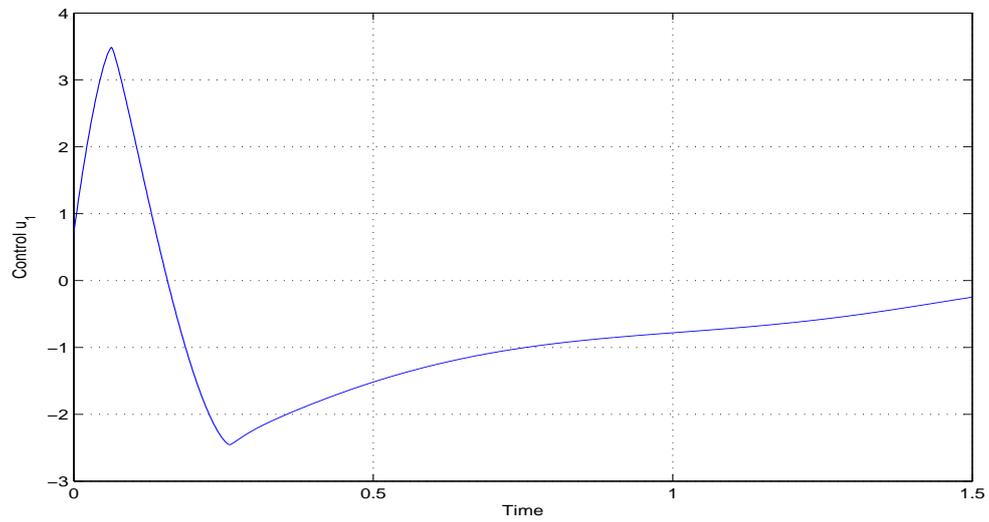


Figura 6.5: Control Min-Max u_1 .

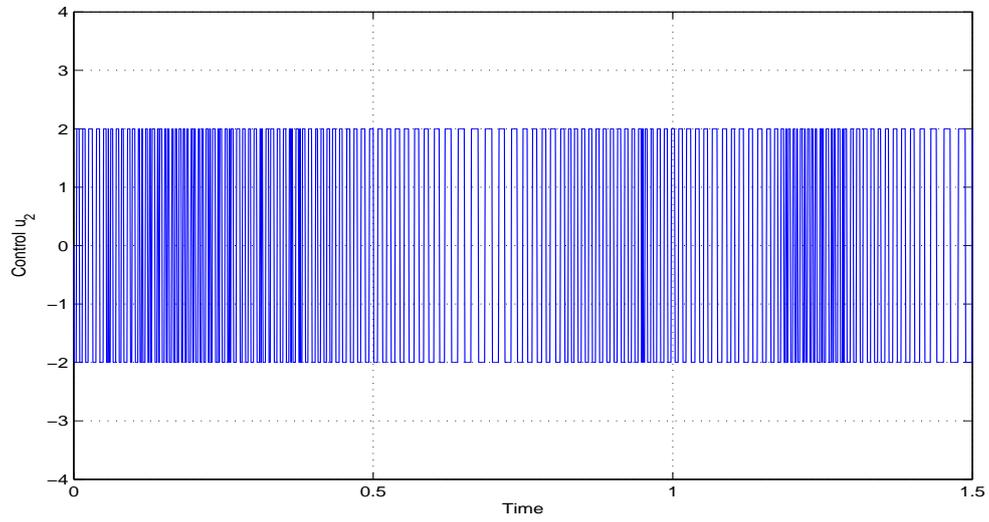


Figura 6.6: Control por modos deslizantes integrales u_2 .

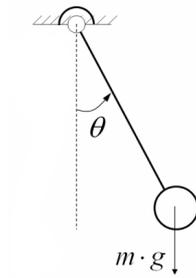


Figura 6.7: Péndulo Simple.

El modelos del sistema de dos escenarios se convierte en:

$$\begin{aligned}\dot{x}_1(t, 1) &= x_2(t, 1), \\ \dot{x}_2(t, 1) &= -3 * (x_1(t, 1) - \frac{1}{6}x_1^3(t, 1)) + u + \text{sen}(1000t), \\ \dot{x}_1(t, 2) &= x_2(t, 2), \\ \dot{x}_2(t, 2) &= -2,95 * (x_1(t, 2) - \frac{1}{6}x_1^3(t, 2)) + u + \text{sen}(1000t),\end{aligned}$$

con la ecuación de salida:

$$\begin{aligned}y(t, 1) &= x_1(t, 1), \\ y(t, 2) &= x_1(t, 2),\end{aligned}$$

y el índice de desempeño:

$$g(x(t; \alpha), u, \alpha) = x_1^2(T, \alpha) + x_2^2(T, \alpha) + \int_0^T (x_1^2(t, \alpha) + x_2^2(t, \alpha) + u^2(t)) dt.$$

Finalmente, las ecuaciones del observador son:

$$\begin{aligned}\dot{\hat{x}}_1 &= w_1 = \hat{x}_2 - \alpha_3 M^{1/3} |\hat{y}_1 - y_1|^{2/3} \text{sign}(\hat{y}_1 - y_1), \\ \dot{\hat{x}}_2 &= w_2 = \hat{x}_3 - \alpha_2 M^{1/2} |\hat{x}_2 - w_1|^{1/2} \text{sign}(\hat{x}_2 - w_1), \\ \dot{\hat{x}}_3 &= \alpha_1 M \text{sign}(\hat{x}_3 - w_2), \\ \hat{y}_1 &= \hat{x}_1, \\ \\ \dot{\hat{x}}_4 &= w_4 = \hat{x}_5 - \alpha_6 M^{1/3} |\hat{y}_2 - y_2|^{2/3} \text{sign}(\hat{y}_2 - y_2), \\ \dot{\hat{x}}_5 &= w_5 = \hat{x}_6 - \alpha_5 M^{1/2} |\hat{x}_5 - w_4|^{1/2} \text{sign}(\hat{x}_5 - w_4), \\ \dot{\hat{x}}_6 &= \alpha_4 M \text{sign}(\hat{x}_6 - w_5), \\ \hat{y}_2 &= \hat{x}_2,\end{aligned}$$

donde la condición inicial para las variables del observador es fijado como cero. La ley de control toma la forma:

$$u = (P_{21} + P_{41})\hat{x}_1 + (P_{22} + P_{42})\hat{x}_2 + (P_{23} + P_{43})\hat{x}_3 + (P_{24} + P_{44})\hat{x}_4.$$

En este ejemplo, las figuras son organizadas como sigue: La figura 8 y 9 muestran las variables de estado y los estimados del modelos 1 y 2. Como en el ejemplo anterior, ambas plantas muestran el mismo comportamiento, cuando el control robusto es aplicado, demostrando así que el diseño robusto trabaja bien para cualquier de las dos plantas. Los estimados también parecen casi iguales todo el tiempo. Para este caso, el valor óptimo es $\lambda = 0,8779$. La figura 10 muestra el estimado del error. Finalmente, La figura 11 muestra el control por modos deslizante integrales basado en la salida. De nuevo, se puede observar que con la aplicación del modo deslizante integral basado en la salida la perturbación no afecta la dinámica de ninguna de las plantas, resultando en un buen desempeño a través de combinación de las dos leyes de control.

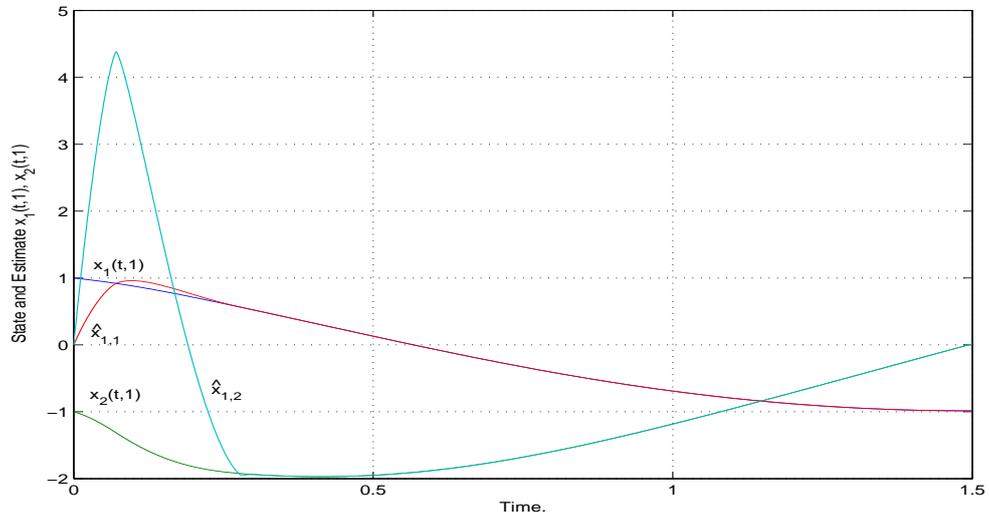


Figura 6.8: Estado y estimado. Péndulo 1.

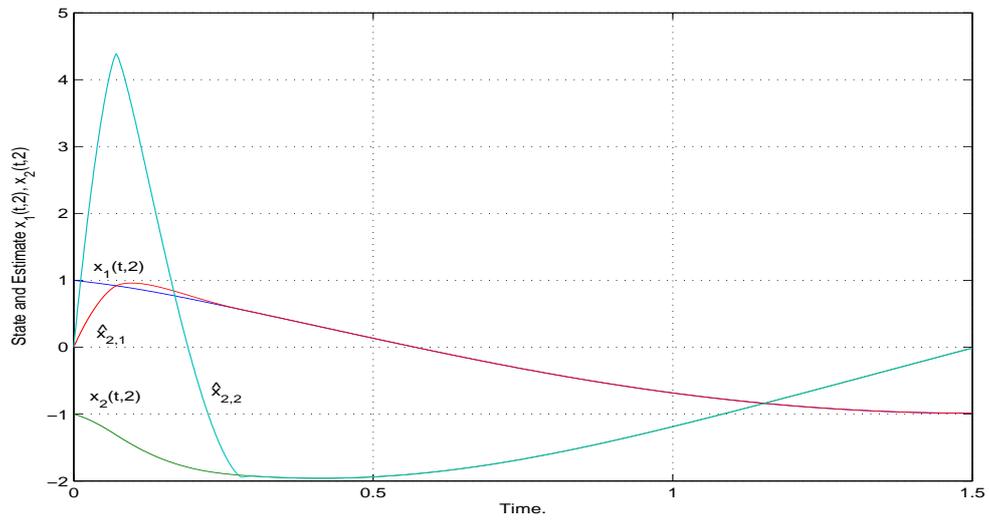


Figura 6.9: Estado y estimado. Péndulo 2.

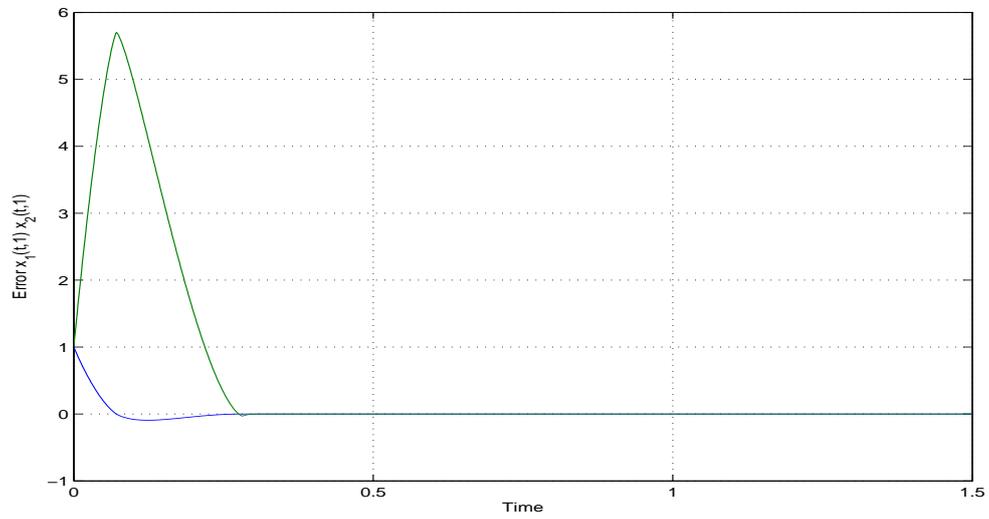


Figura 6.10: Error de Estimación.

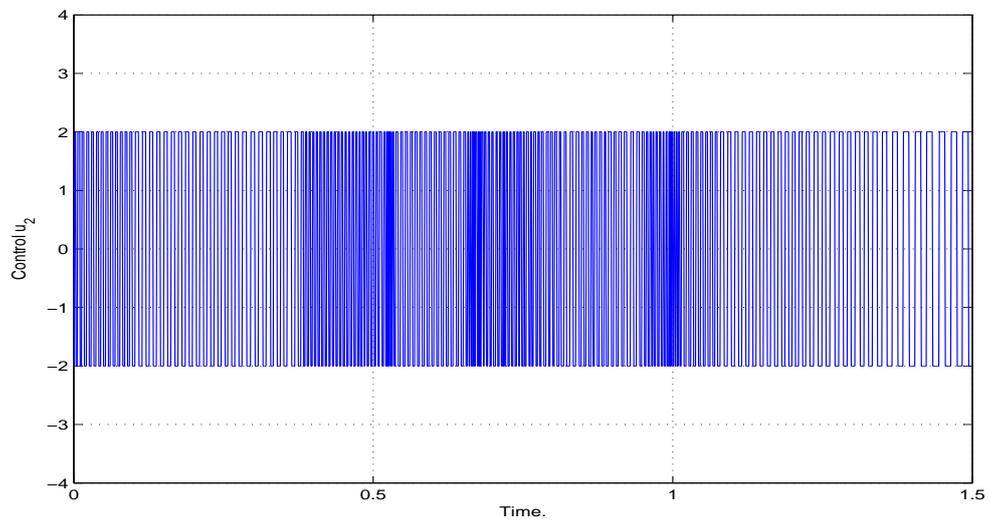


Figura 6.11: Control por modos deslizantes integrales u_2 .

7

ESTRATEGIAS DE STACKELBERG CON INCENTIVOS MINI-MAX PARA JUEGOS DIFERENCIALES LINEAL-CUADRATICOS CON INCERTIDUMBRES.

7.1. DESCRIPCIÓN DEL PROBLEMA Y SUPOSICIONES BÁSICAS

Considere el juego diferencial lineal cuadrático de 2 jugadores (LQDG) donde la dinámica de los jugadores esta dada por la siguiente ecuación diferencial ordinaria (ver [8] y[4]):

$$\dot{x}(t) = A(t)x(t) + \sum_{n=1}^2 B^j(t)u^j(t); x(t_0) = x_0, \quad (7.1)$$

donde $x(t) \in \mathbb{R}^n$ es el vector de estado en tiempo $t \in [0, T]$, el índice j denota el número del jugador ($j = 1, 2$), en este caso se selecciona el jugador 1 como el líder y el jugador 2 como el seguidor, $u^j(t) \in \mathbb{R}^{m_j}$ son las acciones de control correspondientes a ambos jugadores. Las matrices $A(t) \in \mathbb{R}^{n \times n}$ y $B^j(t) \in \mathbb{R}^{n \times m_j}$ son las matrices del sistema y del control con entradas continuas en el intervalo $[0, T]$. Toda esta información y el estado inicial (Acotado) x_0 se asume que se conoce por ambos jugadores. En esta sección suponemos que el tiempo T esta fijo, después se abandona esta suposición para considerar el caso de horizonte de tiempo infinito. También se supone lo siguiente:

1. Las matrices $A(t)$ así como la matriz del control $B^j(t)$ ($j = 1, 2$) pertenecen a un conjunto finito fijo de

familia de funciones matriciales conocidas a priori, que caracterizan la dinámica posible, esto es:

$$\begin{aligned} A(t) &\in \{A(t, 1), A(t, 2), \dots, A(t, M)\}, \\ B^j(t) &\in \{B^j(t, 1), B^j(t, 2), \dots, B^j(t, M)\}, \\ j &= 1, 2, \end{aligned} \quad (7.2)$$

aquí el numero M es el numero finito posible del conjunto de parámetros, usando la notación $A(t, \alpha)$ y el parámetro $\alpha \in \mathcal{M} = \{1, \dots, M\}$, representando la incertidumbre sobre la dinámica del juego

2. La información completa del estado esta disponible para ambos jugadores de modo que pueden utilizarla en la selección de sus acciones de control.

Por lo tanto, para cada parámetro fijo α tenemos una realización concreta de la dinámica, el siguiente Multi-Modelo es obtenido:

$$\dot{x}(t, \alpha) = A(t, \alpha)x(t, \alpha) + \sum_{j=1}^2 B^j(t, \alpha)u^j(t); x(t_0, \alpha) = x_0, \quad (7.3)$$

Para un conjunto dado de acciones de control (u^1, u^2) , nos interesamos en la trayectoria correspondiente a partir del punto inicial x_0 . Pero el valor posible realizado de α es a priori desconocido. Esto es porque, es considerado el conjunto de las familias de trayectorias $x(t, \alpha)$ con insuficiente información acerca de la trayectorias realizadas. Cada jugador tiene un índice de desempeño cuadrático $h^i(u^1, u^2, \alpha)$ (que también depende de α), que esta dado en la forma estándar de Bolza

$$h^i(u^1, u^2, \alpha) = \frac{1}{2}x^\top(T, \alpha)Q_f^i x(T, \alpha) + \frac{1}{2} \int_{t=t_0}^T \left[x^\top(t, \alpha)Q^i x(t, \alpha) + \sum_{j=1}^2 u^{j\top} R^{ij} u^j \right] dt, \quad (7.4)$$

con la siguiente restricción en las matrices de peso:

$$\begin{aligned} Q^i &= Q^{i\top} \geq 0, \quad R^{ii} = R^{ii\top} > 0, \quad R^{12} = R^{12\top} > 0, \\ Q_f^i &= Q_f^{i\top} \geq 0, \quad R^{21} = R^{21} \geq 0, \end{aligned} \quad (7.5)$$

Dado que el valor del parámetro obtenido α es desconocido, el peor caso con respecto a los escenarios posibles es definido en el funcional de costo J^i para cada jugador bajo controles admisibles fijos $u^1 \in U^1$ y $u^2 \in U^2$ se puede definir como sigue:

$$J^i(u^1, u^2) := \max_{\alpha} h^i(u^1, u^2, \alpha). \quad (7.6)$$

Dentro de este marco de incertidumbres nuestro objetivo es encontrar un control de incentivos que proporcionara el costo mas favorable para el líder, por los ajustes de los multi-parámetros. Esto representa un problema debido a la imposibilidad de encontrar cual es el mejor costo que el líder puede obtener en un problema multi escenario. Si nos fijamos en la teoría estándar existente la mayoría de los esquemas de incentivos en la literatura se basan de una o de otra manera en la llamada solución en equipo optima, que es la determinación de un limite inferior para la función de costo del líder. Este limite se encuentra asumiendo la situación ideal de ambos jugadores que persiguen la minimization del funcional del líder. Una vez que el líder obtuvo tal solución el siguiente problema

es diseñar una estrategia adecuada que realiza tal objetivo tomando en consideración la posible reacción racional del seguidor. Por lo tanto, nuestro enfoque a este problema es el siguiente: Considerando que la actual realización del parámetro no es conocida, por consiguiente, la solución en equipo optima tradicional no esta disponible. Sin embargo, podemos encontrar la solución mini-max para un problema de equipo optima y a continuación, se intenta diseñar un control que ajuste al juego (en el sentido mini-max) a la mejor solución. En el momento que el líder a determinado su estrategia de incentivos y la anuncia antes de tiempo, el problema de control mini-max que se enfrenta el seguidor, consiste en encontrar un control admisible $\{u^2\}_{t \in [0, T]}$ que para una condición inicial dada $x(t_0, \alpha) = x_0$, proporciona las siguientes propiedades de optimalidad:

$$\{u^2\}_{t \in [0, T]} \in Arg \min_{\text{admissible}\{u^2\}_{t \in [0, T]}} \max_{\alpha} h^2(u^1, u^2, \alpha).$$

7.1.1. BREVE RESEÑA DEL FORMATO DE INCENTIVO.

Fundamentalmente, los diferentes tipos de incentivos están relacionados con el tipo de estructura de información que los jugadores tienen disponible o usan, por ejemplo en el artículo [46] los autores proponen como forma de incentivo $u_1(t) = u_1^d(t) + P[u_2(t) - u_2^d(t)]$, donde $u_1(t)$ es la ley de control del líder y $u_2(t)$ es el control de seguidor, P es la denominada matriz de incentivo, el superíndice d representa el conjunto ideal de estrategias que el líder puede esperar, tal conjunto de estrategias por lo general se eligen para ser los mismos como los valores en equipo óptimo, que como hemos visto se asume que todos los participantes minimizan el costo del líder. Por lo tanto, en [46] el líder debe tener acceso al control del seguidor que representa un problema (para muchas aplicaciones). La aproximación de [45], para sistemas de tiempo discreto, hace uso de un paso de retardo en la representación del estado como $u_1(k) = u_1^t(k) + P[x(k) - \bar{x}(k-1)]$, donde $\bar{x}(k-1) = F(k-1)x(k-1)$ denota el valor en equipo óptimo de $x(k)$ y $F(k-1)$ es matriz en lazo cerrado. Este enfoque ha demostrado que funciona bien, pero los problemas de estabilidad no se aclaran del todo. Por supuesto, la información disponible requerida supone un retraso en el estado, este enfoque también se puede extender a múltiples retrasos. El trabajo de [48], aunque trabaja en tiempo continuo, propone un control puro del tipo retroalimentado de la forma: $u_1(t) = u_1^{te}(t) + P^T[x(t) - x^{te}(t)]$, el diseño propuesto para la matriz de incentivos P , que garantiza la estabilidad del sistema en lazo cerrado esta dada, sin embargo, sin demostración. En el trabajo de ([47]), los autores obtienen las estrategias de equipo óptimo retroalimentada del estado de Stackelberg (sin retardo) para una clase de juegos en tiempo discreto, y la condición suficiente para encontrar la matriz P esta dada a través de la solución de un conjunto de ecuaciones matriciales.

Como se menciona en este trabajo se sigue la forma de incentivos dada en [48], el líder anuncia su estrategia del tipo retroalimentado sin retardos:

$$\begin{aligned} u_1(t) &= u_1^{te}(t) + P^T[x(t) - x^{te}(t)], \\ P &\in \mathbb{R}^{n \times m}, \end{aligned} \tag{7.7}$$

donde P , la matriz de incentivos, es necesario encontrarla. El reto con esta matriz es que debe asegurarse de que la solución en equipo óptima del líder se logra, y esta matriz puede no ser única.

7.2. SOLUCIÓN EN EQUIPO OPTIMA MINI-MAX.

Considere que se toma el peor de los escenarios con respecto a al conjunto de parámetros de la función de costo del líder usando (7.6), el problema de equipo óptimo es calculado asumiendo que ambos jugadores persiguen la minimización en el peor caso del funcional del líder que es expresado como:

$$\underset{u^1 \in U^1, u^2 \in U^2}{\text{minimize}} \quad J^1(u_1, u_2), \quad (7.8)$$

Como hemos dicho antes, la solución de la minimización conjunta del funcional $J^1(u_1, u_2)$ sobre $u^1 \in U^1$, $u^2 \in U^2$ es lo suficientemente buena, permitiendo que el líder pueda lograrla. En este caso llamamos tal solución como una solución mini-max o una solución en equipo optima robusta. Siguiendo las técnicas conocidas como el Principio del Máximo Robusto ([30], [27]) considere la siguiente función Hamiltoniana generalizada para el problema de equipo óptimo:

$$\mathcal{H}^\diamond = \sum_{\alpha} \left[\frac{1}{2} \psi_{n+1}(t, \alpha) (x^\top(t, \alpha) Q^1 x(t, \alpha) + u^{1\top} R^{11} u^1 + u^{2\top} R^{12} u^2) + \psi^\top(t, \alpha) (A(t, \alpha) x(t, \alpha) + B^1(t, \alpha) u^1 + B^2(t, \alpha) u^2) \right], \quad (7.9)$$

donde $\psi(t, \alpha)$, $\alpha \in \mathcal{M} = \{1, \dots, M\}$, son los vectores de co-estado de las dinámicas correspondientes. Para cada valor fijo del parámetro incierto α la función Hamiltoniana utilizada en la construcción de la estrategia optima robusta del seguidor, representa las funciones Hamiltonianas estándar. Usando vectores extendidos como en [30], [27], considere la siguiente representación de los juegos dinámicos dada por (fuente en negrita para los vectores y matrices extendidos) $\dot{\mathbf{x}}(t) = \mathbf{A}(t) + \mathbf{B}^1(t) u^1(t) + \mathbf{B}^2(t) u^2(t)$, aquí los vectores y las matrices extendidas son:

$$\mathbf{A} := \begin{bmatrix} A(t, 1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A(t, M) \end{bmatrix}, \quad \mathbf{Q}^i := \begin{bmatrix} Q^i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Q^i \end{bmatrix}, \quad (7.10)$$

$$\mathbf{Q}_f^i := \begin{bmatrix} Q_f^i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Q_f^i \end{bmatrix}, \quad \mathbf{\Lambda} := \begin{bmatrix} \lambda(1)I & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda(M)I \end{bmatrix},$$

$$\mathbf{B}^{i\top} := [B^{i\top}(t, 1) \cdots B^{i\top}(t, M)], \quad I \in \mathbb{R}^{n \times n} \quad i = 1, 2$$

$$\mathbf{x}^{te\top} := (x^{te\top}(t, 1), \dots, x^{te\top}(t, M))^\top \in \mathbb{R}^{n \cdot M}$$

Define también para cualquier vector $\lambda = (\lambda(1), \dots, \lambda(M))^\top$

$$\lambda \in S^M := \left\{ \lambda \in \mathbb{R}^M, \lambda(\alpha) \geq 0, \sum_{\alpha} \lambda(\alpha) = 1 \right\}. \quad (7.11)$$

La solución para el problema de parámetros conjuntos en el peor caso (7.8) esta dado en el siguiente teorema.

Teorema 7.1 Supongamos que las condiciones (7.5) para el juego lineal cuadrático de las dos personas son satisfechas, entonces el conjunto de controles mini-max realiza el problema de optimización conjunta en el peor

caso (7.8) están dados por

$$\begin{aligned} u^{1te} &= - (R^{11})^{-1} \mathbf{B}^{1\top} \mathbf{S}_\lambda x^{te}, \\ u^{2te} &= - (R^{12})^{-1} \mathbf{B}^{2\top} \mathbf{S}_\lambda x^{te}, \end{aligned} \quad (7.12)$$

donde la matriz $\mathbf{S}_\lambda \in \mathbb{R}^{n \cdot M \times n \cdot M}$ es la solución de la siguiente ecuación diferencial parametrizada (backward-forward)

$$\begin{aligned} -\mathbf{S}_\lambda \mathbf{A} - \mathbf{A}^\top \mathbf{S}_\lambda - \Lambda \mathbf{Q}^1 + \mathbf{S}_\lambda \mathbf{B}^2 (R^{12})^{-1} \mathbf{B}^{2\top} \mathbf{S}_\lambda + \\ \mathbf{S}_\lambda \mathbf{B}^1 (R^{11})^{-1} \mathbf{B}^{1\top} \mathbf{S}_\lambda = \dot{\mathbf{S}}_\lambda; \quad \mathbf{S}_\lambda(T) = \Lambda \mathbf{Q}_f^1, \end{aligned} \quad (7.13)$$

la matriz Λ es una matriz de peso que contiene el vector $\lambda = \lambda^*$ que es la solución para el siguiente problema de optimización de dimension finita:

$$\lambda^* = \arg \min_{\lambda \in S^M} J^1(\lambda). \quad (7.14)$$

Demostración Se sigue de [30].

Tenga en cuenta que las estrategias de equipo óptimo mini-max son actualmente una suma ponderada de las estrategias que son el equipo óptimo para cada conjunto de parámetros en el juego diferencial (7.3), y pueden ser expresados como la suma ponderada: $u^{ite} = - \sum_{\alpha} (R^{1i})^{-1} B^{i\top}(t, \alpha) S_\lambda^\alpha x^{te}(t, \alpha)$, donde la parametrización de la matriz S_λ^α viene de la matriz Λ en (7.13).

7.2.1. DISEÑO DE INCENTIVOS PARA HACER CUMPLIR LA SOLUCIÓN DE EQUIPO ÓPTIMO MINI-MAX DEL LÍDER

Una vez que la mejor posible minimización para el índice de desempeño del líder a sido obtenida, el siguiente problema que enfrenta el líder es diseñar una estrategia del incentivo para forzar su solución minimax. Para lograr esto, considere la siguiente suma de estrategia del incentivo para el control del líder:

$$\begin{aligned} u^1 = \sum_{\alpha} \left[- (R^{11})^{-1} B^{1\top}(t, \alpha) S_\lambda^\alpha x(t, \alpha) + P^\top(t, \alpha) [x(t, \alpha) - x^{te}(t, \alpha)] \right], \\ P(t, \alpha) \in \mathbb{R}^{n \times m}. \end{aligned} \quad (7.15)$$

El problema ahora es encontrar un conjunto de matrices $P(t, \alpha)$ ($\alpha \in \mathcal{M} = \{1, \dots, M\}$), tal que la solución en equipo optima mini-max para el líder sea obtenida. El seguidor a su vez esta optimizando en peor caso su función objetivo $h^2(u^1, u^2, \alpha)$:

$$h^2(u^1, u^2, \alpha) = \frac{1}{2} x^\top(t, \alpha) (T) Q_f^2 x(t, \alpha) (T) + \frac{1}{2} \int_{t=t_0}^T [x^\top(t, \alpha) Q^2 x(t, \alpha) + u^{1te\top} R^{21} u^{1te} + u^{2\top} R^{22} u^2] dt, \quad (7.16)$$

Con el fin de resolver este problema considere de nuevo la aproximación del Hamiltoniano generalizado para la optimización en el peor caso del costo del seguidor:

$$\begin{aligned} \mathcal{H}_2^\diamond = \sum_{\alpha} \mathcal{H}_2^\alpha \\ \mathcal{H}_2^\alpha = \left[\frac{1}{2} \mu_{n+1}(t, \alpha) (x^\top(t, \alpha) Q^2 x(t, \alpha) + u^{1te\top} R^{21} u^{1te} + u^{2\top} R^{22} u^2) + \right. \\ \left. \mu^\top(t, \alpha) (A(t, \alpha) x(t, \alpha) + B^1(t, \alpha) u^1 + B^2(t, \alpha) u^2) \right]. \end{aligned} \quad (7.17)$$

Después de las condiciones generales necesarias mini-max dadas en el trabajo mencionado anteriormente, las variables adjuntas $\mu(t, \alpha)$ para cada parámetro fijo α , satisfacen la propiedad:

$$\begin{aligned} \dot{\mu}(t, \alpha) = & -\frac{\partial}{\partial x(T, \alpha)} \mathcal{H}_2^\alpha = -\left\{ \mu_{n+1}(T, \alpha) Q^2 x(T, \alpha) + \right. \\ & \left(\sum_{\alpha} P^\top(T, \alpha) - \sum_{\alpha} (R^{11})^{-1} B^{1\top}(t, \alpha) S_\lambda^\alpha \right)^\top R^{21} \left(\sum_{\alpha} P^\top(t, \alpha) - \sum_{\alpha} (R^{11})^{-1} B^{1\top}(t, \alpha) S_\lambda^\alpha \right) x^\alpha + \\ & \left(\sum_{\alpha} P^\top(t, \alpha) - \sum_{\alpha} (R^{11})^{-1} B^{1\top}(t, \alpha) S_\lambda^\alpha \right)^\top R^{21} \sum_{\alpha} P^\top(t, \alpha) x^{te}(t, \alpha) - \\ & \left. \mu^\top(t, \alpha) \left(A(t, \alpha) - B^1(t, \alpha) \left(\sum_{\alpha} (R^{11})^{-1} B^{1\top}(t, \alpha) S_\lambda^\alpha + \sum_{\alpha} P^\top(t, \alpha) \right) \right) \right\}, \end{aligned} \quad (7.18)$$

para la ultima coordenada la derivada temporal es:

$$\dot{\mu}_{n+1}(t, \alpha) = 0,$$

con la condición de transversalidad:

$$\begin{aligned} \mu(T, \alpha) &= -\nu(\alpha) \text{grad } h^2(x(T, \alpha), \alpha) \\ &= -\nu(\alpha) Q_f^2 x(T, \alpha), \\ \mu_{n+1}(T, \alpha) &= -\nu(\alpha), \end{aligned}$$

donde la constante $\nu(\alpha)$ es un valor real no negativo que aparece en el caso general de las condiciones de transversalidad (Teorema 4 Condición iii) [30]). Tomando la suma de la función Hamiltoniana individual para cada α^i , ($\alpha = 1, \dots, M$), y normalizando las variables adjuntas. El control robusto que satisface la condición de maximalidad conduce a:

$$u^2 = (R^{22})^{-1} \sum_{\alpha} \theta^\alpha B^{2\top}(t, \alpha) \mu(t, \alpha),$$

donde el vector $\theta = (\theta(1), \dots, \theta(M))^\top$ pertenece al simplex:

$$\theta \in S^M := \left\{ \theta \in \mathbb{R}^M, \theta(\alpha) \geq 0, \sum_{\alpha} \theta(\alpha) = 1 \right\}. \quad (7.19)$$

Uno de los posibles procedimientos para tratar de encontrar un conjunto de matrices de incentivos es, primero asumir que tal conjunto de matrices existe de modo que la solución de equipo optima minimax es logrado. En ese caso se obtiene la siguiente condición correspondiente para el control del seguidor:

$$(R^{22})^{-1} \sum_{\alpha} \theta^\alpha B^{2\top}(t, \alpha) \mu(t, \alpha) = (R^{12})^{-1} \sum_{\alpha} B^{2\top}(t, \alpha) S_\lambda^\alpha x^{te}(t, \alpha). \quad (7.20)$$

Se definen ahora los siguientes vectores y matrices extendidos:

$$\begin{aligned} \mathbf{P} &:= \begin{bmatrix} P(t, 1) & \dots & P(t, M) \end{bmatrix} \in \mathbb{R}^{n \times M \cdot m}; \quad \Theta := \begin{bmatrix} \theta(1)I & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \theta(M)I \end{bmatrix} \\ \mu_k^\top &:= (\mu_1^\top(t, 1), \dots, \mu_n^\top(t, 1); \dots; \mu_1^\top(t, M), \dots, \mu_n^\top(t, M)) \in \mathbb{R}^{1 \times nM} \\ \mathbf{x}^{te\top} &:= (x^\top(t, 1), \dots, x^\top(t, M))^\top \in \mathbb{R}^{n \cdot M} \end{aligned} \quad (7.21)$$

con esta definición del vector y matrices extendidas, considere la hipótesis de la existencia del conjunto de matrices de incentivos \mathbf{P} , entonces el conjunto de trayectorias generadas en la solución en equipo optima $x^{te}(t, \alpha)$ coincidirá con el conjunto de trayectorias de este problema de optimización $x(t, \alpha)$. Por lo tanto, se obtiene:

$$\dot{\mu} = \left(\Theta \mathbf{Q}^2 + \left((R^{11})^{-1} \mathbf{B}^{1\tau} \mathbf{S}_\lambda - \mathbf{P}^\tau \right)^\top R^{21} (R^{11})^{-1} \mathbf{B}^{1\tau} \mathbf{S}_\lambda \right) \mathbf{x}^{te} - \left(\mathbf{A} - \mathbf{B}^1 (R^{11})^{-1} \mathbf{B}^{1\tau} \mathbf{S}_\lambda + \mathbf{B}^1 \mathbf{P}^\tau \right)^\top \mu; \quad \mu(T) = \Theta \mathbf{Q}_f^2, \quad (7.22)$$

y el control del seguidor toma la forma:

$$u^2 = (R^{22})^{-1} \mathbf{B}^{2\tau} \Theta \mu. \quad (7.23)$$

Note que la dependencia en el parámetro incierto α ha desaparecido, la nueva dinámica definida incluye la familia completa de plantas pero el control es el mismo para todas las plantas. Esto significa que el mismo control se aplica simultáneamente para todas las dinámicas. Es posible presentar ahora el siguiente resultado

Teorema 7.2 Si para el juego lineal cuadrático de Stackelberg (7.3), existe un conjunto de matrices de incentivos $P(t, \alpha)$ ($\alpha = 1, \dots, M$), tal que la solución para la siguiente ecuación diferencial lineal matricial para la matriz $\mathbf{M}_\theta \in \mathbb{R}^{n \cdot M \times n \cdot M}$ se satisface:

$$\begin{aligned} & \mathbf{M}_\theta \mathbf{X} + \mathbf{X}^\top \mathbf{M}_\theta + \mathbf{P} \left(\mathbf{B}^{1\tau} \mathbf{M}_\theta - R^{21} (R^{11})^{-1} \mathbf{B}^{1\tau} \mathbf{S}_\lambda \right) + \\ & \Theta \mathbf{Q}^2 + \mathbf{S}_\lambda \left(\mathbf{B}^1 (R^{11})^{-1} R^{21} (R^{11})^{-1} \mathbf{B}^{1\tau} + \right. \\ & \left. \mathbf{B}^2 (R^{12})^{-1} R^{22} (R^{12})^{-1} \mathbf{B}^{2\tau} \right) \mathbf{S}_\lambda = -\dot{\mathbf{M}}_\theta, \quad \mu(T) = \Theta \mathbf{Q}_f^2, \end{aligned} \quad (7.24)$$

donde:

$$\mathbf{X} = \mathbf{A} - \left(\mathbf{B}^1 (R^{11})^{-1} \mathbf{B}^{1\tau} + \mathbf{B}^2 (R^{12})^{-1} \mathbf{B}^{2\tau} \right) \mathbf{S}_\lambda, \quad (7.25)$$

con la condición de restricción adicional siguiente:

$$(R^{22})^{-1} (t) \mathbf{B}^{2\tau} (t) \mathbf{M}_\theta (t) = (R^{12})^{-1} (t) \mathbf{B}^{2\tau} (t) \mathbf{S}_\lambda (t), \quad (7.26)$$

donde la matriz Θ es una matriz de peso que contiene el vector $\theta = \theta^*$ que es la solución para el siguiente problema de optimización de dimension finita:

$$\theta^* = \arg \min_{\theta \in S^M} J^2(\theta). \quad (7.27)$$

Entonces, el control de incentivos para que el líder realice la solución de equipo optima mini-max (7.12) y el control óptimo mini-max de la reacción del seguidor, están dados por:

$$\begin{aligned} u^1 &= - (R^{11})^{-1} \mathbf{B}^{1\tau} \mathbf{S}_\lambda \mathbf{x} + \mathbf{P}^\tau [\mathbf{x} - \mathbf{x}^{te}], \\ u^2 &= - (R^{12})^{-1} \mathbf{B}^{2\tau} \mathbf{S}_\lambda \mathbf{x}^{te}, \end{aligned}$$

donde las matrices $\Lambda(\lambda^*)$ y $\Theta(\theta^*)$ definidas en (7.10) y (7.21) son solución de los problemas (7.14) y (7.27).

Demostracion Considerando (7.23), vamos a tratar de representar $\Theta \mu(t) = \mathbf{M}_\theta \mathbf{x}^{te}$ en (7.22), donde \mathbf{M}_θ es una $\mathbb{R}^{n \cdot M \times n \cdot M}$ función matricial. Sustituyendo la representación propuesta por $\Theta \mu(t)$ en (7.22) se sigue el resultado.

7.3. INCENTIVO MINI-MAX PARA UN JUEGO ESTACIONARIO DE STACKELBERG CON HORIZONTE DE TIEMPO INFINITO.

Esta sección presenta la solución de estado estacionario de las ecuaciones mencionada anteriormente. La versión estacionaria del problema anterior se ve mucho más fácil de resolver (que en el caso de tiempo finito) en una simulación numérica. Para ello, considere que la variable T tiende a infinito y las matrices del juego son estacionarias, es decir:

$$A(t, \alpha) = A(\alpha), \quad B^j(t, \alpha) = B^j(\alpha), \\ j = 1, 2$$

Para la solución de equipo óptimo se obtiene el siguiente conjunto de controles:

$$u^{1te} = - (R^{11})^{-1} \mathbf{B}^{1\top} \mathbf{S}_\lambda x^{te}, \\ u^{2te} = - (R^{12})^{-1} \mathbf{B}^{2\top} \mathbf{S}_\lambda x^{te}, \quad (7.28)$$

con \mathbf{S}_λ como la solución de

$$-\mathbf{S}_\lambda \mathbf{A} - \mathbf{A}^\top \mathbf{S}_\lambda - \Lambda \mathbf{Q}^1 + \mathbf{S}_\lambda \mathbf{B}^2 (R^{12})^{-1} \mathbf{B}^{2\top} \mathbf{S}_\lambda + \\ \mathbf{S}_\lambda \mathbf{B}^1 (R^{11})^{-1} \mathbf{B}^{1\top} \mathbf{S}_\lambda = \mathbf{0} \quad (7.29)$$

donde el vector λ resuelve el siguiente problema de optimización:

$$\lambda^* = \arg \min_{\lambda \in SM} J_\infty^1(\lambda), \quad (7.30)$$

Aquí asumimos que la condición de estabilidad para el sistema extendido se mantiene, es decir, $\{\mathbf{A}, (\mathbf{B}^1, \mathbf{B}^2)\}$ es estabilizable, de modo que la matriz extendida $\mathbf{X} = \mathbf{A} - (\mathbf{B}^1 (R^{11})^{-1} \mathbf{B}^{1\top} + \mathbf{B}^2 (R^{12})^{-1} \mathbf{B}^{2\top}) \mathbf{S}_\lambda$ es estable. El siguiente paso es diseñar un control de incentivos que se espera que obtenga la solución de equipo óptima en tiempo infinito. Al igual que en la última sección el control propuesto es:

$$u^1 = \sum_{\alpha} \left[- (R^{11})^{-1} \mathbf{B}^{1\top}(\alpha) \mathbf{S}_\lambda^\alpha x(t, \alpha) + P^\top(\alpha) [x(t, \alpha) - x^{te}(t, \alpha)] \right], \\ P(\alpha) \in \mathbb{R}^{n \times m} \quad (7.31)$$

Para el conjunto de matrices estacionarias $P(\alpha)$ ($\alpha = 1, \dots, M$), el problema es estudiar la existencia de dichas matrices, para que la reacción óptima mini-max del seguidor coincide con (7.28). Procediendo de la misma manera que en el horizonte de tiempo finito obtenemos el siguiente corolario:

Corolario 7.1 Si para el juego lineal cuadrático Stackelberg (7.3) formulada para una clase de modelos estacionarios multi-lineales con un índice cuadrático (7.4) dentro del horizonte de tiempo infinito, existe un conjunto de matrices $P(\alpha)$ ($\alpha = 1, \dots, M$), tal que la única solución de la ecuación lineal matricial extendida:

$$\mathbf{M}_\theta \mathbf{X} + \mathbf{X}^\top \mathbf{M}_\theta + \mathbf{P} \left(\mathbf{B}^{1\top} \mathbf{M}_\theta - R^{21} (R^{11})^{-1} \mathbf{B}^{1\top} \mathbf{S}_\lambda \right) + \\ \Theta \mathbf{Q}^2 + \mathbf{S}_\lambda \left(\mathbf{B}^1 (R^{11})^{-1} R^{21} (R^{11})^{-1} \mathbf{B}^{1\top} + \right. \\ \left. \mathbf{B}^2 (R^{12})^{-1} R^{22} (R^{12})^{-1} \mathbf{B}^{2\top} \right) \mathbf{S}_\lambda = \mathbf{0}, \quad (7.32)$$

satisface las restricciones:

$$(R^{22})^{-1} \mathbf{B}^{2\tau} \mathbf{M}_\theta = (R^{12})^{-1} \mathbf{B}^{2\tau} \mathbf{S}_\lambda, \quad (7.33)$$

donde el vector $\theta = \theta^*$ resuelve el problema de optimización:

$$\theta^* = \arg \min_{\theta \in SM} J_\infty^2(\theta). \quad (7.34)$$

Entonces, para el juego de Stackelberg en horizonte de tiempo infinito, el control de incentivo para que el líder realice la solución de equipo optima mini-max (7.28) y la reacción de control optima mini-max del seguidor, esta dada por:

$$\begin{aligned} u^1 &= - (R^{11})^{-1} \mathbf{B}^{1\tau} \mathbf{S}_\lambda \mathbf{x} + \mathbf{P}^\tau [\mathbf{x}^\alpha - \mathbf{x}^{te, \alpha}], \\ u^2 &= - (R^{12})^{-1} \mathbf{B}^{2\tau} \mathbf{S}_\lambda \mathbf{x}^{te}, \end{aligned} \quad (7.35)$$

donde las matrices $\Lambda(\lambda^*)$ y $\Theta(\theta^*)$ definidas en (7.10) y (7.21) son solución de los problemas (7.30) y (7.34) y la matriz \mathbf{S}_λ es una solución de (7.29).

Como se demostró en [49], es posible obtener un conjunto de ecuaciones mas convenientes para resolver el caso de tiempo infinito. Esto se hace de la siguiente manera, sumar y restar de (7.29) el termino, $\mathbf{S}_\lambda [\mathbf{B}^1 (R^{11})^{-1} \mathbf{B}^{1,T} + \mathbf{B}^2 (R^{12})^{-1} \mathbf{B}^{2,T}] \mathbf{S}_\lambda$ obteniendo

$$-\mathbf{S}_\lambda [\mathbf{B}^1 R_{11}^{-1} \mathbf{B}^{1\tau} + \mathbf{B}^2 (R^{12})^{-1} \mathbf{B}^{2\tau}] \mathbf{S}_\lambda - \mathbf{S}_\lambda \mathbf{X} - \mathbf{F}^T \mathbf{S}_\lambda - \Lambda \mathbf{Q}^1 = 0.$$

Sumando y restando esta ecuación a (7.32) obtenemos:

$$\begin{aligned} \mathbf{Z}\mathbf{X} + \mathbf{X}^\tau \mathbf{Z} + \Theta \mathbf{Q}^2 - \Lambda \mathbf{Q}^1 + \mathbf{P}\mathbf{B}^{1\tau} \mathbf{Z} + \mathbf{P}[I - R^{12} (R^{11})^{-1}] \mathbf{B}^{1\tau} \mathbf{S}_\lambda \\ + \mathbf{S}_\lambda \mathbf{B}^1 (R^{11})^{-1} [R^{21} (R^{11})^{-1} - I] \mathbf{B}^{1\tau} \mathbf{S}_\lambda + \\ \mathbf{S}_\lambda \mathbf{B}^2 (R^{12})^{-1} [R^{22} (R^{12})^{-1} - I] \mathbf{B}^{2\tau} \mathbf{S}_\lambda = 0, \end{aligned}$$

donde $\mathbf{Z} = (\mathbf{M}_\theta - \mathbf{S}_\lambda)$, con la restricción:

$$(R^{12})^{-1} \mathbf{B}^{2,T} \mathbf{S}_\lambda = (R^{22})^{-1} \mathbf{B}^{2,T} \mathbf{M}_\theta,$$

si consideramos el caso cuando $R^{22} = R^{12}$ obtenemos la ecuación:

$$\begin{aligned} \mathbf{Z}\mathbf{X} + \mathbf{X}^\tau \mathbf{Z} + \Theta \mathbf{Q}^2 - \Lambda \mathbf{Q}^1 + \mathbf{P}\mathbf{B}^{1\tau} \mathbf{Z} + \mathbf{P}[I - R^{12} (R^{11})^{-1}] \mathbf{B}^{1\tau} \mathbf{S}_\lambda \\ + \mathbf{S}_\lambda \mathbf{B}^1 (R^{11})^{-1} [R^{21} (R^{11})^{-1} - I] \mathbf{B}^{1\tau} \mathbf{S}_\lambda = 0, \end{aligned} \quad (7.36)$$

con la restricción:

$$\mathbf{B}^{2\tau} \mathbf{Z} = \mathbf{0}, \quad (7.37)$$

este conjunto de ecuaciones se puede resolver como es demostrado en el ejemplo y ayuda a encontrar la matriz de incentivos. A continuación describimos un procedimiento para encontrar las matrices de incentivos mini-max para el juego con 2 jugadores.

Observación 7.1 El numero de variables desconocidas en (7.36) son $Mn \times Mn + Mn + m$ y el numero total de ecuaciones son $Mn \times Mn$, así que usamos la restricción (7.37) que nos dan $m \times Mn$ ecuaciones adicionales que junto con las ecuaciones anteriores se pueden resolver.

7.3.1. ALGORITMO 1 SOLUCIÓN DEL PROBLEMA DE EQUIPO ÓPTIMO MINI-MAX

El algoritmo para encontrar los controles en equipo óptimo mini-max se resume de la siguiente manera:

1. Se selecciona una condición inicial para los pesos en equipo óptimo.
2. Se aplican las acciones de control igual a la combinación ponderada de las estrategias en equipo optima para calcular los correspondientes funcionales de costo.
3. Correr por los valores de lambda ($\lambda \in [0, 1]$) para encontrar el valor (7.11) para que el problema $\lambda^* = \arg \min_{\lambda \in S^M} J_{\infty}^1(\lambda)$ sea resuelto.
4. Se encuentra la matriz S_{λ} (7.29).

7.3.2. ALGORITMO 2 ENCONTRAR LA ESTRATEGIA DE INCENTIVOS MINI-MAX.

El algoritmo para encontrar la matriz de incentivos mini-max se resume de la siguiente manera: (7.36). Usando

1. Se aplica la restricción (7.37) para reducir el numero de variables de Z .
2. Se aplica el conjunto de ecuaciones obtenido a (7.36).
3. Se resuelve para Z y P usando métodos algebraicos computacionales.
4. Se encuentra la matriz $M_{\theta} = Z + S_{\lambda}$.

7.4. EJEMPLOS NUMÉRICOS

7.4.1. EJEMPLO ESCALAR.

Considere el siguiente juego multi-modelo de dos jugadores

$$\begin{aligned}\dot{x}(t, \alpha) &= A(\alpha)x(t, \alpha) + B^1(\alpha)u^1(t) + B^2(\alpha)u^2(t), \\ \alpha &= 1, 2; \quad x \in \mathbb{R}, \quad u^1, u^2 \in \mathbb{R}\end{aligned}$$

donde los coeficientes están dados por: $A(1) = 2$, $A(2) = 1,8$, $B^1(1) = 1,9$, $B^2(1) = 1,8$, $B^1(2) = 2$, $B^2(2) = 1,5$. con un costo cuadrático:

$$\begin{aligned}h^i(u^1, u^2, \alpha) &= \frac{1}{2} \int_0^{\infty} \left[x^{\top}(t, \alpha) Q^i x(t, \alpha) + \sum_{j=1}^2 u^{j\top} R^{ij} u^j \right] dt, \\ i &= 1, 2; \quad \alpha = 1, 2\end{aligned}$$

donde $Q^1 = Q^2 = R^{12} = R^{21} = R^{22} = 1$. El sistema extendido esta dado por:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}^1 u_1(t) + \mathbf{B}^2 u_2(t),$$

donde $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 1,8 \end{pmatrix}$, $\mathbf{B}^1 = \begin{pmatrix} 1,9 \\ 2 \end{pmatrix}$, $\mathbf{B}^2 = \begin{pmatrix} 1,8 \\ 1,5 \end{pmatrix}$.

ALGORITMO 1: Ejecutando el algoritmo 1 la solución al problema (7.30), el parámetro obtenido es $\lambda = 0,0521$ y la solución para la matriz \mathbf{S}_λ es:

$$\mathbf{S}_\lambda = \begin{pmatrix} 39,77115300 & -39,72347798 \\ -39,72347798 & 40,45673335 \end{pmatrix},$$

ALGORITMO 2: Aplicando la restricción (7.37) y sustituyendo en (7.36) obtenemos el siguiente conjunto de ecuaciones no lineales:

$$\begin{aligned} 15,016 + 2\theta_1 + p_1(3,88177 + 0,416667z_{21}) + 10,1422z_{21} + 8,53396z_{22} &= 0, \\ -21,1124 - 7,61421z_{21} + p_1(-5,43886 + 0,416667z_{22}) - 6,00595z_{22} &= 0, \\ -21,1124 + p_2(3,88177 + 0,416667z_{21}) - 12,696z_{21} - 10,2408z_{22} &= 0, \\ 30,6333 - 2\theta_2 + 9,13706z_{21} + p_2(-5,43886 + 0,416667z_{22}) + 6,68183z_{22} &= 0, \end{aligned}$$

tenga en cuenta que es un conjunto de ecuaciones consistente para θ un parámetro dado. Este sistema se resuelve mediante rutina numérica de encontrar raíces de Mathematica basada en los métodos de Newton para encontrar las soluciones. La solución para \mathbf{M}_θ y \mathbf{P} se puede calcular para todos los parámetros $\theta_1 \in [0, 1]$ y $\theta_2 = 1 - \theta_1$ y esta dado por:

θ_1	θ_2	p_1	p_2	M_{11}	M_{12}	M_{21}	M_{22}
0	1	-4.12229249	6.34805647	73.1559	-72.7767	-79.7852	80.1206
1	0	-36.06984460	40.94384724	45.2014	-48.6246	-46.2398	51.1381
.5	.5	0.97618174	0.80513753	55.9188	-57.7498	-59.1006	62.0884
.3	.7	-2.04453584	4.08294239	61.2061	-62.3083	-65.4454	67.5585
.7	.3	10.47702536	-9.49328814	51.363	-53.8552	-53.6337	57.4148

Table 1.

En la tabla 1 se observa que para cualquier valor del parámetro θ podemos encontrar la matriz del incentivo P para que el líder siempre obtenga el costo mini-max previamente calculado. El valor numérico del índice del seguidor es $J^2 = 0,77848$, como se muestra en la figura 5. La figura 1 muestra las trayectorias del estado mini-max después de que el líder ha implementado el incentivo, la figura 2 los controles para cada jugador. La figura 3 representa el criterio y la figura 4 los valores de J^1 como una función del parámetro λ . Finalmente la figura 5 muestra el valor del criterio para el seguidor, observamos que para cualquier $\theta_1 \in [0, 1]$ y $\theta_2 = 1 - \theta_1$, se encuentra el mismo valor.

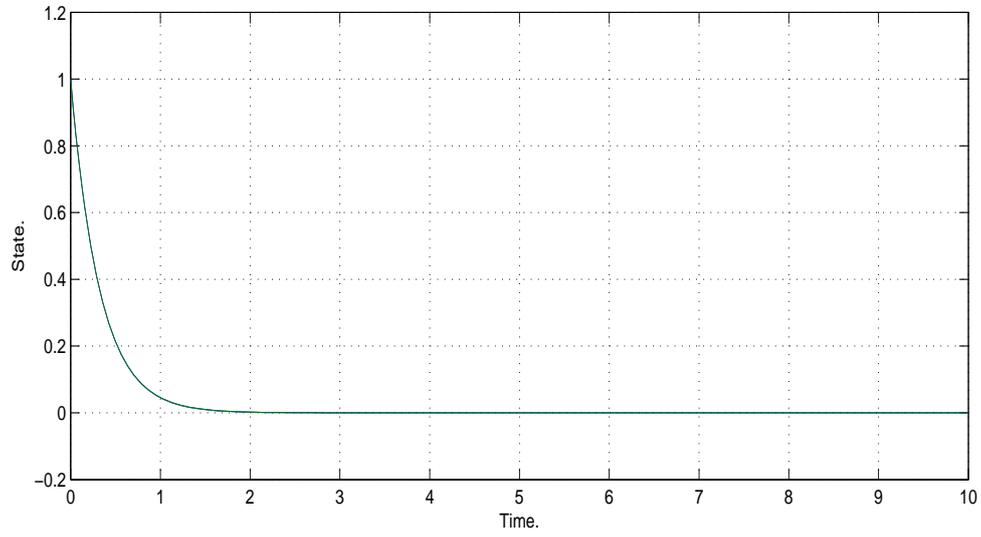


Figura 7.1: Trayectorias del estado del juego.

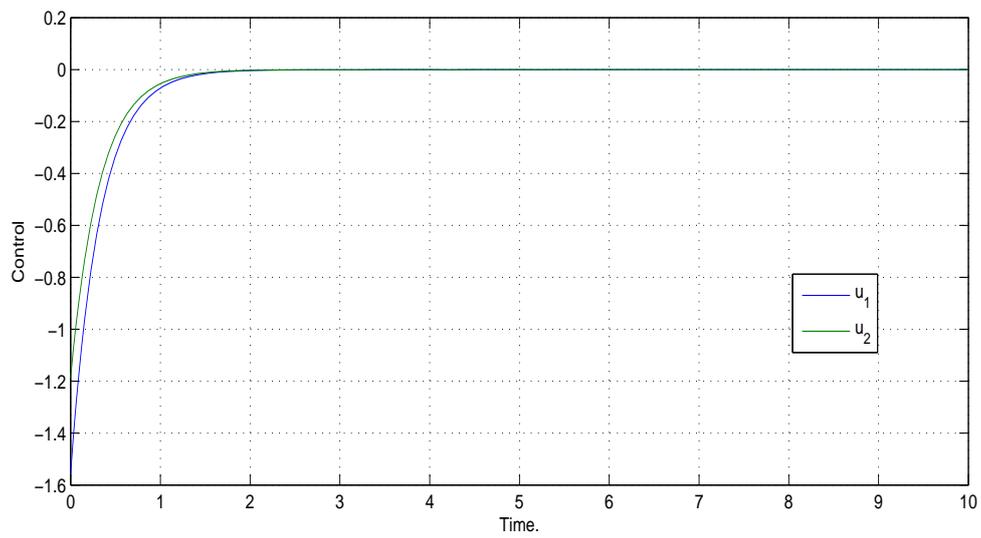


Figura 7.2: Control de los dos jugadores.

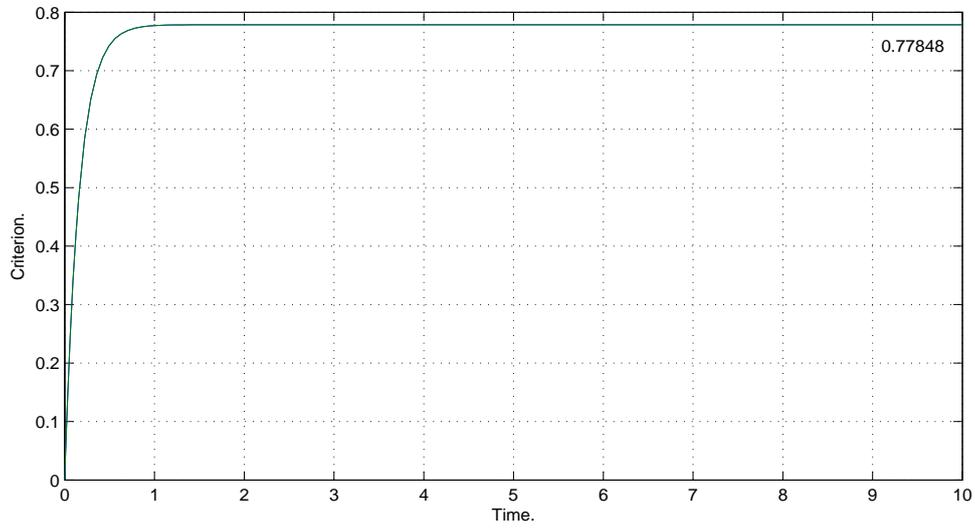


Figura 7.3:

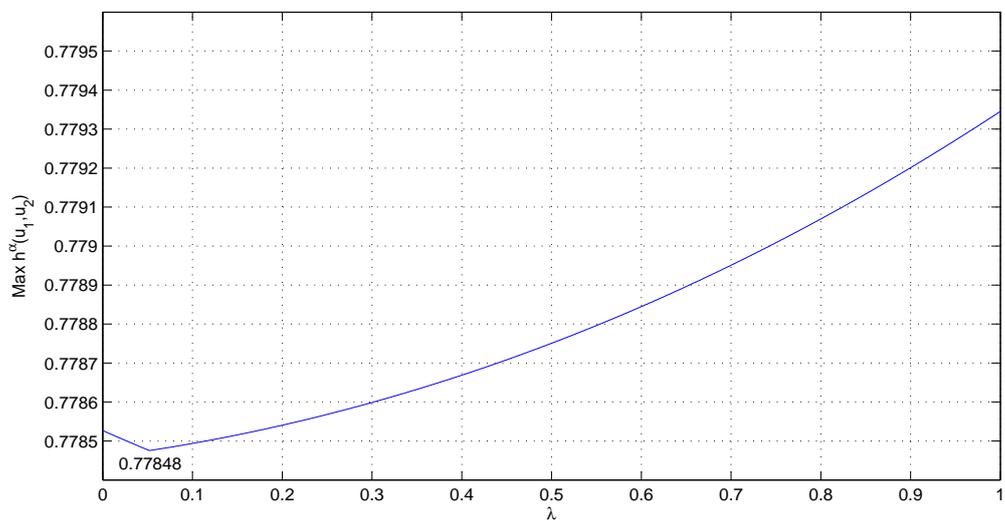


Figura 7.4: Índice Mini-max como función de λ ($\lambda^* = 0,0521$).

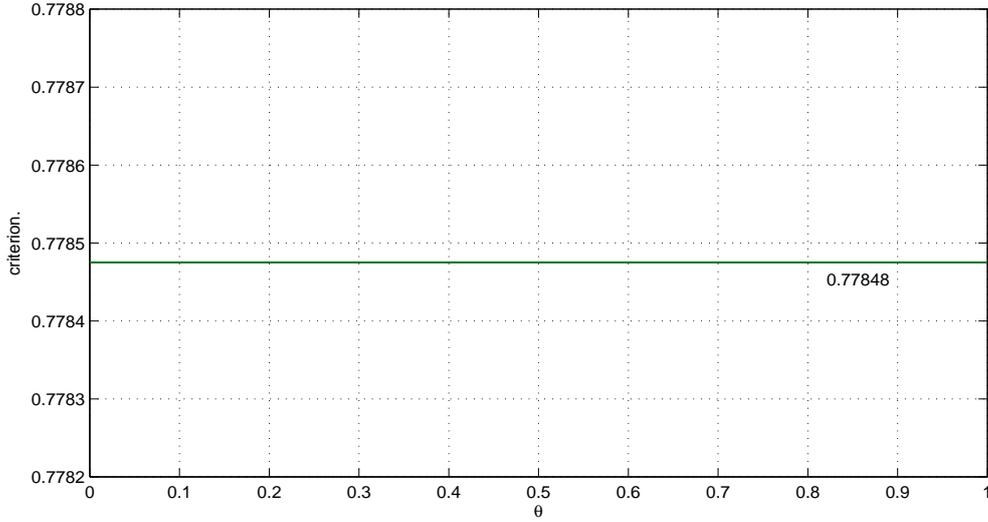


Figura 7.5: Criterio del seguidor J^2 como función de θ .

7.4.2. EJEMPLO VECTORIAL.

Considere el juego de dos estados multi-planta dado por:

$$\begin{aligned} \dot{x}(t, \alpha) &= A(\alpha)x(t, \alpha) + B^1(\alpha)u_1(t) + B^2(\alpha)u_2(t), \\ \alpha &= 1, 2; \quad x \in \mathbb{R}^2, \quad u^1, u^2 \in \mathbb{R}^2 \end{aligned}$$

donde las matrices involucradas son: $A(1) = \begin{pmatrix} 0 & 1 \\ 0,5 & 1,3 \end{pmatrix}$, $A(2) = \begin{pmatrix} 0 & 1,1 \\ 1 & 1,5 \end{pmatrix}$, $B^1(1) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$,
 $B^2(1) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $B^1(2) = \begin{pmatrix} 1,1 & ,9 \\ 1,3 & 0 \end{pmatrix}$, $B^2(2) = \begin{pmatrix} 0,9 & 1,1 \\ 0,9 & 0 \end{pmatrix}$. Con el siguiente criterio cuadrático:

$$\begin{aligned} h^i(u^1, u^2, \alpha) &= \frac{1}{2} \int_0^{\infty} \left[x^\top(t, \alpha) Q^i x(t, \alpha) + \sum_{j=1}^2 u^{j\top} R^{ij} u^j \right] dt, \\ i &= 1, 2; \quad \alpha = 1, 2 \end{aligned}$$

donde $Q^1 = Q^2 = R^{11} = R^{12} = R^{21} = R^{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. El sistema extendido tiene matrices definidas como:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0,5 & 1,3 & 0 & 0 \\ 0 & 0 & 0 & 1,1 \\ 0 & 0 & 1 & 1,5 \end{pmatrix}, \quad \mathbf{B}^1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1,1 & ,9 \\ 1,3 & 0 \end{pmatrix}, \quad \mathbf{B}^2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0,9 & 1,1 \\ 0,9 & 0 \end{pmatrix},$$

ALGORITMO 1: El algoritmo 1 ofrece la solución al problema (7.30), el parámetro obtenido es $\lambda = 0,663$ y la solución para la matriz S_λ

$$S_\lambda = \begin{pmatrix} 3,32038952 & 8,30444247 & -3,82944445 & -7,66613310 \\ 8,30444247 & 26,10320776 & -10,96514265 & -22,73269429 \\ -3,82944445 & -10,96514265 & 5,11983685 & 10,05843260 \\ -7,66613310 & -22,73269429 & 10,05843260 & 21,18613248 \end{pmatrix},$$

ALGORITHM 2: Aplicando la restricción (7.37) y sustituyendo en (7.36) obtenemos el sistema de ecuaciones no lineales que puede ser resuelto (pero es demasiado largo para mostrarlo aquí). De nuevo se utiliza la rutina de encontrar raíces de Mathematica para encontrar las soluciones. La solución de la matriz M_θ y la matriz de incentivos P depende de los valores del parámetro $\theta_1 \in [0, 1]$ y $\theta_2 = 1 - \theta_1$. Es posible obtener la matriz de incentivos para cualquier valor del parámetro θ la siguiente tabla muestra para cuatro puntos:

θ_1	θ_2	P	M_θ	$J^2(u_1, u_2)$
0	1	$\begin{pmatrix} 1,07035 & 5,2015 \\ -8,71166 & -5,15142 \\ 0,148545 & -3,67385 \\ 9,17596 & 4,50658 \end{pmatrix}$	$\begin{pmatrix} 6,39912 & 12,2472 & -9,69466 & -13,0314 \\ 4,33416 & 19,8332 & -5,95534 & -15,2323 \\ -6,62829 & -14,5495 & 10,4519 & 14,936 \\ -3,87667 & -16,5626 & 5,67687 & 13,9362 \end{pmatrix}$	2.0648
.1	.9	$\begin{pmatrix} 1,04175 & 5,19136 \\ -8,73403 & -5,12626 \\ 0,236086 & -3,75641 \\ 9,19588 & 4,51023 \end{pmatrix}$	$\begin{pmatrix} 5,95267 & 12,175 & -9,32277 & -12,6851 \\ 4,88568 & 20,5594 & -6,43698 & -16,1182 \\ -6,22242 & -14,4838 & 10,1138 & 14,6211 \\ -4,39928 & -17,3548 & 6,1369 & 14,8506 \end{pmatrix}$	2.0648
.3	.7	$\begin{pmatrix} 0,994946 & 5,26048 \\ -8,77116 & -5,10871 \\ 0,416953 & -4,02996 \\ 9,22385 & 4,54418 \end{pmatrix}$	$\begin{pmatrix} 4,93123 & 12,3476 & -8,58126 & -12,1117 \\ 6,07793 & 21,8629 & -7,42653 & -17,8724 \\ -5,29384 & -14,6408 & 9,43967 & 14,0999 \\ -5,51765 & -18,8381 & 7,0866 & 16,6839 \end{pmatrix}$	2.0648
.9	.1	$\begin{pmatrix} -3,14126 & 3,88979 \\ -6,54864 & -3,92493 \\ 4,73952 & -3,53303 \\ 7,53125 & 3,59013 \end{pmatrix}$	$\begin{pmatrix} 0,33541 & 13,3949 & -2,9971 & -9,08903 \\ 10,9347 & 24,8836 & -12,488 & -23,8942 \\ -1,11583 & -15,5928 & 4,36316 & 11,352 \\ -9,98564 & -22,4059 & 11,5823 & 22,7642 \end{pmatrix}$	2.0648

Table 2.

Observe que para todos los valores de θ podemos encontrar una matriz de incentivos que lleva al índice del seguidor al mismo valor.

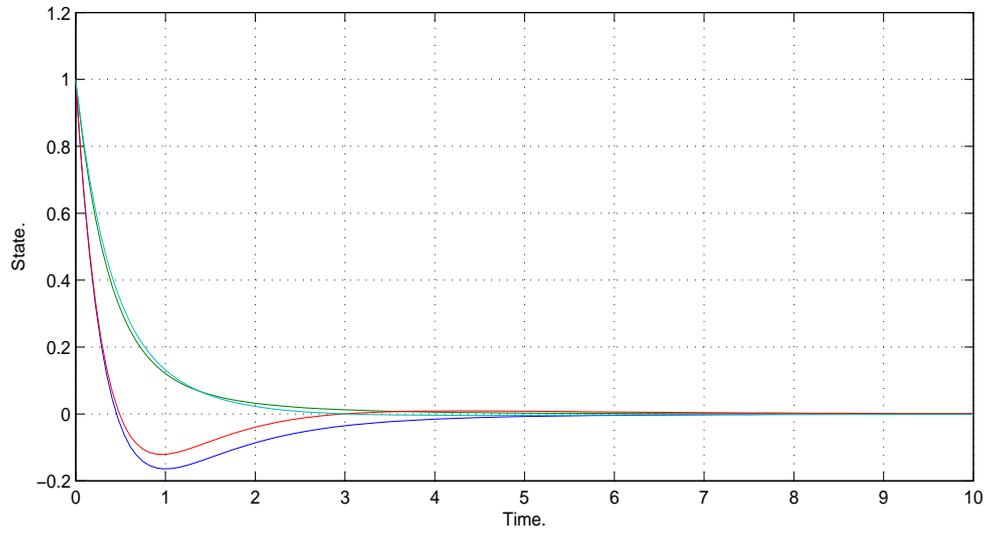


Figura 7.6: Trayectorias de los estados del juego.

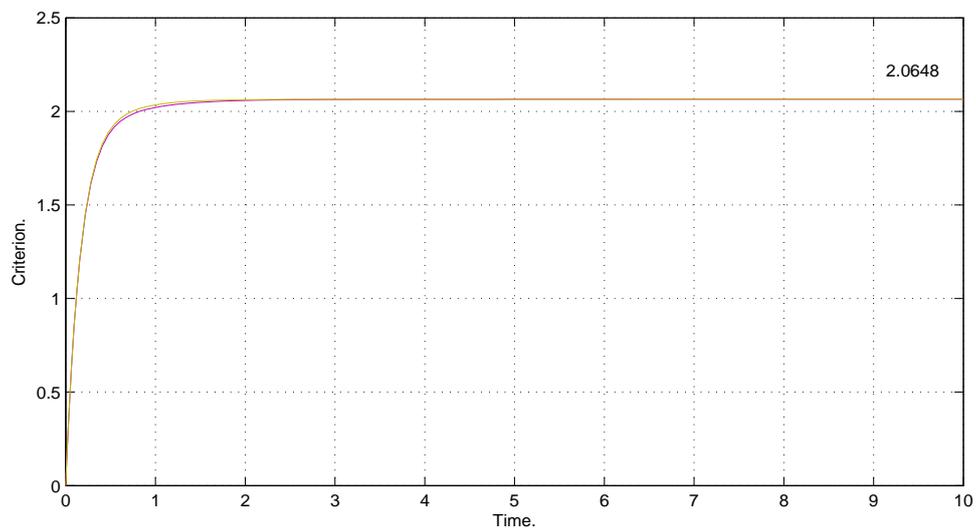


Figura 7.7: Criterios de los dos jugadores.

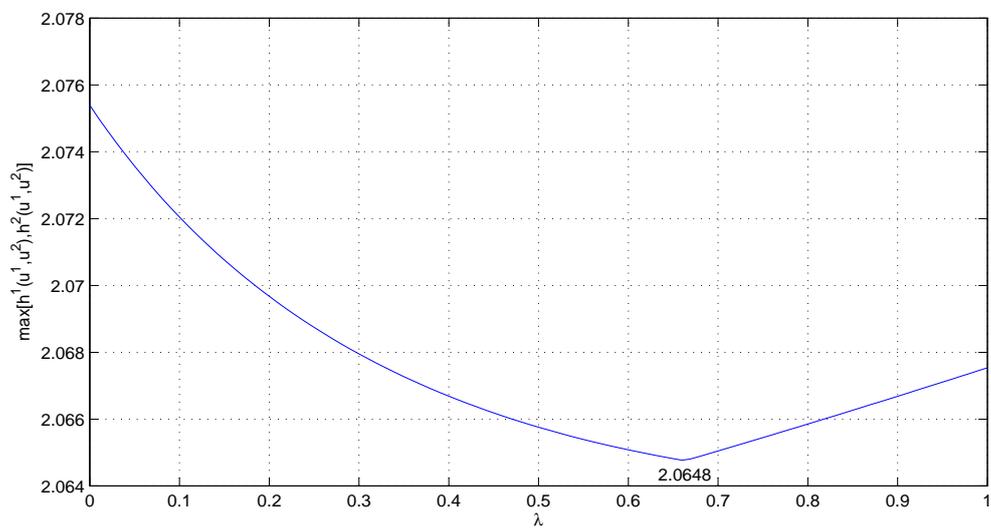


Figura 7.8: Índice Mini-max como función de λ ($\lambda^* = 0,663$).

8

CONCLUSION Y TRABAJO FUTURO.

8.1. CONCLUSIONES.

En esta tesis se diseñan las estrategias para juegos diferenciales polinomiales de suma no cero, también se diseñaron las estrategias para el equilibrio ε -Nash para juegos diferenciales polinomiales estocásticos con información incompleta sobre observaciones lineales ambos diseños en horizonte de tiempo finito e infinito. En ambos casos se demuestra que la llamada ecuación tipo Riccati dependiente del estado provee una técnica para buscar un conjunto de estrategias que garantiza el equilibrio de Nash para esta clase particular de juegos no lineales. Por otra parte de desarrollo una solución para el problema del regulador óptimo robusto para sistemas polinomiales inciertos basados en información de salida y por ultimo se diseño unas estrategias de incentivos para juegos diferenciales de dos jugadores para una familia de plantas sin información de la realización de la planta. Para el ultimo problema se plantea unos incentivos para el juego de Stackelberg en que los parámetros que describen la dinámica del juego dependen de un vector desconocido contenido en un conjunto de parámetros finitos y la solución de la estrategia de incentivos esta dada en términos del peor de los casos.

8.2. TRABAJO FUTURO.

Existen varios trabajos en los cuales se podría extender las herramientas desarrolladas en este trabajo, como por ejemplo:

- Estrategias con Incentivos Min-Max para Juegos Diferenciales con Información Incompleta.
- Estrategias con Incentivos Min-Max con Múltiples seguidores para juegos Diferenciales.
- Controlador Min-Max para Sistemas Polinomiales Con Información Incompleta.
- Equilibrio de Nash en lazo abierto via control Min-Max para Juegos Diferenciales Polinomiales.

8.3. ARTÍCULOS.

Capítulo 3:

- Optimal Control for a Polynomial System with a Quadratic Criterion over Infinite Horizon, International Journal of Systems Science, Publicado, 10.1080/00207721.2013.823528.
- Optimal Control for a Polynomial System with a Quadratic Criterion over Infinite Horizon, 2013 American Control Conference, June 17-June 19, 2013, Washington, DC, USA.

Capítulo 4:

- Open-Loop Nash Equilibrium in Polynomial Differential Games via State-Dependent Riccati Equation, Automatica, Sometido.
- Open-Loop Nash Equilibrium in Polynomial Differential Games via State-Dependent Riccati Equation, 2013 American Control Conference, June 17-June 19, 2013, Washington, DC, USA

Capítulo 5:

- Open-Loop Nash Equilibrium Problem in Polynomial Stochastic Differential Games with Incomplete Information, 2013 Asian Control Conference, June 23-June 26, 2013, Istanbul, Turkey.

Capítulo 6:

- Output mini-max control for polynomial systems: analysis and applications, International Journal of Systems Science, Publicado, <http://dx.doi.org/10.1080/00207721.2012.757385>.
- HOSM observer for robust output regulator in uncertain nonlinear polynomial systems, International Workshop on Variable Structure Systems - VSS , pp. 373-378, 2012.
- Output Mini-Max Control for Polynomial Systems: Analysis and Application, 2012 Conference on Decision and Control, December 10-13, 2012, Maui, Hawaii, USA.

Capítulo 7:

- Mini-Max Incentive Strategy For LQ Stackelberg Differential Games With Uncertainties, Journal of The Franklin Institute, Sometido.

BIBLIOGRAFÍA

- [1] R. Isaacs, *Differential Games*. John Wiley and Sons, Inc., 1965.
- [2] A. Starr and Y. Ho, “Nonzero-sum differential games,” *Journal of Optimization Theory and Applications*, vol. 3, no. 3, pp. 184–206, 1969.
- [3] —, “Further properties of nonzero-sum differential games,” *Journal of Optimization Theory and Applications*, vol. 3, no. 4, pp. 207–219, 1969.
- [4] J. Engwerda, *Lq dynamic optimization and differential games*. Wiley, 2005.
- [5] J. Nash, “Equilibrium points in n-person games,” *Proc. Nat. Acad. U.S.A.*, vol. 36, pp. 48–49, 1950.
- [6] K. Tanaka and K. Yokoyama, “On ε -equilibrium point in a noncooperative N-person game,” *Journal of Mathematical Analysis and Applications*, 160, 413-423, vol. 160, pp. 413–423, 1991.
- [7] M. Jimenez and A. Poznyak, “ ε -equilibrium in lq differential games with bounded uncertain disturbances: robustness of standard strategies and new strategies with adaptation,” *International Journal of Control*, vol. 79, no. 7, pp. 786–797, 2006.
- [8] T. Basar and G. Olsder, *Dynamic Noncooperative Game Theory*. Academic Press, New York, 1982.
- [9] G. Kossiorisa, M. Plexousakis, A. Xepapadeas, A. de Zeeuwe, and K. Maler, “Feedback Nash equilibria for non-linear differential games in pollution control,” *Journal of Economic Dynamics & Control*, vol. 32, pp. 1312–1331, 2008.
- [10] G. Sorger, “Discrete-time dynamic game models for advertising competition in a duopoly,” *Optim Control Methods Appli*, vol. 16, pp. 175–188, 1995.
- [11] H. Zhanga, Q. Weib, and D. Liu, “An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games,” *Automatica*, vol. 47, pp. 207–214, 2011.
- [12] K. M. Ramachandran and C. P. Tsokos, *Stochastic Differential Games. Theory and Applications*, Springer, Ed., 2012.
- [13] R. Kalman, “Contributions to the theory of optimal control,” *Boletín de la Sociedad Matemática Mexicana*, vol. 5, pp. 102–119, 1960.

- [14] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. Wiley-Interscience, New York, 1972.
- [15] R. Bellman, *Dynamic Programming*. Princeton University Press, Princeton, 1957.
- [16] E. Al'brekht, "On the optimal stabilization of nonlinear systems," *J. Appl. Math. Mech*, vol. 25, pp. 1254–1266, 1961.
- [17] T. Yoshida and K. Loparo, "Quadratic regulatory theory for analytic nonlinear systems with additive controls," *Automatica*, vol. 25, pp. 531–544, 1989.
- [18] W. Wonham, "Some applications of stochastic differential equations to nonlinear filtering," *SIAM J. Control*, vol. 2, pp. 347–369, 1965.
- [19] V. Benes, "Exact finite-dimensional filters for certain diffusions with nonlinear drift," *Stochastics*, vol. 5, pp. 65–92, 1981.
- [20] S. Yau, "Finite-dimensional filters with nonlinear drift. I: A class of filters including both Kalman-Bucy and Benes filters," *J. Math. Systems, Estimation, and Control*, vol. 4, pp. 181–203, 1994.
- [21] X. Yan, J. Lam, and L. Xie, "Robust observer design for nonlinear interconnected systems using structural characteristics," *International Journal of Control*, vol. 76, pp. 741–746, 2003.
- [22] M. Basin, "On optimal filtering for polynomial system states," *ASME Trans. J. Dynamic Systems, Measurement, and Control*, vol. 125, pp. 123–125, 2003.
- [23] M. Basin, D. Calderon-Alvarez, and M. Skliar, "Optimal filtering for incompletely measured polynomial states over linear observations," *International J. Adaptive Control and Signal Processing*, vol. 22, pp. 482–494, 2008.
- [24] P. Shi and S. Shue, " H^∞ control of interconnected nonlinear sampled-data systems with parametric uncertainty," *International Journal of System Science*, vol. 28, pp. 961–976, 1997.
- [25] P. Shi, R. Agarwal, E. Boukas, and S. Shue, "Robust H^∞ state feedback control of discrete time-delay linear systems with norm-bounded uncertainty," *International Journal of System Science*, vol. 42, pp. 409–415, 2011.
- [26] X. Guan, P. Shi, and Y. Liu, "Robust output feedback control for uncertain discrete time delay systems," *Systems Analysis Modelling Simulation*, vol. 42, pp. 1829–1840, 2002.
- [27] V. Boltyanski and A. Poznyak, "Robust maximum principle in minimax control," *International Journal of Control*, vol. 72, no. 4, pp. 305–314, 1999.
- [28] —, "Linear multi-model time optimization." *Optimal Control Applications and Methods*, vol. 23, pp. 141–161, 2002.

- [29] M. Jimenez-Lizarraga and A. Poznyak, "Robust nash equilibrium in multi-model LQ differential games: Analysis and extraproximal numerical procedure," *Optimal Control Applications and Methods*, vol. 8, no. 2, pp. 117–141, 2007.
- [30] A. Poznyak, T. Duncan, B. Pasik-Duncan, and V. Boltyanski, "Robust maximum principle for multi-model LQ-problem," *International Journal of Control*, vol. 75, no. 15, pp. 1770–1777, 2002.
- [31] A. Levant, "Robust exact differentiation via sliding mode technique," *Automatica*, vol. 34, pp. 379–384, 1998.
- [32] ———, "Robust exact differentiation via siliding modes technique," *Automatica*, pp. 379–384, 1998.
- [33] ———, "Sliding order and sliding accuracy in sliding mode control," *International Journal of Control*, vol. 58, no. 6, pp. 1247–1263, 1993.
- [34] L. Fridman, Y. Shtessel, C. Edwards, and X. Yan, "High-order sliding-mode observer for state estimation and input reconstruction in nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 18, pp. 399–413, 2008.
- [35] L. Fridman, A. Levant, and J. Davila, "Observation of linear system with unknown inputs via high-order sliding-mode," *International Journal of System Science*, vol. 38, pp. 773–791, 2007.
- [36] J. Bejarano, L. Fridman, and A. Poznyak, "Hierarchical second-order sliding-mode observer for linear time invariant systems with unknown inputs," *International Journal of Systems Science*, vol. 38, pp. 793–802, 2007.
- [37] C. Chen and J. B. Cruz, Jr., "Stackelberg solution for a two person games with biased information patterns," *IEEE Transactions of Automatic Control*, vol. 17, no. 6, pp. 791–798, 1972.
- [38] M. Simman and J. B. Cruz, Jr., "On the stackerberg strategy in nonzero games," *Journal of Optimization Theory and Applications*, vol. 11, no. 5, 1973.
- [39] F. Kyndland, "Noncooperative and dominant player solutions in discrete dynamic games," *International Economic Review*, vol. 16, no. 2, pp. 321–335, 1975.
- [40] F. Kyndland, "Equilibrium solution in dynamic dominant player models," *Journal of Economic Theory*, vol. 15, pp. 307–325, 1977.
- [41] A. Bagchi, *Stackelberg Differential Games in Economic Models*. Springer-Verlag, New York, NY, 1984.
- [42] X. He, A. Prasad, S. Sethi, and G. Gutierrez, "A survey of stackelberg differential game models in supply and marketing channels," *J Syst Sci Syst Eng*, vol. 16, pp. 385–413, 2007.

- [43] C. Harris and J. Vickers, "Innovation and natural resources: a dynamic game with uncertainty," *RAND Journal of Economics*, vol. 26, pp. 418–430, 1995.
- [44] H. Von Stackelberg, *The theory of the market economy*. Oxford University Press, 1952.
- [45] T. Basar and H. Selbuz, "Closed-loop stackelberg strategies with applications in the optimal control of multilevel systems," *IEEE Transactions on Automatic Control*, vol. AC-24, pp. 166–178, 1979.
- [46] Y. Ho, P. Luh, and G. Olsder, "A control-theoretic view of incentives," *Automatica*, vol. 18, pp. 167–180, 1982.
- [47] M. Li, J. B. Cruz, and M. Simman, "An approach to discrete-time incentive feedback stackelberg games," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 32, pp. 472–781, 2002.
- [48] M. Salman and J. B. Cruz, "Team-optimal closed-loop stackelberg strategies for systems with slow and fast modes," *International Journal of Control*, vol. 37, pp. 1401–1416, 1983.
- [49] T. Basar and G. Olsder, "Team-optimal closed-loop stackelberg strategies in hierarchical control problems," *Automatica*, vol. 16, pp. 409–414, 1980.
- [50] M. Basin, J. Perez, and M. Skliar, "Optimal filtering for polynomial system states with polynomial multiplicative noise," *Int. J. Robust Nonlin. Contr.*, vol. 16, pp. 287–298, 2006.
- [51] S. Bittanti, A. Laub, and J. Willems, *The Riccati Equation*. Springer, 1991.
- [52] M. Basin, J. Perez, and M. Skliar, "Optimal filtering for polynomial system states with polynomial multiplicative noise," *International Journal of Robust and Nonlinear Control*, vol. 16, pp. 287–298, 2006.
- [53] C. P. Mracek and J. R. Cloutier, "Control designs for the nonlinear bench mark problem via the state-dependent Riccati equation method," *International Journal of Robust and Nonlinear Control*, vol. 8, pp. 401–433, 1998.
- [54] T. Çimen, "State-dependent Riccati equation (SDRE) control: A survey," *Proceedings of the 17th World Congress The International Federation of Automatic Control Seoul, Korea.*, pp. 3761–3775, 2008.
- [55] M. Basin and D. Calderon-Alvarez, "Optimal controller for uncertain stochastic polynomial systems with deterministic disturbances," *International Journal of Control*, vol. 82, pp. 1435–1447, 2009.
- [56] G. Bryson and Y. Ho, *Applied Optimal Control*. Hemisphere, Washington, DC, 1975.
- [57] H. Abou-Kandil, G. Freiling, V. Ionescu, and G. Jank, *Matrix Riccati Equations in Control and System Theory*. Birkhauser, 2003.
- [58] V. S. Pugachev and I. N. Sinitsyn, *Stochastic Systems: Theory and Applications*. World Scientific, 2001.

- [59] K. J. Åström, *Introduction to Stochastic Control Theory*. Academic Press: New York, 1970.
- [60] M. V. Basin, *New Trends in Optimal Filtering and Control for Polynomial and Time-Delay Systems*. Heidelberg: Springer-Verlag, 2008.
- [61] A. Isidori, *Nonlinear Systems*. Berlin: Springer., 2001.
- [62] M. Jimenez-Lizarraga, M. Basin, and P. Rodriguez-Ramirez, “Robust mini-max regulator for uncertain nonlinear polynomial systems,” *Proc. of the IEEE Multi-Conference on Systems and Control Denver, CO*, pp. 1534–1539, 2011.
- [63] A. Loria, E. Panteley, and H. Nijmeijer, “Control of the chaotic Duffing equation with uncertainty in all parameters,” *IEEE Transactions on Circuits and Systems-I Fundamental Theory and Applications*, vol. 45, pp. 1252–1255, 1998.
- [64] H. Nijmeijer and B. Harry, “On Lyapunov control of the Duffing equation,” *IEEE Transactions on Circuits and Systems-I Fundamental Theory and Applications*, vol. 42, pp. 473–477, 1995.
- [65] J. Aubin, *Mathematical Methods of Game and Economic Theory*. North-Holland, 1979.
- [66] F. Amato, M. Mattei, and A. Pironti, “A note on singular zero-sum linear quadratic differential games,” *Proceedings of the 33rd IEEE Conference on Decision and Control*, vol. 2, pp. 1533 – 1535, Dec 1994, lake Buena Vista Florida.
- [67] D. Skataric and B. Petrovic, “Recursive approach to nash games of quasi-singularly perturbed linear systems,” *Proceedings of the 1998 American Control Conference*, vol. 3, pp. 1513 – 1517, June 1998, pennsylvania USA.
- [68] P. Zigang and T. Basar, “Multi-time scale zero-sum differential games with perfect state measurements,” *Proceedings of the 32nd IEEE Conference on Decision and Control*, vol. 4, pp. 3366 – 3371, Dec, san Antonio Texas.
- [69] H. Khalil, “Approximation of nash strategies,” *IEEE Transactions on Automatic Control*, vol. AC-25, no. 2, 1980.
- [70] J. Engwerda, B. van Aarle, and J. Plasmans, “The (in)finite horizon open-loop Nash LQ game: An application to EMU,” *Annals o Operational Research*, vol. 88, pp. 251–273, 1999.
- [71] W. Ma and H. Peng, “Solution of continuous-time preview two-player differential games,” *ASME J. of Dynamic System, Measurement and Control*, vol. 121, no. 2, pp. 326–330, 1999.
- [72] R. Hamalainen and H. Ehtamo, *Dynamic Games in Economic Analysis*, ser. Lectures Notes in Control and Information Sciences 157. Springer-Verlang, 1991.

- [73] J. Mertens, “The value of two-person zero-sum repeated games with lack of information on both sides,” *Int. Journal of Game Theory*, vol. 1, no. 1, 1971.
- [74] J. Case and G. Kimeldorf, “On nash points and Games of imperfect information,” *Journal of Optimization Theory and Applications*, vol. 9, no. 5, pp. 302–323, 1972.
- [75] J. Ponssard and S. Zamir, “Zero-sum sequential games with incomplete information,” *Journal of Game Theory*, vol. 2, no. 2, 1973.
- [76] G. Hexner, “A differential game of incomplete information,” *Journal of Optimization Theory and Applications*, vol. 28, no. 2, pp. 213–232, June 1979.
- [77] I. Rhodes and D. Luenberger, “Stochastic differential games with constrained state estimators,” *IEEE Transactions on Automatic Control*, vol. AC-14, no. 5, pp. 476–481, October 1969.
- [78] ———, “Differential games with imperfect state information,” *IEEE Transactions on Automatic Control*, vol. AC-14, no. 1, pp. 29–38, February 1969.
- [79] W. van den Broek, J. Engwerda, and J. Schumacher, “Robust equilibria in indefinite linear-quadratic differential games,” *Journal of Optimization Theory and Applications*, vol. 119, no. 3, pp. 565–595, 2003.
- [80] J. Engwerda, “On the open-loop nash equilibrium in lq-games,” *Journal of Economic Dynamics and Control*, vol. 22, pp. 729–762, 1998.
- [81] J. Engwerda, W. van de Broek, and Schumacher, “Feedback nash equilibria in uncertain infinite time horizon differential games,” <http://arno.uvt.nl/show.cgi?fid=4752>, 2005.
- [82] M. Jungers and H. Abou-Kandil, “Properties of coupled riccati equations in stackelberg games with time preference rate,” in *Proceedings of the 2nd Symposium on System, Structure and Control.*, 2004.
- [83] H. Abou-Kandil, G. Freiling, and G. Jank, “Necessary conditions for constant solutions of coupled riccati equations in nash games.” *System and Control Letters*, vol. 21, no. 4, pp. 295–306, 1993.
- [84] J. Papavassilopoulos and J. Cruz, “On the existence of solutions to coupled matrix riccati differential equations in linear-quadratic nash games.” *Journal of Optimization Theory and Applications*, no. 1, pp. 127–129, 1979.
- [85] H. Khalil, “Multimodel design of a nash strategy,” *Journal of Optimization Theory and Applications*, vol. 31, no. 4, pp. 553–555, 1980.
- [86] H. Mukaidani and H. Xu, “Nash strategies for a large scale interconnected systems,” in *43rd IEEE Conference on Decision and Control, December 14-17, 2004.*

- [87] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1944.
- [88] Y. Ho, “Survey paper: differential games, dynamic optimization, and generalized control theory,” *Journal of Optimization Theory and Applications*, vol. 6, no. 3, pp. 179–207, 1970.
- [89] J. Engwerda, “Uncertainty in a fishery management game. inbook: ”linear quadratic dynamic optimization and differential game theory”,” 2005.
- [90] T. Basar and P. Bernhard, *H^∞ -Optimal Control and Related Minimax Design Problems*. Birkhauser, Boston, 1995.
- [91] E. Dockner, S. Jorgensen, L. N. van, and G. Sorger, *Differential Games in Economics and Management Science*. Cambridge University Press, Cambridge, 2000.
- [92] M. Jungers, A. Franco, E. De Pieri, and Abou-Kandil, “Nash strategy applied to active magnetic bearing control,” *Proceedings of IFAC world congress*, 2005.
- [93] H. Khalil and P. Kokotovic, “Control strategies for decision makers using different models of the same system,” *IEEE Transactions on Automatic Control*, vol. AC-23, no. 2, pp. 289–298, 1978.
- [94] A. Filippov, *Differential equations with discontinuous right hand-sides, Mathematics and its applications*. Kluwer Academic Publisher, 1983.
- [95] M. Jimenez-Lizarraga and A. Poznyak, “Near-nash equilibrium strategies for lq differential game with inaccurate state information,” *Mathematical Problems in Engineering*, vol. Article ID 21509, pp. 1–24, 2006.
- [96] A. Poznyak and C. Jimenez, “Multimodel prey-predator lq differential game,” in *Proceedings of the American Control Conference 2003, Denver, Colorado USA*, 2003, pp. 5369–5374.
- [97] A. Antipin, “An extraproximal method for solving equilibrium programming problems and games,” *Computational Mathematics and Mathematical Physics*, vol. 45, no. 11, pp. 1893–1914, 2005.
- [98] —, “An extraproximal method for solving equilibrium programming problems and games with coupled variables,” *Computational Mathematics and Mathematical Physics*, vol. 45, no. 12, pp. 2020–2029, 2005.
- [99] J. B. Cruz, Jr., “Leader-follower strategies for multilevel systems,” *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 244–255, April 1978.
- [100] —, “Survey of nash and stackelberg equilibrium strategies in dynamic games,” *Annals of Economic and Social Measurement*, vol. 4, no. 4, pp. 339–344, 1975.
- [101] J. Engwerda, *Uncertainty in a fishery management game. Inbook: ”Linear Quadratic Dynamic Optimization and Differential Game Theory”*. John Wiley and Sons, Chichester, 2005.

- [102] G. Freiling, G. Jank, and S. Lee, “Existence and uniqueness of open-loop stackelberg equilibria in linear-quadratic differential games,” *Journal of Optimization Theory and Applications*, vol. 110, no. 3, pp. 515–544, 2001.
- [103] D. Fudenberg and K. Levine, *The Theory of Learning in Games*. Cambridge, MA: MIT Press, 1998.
- [104] V. Y. Glizer, “Asymptotic solution of zero-sum linear-quadratic differential game with cheap control for minimizer,” *Nonlinear Differ. Equ. Appl.*, vol. 7, no. 2, pp. 231–258, 2000.
- [105] A. A. Melikyan, “Singular characteristics of first order pdes in optimal control and differential games,” *J. Mathematical Sciences*, vol. 103, no. 6, pp. 745–755.
- [106] M. Miller and M. Salmon, “Dynamic games and the time inconsistency of the optimal policy in open economies,” *Economic Journal*, vol. 95, pp. 124–137, 1985.
- [107] G. Papavassilopoulos and J. B. Cruz, Jr., “Nonclassical control problems and stackelberg games,” *IEEE Transactions on Automatic Control*, vol. AC-24, pp. 127–129, 1979.
- [108] X. Tan and J. B. Cruz, Jr., *Dynamic Noncooperative Game Models for Deregulated Electricity Markets*. Nova Science Publishers, Inc., 2009.
- [109] B. Tolwinski, “A stackelberg solution of dynamic games,” *IEEE Transactions of Automatic Control*, vol. 28, pp. 85–93, 1983.
- [110] X. Wang and J. B. Cruz, “Asymptotic epsilon-nash equilibrium for 2nd order two-player nonzero-sum singular lq games with decentralized control,” *Proceedings of IFAC world congress Seoul, Korea*, pp. 3970–3975, 2008.
- [111] A. Ferreira de Loza, M. Jimenez-Lizarraga, and L. Fridman, “Robust output nash strategies based on sliding mode observation in a two-player differential game,” *Journal of the Franklin Institute*, vol. to appear, 2011.
- [112] M. Jimenez-Lizarraga and L. Fridman, “Robust nash strategies based on integral sliding mode control for a two players uncertain linear affine-quadratic game,” *International Journal of Innovative Computing, Information and Control*, vol. 42, pp. 117–122, 2009.
- [113] M. Jimenez-Lizarraga and A. Poznyak, “Necessary conditions for robust stackelberg equilibrium in a multi-model differential game,” *Optimal Control Applications and Methods*, vol. to appear, 2011.
- [114] ———, “Robust stackelberg equilibrium for a multi-scenario two players linear affine-quadratic differential game,” *Proceedings of the 17th World Congress The International Federation of Automatic Control*, pp. 3959–3963, 2008.

- [115] ———, “Quasi-equilibrium in lq differential games with bounded uncertain disturbances: Robust and adaptive strategies with pre-identification via sliding mode technique,” *International Journal of Systems Science*, vol. 38, no. 7, pp. 585–599, July 2007.
- [116] K. Tanaka and K. Yokoyama, “On ε -equilibrium point in a noncooperative N-person game,” *Journal of Mathematical Analysis and Applications*, 160, 413-423, vol. 160, pp. 413–423, 1991.
- [117] J. Engwerda, *LQ dynamic optimization and differential games*. John Wiley & Sons, West Sussex, England, 2005.
- [118] M. Jimenez-Lizarraga and L. Fridman, “Robust Nash strategies based on integral sliding mode control for a two players uncertain linear affine-quadratic game,” *International Journal of Innovative Computing Information and Control*, vol. 5, pp. 1349–4198, 2009.
- [119] M. Jimenez-Lizarraga and A. Poznyak, “Robust stackelberg equilibrium for a multi-scenario two players linear affine-quadratic differential game,” in *Proceedings of the 17th World Congress The International Federation of Automatic Control Seoul, Korea, July 6-11, 2008*, 2008, pp. 3959–3963.
- [120] R. Kalman and R. Bucy, “New results in linear filtering and prediction theory,” *ASME Trans., Part D (J. of Basic Engineering)*, vol. 83, pp. 95–108, 1961.
- [121] P. Nie, “Commitment for storable goods under vertical integration,” *Economic Modelling*, vol. 26, no. 2, pp. 414–417, 2009.
- [122] ———, “Dynamic stackelberg games under open-loop complete information,” *Journal of the Franklin Institute-Engineering and Applied Mathematics*, vol. 342, no. 7, pp. 737–748, 2005.
- [123] J. B. Cruz and T. Xiaohuan, *Dynamic Noncooperative Game Models for Deregulated Electricity Markets*. Nova Science Publishers, Inc., 2009.
- [124] J. Engwerda, “Feedback nash equilibria in the scalar infinite horizon lq-game,” *Automatica*, vol. 36, pp. 135–139, 2000.
- [125] D. Gu, “A differential game approach to formation control,” *IEEE Transactions on Control Systems Technology*, vol. 16, pp. 85–93, 2008.
- [126] S. Jorgensen, G. Martin-Herran, and G. Zaccour, “Dynamic games in the economics and management of pollution,” *Environmental Modeling and Assessment*, vol. 15, pp. 433–467, 2010.
- [127] E. Semsar-Kazerooni and K. Khorasani, “Multi-agent team cooperation: A game theory approach,” *Automatica*, vol. 45, pp. 2205–2213, 2009.
- [128] “Optimal consensus algorithms for cooperative team of agents subject to partial information,” *Automatica*, vol. 44, pp. 2766–2777, 2008.

- [129] T. Çimen, “State-dependent riccati equation control: A survey,” *Proc. 17th World Congress. Int. Federation Automat. Contr.*, pp. 3761–3775, 2008.
- [130] J. Cloutier and J. Cockburn, “The state-dependent nonlinear regulator with state constraints,” *Proc. Amer. Contr. Conference*, pp. 390–395, 2001.
- [131] T. Do, H. Choi, and J. Jung, “Sdre-based near optimal control system design for pm synchronous motor,” *IEEE Transactions on Industrial Electronics*, pp. 4063–4074, 2012.
- [132] R. Guo and J. Chen, “Sdre attitude control with global asymptotic stability for an unmanned helicopter,” *2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE)*, pp. 1044–1049, 2011.
- [133] A. Jazwinski, *Stochastic Processes and Filtering Theory*. New York, 1970.
- [134] P. Menon, T. Lam, L. Crawford, and V. Cheng, “Real-time computational methods for sdre nonlinear control of missiles,” *Proc. Amer. Contr. Conference*, pp. 232–237, 2002.
- [135] A. Merttopcuoglu, A. Kahvecioglu, and T. Çimen, “Sdre control of the control actuation system of a guided missile,” *Proc. IFAC Symp. Automat. Contr. Aerospace*, vol. 17, pp. 774–779, 2007.
- [136] C. Mracek and J. Cloutier, “Control designs for the nonlinear benchmark problem via the state-dependent riccati equation method,” *Int. J. Robust Nonlin. Contr.*, vol. 8, pp. 401–433, 1998.
- [137] H. Park, S. Park, and K. Choi, “Satellite formation reconfiguration and station-keeping using state-dependent riccati equation technique,” *Aerospace Science and Technology*, vol. 15, pp. 440–452, 2011.
- [138] L. Peled-Eitan and I. Rusnak, “State dependent difference riccati equation based estimation for maneuvering target,” *2012 15th International Conference on Information Fusion (FUSION)*, pp. 1689–1696, 2012.