

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN CIENTÍFICA

DE LA FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



ALGORITMOS DE FILTRADO ESTOCÁSTICO PARA SISTEMAS POLINOMIALES CON RUIDO DE POISSON

TESIS PRESENTADA POR:

JUAN JOSÉ MALDONADO ORTIZ

PARA OBTENER EL GRADO DE:

DOCTOR EN INGENIERÍA FÍSICA INDUSTRIAL

SAN NICOLÁS DE LOS GARZA, NUEVO LEÓN

AGOSTO 2012

*A mis padres
por apoyarme siempre y
confiar en mis decisiones,
por inspirar mis logros.*

Agradecimientos

Quiero agradecer al grupo de extraordinarias personas que han formado parte de mi equipo de apoyo durante estos años de mi carrera.

A mi asesor, el Doctor Mikhail Valentinovich Basin por su enseñanza, guía y apoyo durante mi formación, a quien admiro por su trayectoria y logros como investigador.

Al Doctor Vadim Azhmyakov Grudnikov por sus observaciones de rigor matemático y por las ideas que aportó a mi trabajo.

Al Doctor Rodolfo Martínez Zúñiga por el apoyo brindado y por sus acertados comentarios y observaciones.

Al Doctor Francisco Javier Almaguer Martínez por sus observaciones en la parte aplicada de mi tesis.

Al Doctor Manuel Alejandro Jiménez Lizárraga por sus comentarios constructivos en la revisión de mi tesis.

Al Doctor José Rubén Morones Ibarra por su apoyo en todo momento y por brindarme su confianza apoyándome en la realización de mi estancia en el extranjero.

A la Licenciada Gabriela Valentina Medellín González por todo su apoyo en la realización de los trámites ante el CONACyT.

A la Doctora Lilia López Vera por darme la oportunidad de impartir clases en la Facultad de Ciencias Físico Matemáticas durante mis estudios de posgrado.

A la Doctora Nataliya Ivanovna Kalashnikova por sus enseñanzas durante mis estudios. Al Doctor Mario César Osorio Abraham por apoyarme siempre, particularmente en los momentos más difíciles de la carrera.

Al M.C. Mauricio Torres Torres por su apoyo en los momentos en que pensaba abandonar mis estudios de posgrado.

Al Lic. Pedro Marcelo Soto Pérez por su compañerismo.

A la M.C. Perla Cecilia Hernández Lara, porque aún en la distancia siempre me ha mostrado su apoyo, siempre tiene las palabras más acertadas en los momentos indicados.

A la M.C. Nidia Lizette Gómez Duarte por confiar en mí y ser una gran fuente de apoyo durante todos estos años.

A la M.C. Mónica Guadalupe Elizondo Amaya, porque charlar con ella una hora me proporciona fuerza espiritual durante varias semanas.

A la M.C. Yadira Isabel Silva Soto por su amistad y apoyo.

A la Lic. Sandra Eguía Eguía y su familia por el invaluable apoyo que me brindaron desde el inicio de mi estancia en Monterrey.

A la Lic. Carolina Lozano Banda y su esposo, el M.C. Bruno Rodrigo Fuentes López por su amistad y apoyo constante durante mis estudios.

Al Lic. Julio Cerrillo por su apoyo en los diversos momentos de dificultad.

A mi gran amiga Jérica Michel Botello Zertuche por creer en mí, apoyarme y motivarme a dar lo mejor siempre.

A todos y cada uno de mis amigos de preparatoria, universidad y posgrado, que han confiado en mí, que han celebrado conmigo los logros y me han brindado su apoyo en los momentos complicados.

A la familia García Molina por su hospitalidad, motivación y apoyo en todo momento. A mi familia por su comprensión y apoyo en la realización de cada uno de mis objetivos. Al CONACyT por el apoyo económico brindado durante mis estudios.

I would like to thank Professor Hamid Reza Karimi for his support in making my research stay in Norway possible.

I would like to express my gratitude to PhD student Knut Berg Kaldestad who always supported me during my research stay in Norway.

Sin el apoyo de cada uno de Ustedes, esto no hubiera sido posible, gracias.

Resumen

Este trabajo de tesis aporta importantes resultados a la teoría de filtrado estocástico, a través del estudio de sistemas dinámicos sujetos a perturbaciones consideradas ruido blanco de Poisson, área que no ha sido tratada consistentemente. De esta manera se extiende el estudio de aquellos casos que se limitan a considerar solamente perturbaciones de naturaleza Gaussiana.

Se desarrollaron cuatro teoremas que proporcionan los algoritmos que permiten la obtención de los filtros e identificadores óptimos y un controlador óptimo para sistemas con ruido de Poisson, detallando el procedimiento de la obtención de los mismos, aportando así las bases teóricas para la aplicación en diversas áreas, tales como la fotografía, centellografía, reconocimiento de objetos en video, telecomunicaciones, etc.

Se obtuvo el filtro óptimo para sistemas no lineales incompletamente medidos sobre observaciones lineales con ruido de Poisson en el estado y en las observaciones, partiendo de la expresión general de la diferencial estocástica Ito del estimado óptimo y la varianza del error, y aplicando propiedades de las variables aleatorias de Poisson con la finalidad de obtener un sistema cerrado de las ecuaciones de filtrado óptimo. Después se diseñó un filtro y un identificador de parámetro para sistemas lineales con ruido de Poisson, introduciendo nuevas variables que permitieron reducir el problema al caso previamente resuelto, mod-

elando los parámetros desconocidos como procesos de Poisson e incorporándolos como variables de estado. Después se diseñó un controlador óptimo para sistemas lineales con ruido de Poisson con parámetros desconocidos, utilizando el principio de separación donde primero se introdujeron nuevas variables para reducir el problema al caso resuelto previamente y así poder aplicar el filtro que ya había sido obtenido, para después aplicar un control al sistema ya filtrado. Finalmente, se diseñó un filtro para sistemas polinomiales con ruido de Poisson donde las observaciones son polinomiales, introduciendo un cambio de variable y aplicando la diferencial estocástica Ito al nuevo proceso estocástico que se introdujo para así reformular el problema original, reduciendo el caso a aquél donde las observaciones son lineales, para de esta manera aplicar uno de los filtros óptimos obtenidos previamente, obteniendo el filtro deseado.

Los algoritmos obtenidos como resultado de los teoremas fueron aplicados en casos particulares para mostrar su rendimiento. Se compararon los resultados obtenidos por la aplicación de los algoritmos diseñados en esta tesis con los resultados obtenidos por la aplicación de los algoritmos óptimos diseñados en contexto Gaussiano, a través de simulaciones computacionales realizadas en MATLAB. Se presentan las gráficas de los errores en cada caso para ilustrar las ventajas que se obtienen al aplicar los algoritmos diseñados en este trabajo de investigación.

Índice general

Dedicatoria	I
Agradecimientos	III
Resumen	V
Lista de figuras	XIII
Notación	XV
1. Introducción	1
1.1. Introducción	1
1.2. Antecedentes	3
1.3. Motivación	6
1.4. Aportaciones	7
1.4.1. Filtrado Óptimo para Estados Polinomiales Incompletamente Med- tibles con Ruido de Poisson	7

1.4.2.	Filtrado Óptimo e Identificador de Parámetro para Sistemas Lineales con Ruidos de Poisson	8
1.4.3.	Controlador Óptimo LQP para Sistemas Estocásticos Lineales con Parámetros Desconocidos y Ruido de Poisson	8
1.4.4.	Filtro Aproximado de Dimensión Finita para Estados Polinomiales con Ruido de Poisson sobre Observaciones Polinomiales	9
1.5.	Organización de la Tesis	10
2.	Marco Teórico	11
2.1.	Probabilidad y Estadística	11
2.1.1.	Variables Aleatorias	11
2.1.2.	Convergencia de Variables Aleatorias	17
2.2.	Procesos Estocásticos	18
2.2.1.	Procesos Estocásticos	18
2.2.2.	Procesos con Incrementos no Correlacionados	20
2.2.3.	Procesos con Incrementos Independientes	22
2.2.4.	Proceso de Wiener	22
2.2.5.	Proceso de Poisson	25
2.3.	Cálculo Estocástico	28
2.3.1.	La Integral de Ito	28
2.3.2.	Ecuaciones Diferenciales Estocásticas	29
2.4.	Teoría de Filtrado Óptimo	35
2.4.1.	Filtro de Wiener	35
2.4.2.	Filtro de Kalman-Bucy (Caso Discreto)	37
2.4.3.	Filtro de Kalman-Bucy (Caso Continuo)	39
2.4.4.	Filtro Óptimo Polinomial	41

2.4.5. Ecuación General de Filtrado Óptimo	43
3. Filtrado Óptimo para Estados Polinomiales Incompletamente Medibles con Ruido de Poisson	53
3.1. Planteamiento del Problema	53
3.2. Diseño del Filtro Óptimo	55
3.2.1. Filtro Óptimo para un Estado Polinomial de Tercer Orden	60
3.3. Ejemplo	63
4. Filtrado Óptimo e Identificación de Parámetro para Sistemas Lineales con Ruidos de Poisson	69
4.1. Planteamiento del Problema	69
4.2. Diseño del Filtro Óptimo e Identificador	71
4.3. Ejemplo	74
5. Controlador Óptimo LQP para Sistemas Estocásticos Lineales con Parámetros Desconocidos y Ruido de Poisson	83
5.1. Planteamiento del Problema	83
5.2. Diseño del Controlador Óptimo	85
5.2.1. Reducción del Problema	85
5.2.2. Principio de Separación	86
5.2.3. Solución del Problema de Control Óptimo	88
5.3. Ejemplo	90
6. Filtro Aproximado de Dimensión Finita para Estados Polinomiales con Ruido de Poisson sobre Observaciones Polinomiales	100
6.1. Planteamiento del Problema	100

6.2. Diseño del Filtro	103
6.3. Ejemplo: Problema de Filtrado para el Sensor de Tercer Grado para un Sistema Cuadrático	107
7. Conclusiones y Trabajos Futuros	114
7.1. Conclusiones	114
7.2. Trabajos Futuros	116
A. Glosario	117
Bibliografía	121

Índice de figuras

3.1.	Gráfica del error entre el estado real $x_1(t)$, que satisface (3.16), y el estimado del filtro óptimo $m_1(t)$, que satisface (3.18), en el intervalo de simulación $[0, 2]$	66
3.2.	Gráfica del error entre el estado real $x_2(t)$, que satisface (3.16), y el estimado del filtro óptimo $m_2(t)$, que satisface (3.18), en el intervalo de simulación $[0, 2]$	67
3.3.	Gráfica del error entre el estado real $x_1(t)$, que satisface (3.16), y el estimado $m_{k1}(t)$, que satisface (3.20), en el intervalo de simulación $[0, 1,7842]$	68
3.4.	Gráfica del error entre el estado real $x_2(t)$, que satisface (3.16), y el estimado $m_{k2}(t)$, que satisface (3.20), en el intervalo de simulación $[0, 1,7842]$	68
4.1.	Gráficas de las variables de estado de referencia $x_1(t)$ y $x_2(t)$ (línea gruesa en la primera y segunda gráfica), estimados del estado óptimo $m_1(t)$ y $m_2(t)$ (línea delgada en la primera y segunda gráfica) y estimado óptimo del parámetro $m_3(t)$ (línea gruesa en la tercera gráfica), que satisfacen (4.10), para el sistema estable (4.7) en el intervalo de simulación $[0, 50]$	79

4.2.	Gráficas de las variables del estado de referencia $x_1(t)$ y $x_2(t)$ (línea gruesa en la primera y segunda gráfica), estimados óptimos del estado $m_1(t)$ y $m_2(t)$ (línea delgada en la primera y segunda gráfica) y el estimado óptimo del parámetro $m_3(t)$ (línea gruesa en la tercera gráfica), que satisfacen (4.10), para el sistema inestable (4.7) en un intervalo de simulación $[0, 10]$.	80
4.3.	Gráficas de las variables del estado de referencia $x_1(t)$ y $x_2(t)$ (línea delgada en la primera y segunda gráfica), estimados del estado $m_{G1}(t)$ y $m_{G2}(t)$ (línea delgada en la primera y segunda gráfica) y el estimado del parámetro $m_{G3}(t)$ (línea gruesa en la tercera gráfica), que satisface (4.12) para el sistema estable (4.7) en el intervalo de simulación $[0, 50]$.	81
4.4.	Gráficas de las variables del estado de referencia $x_1(t)$ y $x_2(t)$ (línea gruesa en la primera y segunda gráfica), estimados óptimos del estado $m_{G1}(t)$ y $m_{G2}(t)$ (línea delgada en la primera y segunda gráfica) y estimado del parámetro $m_{G3}(t)$ (línea gruesa en la tercera gráfica), que satisface (4.12), para el sistema inestable (4.7) en el intervalo de simulación $[0, 10]$.	82
5.1.	Gráfica del control $u(t)$ (5.17), el error entre las variables de estado reales $x_1(t)$ y $x_2(t)$, que satisfacen (5.14), y los estimados óptimos controlados $m_1(t)$ y $m_2(t)$ que satisfacen (5.18), y gráfica del criterio (5.16) $J(t)$ en el intervalo $[0, 1,6]$.	97
5.2.	Gráfica del control $u(t)$ (5.22), el error entre las variables de estado reales $x_1(t)$ y $x_2(t)$, que satisfacen (5.14), y los estimados óptimos controlados $\hat{x}_1(t)$ y $\hat{x}_2(t)$ que satisfacen (5.23)–(5.26), la gráfica del estimado del parámetro $\hat{\theta}$ que satisface (5.24), y la gráfica del criterio (Ec. (5.16)) $J(t)$ en el intervalo $[0, 1,6]$.	98

5.3.	Gráfica del control $u(t)$ (5.22), el error entre las variables de estado reales $x_1(t)$ y $x_2(t)$, que satisfacen (14), y los estimados controlados $\hat{x}_{G1}(t)$ y $\hat{x}_{G2}(t)$ que satisfacen (5.28)–(5.31), la gráfica del estimado del parámetro $\hat{\theta}$ satisfying (5.29) , y la gráfica del criterio (Ec. (16)) $J(t)$ en el intervalo $[0, 1,6]$	99
6.1.	Arriba. Gráfica del error de estimación entre el estado de referencia $x(t)$ (15) y su estimado óptimo $m_2(t)$ (6.20) en el intervalo $[0, 4,0575]$. Abajo. Gráfica del error de estimación entre el estado de referencia $z(t)$ (6.17) y su estimado óptimo $m_1(t)$ (6.19) en el intervalo $[0, 4,0575]$	112
6.2.	Arriba. Gráfica del error de estimación entre el estado de referencia $x(t)$ (6.15) y su estimado $m_{G2}(t)$ (6.25) en el intervalo $[0, 4,0575]$. Abajo. Gráfica del error de estimación entre el estado de referencia $z(t)$ (6.17) y el estimado $m_{G2}(t)$ (6.24) en el intervalo $[0, 4,0575]$	113

Notación

$P(A)$	Probabilidad del evento A
Ω	Espacio de eventos elementales
δ, \mathcal{F}	Conjunto de subconjuntos de Ω
(Ω, δ, P)	Espacio de probabilidad
$P(A B)$	Probabilidad condicional de A relativo a B
$X_t, X(t)$	Proceso estocástico
$F(x)$	Función de distribución de la variable aleatoria X
$f(x)$	Función de densidad de la variable aleatoria X
$g_x(\lambda)$	Función característica de la variable aleatoria X
$E(X), m_x$	Esperanza de la variable aleatoria X
$E(\varphi(X) Y)$	Esperanza condicional de la función $\varphi(X)$ relativo a Y
Γ_x	Momento de segundo orden del vector aleatorio X
K_x	Matriz de covarianza relativa al vector aleatorio X
Γ_{xy}	Momento de segundo orden relativo a X y Y
K_{xy}	Matriz de covarianza cruzada
$\nu(t)$	Intensidad de cierto ruido blanco
$W(t)$	Proceso de Wiener
$N(t)$	Proceso de Poisson

Capítulo 1

Introducción

1.1. Introducción

En el procesamiento de señales, un filtro es una función o procedimiento que remueve las partes no deseadas de una señal, tales partes indeseadas son los ruidos o perturbaciones que afectan el proceso. Estas señales pueden ser modeladas por sistemas dinámicos cuyo comportamiento puede ser, en cierta medida, monitorizado. Dentro de los sistemas estocásticos tienen especial relevancia aquellos sistemas cuyas ecuaciones de estado son lineales, aún así existe un gran número de problemas donde el modelo lineal no representa de forma adecuada la evolución temporal del fenómeno objeto de estudio, siendo preciso recurrir a modelos no lineales, lo que supone una mayor complejidad en el cálculo matemático asociado. Los ruidos que afectan el sistema pueden ser de distintos tipos, en esta tesis se aborda el caso donde el ruido que altera a los sistemas dinámicos es modelado como ruido blanco. El ruido blanco, es una señal aleatoria caracterizada porque sus valores en instantes de tiempo distintos no tienen relación alguna entre sí, es decir, no existe correlación estadística entre sus valores. En la práctica es común encontrar que dichos

ruidos son en su mayoría representados como ruido blanco Gaussiano, que es aquél cuya función de densidad responde a una distribución normal, es común encontrar en la literatura que en la mayoría de los casos se asume que el ruido alterando el sistema dinámico corresponde a ruido Gaussiano, debido a la gran cantidad de fenómenos que presentan este tipo de disturbio. Por otro lado, el ruido blanco de Poisson, que a diferencia del ruido blanco Gaussiano, se presenta únicamente en momentos aleatorios aislados en el tiempo, también está presente en una considerable cantidad de fenómenos en la naturaleza, en procesos industriales, telecomunicaciones, etcétera y aún así no ha sido tratado consistentemente. El número de publicaciones encontrado en la literatura sobre filtrado en promedio cuadrático para sistemas con ruido de Poisson, es relativamente pequeño, en [20, 32, 40, 41, 47, 72, 73] se pueden encontrar algunos resultados obtenidos relacionados con sistemas no lineales perturbados por ruido de Poisson, por lo que en esta tesis se centra el estudio en el desarrollo de algoritmos de filtrado para sistemas que son alterados por ruido blanco de Poisson. El diseño de filtros para sistemas con ruido blanco de Poisson, es de gran importancia tanto en la teoría de filtrado como en la práctica, pues permite estudiar problemas de estimación en promedio cuadrático para sistemas no lineales que son perturbados con ruido blanco no necesariamente Gaussiano. Existen muchas situaciones en la práctica donde los ruidos que perturban los sistemas dinámicos no podrían ser modelados como ruidos blancos Gaussianos, actuando uniformemente, como por ejemplo una serie de impulsos electromagnéticos. Este tipo de ruido blanco podrían ser modelado como ruido blanco de Poisson, actuando de forma aleatoria en ciertos momentos aislados del tiempo.

Por otro lado, en las décadas de 1960 y 1970 se dieron una gran cantidad de aplicaciones de las técnicas de control moderno, en especial el control óptimo, lo cual dio origen a técnicas de diseño sistemáticas como LQR, LQG y LTR, conocidas como las técnicas

posmodernas. Aún así, el problema del controlador óptimo LQP (de sus siglas en inglés, Linear Quadratic Poisson controller, esto último viene de la caracterización estadística del ruido empleado) para sistemas lineales con parámetros desconocidos y ruido de Poisson no ha sido tratado consistentemente. De hecho podría parecer no tener sentido buscar una solución óptima para este problema, pues algunos de los parámetros no son determinados. El planteamiento del problema empieza a tener sentido cuando los parámetros desconocidos son modelados, para esto se toman en cuenta especificaciones estocásticas de Poisson del problema óptimo LQP, así los parámetros desconocidos son representados como procesos de Poisson.

Esta tesis presenta los algoritmos de filtrado en promedio cuadrático óptimo para resolver los problemas de sistemas lineales y no lineales estocásticos con presencia de ruido de Poisson. Se realiza una comparación de los resultados obtenidos con el caso donde los algoritmos son desarrollados considerando que el ruido a filtrar es de tipo Gaussiano. Esta comparación se realiza con el fin de resaltar la importancia del diseño de estos filtros y mostrar que dichos filtros permiten obtener un resultado óptimo.

1.2. Antecedentes

Aproximadamente al mismo tiempo en que la teoría de filtrado se estaba desarrollando usando ecuaciones diferenciales estocásticas lineales, un desarrollo análogo para el control óptimo de sistemas dinámicos lineales con una función de costo cuadrática estaba tomando lugar. Este trabajo fue inspirado por el desarrollo de la Programación Dinámica de Bellman [17] y fue desarrollado por Kalman [37]. El resultado más famoso, el filtro de Kalman Bucy [39], aborda el caso donde el estado y las observaciones son lineales, donde solamente dos momentos, el estimado mismo y su varianza, forman un sistema cerrado de

las ecuaciones de filtrado. Tiempo más tarde, una extensión del problema de control óptimo con función de costo cuadrática para sistemas dinámicos lineales en presencia ruido blanco aditivo llevó a considerar el problema de costo cuadrático para sistemas dinámicos estocásticos lineales. Este problema en algunas situaciones especializadas fue examinado y resuelto por Florentin [25, 26] y por Kushner [43]. Fue Kushner también [42] quien un par de años más tarde dio la solución óptima al problema de filtrado en promedio cuadrático para sistemas no lineales con ecuación de observación no lineales considerando los disturbios como ruido blanco Gaussiano, a través de la llamada ecuación de Kushner, por la densidad condicional de un estado no observado con respecto a las observaciones. Existen muy pocos ejemplos conocidos de sistemas no lineales donde la ecuación de Kushner puede reducirse a un sistema cerrado de dimensión finita de las ecuaciones de filtrado para un cierto número de momentos condicionales. En [1, 18, 24, 61, 64, 65, 70, 71] se pueden encontrar algunos otros filtros en promedio cuadrático no lineales de dimensión finita, así como también existe una considerable bibliografía sobre filtrado robusto para sistemas lineales y no lineales en presencia de ruido blanco Gaussiano (ver por ejemplo [19, 30, 31, 34, 48, 49, 54, 56, 57, 58, 63, 67]). Con un enfoque distinto, también la teoría de filtrado H-Infinito ha sido ampliamente estudiada en una variedad de artículos para sistemas estocásticos lineales [27, 28, 29, 50, 55, 60, 62, 66, 68, 69, 74]. Los métodos de H-Infinito también son utilizados en la teoría de control para lograr un rendimiento robusto o de estabilización en los controladores y aunque el problema del controlador óptimo LQG para sistemas lineales fue resultado en la década de 1960, basada en las soluciones para el problema de filtrado óptimo [39] y el regulador óptimo [23, 44], el problema del controlador óptimo LQP para sistemas lineales con parámetros desconocidos y ruido de Poisson alterando el sistema no ha sido tratado consistentemente. Otro de los problemas correspondientes a la teoría de filtrado, es el llamado problema de identificación de

parámetros, el problema de obtener simultáneamente el estimador óptimo y el identificador de parámetro para sistemas estocásticos Gaussianos con parámetros desconocidos ha sido tratado sistemáticamente partiendo desde el artículo seminal [3]. El resultado óptimo fue obtenido en [3] para un sistema lineal en tiempo discreto con parámetros desconocidos constantes dentro de un intervalo de filtrado en tiempo finito, utilizando el principio de máxima verosimilitud, en vista de un intervalo finito del estado y los valores del parámetro en instantes del tiempo. La aplicación del principio de máxima verosimilitud se extendió después para sistemas lineales en tiempo discreto [22] y sistemas lineales en tiempo continuo en [21]. Sin embargo, el uso del principio de máxima verosimilitud presenta ciertas limitaciones en el resultado final: a) los parámetros desconocidos se consideran constantes para evitar complicaciones en el problema de optimización generado, b) no se pueden obtener ecuaciones dinámicas directas que puedan dar seguimiento tanto al estado y a los estimados del parámetro en la situación general, sin poner restricciones especiales en la estructura del sistema. Por otro lado, el número de publicaciones en el área de filtrado en promedio cuadrático para sistemas con ruido blanco de Poisson es relativamente pequeño. Se sabe que el filtro en promedio cuadrático para sistemas lineales con ruido blanco de Poisson coincide con el filtro de Kalman-Bucy [46, 53]. Aún así, los filtros en promedio cuadrático para sistemas polinomiales con ruido blanco de Poisson, similares a esos obtenidos en [4, 11, 14], no habían sido diseñados hasta ahora. Hasta este momento se ha tratado con sistemas donde la ecuación de estado pueden ser lineales o no lineales, pero la ecuación de observaciones es lineal en ambos casos, recientemente el problema de filtrado para sistemas donde tanto el estado como las observaciones son polinomiales, fue resuelto en [15] pero considera únicamente el caso donde el ruido blanco que altera tanto la ecuación de estado y las observaciones es Gaussiano, por lo que en esta tesis se ampliará la situación al caso en un contexto de sistemas de Poisson.

1.3. Motivación

El número de publicaciones encontradas en la literatura sobre filtrado en promedio cuadrático para sistemas con ruido de Poisson es relativamente pequeño. La principal causa puede ser que, a pesar de que la naturaleza de los ruidos Gaussianos y de Poisson es muy distinta, conforme la intensidad del ruido aumenta, la distribución Gaussiana puede utilizarse para aproximar la distribución de Poisson bajo algunas consideraciones. Aún así existe una gran cantidad de ejemplos y aplicaciones que requieren del diseño de filtros que consideren con rigor la naturaleza de las variables aleatorias y procesos involucrados en el sistema para obtener un resultado lo más parecido al resultado deseado, libre de perturbaciones. En la práctica no suele hacerse una diferencia entre la naturaleza real del ruido que está afectando el sistema, pues teóricamente no hay las suficientes herramientas que lo permitan, por lo que diseñar filtros e identificadores para sistemas con ruido de Poisson representa un gran aporte a la teoría de filtrado estocástico y a la práctica. El desarrollo de teoría en este contexto permitirá mejorar los resultados deseados en aplicaciones de gran importancia tales como las telecomunicaciones, fotografía, centellografía, reconocimiento de objetos en video, entre otras, donde hasta ahora se aplican filtros convencionales por el hecho de que han sido estudiados ampliamente. Sin embargo la aplicación de filtros que consideran las características concretas de la naturaleza del ruido que afecta los sistemas permitirá una mejora en los resultados esperados.

1.4. Aportaciones

1.4.1. Filtrado Óptimo para Estados Polinomiales Incompletamente Medibles con Ruido de Poisson

Este capítulo presenta el diseño del filtro óptimo de dimensión finita para sistemas polinomiales incompletamente medibles que incluyen ruidos de Poisson, sobre observaciones lineales. El problema de filtrado óptimo es tratado partiendo de la expresión general para la diferencial estocástica Ito del estimado óptimo y la varianza del error [53]. Como primer resultado, se derivan las diferenciales Ito y la correspondiente varianza del error del problema planteado. Después se introduce una transformación de la ecuación de observación para reducir el problema original a uno con una matriz de observación invertible. Finalmente se muestra que, utilizando la técnica de representar los momentos de orden superior de una variable aleatoria de Poisson como funciones de su esperanza y su varianza, puede obtenerse un sistema cerrado de dimensión finita de las ecuaciones de filtrado óptimo con respecto a un número finito de variables de filtrado, para una ecuación de estado polinomial y observaciones lineales con una matriz de observación arbitraria. En este caso, se establece el procedimiento correspondiente para el diseño de las ecuaciones de filtrado óptimo. Finalmente, el sistema cerrado de las ecuaciones de filtrado óptimo con respecto a las dos variables, el estimado óptimo y la varianza, se deriva en la forma explícita en el caso particular de una ecuación de estado de tercer grado.

Los resultados de este capítulo han sido publicados en [8].

1.4.2. Filtrado Óptimo e Identificador de Parámetro para Sistemas Lineales con Ruidos de Poisson

Este capítulo presenta el filtro óptimo e identificador de parámetro para sistemas estocásticos lineales con parámetros aditivos y multiplicativos desconocidos sobre observaciones lineales, donde los parámetros desconocidos son modelados como procesos de Poisson. El problema de filtrado se formaliza considerando a los parámetros desconocidos como estados adicionales del sistema que satisfacen ecuaciones diferenciales estocásticas lineales con término drif igual a cero y difusión uno. Así, el problema se reduce al problema de filtrado para sistemas polinomiales medibles incompletamente con ruido de Poisson sobre observaciones lineales, cuya solución es obtenida en [7]. Esto presenta el algoritmo óptimo simultáneamente para el estimado óptimo del estado e identificación del parámetro en sistemas lineales de Poisson con parámetros desconocidos multiplicativos y aditivos sobre observaciones lineales. Los resultados de este capítulo fueron publicados en [12].

1.4.3. Controlador Óptimo LQP para Sistemas Estocásticos Lineales con Parámetros Desconocidos y Ruido de Poisson

Este capítulo presenta la solución al problema del controlador óptimo LQP para sistemas lineales con ruido de Poisson y parámetros desconocidos. El problema primero se reduce al problema del controlador óptimo para un estado bilineal sobre observaciones lineales y un criterio cuadrático. Por el principio de separación para sistemas lineales y un criterio cuadrático, el cual se plantea y fundamenta en el artículo [44] análogamente para sistemas lineales sin incertidumbre, el problema original del controlador óptimo se divide en dos partes, una el problema de filtrado óptimo para estados bilineales incompletamente

medibles sobre observaciones lineales y otra, el problema de control óptimo (regulador) para el estimado del filtro diseñado. Finalmente, puede ser demostrado fácilmente que la solución al problema original del controlador óptimo se compone de las dos soluciones. Los resultados de este capítulo fueron publicados en [10].

1.4.4. Filtro Aproximado de Dimensión Finita para Estados Polinomiales con Ruido de Poisson sobre Observaciones Polinomiales

Este capítulo presenta un filtro aproximado de dimensión finita para estados polinomiales con ruidos de Poisson sobre observaciones polinomiales, continuando la investigación en el área de filtrado en promedio cuadrático para sistemas polinomiales iniciada en ([4, 6, 11, 13, 16]). Se aborda el caso general donde la ecuación de estado y las observaciones son no lineales, y son afectadas con ruido de Poisson. Diseñar un filtro de dimensión finita en forma cerrada para sistemas con ruido de Poisson sobre observaciones polinomiales presenta una significativa ventaja en la teoría de filtrado y en la práctica, pues permite tratar problemas donde tanto la ecuación de estado como las observaciones son de naturaleza no lineal y el ruido aditivo no es necesariamente Gaussiano, tal como el problema del sensor cúbico [33] para varios sistemas polinomiales en contexto de Poisson. Incluso, el resultado principal de este capítulo permite diseñar un filtro subóptimo en promedio cuadrático de dimensión finita para cualquier estado polinomial perturbado por ruido de Poisson sobre observaciones lineales. Por lo tanto, como cualquier función no lineal puede ser aproximada, con cierta precisión, por un polinomio de cierto grado, esto potencialmente permitiría tratar con el diseño de filtros para cualquier estado no lineal con ruido de Poisson con observaciones y término drift no lineales.

Los resultados de este capítulo fueron publicados en [9].

1.5. Organización de la Tesis

La tesis se divide en siete capítulos. En el Capítulo 2 se presenta una síntesis de la teoría de probabilidad y estadística, procesos estocásticos, los procesos de Wiener y de Poisson, cálculo estocástico y teoría de filtrado. En el Capítulo 3 se plantea y se resuelve el problema de filtrado óptimo para estados polinomiales incompletamente medibles con ruido de Poisson. En el Capítulo 4 se plantea y se resuelve el problema de filtrado óptimo e identificador de parámetro para sistemas lineales con ruido de Poisson. En el Capítulo 5 se plantea y se resuelve el problema del controlador óptimo LQG para sistemas estocásticos lineales con parámetros desconocidos y ruido de Poisson. En el Capítulo 6 se desarrolla el filtro aproximado de dimensión finita para estados polinomiales con ruido de Poisson sobre observaciones polinomiales. Finalmente, en el Capítulo 7 se presentan las conclusiones y los trabajos futuros que pueden desarrollarse siguiendo esta línea, así como las recomendaciones para mejorarlo.

Capítulo 2

Marco Teórico

2.1. Probabilidad y Estadística

2.1.1. Variables Aleatorias

Una variable aleatoria X es una función real cuyo dominio es Ω y el cual es δ -medible, esto es, para cada número real x , $\{w \in \Omega / X(w) \leq x\} \in \Omega$. En otra forma, es una variable la cual asume como valores uno de los resultados de una prueba, y es imposible saber que valor tomará y al valor de una variable aleatoria en algún punto dado w del espacio Ω (i.e. el valor el cual se asume cuando aparece un resultado de la prueba y se obtiene el evento w) es llamado una realización de esta variable aleatoria.

Si P_x es la medida de probabilidad de la variable aleatoria X . La función

$F(x) = P(X \in (-\infty, x)) = P_x(-\infty, x)$ es llamada una función de distribución de la variable aleatoria X .

Propiedades de la Función de Distribución

- Una función de distribución es una función no decreciente de x .
- $F(-\infty) = 0; F(\infty) = 1$
- $F(x)$ es continua por la izquierda, es decir $F(x) = F(x - 0)$.
- La función de distribución de una variable aleatoria discreta con saltos en los puntos x_1, x_2, \dots, x_N igual a p_1, p_2, \dots, p_N respectivamente y es constante en algún intervalo el cual no contiene alguno de los valores x_1, x_2, \dots, x_N , como la probabilidad del evento $X < x$ no cambia si x varía en cada intervalo. Entonces la función de distribución de una variable aleatoria discreta, es representada por una función escalón.
- La función de distribución de una variable aleatoria continua es continua y diferenciable a lo largo de eje numérico.
- La función de distribución de una variable aleatoria continua discreta tiene puntos de discontinuidad x_1, x_2, \dots, x_N con saltos p_1, p_2, \dots, p_N respectivamente y es continua y diferenciable en todos los otros puntos del eje numérico.

Diferenciando la fórmula anterior con respecto a x en el caso de una variable aleatoria continua escalar y aplicando el Teorema Fundamental del Cálculo, se obtiene

$$f(x) = F'(x)$$

Así la densidad de una variable aleatoria es la derivada de su función de distribución.

El límite de la proporción de la probabilidad de la ocurrencia de una variable escalar aleatoria X en un intervalo $[x, x + \Delta x)$ con longitud Δx cuando $\Delta x \rightarrow 0$ es llamado

densidad o densidad de probabilidad de la variable aleatoria X en el punto x .

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x \leq X < x + \Delta x)}{\Delta x} = F'(x)$$

Propiedades de la Densidad de Probabilidad

- $f(x) \geq 0$
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- La densidad de $Y = X + a$ está dada por

$$f(y-a) = \lim_{\Delta y \rightarrow 0} \frac{P(y \leq X + a < y + \Delta y)}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{P(y - a \leq X < y - a + \Delta y)}{\Delta y}$$

La probabilidad de ocurrencia de una variable aleatoria X en el dominio A es determinada por la fórmula

$$P(X \in A) = \int_A f(x)dx$$

si la densidad $f(x)$ es continua parte por parte y acotada en el dominio A . La función de distribución de la variable aleatoria X está dada por $F(x) = P(X \leq x)$. Utilizando la densidad,

$$F(x) = P(-\infty < X < x) = \int_{-\infty}^x f(u)du$$

cuando X es una variable aleatoria continua escalar.

El valor esperado de una función $\phi(X)$ de la variable aleatoria continua X con densidad $f(x)$ está dado por

$$E(X) = \int_{-\infty}^{\infty} \phi(x)f(x)dx$$

La varianza de una variable aleatoria escalar es definida como

$$\sigma^2 = E(x - E(x))^2$$

Una probabilidad condicional relativa a la variable aleatoria $X = x(w)$ en el espacio (Ω, S) es llamada probabilidad condicional relativa a la σ -álgebra \mathcal{S}_x inducida por esta variable aleatoria. Sea $Y = y(w)$ una nueva variable aleatoria en el espacio (Y, B) . La función:

$$\mu_{y/x}(B/x) = Q_x(y^{-1}(B)/x), B \in B$$

es llamada distribución condicional de la variable aleatoria Y relativa a la variable aleatoria X .

La esperanza condicional de una función dada $\varphi(X)$ de una variable o vector aleatorio escalar X dado el valor de una variable o vector escalar aleatorio Y , está dada por:

$$E[\varphi(X)|Y] = \int_{-\infty}^{\infty} \varphi(x)f(x|y)dx,$$

donde $f(x|y)$ es la densidad condicional de la variable aleatoria X dado el valor y de la variable aleatoria Y .

Dado un vector aleatorio X , el momento de segundo orden está dado por $\Gamma_x = EXX^T$ y la matriz de covarianza es determinada por la fórmula

$$K_x = EX^0X^{0T}$$

donde $X^0 = X - m_x$ y $m_x = EX$. Además

$$\Gamma_x = K_x + m_x m_x^*.$$

Donde el $*$ indica la transposición de una matriz, cambiando sus elementos complejos por sus conjugados correspondientes. Para dos vectores aleatorios X y Y , la matriz del

momento de segundo orden Γ_{xy} y la matriz de covarianza cruzada K_{xy} están dadas por las fórmulas

$$\Gamma_{xy} = EXY^T,$$

$$K_{xy} = EX^0Y^{0T}$$

Y además

$$\Gamma_{xy} = K_{xy} + m_x m_y^*.$$

Para los momentos centrales e iniciales de órdenes superiores, la fórmula está dada por

$$\begin{aligned} \mu_r &= E(X^0)^r, \alpha_r = EX^r, (r = 1, 2, \dots) \\ \alpha_r &= \alpha_{r_1, \dots, r_n} = EX_1^{r_1} \dots X_n^{r_n}, \\ \mu_r &= \mu_{r_1, \dots, r_n} = E(X_1^0)^{r_1} \dots (X_n^0)^{r_n}, \\ (|r| &= r_1 + \Delta\Delta\Delta + r_n, |r| = 1, 2, \dots). \end{aligned}$$

La función característica de una variable aleatoria X está determinada por el valor esperado de la variable aleatoria $e^{i\lambda^T X}$ [53] y es considerada como una función de la variable real λ . Su fórmula está dada por

$$g(\lambda) = E(e^{i\lambda^T X}) = \int_{-\infty}^{\infty} e^{i\lambda^T x} f(x) dx.$$

La dimensión de la variable λ coincide con la dimensión de la variable X .

Propiedades de la Función Característica

- La función característica es continua y $|g(\lambda)| \leq 1, g(0) = 1, g(-\lambda) = \overline{g(\lambda)}$. Donde indica el conjugado.

- La función característica es positiva definida, esto es: para algunos valores $\lambda_1, \dots, \lambda_N$ de una variable λ y algunos números complejos ξ_1, \dots, ξ_N

$$\sum_{p,q=1}^N g(\lambda_p - \lambda_q) \xi_p \bar{\xi}_q \geq 0.$$

- La función característica $g_2(\mu)$ de una variable aleatoria $Y = AX + a$ obtenida como el resultado de una transformación de una variable aleatoria X es expresada en términos de la función característica $g_1(\lambda)$ de la variable aleatoria X por

$$g_2(\mu) = e^{i\mu^T a} g_1(A^T \mu)$$

- La función característica de la proyección de un vector aleatorio en algún subespacio G es igual a la contracción de su función característica en este espacio. Si $a = 0$ y A es la matriz proyección en G , entonces $A^T = A$, $A\lambda = \lambda$ para algún $\lambda \in G$ y $A\lambda = 0$ para algún vector λ ortogonal a G .
- La función característica $g(\lambda)$ de la suma de variables aleatorias independientes X_1, \dots, X_n es igual al producto de sus funciones características $g_k(\lambda)$, ($k = 1, \dots, n$) :

$$g(\lambda) = \prod_{k=1}^n g_k(\lambda).$$

- Si X_1, \dots, X_n son variables aleatorias independientes, entonces la función característica correspondiente $g(\lambda)$, $\lambda = [\lambda_1, \dots, \lambda_n]^T$ de el vector compuesto aleatorio $X = [X_1^T, \dots, X_n^T]^T$ es igual al producto de las funciones características $g_k(\lambda_k)$, ($k = 1, \dots, n$) de las variables aleatorias X_1, \dots, X_n :

$$g(\lambda) = \prod_{k=1}^n g_k(\lambda_k).$$

El inverso también se cumple.

2.1.2. Convergencia de Variables Aleatorias

Sea $\{X_r\}$, $X_r = x_r(w)$ un conjunto arbitrario de variables aleatorias con valores en un espacio dependiente del parámetro r el cual toma valores de algún conjunto R .

Se dice que la variable aleatoria X_r converge casi seguro (o con probabilidad 1) a la variable aleatoria $X = x(w)$ con $r \rightarrow r_0$ si

$$P(X_r \rightarrow X) = P(w : x_r(w) \rightarrow x(w)) = 1. \quad (2.1)$$

.

Se dice que la variable aleatoria X_r converge en probabilidad a X si para algún $\epsilon > 0$

$$\lim_{r \rightarrow r_0} P(\|X_r - X\| \geq \epsilon) = \lim_{r \rightarrow r_0} P(w : \|x_r(w) - x(w)\| \geq \epsilon) = 0 \quad (2.2)$$

.

Consideremos un conjunto de variables aleatorias $\{X_r\}$, $X_r = x_r(w)$ con valores en el espacio X . Se dice que X_r converge en media cuadrada a la variable aleatoria X , $X_r \xrightarrow{m.s.} X$ si

$$E \|X\|^2, E \|X_r\|^2 < \infty \quad (2.3)$$

y

$$E \|X_r - X\|^2 \rightarrow 0 \quad (2.4)$$

cuando $r \rightarrow r_0$.

La convergencia estocástica es la más débil de las tres anteriores, se puede verificar que se cumplen las siguientes implicaciones:

$$(2,1) \Rightarrow (2,2) \Leftarrow (2,3).$$

Cuando $|X_n| \leq Y$, para algún $Y \in \mathbb{L}^2(\Omega, \mathcal{F}, P)$ (2.24) es el más fuerte de los modos de convergencia y los dos restantes resultan ser equivalentes:

$$(2,1) \Rightarrow (2,2) \Leftrightarrow (2,3) \Leftarrow (2,1).$$

La implicación $(2,1) \Rightarrow (2,3)$, bajo las condiciones dadas, es lo que se conoce como el Teorema de la Convergencia Dominada.

2.2. Procesos Estocásticos

2.2.1. Procesos Estocásticos

Un proceso estocástico es una colección de variables aleatorias, $X_t, t \in T$, definidas sobre el mismo espacio de probabilidad, donde T es un conjunto indizado y X_t es una variable aleatoria para cada t . X_t también se denota por $X(t)$ y los valores que esta variable aleatoria asigna al evento elemental w se denotarán por $X(t, w)$.

Las funciones del momento de segundo orden de la variable aleatoria $X(t)$, $t \in T_1$ (matriz $\Gamma_x(t_1, t_2)$) y la función de covarianza (matriz $K_x(t_1, t_2)$) son determinadas por las siguientes fórmulas

$$\begin{aligned} \Gamma_x(t_1, t_2) &= EX(t_1)X(t_2)^T \\ K_x(t_1, t_2) &= EX^0(t_1)X^0(t_2)^*, \end{aligned} \tag{2.5}$$

además,

$$\Gamma_x(t_1, t_2) = K_x(t_1, t_2) + m_x(t_1)m_y^*(t_2),$$

donde el * indica la transposición de una matriz, cambiando sus elementos por sus conjugados. El momento cruzado de segundo orden (matriz $\Gamma_{xy}(t_1, t_2)$) y la función de co-

varianza cruzada (matriz $K_{xy}(t_1, t_2)$) para dos funciones aleatorias $X(t), Y(t), t \in T_1$ son determinadas por

$$\begin{aligned}\Gamma_x(t_1, t_2) &= EX(t_1)Y(t_2)^T, \\ K_x(t_1, t_2) &= EX^0(t_1)Y^0(t_2)^*,\end{aligned}\tag{2.6}$$

y además,

$$\Gamma_{xy}(t_1, t_2) = K_{xy}(t_1, t_2) + m_x(t_1)m_y^*(t_2).$$

Análogamente, los momentos superiores para funciones reales, escalares aleatorias $X(t), t \in T_1$ son determinados por las siguientes fórmulas

$$\begin{aligned}\alpha_r(t_1, t_2, \dots, t_r) &= EX(t_1)\dots X(t_r) = \\ &\int_{-\infty}^{\infty} \Delta\Delta\Delta \int_{-\infty}^{\infty} x_1\dots x_r f_r(x_1, \dots, x_r; t_1\dots t_r) dx_1\dots dx_r \\ \mu_r(t_1, \dots, t_r) &= EX^0(t_1)\dots X^0(t_r) = \\ &\int_{-\infty}^{\infty} \Delta\Delta\Delta \int_{-\infty}^{\infty} [x_1 - m_x(t_1)]\dots [x_r - m_x(t_r)] f_r(x_1, \dots, x_r; t_1\dots t_r) dx_1\dots dx_r\end{aligned}\tag{2.7}$$

El proceso aleatorio $X(t)$ con esperanza cero y función de covarianza la cual contiene como un múltiplo la función δ ,

$$m_x(t) = 0, K(t_1, t_2) = \nu(t_1)\delta(t_1 - t_2)$$

es llamado ruido blanco, en sentido amplio. Tomando en consideración que $\delta(t_1 - t_2) = 0$ en $t_1 \neq t_2$ el multiplicador $\nu(t_1)$ puede ser reemplazado por el multiplicador $\nu(t_2)$ o por el multiplicador simétrico $\sqrt{\nu(t_1)\nu(t_2)}$. El multiplicador $\nu(t)$ de la función δ es llamado la intensidad del ruido blanco $X(t)$. La intensidad de un ruido blanco escalar es positiva. La intensidad del ruido blanco vector representa una matriz simétrica definida no negativa. Si la variable

$$\frac{K_x(t_1, t_2)}{K_x(t_1, t_1)}$$

para el proceso aleatorio escalar $X(t)$ puede ser asumida prácticamente igual a cero como $|t_1 - t_2| > \tau_k$ y la variable τ_k es suficientemente pequeña entonces el proceso aleatorio $X(t)$ puede ser asumido como un ruido blanco no estacionario con intensidad igual a

$$\nu(t) = \int_{-\infty}^{\infty} K_x(t, t + \tau) d\tau.$$

2.2.2. Procesos con Incrementos no Correlacionados

Se dice que un proceso aleatorio $X(t)$ es llamado proceso con incrementos no correlacionados si para todos los intervalos disjuntos $[t_1, t_2), [t_3, t_4), t_1 < t_2 \leq t_3 < t_4$ los incrementos correspondientes $X_{t_2} - X_{t_1}$, y $X_{t_4} - X_{t_3}$ del proceso $X(t)$ son no correlacionados. De aquí se sigue que los incrementos de cada proceso o algunos intervalos finitos tienen momentos finitos de segundo orden y consecuentemente de primer orden. En este caso, el proceso $X(t)$ puede no tener media ni momento de segundo orden. Pero si en algún instante t_0 el valor del proceso con incrementos no correlacionados $X(t)$ es igual a cero con probabilidad 1, entonces el proceso $X(t)$ tiene esperanza finita y momento de segundo orden como el valor X_t que en algún instante coincide con su incremento en el intervalo $[t_0, t)$ con $t > t_0$ y con su incremento en el intervalo $[t, t_0)$ tomado en forma inversa si $t < t_0$.

El proceso aleatorio $Y(t) = X(t) - X_{t_0}$ representa un proceso con incrementos no correlacionados los cuales tienen la propiedad $Y_{t_0} = 0$. Además, si $X_{t_0} = 0$ (con probabilidad 1) entonces el valor X_t del proceso $X(t)$ en algún instante t es no correlacionado con sus futuros incrementos en los intervalos en los cuales sigue el instante t_0 y con sus previos

incrementos en el intervalo el cual precede al instante t_0 : $EX_t^0(X_{t_2}^{0*} - X_{t_1}^{0*}) = 0$ con $t \leq t_1 < t_2, t_1 \geq t_0$ o $t_1 < t_2 \leq t, t_2 \leq t_0$, donde $X_t^0 = X_t - m_t; (m_t = EX_t)$.

Consideremos la siguiente fórmula:

$$\begin{aligned} k(t) &= EX_t^0 X_t^{0*}; t > t_0, \\ &= 0; t = t_0, \\ &= -EX_t^0 X_t^{0*}; t < t_0. \end{aligned} \tag{2.8}$$

Donde el asterisco indica el vector transpuesto cuyos componentes son los conjugados. La matriz de covarianza del incremento $X_{t_2} - X_{t_1}$ del proceso $X(t)$ en algún intervalo y la función de covarianza del proceso $X(t)$ son determinadas por las siguientes fórmulas:

$$\begin{aligned} E(X_{t_2}^0 - X_{t_1}^0)(X_{t_2}^{0*} X_{t_1}^{0*}) &= k(t_2) - k(t_1), \\ K_x(t_1, t_2) &= k(\max(t_1, t_2)); t_1, t_2 > t_0, \\ &= 0; t \leq t_0 \leq t_2, t_2 \leq t_0 \leq t_1, \\ &= -k(\min(t_1, t_2)); t_1, t_2 < 0. \end{aligned} \tag{2.9}$$

El proceso aleatorio con incrementos no correlacionados es continuo por el criterio de media cuadrada si y sólo si la función $k(t)$ es continua.

Por otro lado, si $k(t)$ es no sólo continua, sino además diferenciable, en este caso la fórmula (2.10) puede ser escrita (con $t_1, t_2 > t_0$) en la forma:

$$k(t) = \int_{t_0}^t \nu(\tau) d\tau, \tag{2.10}$$

donde $\nu(t)$ es una función no negativa la cual es llamada la intensidad de un proceso $X(t)$ con incrementos no correlacionados.

2.2.3. Procesos con Incrementos Independientes

Se dice que un proceso aleatorio $X(t)$ es un proceso con incrementos independientes si para cualquier $\mathcal{N}, t_0 < t_1 < \dots < t_{\mathcal{N}}$ las variables aleatorias $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_{\mathcal{N}}} - X_{t_{\mathcal{N}-1}}$, son independientes.

Teorema. La función característica del incremento de un proceso con incrementos independientes está completamente determinada por su función característica en una dimensión, i.e. su distribución uni-dimensional (ver demostración en [53]).

Todo proceso con incrementos independientes y con función de covarianza diferenciable tiene una derivada débil (en el criterio media cuadrada), la cual representa un ruido blanco. El ruido blanco obtenido por diferenciación de un proceso con incrementos independientes es llamado un ruido blanco en el sentido estricto.

2.2.4. Proceso de Wiener

Un escalar o un proceso aleatorio real con incrementos independientes $W(t), t > 0$, es llamado un proceso de Wiener si satisface las siguientes condiciones:

- Todas las realizaciones $w(t)$ del proceso $W(t)$ son continuas y $w(0) = 0$.
- La distribución uni-dimensional del proceso $W(t)$ es normal.
- La esperanza del proceso $W(t)$ es cero y su función de covarianza es determinada por la fórmula

$$K_w(t_1, t_2) = \int_0^{\min(t_1, t_2)} \nu(\tau) d\tau,$$

donde $\nu(t)$ es una función no negativa la cual representa la intensidad del proceso de Wiener $W(t)$. El ruido blanco representando la derivada débil por media cuadrática de un proceso de Wiener es llamado ruido blanco normalmente distribuido. Un proceso de Wiener como un proceso con incrementos independientes el cual posee media cero y momento de segundo orden finito $k(t)$ para cada instante t , genera una medida estocástica en el eje real con valores independientes en intervalos disjuntos. Esta medida estocástica es determinada por la fórmula $Z((t_1, t_2]) = W(t_2) - W(t_1)$.

Fórmula para la Función $\chi(\mu, t)$ para el Proceso de Wiener

Una forma específica para la función $\chi = \chi(\mu; t)$ en las ecuaciones obtenidas para las funciones características es determinada por el carácter del proceso con incrementos independientes $W(t)$. Aquí es obtenida cuando $W(t)$ es un proceso de Wiener, pero se puede calcular para cualquier proceso, aunque este cálculo puede ser más complicado. Si $W(t)$ es el proceso de Wiener, entonces su función característica uni-dimensional $h_1(\mu, t)$ es determinada por la fórmula

$$h_1(\mu, t) = \exp\left\{-\frac{1}{2}\mu^T \int_0^t \nu(\tau) d\tau \mu\right\} \quad (2.11)$$

Ya que la función $\chi(\mu; t)$ representa la derivada logarítmica de la función característica $h_1(\mu; t)$ con respecto a t : $\chi(\mu; t) = \frac{\partial}{\partial t}[\ln h_1(\mu; t)]$. Sustituyendo aquí la expresión de la función $h_1(\mu; t)$, se obtiene

$$\chi(\mu; t) = -\mu^T \nu(t) \mu / 2 \quad (2.12)$$

Ecuaciones para las Densidades en el Caso Multi-Dimensional del Proceso de Wiener

En el caso de un proceso de Wiener $W(t)$, de acuerdo con (2.12), es obtenida la siguiente ecuación

$$\chi(b(\eta, t)^T \lambda; t) = -\frac{1}{2} \lambda^T b(\eta, t) \nu(t) b(\eta, t)^T \lambda. \quad (2.13)$$

La ecuación (2.12) toma la forma

$$\begin{aligned} \frac{\partial f_1(y, t)}{\partial t} &= \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [i\lambda^T a(\eta, t) - \frac{1}{2} \lambda^T b(\eta, t) \nu(t) b(\eta, t)^T] \\ &\quad \times e^{i\lambda^T(\eta, y)} f_1(\eta; t) d\eta d\lambda. \end{aligned} \quad (2.14)$$

Tomando en cuenta que $u^T A u = \text{tr}(u u^T A)$ para un n-vector u y una matriz $A_{n \times n}$, es obtenida la siguiente fórmula

$$\begin{aligned} \frac{\partial f_1(y, t)}{\partial t} &= \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i\lambda^T a(\eta, t) - \frac{1}{2} \text{tr}[\lambda \lambda^T b(\eta, t) \nu(t) b(\eta, t)^T] \\ &\quad \times e^{i\lambda^T(\eta, y)} f_1(\eta; t) d\eta d\lambda. \end{aligned} \quad (2.15)$$

Utilizando la fórmula integral de la función delta

$$\frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} e^{i\lambda^T(\eta, y)} d\lambda = \delta(\eta - y) \quad (2.16)$$

Diferenciando (2.16) con respecto a η son obtenidas las siguientes fórmulas

$$\begin{aligned} \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} i\lambda e^{i\lambda^T(\eta, y)} d\lambda &= \frac{\partial}{\partial \eta} \delta(\eta - y) = \delta'(\eta - y), \\ -\frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \lambda \lambda^T e^{i\lambda^T(\eta, y)} d\lambda &= \frac{\partial}{\partial \eta} \frac{\partial^T}{\partial \eta} \delta(\eta - y) = \delta''(\eta - y). \end{aligned} \quad (2.17)$$

En base con las últimas tres fórmulas anteriores, es posible escribir

$$\begin{aligned}
\frac{\partial f_1(y, t)}{\partial t} &= \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i\lambda^T a(\eta, t) e^{i\lambda^T(\eta-y)} f_1(\eta; t) d\eta d\lambda = \quad (2.18) \\
&\int_{-\infty}^{\infty} \dot{\delta}(\eta - y)^T a(\eta, t) f_1(\eta; t) d\eta = -\frac{\partial^T}{\partial y} [a(y, t) f_1(y, t)], \\
&-\frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \lambda \lambda^T b(\eta; t) \nu(t) b(\eta; t)^T e^{i\lambda^T(\eta-y)} f_1(\eta; t) d\eta d\lambda = \\
&\int_{-\infty}^{\infty} \delta''(\eta - y) b(\eta; t) \nu(t) b(\eta; t)^T f_1(\eta; t) d\eta = \frac{\partial}{\partial y} \frac{\partial^T}{\partial y} [b(\eta; t) \nu(t) b(\eta; t)^T f_1(\eta; t)]
\end{aligned}$$

Usando las fórmulas (2.14), es posible representar la ecuación (2.15) en la forma

$$\frac{\partial f_1(y, t)}{\partial t} = \frac{\partial^T}{\partial y} [a(y, t) f_1(y, t)] + \frac{1}{2} tr \frac{\partial}{\partial y} \frac{\partial^T}{\partial y} [b(y, t) \nu(t) b(y; t)^T f_1(y, t)] \quad (2.19)$$

Esta ecuación (2.19) es llamada la ecuación de Fokker-Planck-Kolgomorov. La cual es obtenida sólo para el proceso de Wiener. La ecuación (2.19) también es válida para para la densidad n-dimensional $f_n(y_1, \dots, y_n; t_1, \dots, t_n)$ si la diferenciación con respecto a t y a y es considerada como la diferenciación respecto a t_n y y_n .

2.2.5. Proceso de Poisson

Distribución de Poisson

En la práctica ocurren eventos en instancias de tiempo aleatorias. Estas forman una secuencia de eventos que usualmente llamamos flujo de eventos. Frecuentemente podemos asumir que un flujo de eventos satisface las siguientes condiciones:

1. Para cualesquier dos intervalos de tiempo que no se intersectan, la probabilidad de cualquier número de eventos durante uno de ellos es independiente del número de eventos durante el otro.
2. La probabilidad de la aparición de un evento durante un intervalo infinitesimal de tiempos $(t, t+\Delta t)$ es un infinitesimal de orden Δt .

3. La probabilidad de la aparición de más de un evento durante el intervalo de tiempo $(t, t+\Delta t)$ es un infinitesimal de orden Δt .

Denotemos como $p_m(t_1, t_2)$ a la probabilidad de aparición de m eventos en el intervalo de tiempo (t_1, t_2) . Entonces las condiciones 2 y 3 toman la forma

$$p_1(t, t + \Delta t) = \lambda(t)\Delta t + O(\Delta t) \quad (2.20)$$

$$\sum_{k=2}^{\infty} p_k(t, t + \Delta t) = O(\Delta t) \quad (2.21)$$

donde $\lambda(t)$ es una función no negativa.

Si tenemos el caso de un flujo de eventos que satisfacen las condiciones anteriores 1, 2 y 3 consideremos la probabilidad de que en un intervalo de tiempo dado (t_0, t) aparecerán los eventos $(x = 0, 1, 2, \dots)$. Si consideramos el instante t_0 fijo denotamos las probabilidades desconocidas p_x , $(x = 0, 1, 2, \dots)$.

Para evaluar $p_0(t)$ notemos que $p_0(t, t + \Delta t)$ es la probabilidad de la intersección de dos eventos: ninguno de los eventos ocurren en el intervalo (t_0, t) y ninguno de los eventos ocurre en el intervalo $(t, t + \Delta t)$. De acuerdo a la condición 1 mencionada en 2.2, estos eventos son independientes. Consecuentemente

$$p_0(t + \Delta t) = p_0(t)p_0(t, t + \Delta t) \quad (2.22)$$

de esto tenemos:

$$p_0(t, t + \Delta t) = 1 - \sum_{k=1}^{\infty} p_k(t, t + \Delta t) = 1 - \lambda(t)\Delta t + O(\Delta t) \quad (2.23)$$

Para un intervalo de tiempo dado (t_0, t) tenemos el conjunto contable de eventos elementales: ninguno en el intervalo, un evento, dos eventos, etc., y las probabilidades de

estos eventos están determinadas por la fórmula

$$p(x; \lambda) = \frac{\lambda^x \exp^{-\lambda}}{x!}$$
$$x = 0, 1, 2, \dots$$

Por lo tanto esta fórmula, es llamada *Distribución de Poisson*. Por esta razón el flujo de eventos que satisface las condiciones mencionadas anteriormente es llamado *Flujo de Poisson*. El parámetro λ de la distribución de Poisson es el número promedio de eventos que ocurren en el intervalo (t_0, t) .

Proceso de Poisson

Un *Proceso de Poisson*, denotado por $N(t)$, es un proceso con incrementos independientes y su función de distribución unidimensional es la distribución de Poisson. El proceso de Poisson simple $P(t)$ es un proceso aleatorio escalonado con incrementos por unidades en cualquier momento de la ocurrencia de un evento de un flujo de Poisson, donde se tiene que para cualquier t , el valor del proceso $P(t)$ es igual al número de eventos de un flujo de Poisson en el intervalo de tiempo $[0, t]$. Debido a esto el proceso de Poisson simple puede ser llamado como el contador de eventos del flujo de Poisson.

Sea $P(t)$ un proceso simple de Poisson. El proceso aleatorio

$$N(t) = \sum_{k=1}^{P(t)} X_k$$

donde X_k es una secuencia de variables aleatorias independientes igualmente distribuidas del proceso general de Poisson. Este es un proceso con incrementos independientes.

Su derivada es considerada como un ruido blanco. Debido a que la representación gráfica de un proceso de Poisson es una función escalonada, centralizando, es decir, restando su esperanza, es posible obtener una representación gráfica del ruido blanco de Poisson, (ya que la esperanza del proceso de Poisson simple es $\lambda \neq 0$).

2.3. Cálculo Estocástico

2.3.1. La Integral de Ito

Sea $W(t)$ un proceso escalar con incrementos independientes con esperanza cero y momento de segundo orden $k(t)$; sea $X(t)$ un proceso aleatorio continuo por media cuadrada (m.s.), escalar con momento de segundo orden finito cuyo valor en todo t es independiente de los incrementos futuros del proceso $W(t)$, $W(t_2) - W(t_1)$, $t \leq t_1 < t_2$; sea P_n una secuencia de particiones del intervalo $(a, b]$,

$$P_n : (a, b] = \bigcup_{k=1}^{N_n} (t_{k-1}^{(n)}, t_k^{(n)}], t_0^{(n)} = a, t_{N_n}^{(n)} = b,$$

tal que $\max_k (t_k^{(n)} - t_{k-1}^{(n)}) \rightarrow 0$ cuando $n \rightarrow \infty$.

El límite por media cuadrática de la secuencia de sumas integrales $\{Y_n\}$,

$$Y_n = \sum_{k=1}^{N_n} X(t_{k-1}^{(n)}) [W(t_k^{(n)}) - W(t_{k-1}^{(n)})],$$

si este límite existe, es llamado la Integral estocástica de Ito de la función aleatoria $X(t)$ con respecto al proceso con incrementos independientes $W(t)$ sobre el intervalo $(a, b]$:

$$Y = \int_a^b X(t) dW(t) = \lim_{n \rightarrow \infty} Y_n$$

Teorema La integral estocástica de Ito existe si y sólo si la integral

$$\int_a^b E|X(t)|^2 \nu(t) dt = E|Y|^2 = DY$$

existe en cuyo caso esta es igual a la varianza DY de la integral de Ito Y .

Definiendo una θ -integral estocástica para algún $\theta \in [0, 1]$ por la siguiente fórmula:

$$Y_\theta = \int_a^b X(\tau) d_\theta W(\tau) = (1 - \theta)Y + \theta Y_1$$

siendo un caso particular de la integral de Ito cuando $\theta = 0$. Cuando $\theta = 1/2$ la integral estocástica anterior representa la integral estocástica de Stratonovich.

2.3.2. Ecuaciones Diferenciales Estocásticas

La ecuación diferencial

$$\dot{X} = a(X, t) + b(X, t)V \quad (2.24)$$

es llamada ecuación diferencial estocástica si la función aleatoria (generalizada) $V = V(t)$ representa un ruido blanco en el sentido estricto. Sea X_0 un vector aleatorio de la misma dimensión que $X(t)$. La ecuación (2.24) con la condición inicial $X(t_0) = X_0$ determina el proceso aleatorio $X(t)$. La ecuación anterior también puede escribirse como

$$X(t) = X_0 + \int_{t_0}^t a(X(\tau), \tau)d\tau + \int_{t_0}^t b(X(\tau), \tau)V(\tau)d\tau \quad (2.25)$$

donde la primera integral existe por el criterio de media cuadrática. Introduciendo el proceso con incrementos independientes $W(t)$ cuyas derivadas son un ruido blanco $V(t)$ la ecuación anterior se puede reescribir como

$$X(t) = X_0 + \int_{t_0}^t a(X(\tau), \tau)d\tau + \int_{t_0}^t b(X(\tau), \tau)dW(\tau). \quad (2.26)$$

La ecuación (2.26) tiene un sentido exacto. La ecuación (2.24) con la condición inicial $X(t_0) = X_0$ representa una forma concisa de la ecuación (2.26). La ecuación (2.26) en la cual la segunda integral una integral estocástica de Ito, es llamada la ecuación Integral estocástica de Ito, y la ecuación (2.24) y la que se forma sustituyendo dW por V en (2.24) son llamadas ecuaciones diferenciales estocásticas de Ito. Un proceso aleatorio $X(t)$ el cual satisface la ecuación (2.26) en el cual las integrales representan límites por el criterio de media cuadrática de la correspondiente suma de integrales, es llamado solución de la ecuación integral estocástica (2.26) y de la ecuación diferencial estocástica correspondiente (2.24) con la condición inicial $X(t_0) = X_0$.

Momentos

Consideremos el sistema lineal

$$\dot{Y} = aY + a_0 + bV, \quad (2.27)$$

donde $a = a(t)$, $a_0 = a_0(t)$, $b = b(t)$ pueden ser funciones en el tiempo t , y V es un ruido blanco cuya intensidad ν puede ser una función de tiempo t . Resolviendo la ecuación (2.27) es decir, obtenido el vector Y , el cual está dado por la fórmula

$$Y(t) = u(t, t_0)Y_0 + \int_{t_0}^t u(t, \tau)b(\tau)V(\tau)d\tau + \int u(t, \tau)a_0(\tau)d\tau, \quad (2.28)$$

donde $u(t, \tau)$ es la matriz determinada como una función de t por la ecuación diferencial homogénea $\frac{du}{dt} = a(t)u$ y la condición inicial $u(\tau, \tau) = I$.

Momentos de Segundo Orden

Tomando en cuenta que la esperanza del un ruido blanco es igual a cero, en virtud de (2.28), es encontrada la siguiente fórmula para la esperanza del vector estado del sistema $Y(t)$

$$m(t) = u(t, t_0)m_0 + \int_{t_0}^t u(t, \tau)a_0(\tau)d\tau \quad (2.29)$$

donde m_0 es la esperanza del valor inicial Y_0 del vector de estado Y . La función de covarianza del vector de estado Y es determinada por la fórmula

$$K(t_1, t_2) = u(t_1, t_0)K_0u(t_2, t_0) + \int_{t_0}^{\min(t_1, t_2)} u(t_1, \tau)b(\tau)\nu(\tau)b(\tau)^T u(t_2, \tau)^* d\tau \quad (2.30)$$

donde K_0 es la matriz de covarianza del valor inicial Y_0 del vector de estado Y . Dadas las funciones de valores reales $m(t)$ y $K(t_1, t_2)$, sigue la fórmula para el momento de segundo

orden.

$$\Gamma(t_1, t_2) = K(t_1, t_2) + m(t_1)m(t_2)^T. \quad (2.31)$$

La ecuación diferencial para la esperanza del vector Y es obtenida diferenciando (2.29):

$$\begin{aligned} \dot{m}(t) &= u_t(t, t_0)m_0 + \int_{t_0}^t u_t(t, \tau)a_0(\tau)d\tau + a_0(t) \\ &= a(t)[u(t, t_0)m_0 + \int_{t_0}^t u_t(t, \tau)a_0(\tau)d\tau] + a_0(t) \end{aligned} \quad (2.32)$$

pero la expresión entre corchetes es igual a $m = m(t)$ por(2.29). Por lo tanto la ecuación (2.32) se puede escribir en la forma

$$\dot{m} = am + a_0. \quad (2.33)$$

Integrando la ecuación (2.33) con la condición inicial $m(t_0) = m_0$ se puede calcular la esperanza del vector aleatorio Y en el sistema lineal estocástico (2.27).

La ecuación para la matriz de la varianza $K(t)$ del vector Y en (2.30) $t_1 = t_2 = t$:

$$\begin{aligned} K(t) &= K(t, t) = u(t, t_0)K_0u(t, t_0)^* + \\ &\quad \int_{t_0}^t u(t, \tau)b(\tau)\nu(\tau)b(\tau)^T u(t, \tau)^* d\tau \end{aligned} \quad (2.34)$$

Diferenciando esta fórmula respecto a t y sustituyendo $u_t(t, \tau) = a(t)u(t, \tau)$, $u_t(t, \tau)^* = u(t, \tau)^*a(t)^T$ se obtiene

$$\begin{aligned} \dot{K}(t) &= a(t)[u(t, t_0)K_0u(t, t_0)^* + \\ &\quad \int_{t_0}^t u(t, \tau)b(\tau)\nu(\tau)b(\tau)^T u(t, \tau)^* d\tau] + [u(t, t_0)K_0u(t, t_0)^* \\ &\quad \int_{t_0}^t u(t, \tau)b(\tau)\nu(\tau)b(\tau)^T u(t, \tau)^* d\tau]a(t)^T + b(t)\nu(t)b(t)^T. \end{aligned} \quad (2.35)$$

Dado que la expresión en corchetes es igual a $K = K(t)$ entonces

$$\dot{K} = aK + Ka^T + b\nu b^T. \quad (2.36)$$

Integrando la ecuación (2.36) con la condición inicial $K(t_0) = K_0$ se puede calcular la matriz de la varianza del vector aleatorio Y para el sistema lineal estocástico (2.27). La ecuación diferencial para el momento de segundo orden $\Gamma(t)$ del vector Y con $t_1 = t_2 = t$ se puede obtener en base a la fórmula

$$\Gamma(t) = K(t) + m(t)m(t)^T, \quad (2.37)$$

diferenciando la fórmula anterior se obtiene

$$\dot{\Gamma} = \dot{K} + \dot{m}m^T + m\dot{m}^T. \quad (2.38)$$

Sustituyendo aquí las expresiones para \dot{m} y \dot{K} para las ecuaciones (2.33) y (2.36) y usando la fórmula (2.37) se llega a

$$\dot{\Gamma} = a\Gamma + \Gamma a^T + b\nu b^T + a_0 m^T + m a_0^T. \quad (2.39)$$

Integrando la ecuación (2.39) después la ecuación (2.33) la cual determina la esperanza m con la condición inicial $\Gamma(t_0) = \Gamma_0 = K_0 + m_0 m_0^T$ se puede calcular el momento inicial de segundo orden del vector aleatorio Y en el sistema lineal estocástico (2.27). La ecuación para la función de covarianza $K(t_1, t_2)$ del proceso aleatorio Y considerada como una función de t_1 y algún t_1 fijo, con el caso $t_1 < t_2$:

$$K(t_1, t_2) = u(t_1, t_0)K_0 u(t_2, t_0)^* + \int_{t_0}^{t_1} u(t_1, \tau)b(\tau)\nu(\tau)b(\tau)^T u(t_2, \tau)^* d\tau \quad (2.40)$$

Diferenciando la fórmula anterior respecto a t_2 :

$$\begin{aligned} \frac{\partial K(t_1, t_2)}{\partial t_2} &= u(t_1, t_0)K_0 u_{t_2}(t_2, t_0)^* + \\ &\int_{t_0}^{t_1} u(t_1, \tau)b(\tau)\nu(\tau)b(\tau)^T u(t_2, \tau)^* d\tau = u(t_1, t_0)K_0 u(t_2, t_0)^* a^T(t_2) + \\ &\int_{t_0}^{t_1} u(t_1, \tau)b(\tau)\nu(\tau)b(\tau)^T u(t_2, \tau)^* a^T(t_2) d\tau, \\ &= K(t_1, t_2)a(t_2)^T, t_1 < t_2. \end{aligned} \quad (2.41)$$

con la condición inicial $K(t_1, t_1) = K(t_1)$.

Integrando la ecuación (2.41) con varios valores de t_1 se obtiene el número de secciones de la matriz de covarianza $K(t_1, t_2)$ con $t_1 < t_2$. Al obtener $K(t_1, t_2)$ con $t_2 < t_1$ se usa

$$K(t_1, t_2) = K(t_2, t_1)^T$$

Función Característica Uni-dimensional

Considere el sistema cuyo vector de estado es descrito por la ecuación estocástica diferencial de Ito

$$\dot{Y} = a(Y, t) + b(Y, t)V, \quad (2.42)$$

donde V es un ruido blanco en el sentido estricto. El problema es encontrar la distribución multidimensional del estado del sistema $Y(t)$, suponiendo que la distribución uni-dimensional del proceso con incrementos independientes

$$W(t) = W(t_0) + \int_{t_0}^t V(\tau)d\tau \quad (2.43)$$

es conocida. La ecuación de la función característica uni-dimensional está dada por

$$\frac{\partial g_1(\lambda, t)}{\partial t} = E\{i\lambda^T a(Y, t) + \chi(b(Y, t)^T \lambda; t)\}e^{i\lambda^T Y} \quad (2.44)$$

La ecuación multidimensional de la función característica de un vector estado Y de un sistema está dada por

$$g_n(\lambda_1, \dots, \lambda_n; t_1, \dots, t_n) = E \exp\{i \sum_{k=1}^n \lambda_k^T Y(t_k)\} \quad (2.45)$$

Supongamos que la densidad uni-dimensional $f_1(y, t)$ para el vector estado del sistema existe. Entonces la ecuación (2.44) puede escribirse como

$$\frac{\partial g_1(\lambda, t)}{\partial t} = \int_{-\infty}^{\infty} [i\lambda^T a(y, t) + \chi(b(y, t)^T \lambda; t)]e^{i\lambda^T y} f_1(y; t) dy \quad (2.46)$$

Por la Transformada de Fourier

$$f_1(y; t) = \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} e^{-i\mu^T y} g_1(\mu, t) d\mu \quad (2.47)$$

donde p es la dimensión del vector de estado Y , y la integral con respecto a todos los componentes del vector p -dimensional μ es asumida como el valor principal de la integral en el sentido de Cauchy si $g_1(\mu, t)$ es no integrable absolutamente. Sustituyendo la ecuación (2.47) en la ecuación (2.46) es obtenida la ecuación integro diferencial lineal

$$\frac{\partial g_1(\lambda, t)}{\partial t} = \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [i\lambda^T a(y, t) + \chi(b(y, t)^T \lambda; t)] \times \quad (2.48)$$

$$e^{i(\lambda^T - \mu^T)y} g_1(\mu; t) d\mu dy$$

Análogamente, suponiendo que la densidad multidimensional del proceso $Y(t)$ existe, es obtenida la ecuación integro diferencial relativa a $g_n(\lambda_1, \dots, \lambda_n; t_1, \dots, t_n)$:

$$\frac{\partial}{\partial t_n} g_n(\lambda_1, \dots, \lambda_n; t_1, \dots, t_n) \quad (2.49)$$

$$= \frac{1}{(2\pi)^{np}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [i\lambda^T a(y_n, t_n) + \chi(b(y_n, t_n)^T \lambda_n; t_n)]$$

$$\times \exp\left\{i \sum_{k=1}^n (\lambda_k^T - \mu_k^T) y_k\right\} g_n(\mu_1, \dots, \mu_n; t_1, \dots, t_n)$$

$$\times d\mu_1 \Delta \Delta \Delta d\mu_n dy_1 \Delta \Delta \Delta dy_n$$

Ecuaciones Uni-Dimensionales para las Densidades

Reemplazando en la ecuación (2.46) la variable de integración y por η , multiplicando esta ecuación por $(2\pi)^{-p} e^{-i\lambda^T y}$ e integrando esto con respecto a λ es obtenida la ecuación integro-diferencial para la densidad de una dimensión $f_1(y, t)$

$$\frac{\partial f_1(y, t)}{\partial t} = \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [i\lambda^T a(\eta, t) + \chi(b(\eta, t)^T \lambda; t)] e^{i\lambda^T (\eta - y)} f_1(\eta; t) d\eta d\lambda \quad (2.50)$$

2.4. Teoría de Filtrado Óptimo

2.4.1. Filtro de Wiener

Planteamiento del Problema

El filtro de Wiener probablemente representa la primera presentación de terminología en el cual dos importantes ideas han sido rescatadas: sistemas dinámicos y estimación óptima en la presencia de ruido. Se considera una señal $y(\cdot)$ la cual contiene un ruido $v(\cdot)$ y una medida $z(\cdot)$. $y(\cdot)$, $v(\cdot)$, y $z(\cdot)$ pueden originar un problema del tipo continuo, o discreto en el tiempo dependiendo de la naturaleza de los mismos. Las señales de tiempo son consideradas como escalares continuos definidos en el intervalo $(-\infty, \infty)$ solamente. Se asume que $y(\cdot)$, y $v(\cdot)$, son funciones simples de procesos aleatorios estacionarios. Normalmente ellos son independientes y tienen media cero. Posteriormente ellos son considerados para la obtención de $\phi_{yy}(jw)$ y $\phi_{vv}(jw)$, $w \in \mathbb{R}$ -espectra. La tarea del filtro de Wiener es utilizar las mediciones $z(\cdot)$ para estimar $y(\cdot)$. Más precisamente, se requiere que la estimación sea causal, en línea y óptima. Causal significa que $y(t)$ va a ser estimada usando $z(s)$ para algún $s < t$; en línea significa que al tiempo t el estimado de $y(t)$ debería de desempeñarse óptimamente. Óptimo significa que $\hat{y}(t)$, debería de presentar un error cuadrado mínimo, i.e. $E[y(t) - \hat{y}(t)]^2$ el cual debe de ser minimizado. Si $y(\cdot)$, y $v(\cdot)$ son Gaussianos, esto significa que $\hat{y}(t)$ es el estimado condicional, $E[y(t)/z(s), s \leq t]$.

Solución

La solución a este problema está dada en la siguiente explicación: El filtro de Wiener es un sistema lineal, invariante en el tiempo, causal, estable, cuya relación entrada-salida

está dada por una función de transferencia $h(\cdot)$:

$$\hat{y}(t) = \int_{-\infty}^t h(t-s)z(s)ds \quad (2.51)$$

La señal $y(\cdot)$ y el ruido $v(\cdot)$ son representados como la salida de un sistema lineal excitado por ruido blanco. Si $\varepsilon_y(\cdot), \varepsilon_v(\cdot)$ son ruidos blancos con media cero y la intensidad de la varianza 1, entonces

$$E[\varepsilon_y(t)\varepsilon_y(s)] = E[\varepsilon_v(t)\varepsilon_v(s)] = \delta(t-s), \quad (2.52)$$

y por lo tanto

$$\phi_{yy}(jw) = /W_y(jw)/^2, \phi_{vv}(jw) = /W_v(jw)/^2 \quad (2.53)$$

La clave del problema es la obtención de $\phi_{yy}(jw), \phi_{vv}(jw)$ para la función de respuesta al impulso $h(t)$ o su función de transferencia $H(jw)$. El paso crucial es la técnica de factorización espectral. El espectro de $z(\cdot)$ cuando $y(\cdot)$ y $v(\cdot)$ son independientes está dado por

$$\phi_{zz}(jw) = \phi_{yy}(jw) + \phi_{vv}(jw) \quad (2.54)$$

La factorización espectral requiere de la determinación de una función de transferencia $W_z(jw)$ tal que $W_z(s)$ y $W_z^{-1}(s)$ son analíticas en $\mathbb{R}, s \geq 0$ y tal que

$$\phi_{zz}(jw) = /W_z(jw)/^2, \quad (2.55)$$

En [38] esta operación de factorización espectral es presentada como un paso crucial en la obtención de $H(\cdot)$, la cual en [36] es la clave para la determinación del filtro óptimo. A continuación se procede de la siguiente manera. Se define una señal $\varepsilon_z(\cdot)$ como la salida de un sistema lineal de una función de transferencia $W_z^{-1}(jw)$ conducida por $z(\cdot)$. Si existe

$W_z^{-1}(\cdot)$ entonces $\varepsilon_z(\cdot)$ es equivalente a $z(\cdot)$, es decir la estimación de $y(t)$ usando $\varepsilon_z(s)$ para $s < t$ debería dar el mismo resultado como estimación de $y(t)$ usando $z(s)$ para $s < t$ y además $\varepsilon_z(\cdot)$ es un ruido blanco. Esta simplificación es muy importante y es utilizada para la obtención del filtro óptimo en [45].

Además, es notable que la construcción de $W_z(\cdot)$ satisface las condiciones de estabilidad y (2.55) y es un paso importante para la construcción de $H(\cdot)$. La pregunta es ¿cómo puede ser hecho esto si $\phi_{zz}(\cdot)$ es racional?, la clave es la factorización polinomial. En otro caso, utilizar:

$$W_z(jw_0) = \min_{\varepsilon \rightarrow 0} \exp\left\{\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\log \phi_{zz}(jw)}{-j(w - w_0) - \varepsilon} dw\right\} \quad (2.56)$$

Otra forma de resolver el problema de filtrado, en el dominio del tiempo, es utilizando la función de respuesta al impulso $h(t)$, la cual corresponde a la transformada inversa de Laplace de la función $H(jw)$, mediante la ecuación

$$h(t) + \int_0^t h(\tau)K(\tau - s)ds = K(t), t \geq 0, \quad (2.57)$$

donde $K(\tau)$ es la función de covarianza de $y(t)$. Esta ecuación es conocida como la ecuación de Wiener-Hopf.

2.4.2. Filtro de Kalman-Bucy (Caso Discreto)

Prácticamente, todo lo establecido para el filtro de Kalman en el tiempo continuo, se traslada al caso del filtro con tiempo discreto. La teoría en el caso continuo se aprecia más transparente que en el caso discreto, ya que presenta aplicabilidad a más problemas.

Planteamiento del problema

El modelo está dado por

$$\begin{aligned}x_{k+1} &= F_k x_k + G_k w_k \\z_k &= H^T k x_k + v_k\end{aligned}\tag{2.58}$$

con

$$E \begin{bmatrix} w_k \\ v_k \end{bmatrix} \begin{bmatrix} w_l^T & v_l^T \end{bmatrix} = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix} \delta_{kl}$$

y $\{w_k\}, \{v_k\}$ son secuencias con media cero. Por convencionalismo, se considera el tiempo inicial $k = 0$. Agregando que la media \bar{x}_0 y la varianza P_0 de x_0 , son independientes de $\{w_k\}, \{v_k\}$. Todas las variables son Gaussianas. La idea principal es distinguir el efecto de dinámicas y mediciones en el filtro. Más precisamente, sea $\hat{x}_{k/k}$ el estimado óptimo, una media estimada de x_k dada $z_l, l \leq k$, y sea $\hat{x}_{\frac{k+1}{k}}$ dado por $E[x_{k+1}/z_l, l \leq k]$, el primer paso en la predicción del estimado.

Solución

Dado que w_k es independiente de z_l para $l \leq k$, se tiene

$$\hat{x}_{\frac{k+1}{k}} = F_k \hat{x}_{\frac{k}{k}}\tag{2.59}$$

Esto demuestra como actualizar un estimado como resultado de sistemas dinámicos, cuando no aparecen mediciones extras. (2.59) se apoya en

$$\mathbb{V}_{\frac{k+1}{k}} = F_k \mathbb{V}_{\frac{k}{k}} F_k^T + G_k Q_k G_k^T\tag{2.60}$$

Aquí $\mathbb{V}_{\frac{k+1}{k}}$, y $\mathbb{V}_{k/k}$ son las covarianzas del error asociadas con $\hat{x}_{\frac{k}{k}}$ y $\hat{x}_{\frac{k+1}{k}}$. Actualizar las ecuaciones de los estimados equivale a pasar de $\hat{x}_{\frac{k+1}{k}}$ y $\mathbb{V}_{\frac{k+1}{k}}$ a $\hat{x}_{\frac{k+1}{k+1}}$ y $\mathbb{V}_{\frac{k+1}{k+1}}$. Y esto se

muestra a continuación

$$\begin{aligned}\widehat{x}_{\frac{k+1}{k+1}} &= \widehat{x}_{\frac{k+1}{k}} + \bigvee_{\frac{k+1}{k}} H_{k+1} [H_{k+1}^T \bigvee_{\frac{k+1}{k}} H_{k+1} + R_{k+1}]^{-1} \\ &\quad \times [z_{k+1} - H_{k+1}^T \widehat{x}_{\frac{k+1}{k}}] \\ \bigvee_{\frac{k+1}{k+1}} &= \bigvee_{\frac{k+1}{k}} - \bigvee_{\frac{k+1}{k}} H_{k+1} [H_{k+1}^T \bigvee_{\frac{k+1}{k}} H_{k+1} + R_{k+1}]^{-1} H_{k+1}^T \bigvee_{\frac{k+1}{k}}\end{aligned}\tag{2.61}$$

2.4.3. Filtro de Kalman-Bucy (Caso Continuo)

Planteamiento del Problema

La representación del modelo está dada por

$$\frac{dx(t)}{dt} = F(t)x(t) + G(t)w(t)\tag{2.62}$$

$$z(t) = H^T(t)x(t) + v(t)\tag{2.63}$$

en el cual F, G, H son matrices $n \times n, n \times m$, y $n \times p$ respectivamente. El proceso $w(\cdot)$ y $v(\cdot)$ son ruidos blancos Gaussianos con media cero tales que

$$E \begin{bmatrix} w(t) \\ v(t) \end{bmatrix} \begin{bmatrix} w^T(s) & v^T(s) \end{bmatrix} = \begin{bmatrix} Q(t) & S(t) \\ S^T(t) & R(t) \end{bmatrix} \delta(t-s)$$

con $R(t) = R'(t) > 0$ para toda t . Muy frecuentemente, $S(t) \equiv 0$, i.e. $w(\cdot)$ y $v(\cdot)$ son independientes. Lo cual es supuesto. Entonces $Q(t) = Q^T(t) \geq 0$. Se asume un tiempo inicial finito t_0 . Por otro lado $x(t_0)$ será asumida como variable aleatoria Gaussiana con media x_0 y varianza p_0 . La tarea de la estimación es usar mediciones de $z(s)$ para $s < t$ para estimar $x(t)$, este estimado es llamado $\widehat{x}(t)$, el cual minimiza $E[\|x(t) - \widehat{x}(t)\|^2]$. Esto significa que $\widehat{x}(t)$ es necesariamente una estimación de la media condicional, con respecto a las observaciones.

Solución

La solución es obtenida de la siguiente manera. Definamos $P(t) = P^T(t) \geq 0$ como la solución de

$$\dot{P} = PF^T + FP - PHR^{-1}H^T P + GQG^T, P(t_0) = P_0 \quad (2.64)$$

Y $\hat{x}(t)$ es la solución de

$$\frac{d\hat{x}}{dt} = F(t)\hat{x}(t) + P(t)H(t)R^{-1}(t)[z(t) - H^T(t)\hat{x}(t)] \quad (2.65)$$

Donde $P(t)H(t)R^{-1}(t)$ denota la ganancia de Kalman. $E[x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]^T = P(t)$. La efectividad del estimador óptimo es medida por la covarianza del error, la cual es dada por la solución de la ecuación (2.64), y la existencia de la solución a esta ecuación en (t_0, ∞) está garantizada.

Algunas diferencias del filtro de Kalman con respecto al de Wiener son dadas en la siguiente tabla.

Filtro de Wiener	Filtro de Kalman
$t_0 = -\infty$	$t_0 \geq -\infty$
Estacionario	Acepta no estacionario.
Infinito dimensional	Finito dimensional
Ruido no necesariamente blanco	Ruido blanco
Factorización espectral	Solución de la ecuación de Riccati
Estimación de la señal	Estimación del estado

El problema de predicción es resuelto por la teoría de filtrado. Esto consiste en calcular $x(t + \Delta)$ para algún Δ positivo, dado $z(s)$ para $s < t$, esto es:

$$\hat{x}(t + \Delta) = \Phi(t + \Delta)\hat{x}(t) \quad (2.66)$$

2.4.4. Filtro Óptimo Polinomial

Sea (Ω, F, P) un espacio completo de probabilidad con una familia creciente y continua por la derecha de σ -álgebras $F_t, t \geq t_0$, y sean $(W_1(t), F_t, t \geq t_0)$ y $(W_2(t), F_t, t \geq t_0)$ procesos independientes de Wiener. El F_t medible proceso aleatorio $(x(t), y(t))$ es descrito por una ecuación diferencial no lineal con un término drift polinomial para el estado del sistema

$$dx(t) = f(x, t)dt + b(t)dW_1(t), \quad x(t_0) = x_0, \quad (2.67)$$

y una ecuación diferencial lineal para el proceso de observación

$$dy(t) = (A_0(t) + A(t)x(t))dt + B(t)dW_2(t). \quad (2.68)$$

Aquí, $x(t) \in R^n$ es el vector de estado y $y(t) \in R^m$ es el vector de observación lineal, $m \leq n$. 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. Donde la matriz de observación $A(t) \in R^{m \times n}$ no se supone que sea invertible o incluso cuadrada. se asume que $B(t)B^T(t)$ es una matriz definida positiva, por lo tanto, $m \leq q$. Todos los coeficientes en (2.67)-(2.68) son funciones determinísticas de dimensiones apropiadas.

Las ecuaciones de filtrado óptimo pueden ser obtenidas usando la fórmula para la diferencial de Ito de la esperanza condicional $m(t) = E(x(t) | F_t^Y)$

$$dm(t) = E(f(x, t) | F_t^Y)dt + E(x(t)[A(t)(x(t) - m(t))]^T | F_t^Y) \times \\ (B(t)B^T(t))^{-1}(dy(t) - (A_0(t) + A(t)m(t))), \quad (2.69)$$

Donde $f(x, t)$ es el término del drift polinomial en la ecuación de estado.

$$dm(t) = E(f(x, t) | F_t^Y)dt + P(t)A^T(t)(B(t)B^T(t))^{-1} \times$$

$$(dy(t) - (A_0(t) + A(t)m(t))dt), \quad (2.70)$$

$$dP(t) = (E((x(t) - m(t))(f(x, t))^T | F_t^Y) + E(f(x, t)(x(t) - m(t))^T | F_t^Y) + b(t)b^T(t) - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t))dt, \quad (2.71)$$

Con las condiciones iniciales $m(t_0) = E(x(t_0) | F_{t_0}^Y)$ y $P(t_0) = E[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y]$.

En el caso particular donde $(f(x, t))$ toma la forma

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

Las siguientes ecuaciones de filtrado óptimo en forma cerrada son obtenidas en [11]:

$$\begin{aligned} dm(t) &= (a_0(t) + a_1(t)m(t) + a_2(t)m(t)m^T(t) + a_2(t)P(t) + \\ &\quad 3a_3(t)m(t)P(t) + a_3(t)m(t)m(t)m^T(t))dt + \\ &\quad P(t)A^T(t)(B(t)B^T(t))^{-1}[dy(t) - (A_0(t) + A(t)m(t))dt], \\ m(t_0) &= E(x(t_0) | F_{t_0}^Y), \\ dP(t) &= (a_1(t)P(t) + P(t)a_1^T(t) + 2a_2(t)m(t)P(t) + (2a_2(t)m(t)P(t))^T + \\ &\quad 3(a_3(t)[P(t)P(t) + m(t)m^T(t)P(t)] + \\ &\quad 3(a_3(t)[P(t)P(t) + m(t)m^T(t)P(t)]^T + \\ &\quad b(t)b^T(t))dt - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t)dt. \\ P(t_0) &= E((x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y). \end{aligned}$$

2.4.5. Ecuación General de Filtrado Óptimo

Considere el proceso continuo estocástico descrito por la ecuación

$$\dot{X} = \varphi(X, t) + \psi(X, t)V \quad (2.72)$$

donde X es el vector de estado n -dimensional del sistema, V es un vector r -dimensional que representa el ruido blanco Gaussiano, y $\varphi(X, t), \psi(X, t)$ son funciones conocidas del estado del sistema y del tiempo. Los valores de la función $\varphi(X, t)$ son vectores n -dimensionales y los valores de la función $\psi(X, t)$ son matrices $n \times r$. Si el vector de estado del sistema X es medido continuamente, entonces el proceso aleatorio n -dimensional $Y(t) = X(t) + U(t)$ sería el resultado de las mediciones, donde $U(t)$ es el error de la medición, el cual representa usualmente una función aleatoria del tiempo. Por otro lado, si esto no se cumple con el vector de estado, pero si algunas funciones del vector de estado son medidas por alguno de los componentes del vector de observación, el resultado de las mediciones es determinado en forma general por la fórmula

$$Y = Y(t) = \varphi_0(X, U, t), \quad (2.73)$$

donde Y es un vector n_1 -dimensional, U es el error de la medición, representando una función vectorial aleatoria de tiempo de dimensión $r \geq n_1$ y $\varphi_0(x, u, t)$ es una función conocida del estado del sistema, medición del error y del tiempo. El modelo general de mediciones las cuales son llevadas a cabo en un sistema puede ser descrito por la ecuación diferencial

$$\dot{Y} = \varphi_1(Y, X, U, t). \quad (2.74)$$

El resultado de las mediciones representa el proceso aleatorio Y . El problema de filtrado es planteado para el vector de estado del sistema X en cada instante $t > t_0$, usando los

resultados de mediciones continuas del proceso Y determinado por la ecuación (2.74) en el intervalo de tiempo $[t_0, t]$.

Sea un vector aleatorio de un proceso $[Y^T X^T]^T$ determinado por las ecuaciones diferenciales estocásticas de Ito

$$\begin{aligned} dY &= \varphi_1(Y, X, t)dt + \psi_1(Y, X, t)dW, \\ dX &= \varphi(Y, X, t)dt + \psi(Y, X, t)dW, \end{aligned} \tag{2.75}$$

donde Y es un proceso aleatorio n_1 -dimensional, X es un proceso n -dimensional, W es un proceso n -dimensional, $\varphi_1(y, x, t)$ y $\varphi(y, x, t)$ son funciones vectoriales que mapean el espacio $\mathbb{R}^{n_1} \times \mathbb{R}^n \times \mathbb{R}$ en los espacios \mathbb{R}^{n_1} y \mathbb{R}^n respectivamente y $\psi_1(y, x, t)$ y $\psi(y, x, t)$ son matrices de funciones conocidas que mapean $\mathbb{R}^{n_1} \times \mathbb{R}^n \times \mathbb{R}$ en $\mathbb{R}^{n_1 r}$ y \mathbb{R}^{nr} respectivamente. Esto constituye el planteamiento del problema de filtrado para el vector estado del sistema en algún instante $t > t_0$ usando los resultados de mediciones continuas del proceso Y en el intervalo de tiempo $[t_0, t]$.

La solución general al problema de filtrado óptimo se obtiene de la siguiente propiedad para los momentos de segundo orden: el menor de todos los momentos de segundo orden de una variable aleatoria escalar es su varianza. De aquí se sigue que la mejor aproximación de una variable aleatoria por una variable no aleatoria mediante el criterio de media cuadrada es dada por su esperanza condicional respecto a las observaciones. Sea $Y_{t_0}^t$ el conjunto de valores del proceso medido en el intervalo de tiempo $[t_0, t]$, $Y_{t_0}^t = \{Y(\tau) : \tau \in [t_0, t]\}$. Entonces el estimado óptimo del vector $X_u = X(u)$, el cual da la solución del problema para $u = t$ es determinado por la fórmula

$$\hat{X}_u = E[X_u / Y_{t_0}^t] \tag{2.76}$$

Esta fórmula determina el estimado óptimo del valor X_u para alguna función aleatoria $X(u)$ usando los resultados de las mediciones de otra función aleatoria $Y(t)$ en el intervalo

$[t_0, t]$. También es válida para el caso de un vector con argumento t y la medición de la función $Y(t)$ en algún conjunto T de valores de t . La aplicación de la fórmula (2.76) es necesaria para encontrar la distribución condicional de X_u . Este puede ser un problema que en ocasiones no se pueda resolver. En el caso particular en el que $Y(t)$ y $X(t)$ son determinados por las ecuaciones (2.75), este puede ser resuelto bajo algunas restricciones adicionales. La fórmula general para el diferencial estocástico del estimado óptimo de una función del vector de estado dado es la base de la teoría de filtrado óptimo. Sea $f(X_t, t)$ alguna función escalar del vector de estado n -dimensional de un sistema y de tiempo. Su estimado óptimo usando los resultados de observación $Y_{t_0}^t$ de acuerdo con (2.76) es determinado por la fórmula

$$\hat{f}(t) = E[f(X_t, t)/Y_{t_0}^t]. \quad (2.77)$$

Este estimado representa un funcional del proceso aleatorio $Y(t)$ en el intervalo de tiempo $[t_0, t]$, y consecuentemente es por si mismo una función de t . Un problema matemático que sirve de ayuda es encontrar la diferencial estocástica de Ito de este proceso aleatorio. Este problema puede ser resuelto bajo la condición que $W(t)$ en las ecuaciones (2.75) representa el proceso de Wiener cuya dimensión r es no menor que n_1 que es la dimensión del proceso de medición $Y(t)$, y que la función φ_1 en las ecuaciones (2.75) no depende de X . El sistema (2.75) toma la forma

$$\begin{aligned} dY &= \varphi_1(Y, X, t)dt + \psi_1(Y, t)dW, \\ dX &= \varphi(Y, X, t)dt + \psi(Y, X, t)dW, \end{aligned} \quad (2.78)$$

Diferencial de Ito para una Función del Estimado Óptimo

La ecuación diferencial estocástica del estimado óptimo de la variable aleatoria $f(X_t, t)$ para las ecuaciones (2.75) está dada por la fórmula

$$\begin{aligned} d\hat{f} = & E[f_t(X, t) + f_x(X, t)^T \varphi(Y, X, t)] \\ & + \frac{1}{2} tr\{f_{xx}(X, t)(\psi\nu\psi^T)(Y, X, t)\}/Y_{t_0}^t dt + E[f(X, t)\{\varphi_1(Y, X, t)^T - \hat{\varphi}_1^T\}] \\ & + f_x(X, t)^T(\psi\nu\psi_1^T)(Y, X, t)/Y_{t_0}^t(\psi_1\nu\psi_1^T)^{-1}(Y, t)(dY - \hat{\varphi}_1 dt), \end{aligned} \quad (2.79)$$

donde

$$\begin{aligned} (\psi\nu\psi^T)(x, y, t) &= \psi(y, x, t)\nu(t)\psi(y, x, t)^T, \\ (\psi\nu\psi_1^T)(y, x, t) &= \psi(y, x, t)\nu(t)\psi_1(y, t)^T, \\ (\psi_1\nu\psi_1^T)^{-1}(y, t) &= [\psi_1(y, t)\nu(t)\psi_1(y, t)^T]^{-1}, \\ \hat{\varphi}_1 &= \int_{-\infty}^{\infty} \varphi_1 p_t(x) dx = E[\varphi_1(X_t, Y_t, t/Y_{t_0}^t)], \end{aligned} \quad (2.80)$$

$p_t(x)$ es la densidad condicional de X_t relativa a $Y_{t_0}^t$; las derivadas f_t, f_x, f_{xx} y todas las esperanzas condicionales del lado derecho existen.

Ecuación para la Función Característica

Sustituyendo en la ecuación (2.79) $f(x, t) = e^{i\lambda^T X_t}$ se obtendrá la ecuación estocástica para la función condicional característica del vector aleatorio X_t :

$$g_t(\lambda) = E[e^{i\lambda^T X_t} / Y_{t_0}^t]. \quad (2.81)$$

Haciendo las sustituciones

$$\begin{aligned} f_t &= 0, f_x = i\lambda e^{i\lambda^T x}, f_{xx} = -\lambda\lambda^T e^{i\lambda^T x}, \\ tr\{\lambda\lambda^T(\psi\nu\psi^T)(y, x, t)\} &= \lambda^T(\psi\nu\psi^T)(y, x, t)\lambda, \end{aligned} \quad (2.82)$$

de la ecuación (2.79) se obtiene

$$\begin{aligned}
dg_t(\lambda) &= E[i\lambda^T \varphi(Y, X, t) - \frac{1}{2}(\psi\nu\psi^T)(Y, X, t)\lambda]e^{i\lambda^T X/Y_{t_0}^t}dt \\
&+ E[\{\varphi_1(Y, X, t)^T - \widehat{\varphi}_1^T + i\lambda^T(\psi\nu\psi_1^T)(Y, X, t)\}] \\
&\times e^{i\lambda^T X/Y_{t_0}^t}(\psi_1\nu\psi_1^T)^{-1}(Y, t)(dY - \widehat{\varphi}_1 dt).
\end{aligned} \tag{2.83}$$

El lado derecho representa una función de λ . La distribución condicional del vector aleatorio X es completa y únicamente determinada por su función característica. Resolviendo la ecuación (2.83) es posible evaluar el estimado óptimo \widehat{X}_t del vector de estado X_t determinado por la fórmula (2.76). Mediante estas fórmulas es posible obtener la expresión para la esperanza en términos de la función característica.

$$\widehat{X}_t = E[X_t/Y_{t_0}^t] = \left[\frac{\partial g_t(\lambda)}{\partial \lambda}\right]_{\lambda=0} \tag{2.84}$$

Ecuación para la Densidad Condicional

La ecuación estocástica para la densidad condicional $p_t(x)$ del vector aleatorio X_t es derivada a continuación

$$\begin{aligned}
dp_t(x) &= -\frac{\partial^T}{\partial x}[\varphi(Y, x, t)p_t(x)]dt \\
&+ \frac{1}{2}tr\left\{\frac{\partial}{\partial x}\frac{\partial^T}{\partial x}[(\psi\nu\psi^T)(Y, X, t)p_t(x)]\right\}(\psi_1\nu\psi_1^T)^{-1}(Y, t)(dY - \widehat{\varphi}_1 dt)
\end{aligned} \tag{2.85}$$

o

$$\begin{aligned}
dp_t(x) &= L^*p_t(x)dt + \{[\psi_1(Y, x, t)^T - \varphi_1^T]p_t(x) \\
&- \frac{\partial^T}{\partial x}[(\psi\nu\psi^T)(Y, x, t)p_t(x)]\}(\psi_1\nu\psi_1^T)^{-1}(Y, t)(dY - \widehat{\varphi}_1 dt),
\end{aligned} \tag{2.86}$$

donde L^* es el operador adjunto del operador

$$L = \varphi(Y, x, t)^T \frac{\partial}{\partial x} + \frac{1}{2}tr[(\psi\nu\psi^T)(Y, x, t) \frac{\partial}{\partial x} \frac{\partial^T}{\partial x}]. \tag{2.87}$$

Observando la última ecuación de (2.80), se concluye que la ecuación (2.85) representa una ecuación integro-diferencial relativa a la densidad condicional $p_t(x)$. Como el momento inicial t_o la función $p_{t_o}(x)$ sirve como la condición inicial para la ecuación (2.85). Después de resolver la ecuación (2.85), se puede encontrar de acuerdo con la fórmula (2.76) el estimado óptimo \hat{X}_t del vector de estado X_t del sistema

$$\hat{X}_t = E[X_t/Y_{t_o}^t] = \int_{-\infty}^{\infty} xp_t(x)dx. \quad (2.88)$$

Como la fórmula (2.75) determina la diferencial estocástica de Ito del proceso aleatorio $\hat{f}(t)$, las ecuaciones (2.83) y (2.85) representan ecuaciones estocásticas de Ito. La ecuación (2.85) fue originalmente obtenida en otra forma y bajo restricciones más rígidas en (*Stratonovich*1961, 1966) referida como la ecuación estocástica de Stratonovich. Al mismo tiempo, la ecuación para p_t en la forma de Ito fue obtenida en (*Kushner*1964a, b, 1967) también bajo restricciones más rígidas. Por lo tanto es usualmente llamada la ecuación de Stratonovich-Kushner.

Diferencial Estocástica de la Esperanza Matemática

La fórmula (2.76) determinó el estimado óptimo como la esperanza condicional de \hat{X} de la variable aleatoria correspondiente X . El estimado óptimo obtenido como resultado de mediciones es caracterizado por la matriz de covarianza condicional R . Estas fórmulas se pueden obtener de la fórmula general (2.79). Como la fórmula (2.79) determina la diferencial estocástica de una función escalar del estado del sistema, es necesario aplicarla para cada elemento de las matrices \hat{X} y R por separado. Sustituyendo en (2.79) $f(X, t) = X_l$, $f_t = 0$, $f_x = [0, \dots, 1, \dots]^T$, $f_{xx} = 0$, y la fórmula (2.79) toma la forma

$$d\hat{X}_l = \hat{\varphi}_l dt + E[X_l(\varphi_l^T - \hat{\varphi}_l^T) + (\psi\nu\psi_l^T)_l/Y_{t_o}^t](\psi_l\nu\psi_l^T)^{-1}(dY - \hat{\varphi}_l dt)(l = 1, \dots, n) \quad (2.89)$$

donde de acuerdo con la última ecuación de (2.79) $\hat{\varphi}_l = E[\varphi_l(Y, X, t)/Y_{t_0}^t]$, $(\psi\nu\psi_l^T)_l$, siendo la l^{era} columna de la matriz $\psi\nu\psi_l^T$ y los argumentos de las funciones φ_l , $\psi\nu\psi_l^T$ y $(\psi_l\nu\psi_l^T)^{-1}$ son omitidos por brevedad. Entonces la matriz para el diferencial estocástico del estimado óptimo \hat{X} del vector de estado del sistema X está dada por

$$d\hat{X} = \hat{\varphi}dt + E[X\{(\varphi_l(Y, X, t)^T - \hat{\varphi}_l^T)\} + (\psi\nu\psi_l^T)(Y, X, t)/Y_{t_0}^t](\psi_l\nu\psi_l^T)^{-1}(Y, t)(dY - \hat{\varphi}_l dt) \quad (2.90)$$

Diferencial Estocástica del Momento Condicional de Segundo Orden

Sustituyendo en (2.79) $f(X, t) = X_k X_l$ con $k < l$, $f_t = 0$, $f_x = [0, \dots, X_l, \dots, X_k, \dots, 0]^T$,

$$f_{xx} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

Siendo las dos columnas y renglones centrales conteniendo unos, los correspondientes a k , y l respectivamente, de la fórmula (2.79) se tiene

$$d\Gamma_{kl} = E[X_k\varphi_l + X_l\varphi_k + (\psi\nu\psi^T)_{kl}/Y_{t_0}^t]dt + E[X_k X_l(\varphi_l^T - \hat{\varphi}_l^T) + X_k(\psi\nu\psi_l^T)_l + X_l(\psi\nu\psi_l^T)_k/Y_{t_0}^t](\psi_l\nu\psi_l^T)^{-1}(dY - \hat{\varphi}_1 dt)(k, l = 1, \dots, n), \quad (2.91)$$

donde $d\Gamma_{kl} = E[X_k X_l \varphi_k / Y_{t_0}^t]$, y $(\psi\nu\psi^T)_{kl}$ es el elemento correspondiente de la matriz

$(\psi\nu\psi^T)$. Re-escribiendo la fórmula (2.91) como

$$\begin{aligned} d\Gamma_{kl} &= E[X_k\varphi_l + X_l\varphi_k + (\psi\nu\psi^T)_{kl}/Y_{t_0}^t]dt \\ &+ \sum_{\rho=1}^r E[X_kX_l a_\rho + X_k b_{lk} + X_l b_{k\rho}/Y_{t_0}^t](dY_\rho - \widehat{\varphi}_{1\rho}dt), \end{aligned} \quad (2.92)$$

donde a_ρ es el ρ -ésimo elemento de la matriz $(\varphi_1^T - \widehat{\varphi}_1^T)(\psi_1\nu\psi_1^T)^{-1}$ y $b_{k\rho}$ es el elemento del k -ésimo renglón y de la ρ -ésima columna de la matriz $\psi\nu\psi_1^T(\psi_1\nu\psi_1^T)^{-1}$. Denotando por b_ρ la ρ -ésima columna de la matriz $\psi\nu\psi_1^T(\psi_1\nu\psi_1^T)^{-1}$, $b_\rho = [b_{1\rho}, \dots, b_{r\rho}]^T$ ($\rho = 1, \dots, r$), se obtiene la siguiente fórmula diferencial estocástica del momento condicional de segundo orden Γ del vector estado del sistema:

$$\begin{aligned} d\Gamma &= E[X\varphi(Y, X, t)^T + \varphi(Y, X, t)X^T \\ &+ (\psi\nu\psi^T)(Y, X, t)/Y_{t_0}^t]dt + \sum_{\rho=1}^r E[XX^T a_\rho(Y, X, t) \\ &+ Xb_\rho(Y, X, t)^T + b_\rho X^T/Y_{t_0}^t](dY_\rho - \widehat{\varphi}_{1\rho}dt), \end{aligned} \quad (2.93)$$

Diferencial Estocástica de la Matriz de Covarianza

Para encontrar la diferencial estocástica de la matriz de covarianza condicional R del vector estado del sistema se usará la fórmula conocida que relaciona la esperanza, el momento de segundo orden, y la matriz de covarianza del vector aleatorio $R = \Gamma - \widehat{X}\widehat{X}^T$, o en la forma escalar $R_{kl} = \Gamma_{kl} - \widehat{X}_k\widehat{X}_l$. Derivando en ambos lados de la última fórmula, se obtiene la expresión $dR_{kl} = d\Gamma_{kl} - d(\widehat{X}_k\widehat{X}_l)$. Para encontrar $d(\widehat{X}_k\widehat{X}_l)$ se utiliza la fórmula

$$d(Z_1Z_2) = Z_1dZ_2 + Z_2dZ_1 + Y_1\nu Y_2^T dt, \quad (2.94)$$

$Z(t) = [Z_1, Z_2]$ es un proceso de Ito, el cual está dado por

$$dZ(t) = x(t)dt + Y(t)dW(t). \quad (2.95)$$

Aquí $t_0 > 0$, $W(t)$ es un proceso de Wiener, donde Y_1 y Y_2 representan la primera y segunda columnas de la matriz $Y = [Y_1, Y_2]$ respectivamente. $X(t)$, $Y_1(t)$, $Y_2(t)$ son funciones aleatorias que satisfacen las condiciones de existencia. Z_1, Z_2 son los componentes del vector aleatorio $Z(t)$. De acuerdo con (2.90)

$$\begin{aligned} E[X_k(\varphi_1^T - \widehat{\varphi}_1^T) + (\psi\nu\psi_1^T)_k/Y_{t_0}^t](\psi_1\nu\psi_1^T)^{-1}\psi_1, \\ E[X_l(\varphi_1^T - \widehat{\varphi}_1^T) + (\psi\nu\psi_1^T)_l/Y_{t_0}^t](\psi_1\nu\psi_1^T)^{-1}\psi_1, \end{aligned} \quad (2.96)$$

juegan el rol de los renglones Y_1, Y_2 de la matriz, en este caso se llega a

$$\begin{aligned} d(\widehat{X}_k\widehat{X}_l) &= \widehat{X}_k d\widehat{X}_l + \widehat{X}_l d\widehat{X}_k \\ &E[X_k(\varphi_1^T - \widehat{\varphi}_1^T) + (\psi\nu\psi_1^T)_k/Y_{t_0}^t](\psi_1\nu\psi_1^T)^{-1}\psi_1\nu\psi_1^T(\psi_1\nu\psi_1^T)^{-1} \\ &\times E[X_l(\varphi_1^T - \widehat{\varphi}_1^T) + (\psi_1\nu\psi_1^T)_l/Y_{t_0}^t]dt. \end{aligned} \quad (2.97)$$

Sustituyendo aquí las expresiones para $d\widehat{X}_k$ y $d\widehat{X}_l$ de la ecuación (2.89), se tiene

$$\begin{aligned} d(\widehat{X}_k\widehat{X}_l) &= \{\widehat{X}_k\widehat{\varphi}_l + \widehat{X}_l\widehat{\varphi}_k \\ &+ E[X_k(\varphi_1^T - \widehat{\varphi}_1^T) + (\psi\nu\psi_1^T)_k/Y_{t_0}^t](\psi_1\nu\psi_1^T)^{-1}E[X_l(\varphi_1^T - \widehat{\varphi}_1^T) + (\psi\nu\psi_1^T)_l^T/Y_{t_0}^t]\}dt \\ &+ E[(\widehat{X}_kX_l + \widehat{X}_lX_k)(\varphi_1^T - \widehat{\varphi}_1^T) + \widehat{X}_k(\psi\nu\psi_1^T)_l \\ &+ \widehat{X}_l(\psi\nu\psi_1^T)_k/Y_{t_0}^t](\psi_1\nu\psi_1^T)^{-1}(dY - \widehat{\varphi}_1 dt). \end{aligned} \quad (2.98)$$

Substrayendo esta fórmula de (2.92) y adicionando el término

$$E[(\widehat{X}_k\widehat{X}_l)(\varphi_1^T - \widehat{\varphi}_1^T)/Y_{t_0}^t] = \widehat{X}_k\widehat{X}_l(\widehat{\varphi}_1^T - \widehat{\varphi}_1^T) = 0, \quad (2.99)$$

se obtiene

$$\begin{aligned}
dR_{kl} = & \{E[(X_k - \widehat{X}_k)\varphi_l + (X_l - \widehat{X}_l)\varphi_k + (\psi\nu\psi_1^T)_{kl}/Y_{t_0}^t] \\
& - E[X_k(\varphi_1^T - \widehat{\varphi}_1^T) + (\psi\nu\psi_1^T)_k/Y_{t_0}^t](\psi_1\nu\psi_1^T)^{-1}E[X_l(\varphi_1 - \widehat{\varphi}_1) \\
& + (\psi\nu\psi_1^T)_l^T/Y_{t_0}^t]\}dt + E[(X_k - \widehat{X}_k)(X_l - \widehat{X}_l)(\varphi_1^T - \widehat{\varphi}_1^T) \\
& + (X_k - \widehat{X}_k)(\psi\nu\psi_1^T)_l + (X_l - \widehat{X}_l)(\psi\nu\psi_1^T)_k/Y_{t_0}^t](\psi_1\nu\psi_1^T)^{-1} \\
& \times (dY - \widehat{\varphi}_l dt)(k, l = 1, \dots, n).
\end{aligned} \tag{2.100}$$

Haciendo algunas transformaciones en la fórmula anterior (2.100), obtenemos la fórmula de la matriz diferencial estocástica para la matriz de covarianza como la solución de

$$\begin{aligned}
dR = & \{E[(X - \widehat{X})\varphi(Y, X, t)^T + \varphi(Y, X, t)(X^T - \widehat{X}^T) - E[X\{(\varphi_1(Y, X, t)^T - \widehat{\varphi}_1^T)\} + \\
& (\psi\nu\psi_1^T)(Y, X, t)/Y_{t_0}^t](\psi_1\nu\psi_1^T)^{-1}(Y, t)E[\{(\varphi_1(Y, X, t) - \widehat{\varphi}_1)\}X^T \\
& + (\psi_1\nu\psi^T)(Y, X, t)/Y_{t_0}^t]\}dt \\
& + \sum_{\rho=1}^r E[(X - \widehat{X})(X^T - \widehat{X}^T)a_\rho(Y, X, t) + (X - \widehat{X})b_\rho(Y, X, t)^T \\
& + (X - \widehat{X})^T/Y_{t_0}^t](dY_\rho - \widehat{\varphi}_{l\rho} dt).
\end{aligned} \tag{2.101}$$

Hasta aquí se ha establecido el planteamiento del problema y solución para el caso de un sistema representado por ecuaciones de estado lineales, y de observaciones lineales, ambas con la presencia de disturbios los cuales se comportan como ruidos blancos Gaussianos, lo cual fue desarrollado por Kalman-Bucy.

Capítulo 3

Filtrado Óptimo para Estados Polinomiales Incompletamente Medibles con Ruido de Poisson

3.1. Planteamiento del Problema

Sea (Ω, F, P) un espacio completo de probabilidad con una familia de σ -álgebras $F_t, t \geq 0$ creciente y continua por la derecha, y sean $(N_1(t), F_t, t \geq 0)$ y $(N_2(t), F_t, t \geq 0)$ dos procesos de Poisson centralizados independientes. El proceso aleatorio F_t -medible $(x(t), y(t))$ es descrito por una ecuación diferencial estocástica no lineal con un término drift polinomial para el estado del sistema

$$dx(t) = f(x, t)dt + b(t)dN_1(t), \quad x(0) = x_0, \quad (3.1)$$

y una ecuación diferencial lineal para el proceso de observación

$$dy(t) = (A_0(t) + A(t)x(t))dt + B(t)dN_2(t). \quad (3.2)$$

Aquí, $x(t) \in R^n$ es el vector de estado y $y(t) \in R^m$ es el vector de observación lineal, de lo que sigue que $m \leq n$. La condición inicial $x_0 \in R^n$ es un vector de Poisson tal que x_0 , $N_1(t) \in R^p$, y $N_2(t) \in R^q$ son independientes entre sí. La matriz de observación $A(t) \in R^{m \times n}$ no es necesariamente invertible, inclusive no se requiere que sea una matriz cuadrada. Se asume que $B(t)B^T(t)$ es una matriz definida positiva, y por lo tanto, $m \leq q$. Todos los coeficientes en las ecuaciones (3.1)–(3.2) son funciones determinísticas de dimensiones apropiadas.

La función no lineal $f(x, t)$ será considerada una función polinomial de n variables, donde las componentes del vector de estado $x(t) \in R^n$, son coeficientes que dependen del tiempo t . Como $x(t) \in R^n$ es un vector, se requiere una definición especial para el polinomio en el caso donde $n > 1$. De acuerdo con [14], un polinomio de grado p de un vector $x(t) \in R^n$ se considera de la forma p -lineal de n componentes de $x(t)$ y se puede expresar de la siguiente manera

$$f(x, t) = a_0(t) + a_1(t)x + a_2(t)xx^T + \dots + a_p(t)x \dots_p \text{ veces} \dots x, \quad (3.3)$$

donde $a_0(t)$ es un vector de dimensión n , a_1 es una matriz de dimensión $n \times n$, a_2 es un 3D tensor de dimensión $n \times n \times n$, a_p es un $(p+1)$ D tensor de dimensión $n \times \dots_{(p+1) \text{ veces}} \dots \times n$, y $x \times \dots_p \text{ veces} \dots \times x$ es un p D tensor de dimensión $n \times \dots_p \text{ veces} \dots \times n$ que se obtiene al multiplicar el vector $x(t)$, p veces por sí mismo. Tal polinomio también puede ser expresado en la forma de la siguiente sumatoria

$$f_k(x, t) = a_{0k}(t) + \sum_i a_{1ki}(t)x_i(t) + \sum_{ij} a_{2kij}(t)x_i(t)x_j(t) + \dots$$

$$+ \sum_{i_1 \dots i_p} a_p \, k i_1 \dots i_p(t) x_{i_1}(t) \dots x_{i_p}(t), \quad k, i, j, i_1 \dots i_p = 1, \dots, n.$$

El problema de estimación consiste en encontrar el estimado óptimo $\hat{x}(t)$ del estado del sistema $x(t)$, basado en el proceso de observación $Y(t) = \{y(s), t_0 \leq s \leq t\}$, que minimice la segunda norma Euclídeana

$$J = E[(x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t)) \mid F_t^Y]$$

para cada t . Aquí, $E[z(t) \mid F_t^Y]$ representa el valor esperado condicional de un proceso estocástico $z(t) = (x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t))$ con respecto a la σ - álgebra F_t^Y generada por el proceso de observación $Y(t)$ en el intervalo de tiempo $[t_0, t]$. Como se sabe de [53], este estimado óptimo está dado por el valor esperado condicional

$$\hat{x}(t) = m(t) = E(x(t) \mid F_t^Y)$$

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

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

es la estimación de la varianza del error de estimación.

La solución propuesta a este problema de filtrado óptimo se basa en las fórmulas de las diferenciales Ito del valor esperado condicional $E(x(t) \mid F_t^Y)$ y de su varianza $P(t)$ (citado después de [53]) y será desarrollado en la siguiente sección.

3.2. Diseño del Filtro Óptimo

Las ecuaciones de filtrado óptimo se obtendrán utilizando la fórmula de la diferencial Ito del valor esperado condicional $m(t) = E(x(t) \mid F_t^Y)$ en el caso del término drift lineal $A_0(t) + A(t)x(t)$ en la ecuación de observación (ver [53])

$$dm(t) = E(f(x, t) | F_t^Y)dt + E(x(t)[A(t)(x(t) - m(t))]^T | F_t^Y) \times \quad (3.4)$$

$$(B(t)B^T(t))^{-1}(dy(t) - (A_0(t) + A(t)m(t)),$$

donde $f(x, t)$ es el término drift polinomial en la ecuación de estado.

La ecuación (3.4) debe complementarse con la condición inicial $m(t_0) = E(x(t_0) | F_{t_0}^Y)$.

Tratando de formar un sistema cerrado de las ecuaciones de filtrado, la ecuación (3.4) podría ser complementada con la ecuación para la varianza del error $P(t)$. Para esto, se puede utilizar la fórmula para la diferencial Ito de la varianza $P(t) = E((x(t) - m(t))(x(t) - m(t))^T | F_t^Y)$ en el caso del término drift lineal $A_0(t) + A(t)x(t)$ en la ecuación de observación (citada después de [53]):

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

Usando la fórmula de la varianza $P(t) = E((x(t) - m(t))x^T(t) | F_t^Y)$, la última ecuación puede representarse como

$$\begin{aligned} dP(t) = & (E((x(t) - m(t))(f(x, t))^T | F_t^Y) + E(f(x, t)(x(t) - m(t))^T | F_t^Y) + \\ & b(t)b^T(t) - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t))dt. \end{aligned} \quad (3.5)$$

La ecuación (3.5) debe ser complementada 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 (3.4) y (3.5) para los estimados óptimos $m(t)$ y $P(t)$ respectivamente, aún no forman un sistema cerrado de las ecuaciones de filtrado para el estado no lineal (3.1) sobre las observaciones lineales (3.2). Esto significa que el sistema (3.4), (3.5) incluye términos que dependen de x , tales como $E(f(x, t) | F_t^Y)$ y $E((x(t) - m(t))f^T(x, t) | F_t^Y)$, los cuales aún no están expresadas como funciones de las variables del sistema, $m(t)$ y $P(t)$.

Como se muestra en [4, 11], en el caso de ruido blanco Gaussiano en el estado y las ecuaciones de observación, es posible obtener un sistema cerrado de las ecuaciones de filtrado para el estado del sistema (3.1) con el término drift polinomial sobre observaciones lineales. En el caso que se está considerando, de ruido blanco de Poisson, se introducirá la siguiente transformación:

Primero, note que siempre se puede asumir que la matriz A es de rango completo e igual a m , que representa la dimensión de las observaciones linealmente independientes $y(t) \in R^m$, de no ser así, deberán eliminarse las filas linealmente dependientes de la matriz A correspondientes a las observaciones excesivas linealmente dependientes. Una vez haciendo esto, el número de procesos de Poisson en las ecuaciones de observación pueden ser reducidas a m , la dimensión de las observaciones linealmente independientes, sumando y reenumerando los procesos de Poisson en cada ecuación de observación (3.2). Por lo tanto, siempre se puede asumir que la matriz B es una matriz cuadrada de dimensión $m \times m$, tal que $B(t)B^T(t)$ es una matriz definida positiva (ver la Sección 3.1 para esta condición). Después, las nuevas matrices $\bar{A}(t)$ y $\bar{B}(t)$ se definen de la siguiente manera: La matriz $\bar{A}(t) \in R^{n \times n}$ se obtiene a partir de la matriz $A(t) \in R^{m \times n}$, agregando $n - m$ renglones linealmente independientes tales que la matriz resultante $\bar{A}(t)$ sea invertible. La matriz $\bar{B}(t) \in R^{n \times n}$ se obtiene a partir de la matriz $B(t) \in R^{m \times m}$, colocando $B(t)$ en la esquina superior izquierda de $\bar{B}(t)$, y definiendo las demás $n - m$

entradas de la diagonal principal de $\bar{B}(t)$ igual a infinito, y haciendo cero todas las demás entradas de $\bar{B}(t)$ fuera de la diagonal principal y de la submatriz $B(t)$. En otras palabras, $\bar{B}(t) = \text{diag}[B(t), \beta I_{(n-m) \times (n-m)}]$, donde $\beta = \infty$, y $I_{(n-m) \times (n-m)}$ es la matriz identidad de dimensión $(n-m) \times (n-m)$. Entonces, la nueva ecuación de observación está dada por

$$\bar{y}(t) = (\bar{A}_0(t) + \bar{A}(t)x(t))dt + \bar{B}(t)dN_2(t), \quad (3.6)$$

donde $\bar{y}(t) \in R^n$, $\bar{A}_0(t) = [A_0^T(t), 0_{n-m}]^T \in R^n$, y 0_{n-m} es un vector de $n-m$ ceros.

El punto clave de la transformación que se realizó es que el nuevo proceso de observación $\bar{y}(t)$ es físicamente equivalente al antiguo proceso $y(t)$, pues las últimas $n-m$ componentes ficticias de $\bar{y}(t)$, las correspondientes $n-m$ ecuaciones, y las primeras m componentes de $\bar{y}(t)$ coinciden con $y(t)$. Adicionalmente, la matriz de observación $\bar{A}(t)$ es invertible, y la matriz $(\bar{B}(t)\bar{B}^T(t))^{-1} \in R^{n \times n}$ existe y es igual a la matriz cuadrada de dimensión $n \times n$, la cual se forma al ocupar la esquina superior izquierda con la submatriz $(B(t)B^T(t))^{-1} \in R^{m \times m}$ y todas las demás entradas son ceros.

En términos de la nueva ecuación de observación (3.6), las ecuaciones de filtrado (3.4) y (3.5) toman la forma

$$dm(t) = E(f(x, t) | F_t^Y)dt + P(t)\bar{A}^T(t)(\bar{B}(t)\bar{B}^T(t))^{-1} \times$$

$$(d\bar{y}(t) - (\bar{A}_0(t) + \bar{A}(t)m(t))dt), \quad (3.7)$$

$$dP(t) = (E((x(t) - m(t))(f(x, t))^T | F_t^Y) + E(f(x, t)(x(t) - m(t))^T | F_t^Y) +$$

$$b(t)b^T(t) - P(t)\bar{A}^T(t)(\bar{B}(t)\bar{B}^T(t))^{-1}\bar{A}(t)P(t)dt. \quad (3.8)$$

con las condiciones iniciales $m(t_0) = E(x(t_0) | F_{t_0}^Y)$ y $P(t_0) = E[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y]$.

Como la nueva matriz $\bar{A}(t)$ es invertible para cualquier $t \geq t_0$, la variable aleatoria $x(t) - m(t)$ es condicionalmente Poisson con respecto al nuevo proceso de observación $\bar{y}(t)$, y por lo tanto con respecto al proceso de observación original $y(t)$, para cualquier $t \geq t_0$ (ver [53]). Por lo tanto, las siguientes consideraciones son aplicables a las ecuaciones de filtrado (3.4), (3.5).

Si la función $f(x, t)$ es una función polinomial del estado x con coeficientes que dependen del tiempo t , entonces las expresiones para los términos $E(f(x, t) | F_t^Y)$ en (3.7) y $E((x(t) - m(t))f^T(x, t) | F_t^Y)$ en (3.8) incluirían sólo términos polinomiales de x . Entonces, estos términos polinomiales pueden ser representados como funciones de $m(t)$ y $P(t)$ usando la siguiente propiedad de una variable aleatoria de Poisson $x(t) - m(t)$: todos los momentos de una variable aleatoria de Poisson pueden ser representados como funciones de la varianza $P(t)$. Por ejemplo, $m_1 = E[(x(t) - m(t)) | Y(t)] = 0$, $m_2 = E[(x(t) - m(t))^2 | Y(t)] = P$, $m_3 = E[(x(t) - m(t))^3 | Y(t)] = P$, $m_4 = E[(x(t) - m(t))^4 | Y(t)] = 3P^2 + P$, etc. Después de representar todos los términos polinomiales en (3.7) y (3.8), es posible obtener una forma cerrada de las ecuaciones de filtrado, que se genera al expresar $E(f(x, t) | F_t^Y)$, y $E((x(t) - m(t))f^T(x, t) | F_t^Y)$ como funciones de $m(t)$ y $P(t)$.

Finalmente, en vista de la definición de las matrices $\bar{A}(t)$ y $\bar{B}(t)$ y del nuevo proceso de observación $\bar{y}(t)$, las ecuaciones de filtrado (3.7),(3.8) pueden reescribirse en términos de la ecuación de observación original (3.2) usando $y(t)$, $A(t)$, y $B(t)$

$$dm(t) = E(f(x, t) | F_t^Y)dt + P(t)A^T(t)(B(t)B^T(t))^{-1} \times (dy(t) - (A_0(t) + A(t)m(t))dt) \quad (3.9)$$

$$dP(t) = (E((x(t) - m(t))(f(x, t))^T | F_t^Y) + E(f(x, t)(x(t) - m(t))^T | F_t^Y) + b(t)b^T(t) - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t))dt, \quad (3.10)$$

con las condiciones iniciales $m(t_0) = E(x(t_0) | F_{t_0}^Y)$ y $P(t_0) = E[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_{t_0}^Y]$.

Además, se obtendrá la forma cerrada de las ecuaciones de filtrado de (3.9) y (3.10) para una función de tercer orden $f(x, t)$ en la ecuación (3.1), como sigue. Nótese, de cualquier manera, que la aplicación del mismo procedimiento resultará en el diseño de un sistema cerrado de las ecuaciones de filtrado para cualquier función polinomial $f(x, t)$ en (3.1).

3.2.1. Filtro Óptimo para un Estado Polinomial de Tercer Orden

Sea

$$f(x, t) = a_0(t) + a_1(t)x + a_2(t)xx^T + a_3(t)xxx^T \quad (3.11)$$

una función polinomial de tercer orden, donde x es un vector de dimensión n , $a_0(t)$ es un vector de dimensión n , $a_1(t)$ es una matriz de dimensión $n \times n$, $a_2(t)$ es un 3D tensor de dimensión $n \times n \times n$, $a_3(t)$ es un 4D tensor de dimensión $n \times n \times n \times n$.

En este caso, la representación para $E(f(x, t) | F_t^Y)$ y $E((x(t) - m(t))(f(x, t))^T | F_t^Y)$ como funciones de $m(t)$ y $P(t)$ se deriva de la siguiente manera

$$E(f(x, t) | F_t^Y) = a_0(t) + a_1(t)m(t) + a_2(t)m(t)m^T(t) + a_2(t)P(t) + \quad (3.12)$$

$$3a_3(t)m(t)P(t) + a_3(t)m(t)m(t)m^T(t) + a_3(t)P(t) * 1,$$

$$E(f(x, t)(x(t) - m(t))^T | F_t^Y) + E((x(t) - m(t))(f(x, t))^T | F_t^Y) = \quad (3.13)$$

$$\begin{aligned} & a_1(t)P(t) + P(t)a_1^T(t) + 2a_2(t)m(t)P(t) + a_2(t)P(t) * 1 + (a_2(t)(2m(t)P(t) + P(t) * 1))^T + \\ & a_3(t)[(P(t) * (1 * 1^T)) + 3P(t)P(t) + 3m(t)m^T(t)P(t) + 3(m(t)P(t)) * 1^T] + \\ & (a_3(t)[(P(t) * (1 * 1^T)) + 3P(t)P(t) + 3m(t)m^T(t)P(t) + 3(m(t)P(t)) * 1^T])^T. \end{aligned}$$

Aquí, el vector 1 representa un vector de dimensión n con todas sus componentes iguales a 1, y el vector $a_3P(t) * 1 \in R^n$ y las matrices $a_3(t)P(t) * 1 * 1^T \in R^{n \times n}$ y $a_3(t)m(t)P(t) * 1^T \in R^{n \times n}$ se definen como

$$(a_3(t)P(t) * 1)_i = \sum_{j,k,l} a_3 \text{ }_{ijkl}(t) P_{jk}(t) 1_l, \quad i = 1, \dots, n,$$

$$(a_3(t)P(t) * 1 * 1^T)_{ij} = \sum_{h,k,l} a_3 \text{ }_{ihkl}(t) P_{hk}(t) 1_l 1_j, \quad i, j = 1, \dots, n,$$

$$(a_3(t)m(t)P(t) * 1^T)_{ij} = \sum_{h,k,l} a_3 \text{ }_{ihkl}(t) m_h(t) P_{kl}(t) 1_j, \quad i, j = 1, \dots, n.$$

Sustituyendo la expresión (3.12) en (3.9) y la expresión (3.13) en (3.10), se obtienen las siguientes ecuaciones de filtrado para el estimado óptimo $m(t)$ y la varianza del error $P(t)$

$$\begin{aligned} dm(t) &= (a_0(t) + a_1(t)m(t) + a_2(t)m(t)m^T(t) + a_2(t)P(t) + \\ & 3a_3(t)m(t)P(t) + a_3(t)m(t)m(t)m^T(t) + a_3(t)P(t) * 1 + \\ & P(t)A^T(t)(B(t)B^T(t))^{-1}[dy(t) - (A_0(t) + A(t)m(t))dt], \end{aligned} \quad (3.14)$$

$$\begin{aligned}
& m(t_0) = E(x(t_0) | F_t^Y), \quad dP(t) = \\
& (a_1(t)P(t) + P(t)a_1^T(t) + 2a_2(t)m(t)P(t) + a_2(t)P(t) * 1 + (a_2(t)(2m(t)P(t) + P(t) * 1))^T + \\
& \quad a_3(t)[(P(t) * (1 * 1^T)) + 3P(t)P(t) + 3m(t)m^T(t)P(t) + 3(m(t)P(t)) * 1^T] + \\
& \quad (a_3(t)[(P(t) * (1 * 1^T)) + 3P(t)P(t) + 3m(t)m^T(t)P(t) + 3(m(t)P(t)) * 1^T])^T + \\
& \quad b(t)b^T(t)dt - P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t)dt. \tag{3.15}
\end{aligned}$$

$$P(t_0) = E((x(t_0) - m(t_0))(x(t_0) - m(t_0))^T | F_t^Y).$$

Por medio de la derivación anterior, se probará el siguiente resultado.

Teorema 3.1. El filtro óptimo de dimensión finita para el estado de tercer orden (3.1), donde la función polinomial $f(x, t)$ de tercer orden definida por (3.11), sobre las observaciones lineales incompletas (3.2), está dado por la ecuación (3.14) para el estimado óptimo $m(t) = E(x(t) | F_t^Y)$ y la ecuación (3.15) para el estimado de la varianza del error $P(t) = E[(x(t) - m(t))(x(t) - m(t))^T | F_t^Y]$.

Así, basado en el sistema general no cerrado de las ecuaciones de filtrado (3.7),(3.8), se prueba que el sistema cerrado de las ecuaciones de filtrado pueden ser obtenidas para cualquier estado polinomial (3.1) sobre observaciones lineales incompletas (3.2). Por lo tanto, se deriva la forma específica (3.14),(3.15) del sistema cerrado de las ecuaciones de filtrado correspondiente a un estado de tercer orden. En la siguiente sección, se verificará el rendimiento del filtro óptimo diseñado para un estado de tercer orden sobre observaciones lineales incompletas en contra de un filtro en promedio cuadrático convencional para sistemas polinomiales estocásticos con ruidos Gaussianos, obtenido en [11].

3.3. Ejemplo

En esta sección se presenta un ejemplo del diseño del filtro óptimo para un estado bidimensional de tercer orden y para observaciones lineales escalares, y se comparará con un filtro en promedio cuadrático convencional para sistemas polinomiales estocásticos con ruidos Gaussianos [11].

Sea $x(t)$ un estado real bidimensional que satisface el siguiente sistema de tercer orden

$$\dot{x}_1(t) = x_2(t), \quad x_1(0) = x_{10}, \quad (3.16)$$

$$\dot{x}_2(t) = 0,1x_2^3(t) + \psi_1(t), \quad x_2(0) = x_{20},$$

y $y(t)$ el proceso escalar de observación dado por la siguiente ecuación lineal

$$y(t) = x_1(t) + \psi_2(t), \quad (3.17)$$

donde $\psi_1(t)$ y $\psi_2(t)$ son ruidos blancos de Poisson, que son las derivadas débiles de dos procesos estándar de Poisson independientes (ver [53]). Las ecuaciones (3.16),(3.17) presentan la forma convencional de las ecuaciones (3.1),(3.2), las cuales, de hecho, son utilizadas en la práctica [2, 35].

El sistema de filtrado (3.16),(3.17) incluye dos componentes del estado $x(t) = [x_1(t), x_2(t)]^T \in R^2$ y solamente un canal de observación $y(t) \in R$, midiendo la componente del estado $x_1(t)$. Por lo tanto, la matriz de observación $A = [1 \ 0] \in R^{(1 \times 2)}$ y es no invertible. Más aún, la componente no lineal del estado $x_2(t)$ no es medible. El problema de filtrado consiste en encontrar el estimado óptimo para el estado de tercer orden (3.16),

usando observaciones lineales incompletas (3.17) perturbadas con disturbios aleatorios, independientes con impulsos aislados modelados como ruidos blancos de Poisson.

Mostraremos cómo calcular los coeficientes del vector polinomial(3.3) para el sistema (3.16). Incluso, los coeficientes de la matriz a_1 es una matriz de dimensión 2×2 , igual a $a_1 = [0 \ 1 \mid 0 \ 0]$, el coeficiente del 3D tensor a_2 consta sólo de ceros, pues no hay términos cuadráticos o bilineales en (3.16), y el coeficiente del 4D tensor a_3 tiene sólo una entrada diferente de cero, $a_3 \ 2222 = 0,1$, y todas las demás entradas son ceros. Por lo tanto, de acuerdo con (3.14),(3.15), este único término diferente de cero debería entrar en la ecuación para m_2 , multiplicado por $3m_2P_{22} + m_2^3 + P_{22}$, la ecuación para $P_{21} = P_{12}$, multiplicado por $3m_2^2P_{21} + 3P_{22}P_{21} + P_{22} + 3m_2P_{22} = 3m_2^2P_{12} + 3P_{22}P_{12} + P_{22} + 3m_2P_{22}$, en vista de la simetría de la matriz de la varianza P , y la ecuación para P_{22} , multiplicada por $2P_{22} + 6P_{22}^2 + 6m_2P_{22} + 6m_2^2P_{22}$.

Como resultado, las ecuaciones de filtrado (3.14),(3.15) toman la siguiente forma particular para el sistema (3.16),(3.17)

$$\dot{m}_1(t) = m_2(t) + P_{11}(t)[y(t) - m_1(t)] \quad (3.18)$$

$$\dot{m}_2(t) = 0,1m_2^3(t) + 0,3P_{22}(t)m_2(t) + 0,1P_{22}(t) + P_{12}(t)[y(t) - m_1(t)],$$

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

$$\dot{P}_{11}(t) = 2P_{12}(t) - P_{11}^2(t), \quad (3.19)$$

$$\dot{P}_{12}(t) = 1,1P_{22}(t) + 0,3m_2^2(t)P_{12}(t) + 0,3m_2(t)P_{22}(t) + 0,3P_{22}(t)P_{12}(t) - P_{11}(t)P_{12}(t),$$

$$\dot{P}_{22}(t) = 1 + 0,2P_{22}(t) + 0,6m_2^2(t)P_{22}(t) + 0,6m_2(t)P_{22}(t) + 0,6P_{22}^2(t) - P_{12}^2(t),$$

con la condición inicial $P(0) = E((x(0) - m(0))(x(0) - m(0))^T | y(0)) = P_0$.

Los estimados obtenidos al resolver las ecuaciones (3.18)–(3.19) son comparados con los estimados que satisfacen las ecuaciones del filtro en promedio cuadrático convencional para el estado de tercer orden (3.16) sobre observaciones lineales incompletas (3.17) (ver [11]):

$$\dot{m}_{k1}(t) = m_{k2}(t) + P_{k11}(t)[y(t) - m_{k1}(t)], \quad (3.20)$$

$$\dot{m}_{k2}(t) = 0,1m_{k2}^3(t) + 0,3P_{k22}(t)m_{k2}(t) + P_{k12}(t)[y(t) - m_{k1}(t)],$$

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

$$\dot{P}_{k11}(t) = 2P_{k12}(t) - P_{k11}^2(t), \quad (3.21)$$

$$\dot{P}_{k12}(t) = P_{k22}(t) + 0,3m_{k2}^2(t)P_{k12}(t) + 0,3P_{k22}(t)P_{k12}(t) - P_{k11}(t)P_{k12}(t),$$

$$\dot{P}_{k22}(t) = 1 + 0,6m_{k2}^2(t)P_{k22}(t) + 0,6P_{k22}^2(t) - P_{k12}^2(t),$$

Los resultados de la simulación numérica son obtenidos resolviendo los sistemas de las ecuaciones de filtrado (3.18)–(3.19), y (3.20)–(3.21). Los valores obtenidos de los estimados $m_1(t)$, $m_2(t)$, $m_{k1}(t)$ y $m_{k2}(t)$ que satisfacen las ecuaciones (3.18) y (3.20), respectivamente, son comparados con los valores reales de las variables de estado $x_1(t)$ y $x_2(t)$ en (3.16). Para cada uno de los dos filtros (3.18)–(3.19) y (3.20)–(3.21), y el sistema de referencia (3.16)–(3.17), envueltos en simulación, se asignaron los siguientes valores iniciales: $x_{10} = -2,5$, $x_{20} = -0,35$, $m_{10} = -14,6$, $m_{20} = -1,38$, $P_{110} = 20$, $P_{120} = 0,9$, $P_{220} = 0,06$. Las realizaciones de los ruidos blancos de Poisson $\psi_1(t)$ y $\psi_2(t)$ en (3.20) se generaron utilizando la tabla en Simulink sugerida en [5].

Se obtuvieron las siguientes gráficas: las gráficas de los errores entre las componentes del estado de referencia $x_1(t)$ y $x_2(t)$, que satisfacen las ecuaciones (3.16), y las componentes del estimado del filtro óptimo $m_1(t)$ y $m_2(t)$, que satisfacen (3.18), se muestran en las Figuras 1 y 2; las gráficas de los errores entre las componentes del estado de referencia $x_1(t)$ y $x_2(t)$, que satisfacen las ecuaciones (3.16), y las componentes del estimado del filtro polinomial en promedio cuadrático convencional $m_{k1}(t)$ y $m_{k2}(t)$, que satisfacen las ecuaciones (3.20), se muestran en las Figuras 3 y 4. Se puede observar que el error de estimación dado por el filtro óptimo rápidamente alcanza y después mantiene sus valores cercanos al cero. Esto presenta una definitiva ventaja del filtro óptimo diseñado. Por el contrario, el error de estimación dado por el filtro polinomial en promedio cuadrático convencional diverge a infinito en el tiempo $T = 1,7842$. Esto lleva a la conclusión bien justificada, de que el filtro polinomial en promedio cuadrático convencional diseñado para sistemas Gaussianos es inaplicable para sistemas polinomiales corrompidos con ruidos blancos de Poisson, en tal caso debería utilizarse el filtro diseñado en este capítulo.

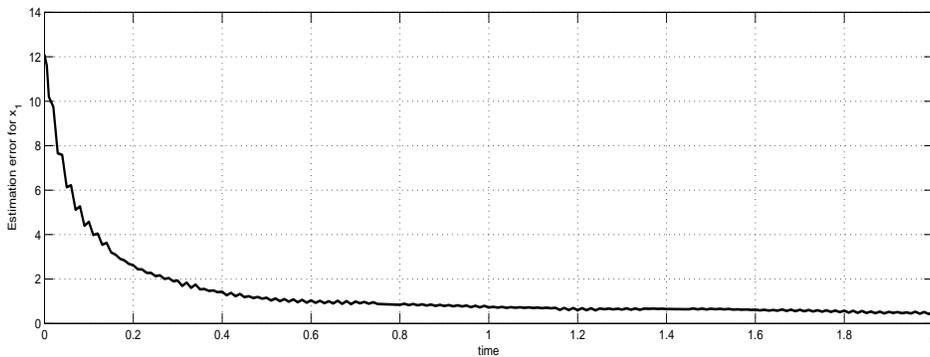


Figura 3.1: Gráfica del error entre el estado real $x_1(t)$, que satisface (3.16), y el estimado del filtro óptimo $m_1(t)$, que satisface (3.18), en el intervalo de simulación $[0, 2]$.

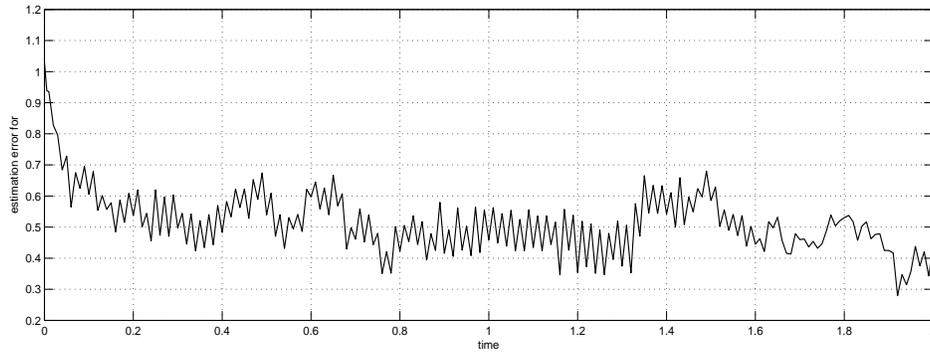


Figura 3.2: Gráfica del error entre el estado real $x_2(t)$, que satisface (3.16), y el estimado del filtro óptimo $m_2(t)$, que satisface (3.18), en el intervalo de simulación $[0, 2]$.

Note que la varianza del error del filtro óptimo $P(t)$ no converge a cero conforme el tiempo tiende al punto de tiempo asintótico, pues la dinámica polinomial de tercer orden es más fuerte que los términos cuadráticos de Ricatti en el lado derecho de las ecuaciones (3.19).

Así, se puede concluir que el filtro óptimo obtenido (3.18)–(3.19) para un estado bidimensional de tercer orden sobre observaciones lineales incompletas definitivamente proporciona mejores estimados que el filtro convencional para sistemas polinomiales con ruidos Gaussianos.

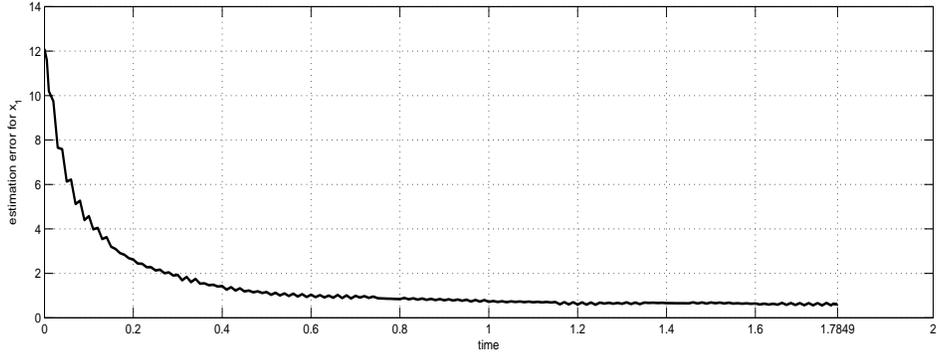


Figura 3.3: Gráfica del error entre el estado real $x_1(t)$, que satisface (3.16), y el estimado $m_{k1}(t)$, que satisface (3.20), en el intervalo de simulación $[0, 1,7842]$.

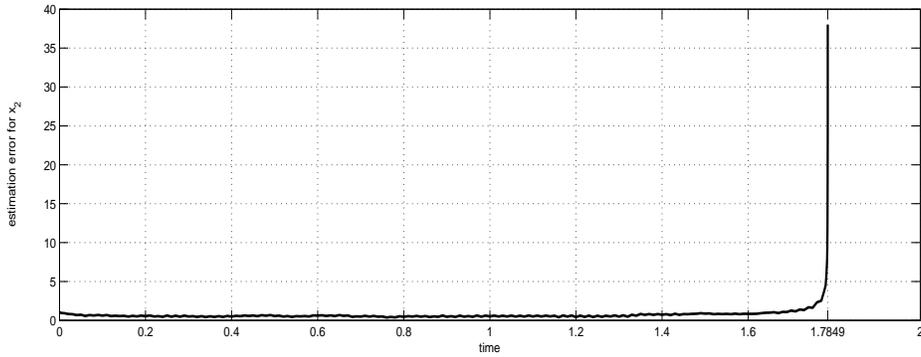


Figura 3.4: Gráfica del error entre el estado real $x_2(t)$, que satisface (3.16), y el estimado $m_{k2}(t)$, que satisface (3.20), en el intervalo de simulación $[0, 1,7842]$.

Capítulo 4

Filtrado Óptimo e Identificación de Parámetro para Sistemas Lineales con Ruidos de Poisson

4.1. Planteamiento del Problema

Sea (Ω, F, P) un espacio completo de probabilidad con una familia de σ -álgebras $F_t, t \geq t_0$ creciente y continua por la derecha, y sean $(N_1(t), F_t, t \geq t_0)$ y $(N_2(t), F_t, t \geq t_0)$ dos procesos de Poisson centralizados independientes. El proceso aleatorio F_t medible $(x(t), y(t))$ es descrito por una ecuación diferencial lineal con un parámetro del vector desconocido $\theta(t)$ para el estado del sistema

$$dx(t) = (a_0(\theta, t) + a(\theta, t)x(t))dt + b(t)dN_1(t), \quad x(t_0) = x_0, \quad (4.1)$$

y una ecuación diferencial lineal para el proceso de observación

$$dy(t) = (A_0(t) + A(t)x(t))dt + B(t)dN_2(t). \quad (4.2)$$

Donde, $x(t) \in R^n$ es el vector de estado, $y(t) \in R^m$ es el proceso de observación, $m \leq n$, y $\theta(t) \in R^p$, $p \leq n \times n + n$, es el vector de entradas desconocidas de la matriz $a(\theta, t)$ y las componentes desconocidas del vector $a_0(\theta, t)$. Esto último significa que ambas estructuras contienen componentes desconocidas $a_{0_i}(t) = \theta_k(t)$, $k = 1, \dots, p_1 \leq n$ y $a_{ij}(t) = \theta_k(t)$, $k = p_1 + 1, \dots, p \leq n \times n + n$, así como componentes conocidas $a_{0_i}(t)$ y $a_{ij}(t)$, cuyos valores son funciones conocidas que dependen del tiempo. La condición inicial $x_0 \in R^n$ es un vector de Poisson tal que x_0 , $N_1(t)$ y $N_2(t)$ son independientes. Se asumirá que $B(t)B^T(t)$ es una matriz definida positiva. Todos los coeficientes en (4.1)–(4.2) son funciones determinísticas del tiempo de dimensiones apropiadas.

Se considerará que no hay información útil sobre los valores de los parámetros desconocidos $\theta_k(t)$, $k = 1, \dots, p$, incluso esta incertidumbre sigue creciendo conforme el tiempo tiende a infinito. En otras palabras, los parámetros desconocidos pueden ser modelados como procesos de Poisson F_t medibles

$$d\theta(t) = dN_3(t), \quad (4.3)$$

con condiciones iniciales desconocidas $\theta(t_0) = \theta_0 \in R^p$, donde $(N_3(t), F_t, t \geq t_0)$ es un proceso de Poisson independiente de x_0 , $N_1(t)$, y $N_2(t)$.

El problema de estimación consiste en encontrar el estimado óptimo $\hat{z}(t) = [\hat{x}(t), \hat{\theta}(t)]$ del vector combinado de los estados del sistema y los parámetros desconocidos $z(t) = [x(t), \theta(t)]$, basado en el proceso de observación $Y(t) = \{y(s), 0 \leq s \leq t\}$. Como se sabe [53], este estimado óptimo está dado por el valor esperado condicional

$$\hat{z}(t) = m(t) = E(z(t) | F_t^Y)$$

del estado del sistema $z(t) = [x(t), \theta(t)]$ con respecto a la σ - álgebra F_t^Y generada por el proceso de observación $Y(t)$ en el intervalo de tiempo $[t_0, t]$. Como es usual, la función matricial

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

es la matriz de la varianza del error .

La solución propuesta a este problema de filtrado óptimo se basa en las ecuaciones de filtrado óptimo para estados polinomiales que no son completamente medibles con ruido de Poisson sobre observaciones lineales , presentadas en el capítulo anterior [8].

4.2. Diseño del Filtro Óptimo e Identificador

Para aplicar las ecuaciones de filtrado óptimo para el vector de estado $z(t) = [x(t), \theta(t)]$, gobernado por las ecuaciones (4.1) y (4.3), sobre las observaciones lineales (4.2) (ver [8]), la ecuación de estado(4.1) debería escribirse en su forma polinomial. Para esto, se introducirá una matriz $a_1(t) \in R^{(n+p) \times (n+p)}$, un tensor cúbico $a_2(t) \in R^{(n+p) \times (n+p) \times (n+p)}$, y un vector $c_0(t) \in R^{(n+p)}$ de la siguiente manera:

La ecuación para la i -ésima componente del vector de estado está dada por

$$dx_i(t) = (a_{0_i}(t) + \sum_{j=1}^n a_{ij}(t)x_j(t))dt + \sum_{j=1}^n b_{ij}(t)dN_{1_j}(t), \quad x_i(t_0) = x_{0_i}.$$

Entonces:

1. Si la variable $a_{0_i}(t)$ es una función conocida, entonces la i -ésima componente del vector $c_0(t)$ será ocupada por esta función, $c_{0_i}(t) = a_{0_i}(t)$. De lo contrario, si la variable $a_{0_i}(t)$ es una función desconocida, entonces la $(i, n + i)$ -ésima entrada de la matriz $a_1(t)$ será ocupada por un 1. Como se ha señalado, el número de componentes desconocidas del vector a_0 es igual a $p_1 \leq n$.

2. Si la variable $a_{ij}(t)$ es una función conocida, entonces la (i, j) -ésima componente de la matriz $a_1(t)$ se ocupará con esta función, $a_{1_{ij}}(t) = a_{ij}(t)$. De lo contrario, si la variable $a_{ij}(t)$ es una función desconocida, entonces la $(i, n + p_1 + k, j)$ -ésima entrada del tensor cúbico $a_2(t)$ contendrá un 1, donde k está contando las entradas desconocidas actuales en la matriz $a_{ij}(t)$ desde la primera hasta la n -ésima entrada en cada renglón. El número de entradas desconocidas en la matriz a_1 es igual a $p - p_1 \leq n^2$, donde $p_1 \leq n$ es el número de componentes desconocidas del vector a_0 , y $p \leq n^2 + n$ es el número total de variables desconocidas.

3. Todas las entradas no asignadas de la matriz $a_1(t)$, el tensor cúbico $a_2(t)$, y del vector $c_0(t)$ serán igualadas a cero.

Usando la notación introducida, las ecuaciones de estado (4.1),(4.3) para el vector $z(t) = [x(t), \theta(t)] \in R^{n+p}$ pueden reescribirse como

$$dz(t) = (c_0(t) + a_1(t)z(t) + a_2(t)z(t)z^T(t))dt + \quad (4.4)$$

$$diag[b(t), I_{p \times p}]d[N_1^T(t), N_3^T(t)]^T, \quad z(t_0) = [x_0, \theta_0],$$

donde la matriz $a_1(t)$, el tensor cúbico $a_2(t)$, y el vector $c_0(t)$ ya han sido definidos, y $I_{p \times p}$ es la matriz identidad de dimensión $p \times p$. La ecuación (4.4) es bilineal con respecto al vector de estado extendido $z(t) = [x(t), \theta(t)]$.

Entonces, el problema de estimación ahora es reformulado como encontrar el estimado óptimo $\hat{z}(t) = m(t) = [\hat{x}(t), \hat{\theta}(t)]$ para el vector de estado $z(t) = [x(t), \theta(t)]$, gobernado por la ecuación bilineal (4.4), basada en el proceso de observación $Y(t) = \{y(s), 0 \leq s \leq t\}$, que satisface la ecuación (4.2). La solución a este problema se obtiene utilizando las ecuaciones de filtrado óptimo para estados lineales-bilineales con la parte lineal medible sólo parcialmente sobre observaciones lineales [8] dada por

$$dm(t) = (c_0(t) + a_1(t)m(t) + a_2(t)m(t)m^T(t) + a_2(t)P(t))dt + \quad (4.5)$$

$$P(t)[A(t), 0_{m \times p}]^T (B(t)B^T(t))^{-1} [dy(t) - (A_0(t) + A(t)m(t))dt],$$

$$m(t_0) = [E(x(t_0) | F_t^Y), E(\theta(t_0) | F_t^Y)],$$

$$dP(t) = (a_1(t)P(t) + P(t)a_1^T(t) + 2a_2(t)m(t)P(t) + a_2(t)P(t) * 1 + \quad (4.6)$$

$$(a_2(t)(2m(t)P(t) + P(t) * 1))^T + (diag[b(t), I_p])(diag[b(t), I_p]^T))dt -$$

$$P(t)[A(t), 0_{m \times p}]^T (B(t)B^T(t))^{-1} [A(t), 0_{m \times p}]P(t)dt,$$

$$P(t_0) = E((z(t_0) - m(t_0))(z(t_0) - m(t_0))^T | F_t^Y),$$

donde el vector 1 es un vector de dimensión n con todas sus componentes iguales a 1; $0_{m \times p}$ es la matriz de ceros de dimensión $m \times p$; $P(t)$ es la varianza condicional del error de estimación $z(t) - m(t)$ con respecto a las observaciones $Y(t)$; y la expresión $a_2(t)P(t) * 1$ se define de la siguiente manera:

$$(a_2(t)P(t) * 1)_{ih} = \sum_{j,k} a_{2\ ij k}(t)P_{jk}(t)1_h$$

Teorema 1. El filtro óptimo de dimensión finita para el vector de estado extendido $z(t) = [x(t), \theta(t)]$, gobernado por la ecuación (4.4), sobre las observaciones lineales (4.2) está dado por la ecuación (4.5) para el estimado óptimo $\hat{z}(t) = m(t) = [\hat{x}(t), \hat{\theta}(t)] = E([x(t), \theta(t)] | F_t^Y)$ y la ecuación (4.6) para la varianza del error de estimación $P(t) = E[(z(t) - m(t))(z(t) - m(t))^T | F_t^Y]$. Además, este filtro aplicado al subvector $\theta(t)$, sirve

como el identificador óptimo para el vector del parámetro desconocido $\theta(t)$ en la ecuación (4.1), produciendo el subvector estimado $\hat{\theta}(t)$ como el estimado del parámetro óptimo.

Demostración. La demostración sigue directamente de los pasos 1-3 para diseñar los coeficientes en la ecuación (4.4), la nueva ecuación del estado extendido (4.4), y las ecuaciones de filtrado óptimo (4.5),(4.6) para estados bilineales incompletamente medibles sobre observaciones lineales, las cuales fueron obtenidas en [8].

Por lo tanto, basado en las ecuaciones de filtrado óptimo para estados bilineales medibles incompletamente sobre observaciones lineales, el filtro de estado óptimo y el identificador de parámetro es obtenido para estado del sistema lineal (4.1) con parámetros desconocidos, modelado por (4.3), sobre las observaciones lineales (4.2). Dado que el problema de identificación original se reduce al problema de filtrado para el estado del sistema extendido incluyendo ambos, el estado y los parámetros, la condición de identificabilidad para el sistema original coincide con la condición de observabilidad para el sistema extendido.

4.3. Ejemplo

En esta sección se presenta un ejemplo de diseño de filtrado óptimo e identificación para un estado bilineal con un parámetro multiplicativo desconocido, basado en mediciones del estado lineal y es comparado con un filtro convencional en promedio cuadrático para sistemas polinomiales estocásticos con ruidos Gaussianos.

Sea $x(t)$ un estado real bidimensional que satisface el siguiente sistema bilineal con un parámetro escalar desconocido $\theta \in R$

$$\dot{x}_1(t) = x_2(t) \tag{4.7}$$

$$\dot{x}_2(t) = \theta x_1(t) + \psi_1(t),$$

y el proceso escalar de observación $y(t) \in R$ está dado por la ecuación lineal

$$y(t) = x_1(t) + \psi_2(t), \quad (4.8)$$

donde $\psi_1(t)$ y $\psi_2(t)$ son ruidos blancos de Poisson, los cuales son las derivadas débiles en promedio cuadrático de procesos estándar de Poisson (ver [53]). Las ecuaciones (4.7),(4.8) presentan la forma convencional para las ecuaciones (4.1)–(4.2), los cuales son utilizados en la práctica [2]. El parámetro θ es modelado como un proceso estándar de Poisson, i.e., satisface la ecuación

$$d\theta(t) = dN_3(t), \quad \theta(0) = \theta_0,$$

la cual también puede escribirse como

$$\dot{\theta}(t) = \psi_3(t), \quad \theta(0) = \theta_0, \quad (4.9)$$

donde $\psi_3(t)$ es un ruido blanco de Poisson. Los ruidos $\psi_1(t)$ y $\psi_3(t)$ en el estado y las ecuaciones del parámetro se asume que son ruidos blancos de Poisson independientes.

El problema de filtrado consiste en encontrar el estimado óptimo para para el estado lineal-bilineal (4.7),(4.9), usando observaciones lineales (4.8) con disturbios aleatorios independientes aislados en el tiempo, modelados como ruidos blancos de Poisson.

Las ecuaciones de filtrado (4.5),(4.6) toma la forma particular para el sistema (4.7)–(4.9)

$$\begin{aligned} \dot{m}_1(t) &= m_2(t) + P_{11}(t)(y(t) - m_1(t)) \\ \dot{m}_2(t) &= P_{13}(t) + m_1(t)m_3(t) + P_{12}(t)(y(t) - m_1(t)) \end{aligned} \quad (4.10)$$

$$\dot{m}_3(t) = P_{13}(t)(y(t) - m_1(t)),$$

con las condiciones iniciales $m_1(0) = E(x_{10} | y(0)) = m_{10}$, $m_2(0) = E(x_{20} | y(0)) = m_{20}$ y $m_3(0) = E(\theta_0 | y(0)) = m_{30}$,

$$\dot{P}_{11}(t) = 2P_{12}(t) - P_{11}^2(t)$$

$$\dot{P}_{12}(t) = P_{13}(t) + P_{22}(t) - P_{11}(t)P_{12}(t) + 2P_{11}(t)m_3(t) \quad (4.11)$$

$$\dot{P}_{13}(t) = P_{23}(t) - P_{11}(t)P_{13}(t) \quad \dot{P}_{22}(t) = 1 + 2P_{13}(t) + 4m_3(t)P_{12}(t) - P_{12}^2(t)$$

$$\dot{P}_{23}(t) = P_{13}(t) - P_{12}(t)P_{13}(t) + 2P_{13}(t)m_3(t)$$

$$\dot{P}_{33}(t) = 1 - P_{13}^2(t),$$

con la condición inicial $P(0) = E((x_0, \theta_0) - m(0))(x_0, \theta_0) - m(0))^T | y(0)) = P_0$.

Los estimados obtenidos al resolver las ecuaciones (4.10),(4.11) serán comparados con los estimados que satisfacen las ecuaciones del filtro polinomial convencional para ruidos Gaussianos para los estados (4.7),(4.9) sobre observaciones lineales incompletas (4.8):

$$\dot{m}_{G1}(t) = m_{G2}(t) + P_{G11}(t)(y(t) - m_{G1}(t))$$

$$\dot{m}_{G2}(t) = P_{G13}(t) + m_{G1}(t)m_{G3}(t) + P_{G12}(t)(y(t) - m_{G1}(t)) \quad (4.12)$$

$$\dot{m}_{G3}(t) = P_{G13}(t)(y(t) - m_{G1}(t)),$$

con las condiciones iniciales $m_{G1}(0) = E(x_{10} | y(0)) = m_{G10}$, $m_{G2}(0) = E(x_{20} | y(0)) = m_{G20}$ y $m_{G3}(0) = E(\theta_0 | y(0)) = m_{G30}$,

$$\dot{P}_{G11}(t) = 2P_{G12}(t) - P_{G11}^2(t)$$

$$\dot{P}_{G12}(t) = P_{G22}(t) - P_{G11}(t)P_{G12}(t) + 2P_{G11}(t)m_{G3}(t) \quad (4.13)$$

$$\dot{P}_{G13}(t) = P_{G23}(t) - P_{G11}(t)P_{G13}(t) \quad \dot{P}_{G22}(t) = 1 + 4m_{G3}(t)P_{G12}(t) - P_{G12}^2(t)$$

$$\dot{P}_{G23}(t) = 2P_{G13}(t)m_{G3}(t) - P_{G12}(t)P_{G13}(t) \quad \dot{P}_{G33}(t) = 1 - P_{G13}^2(t).$$

Los resultados de la simulación numérica se obtienen al resolver los sistemas de ecuaciones de filtrado (4.10),(4.11) y (4.12),(4.13). Los valores obtenidos de $m_1(t)$, $m_2(t)$, $m_3(t)$, $m_{G1}(t)$, $m_{G2}(t)$ y $m_{G3}(t)$ que satisfacen las ecuaciones (4.10), y (4.12), respectivamente, son comparados con los valores reales de las variables de estado $x_1(t)$, $x_2(t)$ y el parámetro $\theta(t)$ en (4.7),(4.9).

Para cada uno de los dos filtros (4.10),(4.11) y (4.12),(4.13), y el sistema de referencia (4.7)–(4.9) se asignaron los siguientes valores iniciales: $x_{10} = 3$, $x_{20} = 3$, $m_{10} = 1$, $m_{20} = 1$, $m_{30} = -0,44$, $P_{110} = P_{220} = P_{330} = 10$, $P_{120} = P_{130} = P_{230} = 1$. El parámetro desconocido θ se asigna como $\theta = -0,1$ en la primer simulación, y como $\theta = 0,1$ en la segunda simulación, entonces considerando el sistema (4.7) estable e inestable, respectivamente. Los ruidos de Poisson $\psi_1(t)$ y $\psi_2(t)$ en (4.9) fueron generados utilizando un diagrama en Simulink sugerido en [5].

Se obtuvieron las siguientes gráficas: Las Figuras 1 y 2 muestran las gráficas de las variables del estado de referencia $x_1(t)$, $x_2(t)$ y los estimados óptimos del estado $m_1(t)$, $m_2(t)$ y el estimado del parámetro $m_3(t)$, que satisface la ecuación (4.10), en el caso estable ($\theta = -0,1$) y en el caso inestable ($\theta = 0,1$) respectivamente; Las Figuras 3 y 4 muestran las gráficas de las variables del estado de referencia $x_1(t)$, $x_2(t)$ y los estimados del filtro convencional Gaussiano $m_{G1}(t)$, $m_{G2}(t)$ y $m_{G3}(t)$, que satisfacen la ecuación (4.12), en el caso estable ($\theta = -0,1$) y en el caso inestable ($\theta = 0,1$) respectivamente. En los casos donde el sistema es estable el tiempo de simulación se fijó en $T = 50$ y para el caso donde el sistema se comporta inestable el tiempo se fijó en $T = 10$ con el fin de apreciar un mejor rendimiento en cada uno de los casos.

Se puede observar que, en ambos casos, los estimados del estado, $m_1(t)$ y $m_2(t)$ las variables de estado reales $x_1(t)$, $x_2(t)$ y el estimado del parámetro $m_3(t)$ convergen al valor real (0.1 or -0.1) del parámetro desconocido $\theta(t)$. Este comportamiento puede clasificarse

como muy confiable, especialmente tomando en cuenta grandes desviaciones en las condiciones iniciales para el estado real y su estimado y grandes valores de la varianza del error inicial.

Otra ventaja que se puede mencionar es que el filtro diseñado y el identificador del parámetro trabajan correctamente en ambos casos sin importar la estabilidad o inestabilidad del sistema, lo cual corresponde a la operación de sistemas bilineales en condiciones nominales bajo la presencia de disturbios externos persistentes., respectivamente. Por el contrario, puede observar que los estimados del estado $m_{G1}(t)$ y $m_{G2}(t)$ ni siquiera se aproximan a las variables de estado reales $x_1(t)$ y $x_2(t)$ en ambos casos, para el sistema estable e inestable (4.7), y el estimado del parámetro $m_{G3}(t)$ no converge al valor real (0,1 or $-0,1$) del parámetro desconocido $\theta(t)$.

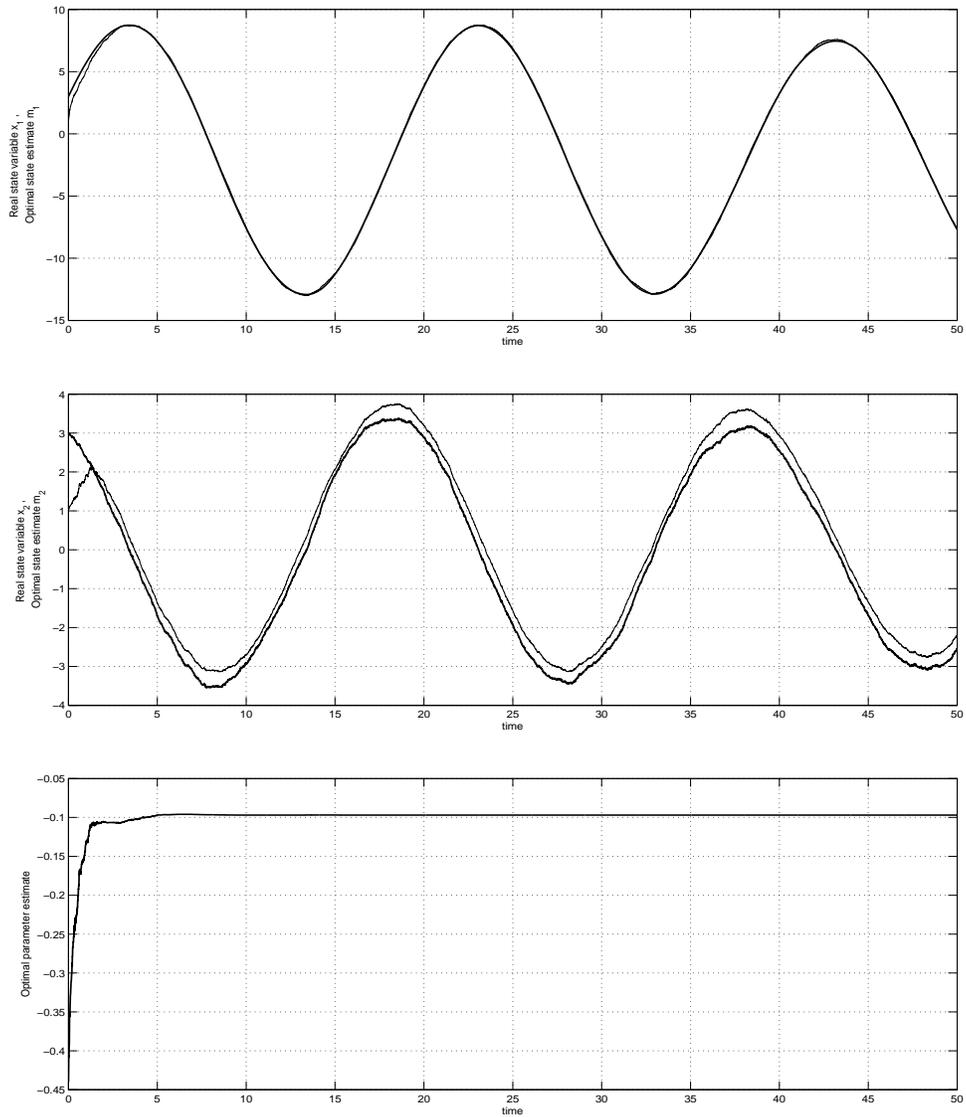


Figura 4.1: Gráficas de las variables de estado de referencia $x_1(t)$ y $x_2(t)$ (línea gruesa en la primera y segunda gráfica), estimados del estado óptimo $m_1(t)$ y $m_2(t)$ (línea delgada en la primera y segunda gráfica) y estimado óptimo del parámetro $m_3(t)$ (línea gruesa en la tercera gráfica), que satisfacen (4.10), para el sistema estable (4.7) en el intervalo de simulación $[0, 50]$.

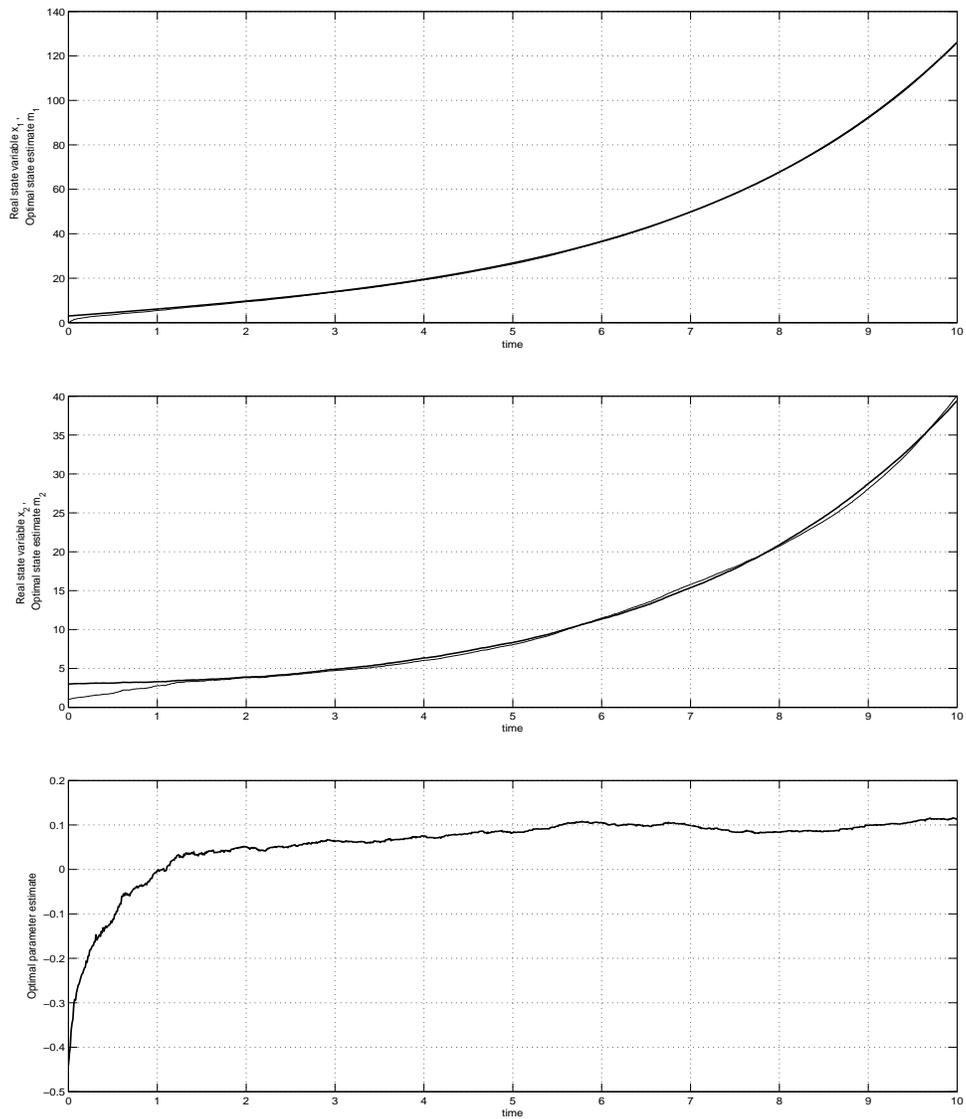


Figura 4.2: Gráficas de las variables del estado de referencia $x_1(t)$ y $x_2(t)$ (línea gruesa en la primera y segunda gráfica), estimados óptimos del estado $m_1(t)$ y $m_2(t)$ (línea delgada en la primera y segunda gráfica) y el estimado óptimo del parámetro $m_3(t)$ (línea gruesa en la tercera gráfica), que satisfacen (4.10), para el sistema inestable (4.7) en un intervalo de simulación $[0, 10]$.

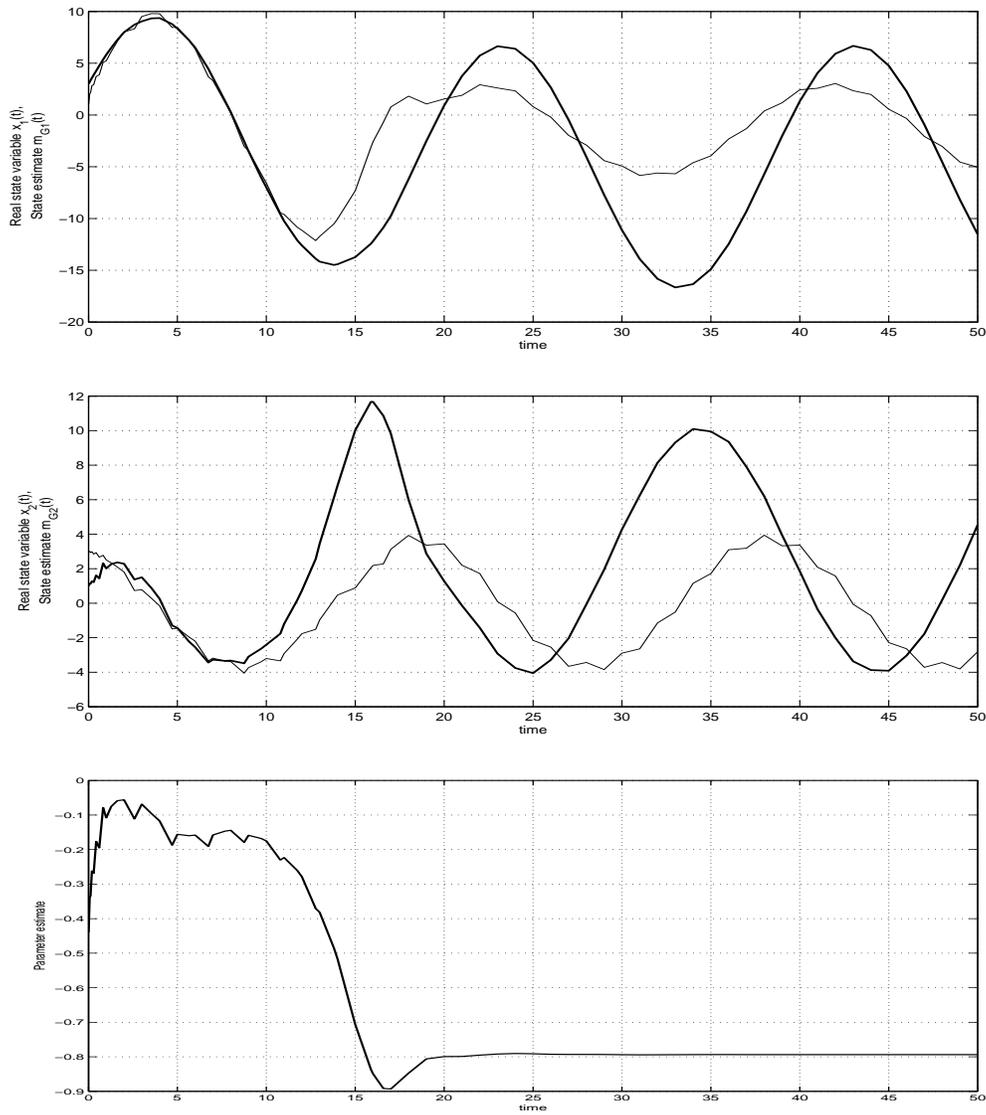


Figura 4.3: Gráficas de las variables del estado de referencia $x_1(t)$ y $x_2(t)$ (línea delgada en la primera y segunda gráfica), estimados del estado $m_{G1}(t)$ y $m_{G1}(t)$ (línea delgada en la primera y segunda gráfica) y el estimado del parámetro $m_{G3}(t)$ (línea gruesa en la tercera gráfica), que satisface (4.12) para el sistema estable (4.7) en el intervalo de simulación $[0, 50]$.

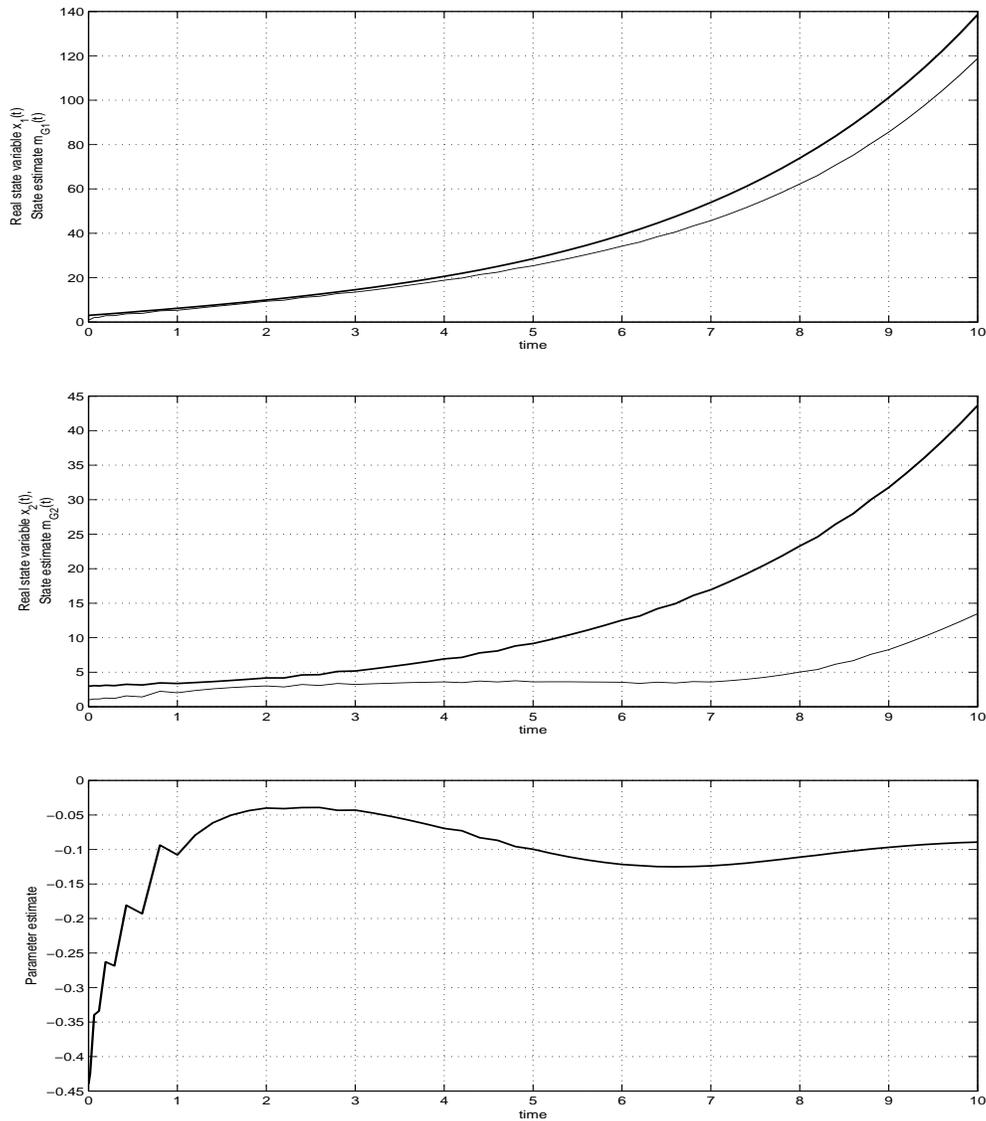


Figura 4.4: Gráficas de las variables del estado de referencia $x_1(t)$ y $x_2(t)$ (línea gruesa en la primera y segunda gráfica), estimados óptimos del estado $m_{G1}(t)$ y $m_{G2}(t)$ (línea delgada en la primera y segunda gráfica) y estimado del parámetro $m_{G3}(t)$ (línea gruesa en la tercera gráfica), que satisface (4.12), para el sistema inestable (4.7) en el intervalo de simulación $[0, 10]$.

Capítulo 5

Controlador Óptimo LQP para Sistemas Estocásticos Lineales con Parámetros Desconocidos y Ruido de Poisson

5.1. Planteamiento del Problema

Sea (Ω, F, P) un espacio completo de probabilidad con una familia creciente de σ -álgebras $F_t, t \geq t_0$, y sean $(N_1(t), F_t, t \geq t_0)$ y $(N_2(t), F_t, t \geq t_0)$ dos procesos independientes de Poisson centralizados. El proceso aleatorio F_t medible $(x(t), y(t))$ es descrito por una ecuación diferencial lineal con un parámetro desconocido $\theta(t)$ para el estado del sistema

$$dx(t) = (a(\theta, t)x(t))dt + B(t)u(t)dt + b(t)dN_1(t), \quad x(t_0) = x_0, \quad (5.1)$$

y una ecuación diferencial para el proceso de observación

$$dy(t) = (A_0(t) + A(t)x(t))dt + G(t)dN_2(t). \quad (5.2)$$

Donde, $x(t) \in R^n$ es el vector de estado, $u(t) \in R^l$ es la entrada del control, $y(t) \in R^m$ es el proceso de observación, y $\theta(t) \in R^p$, $p \leq n \times n$, es el vector de entradas desconocidas de la matriz $a(\theta, t)$. Esto último significa que $a(\theta, t)$ contiene las componentes desconocidas $a_{ij}(t) = \theta_k(t)$, $k = 1, \dots, p \leq n \times n$, así como las componentes conocidas $a_{ij}(t)$, cuyos valores son funciones conocidas que dependen del tiempo. La condición inicial $x_0 \in R^n$ es un vector de Poisson tal que x_0 , $N_1(t)$, y $N_2(t)$ son independientes. Se asume que $G(t)G^T(t)$ es una matriz definida positiva. Todos los coeficientes en (5.1)–(5.2) son funciones determinísticas que dependen del tiempo, de dimensiones apropiadas.

Se considera que no existe información útil sobre los valores de los parámetros desconocidos $\theta_k(t)$, $k = 1, \dots, p$. En otras palabras, los parámetros desconocidos pueden ser modelados como procesos de Poisson F_t medibles

$$d\theta(t) = \beta(t)dN_3(t), \quad (5.3)$$

con condiciones iniciales desconocidas $\theta(t_0) = \theta_0 \in R^p$, donde $(N_3(t), F_t, t \geq t_0)$ es un proceso de Poisson independiente de x_0 , $N_1(t)$, y $N_2(t)$, y $\beta(t) \in R^{p \times p}$ es una función de intensidad.

La función de costo cuadrática J ha de ser minimizada se define de la siguiente manera

$$J = \frac{1}{2}E[x^T(T)\Phi x(T) + \int_{t_0}^T u^T(s)R(s)u(s)ds + \int_{t_0}^T x^T(s)L(s)x(s)ds], \quad (5.4)$$

donde R es definida positiva y Φ, L son matrices simétricas definidas no negativas, $T > t_0$ es un tiempo específico, el símbolo $E[f(x)]$ representa el valor esperado de la función f que depende de una variable aleatoria x , y a^T denota la transpuesta de un vector (matriz) a .

El problema del controlador óptimo consiste en encontrar el control $u^*(t)$, $t \in [t_0, T]$, que minimice el criterio J con la trayectoria no observable $x^*(t)$, $t \in [t_0, T]$, generada al sustituir $u^*(t)$ en la ecuación de estado (5.1).

5.2. Diseño del Controlador Óptimo

5.2.1. Reducción del Problema

Para tratar el problema planteado del controlador, las ecuaciones (5.1) y (5.3) deberían ser reordenadas. Para esto, se introducirán una matriz $a_1(t) \in R^{(n+p) \times (n+p)}$ y un tensor cúbico $a_2(t) \in R^{(n+p) \times (n+p) \times (n+p)}$ de la siguiente manera:

Note que la ecuación para la i -ésima componente del vector de estado (5.1) está dada por

$$dx_i(t) = \left(\sum_{j=1}^n a_{ij}(t)x_j(t) \right) dt + \sum_{j=1}^n b_{ij}(t)dN_{1_j}(t), \quad x_i(t_0) = x_{0_i},$$

entonces, si la variable $a_{ij}(t)$ es una función conocida, entonces la (i, j) -ésima componente de la matriz $a_1(t)$ corresponderá a esta función, $a_{1_{ij}}(t) = a_{ij}(t)$; de lo contrario, si la variable $a_{ij}(t)$ es una función desconocida, entonces la $(i, n+k, j)$ -ésima entrada del tensor cúbico $a_2(t)$ se igualará a 1, donde k es el número actual de entradas desconocidas en la matriz $a_{ij}(t)$, contando el número de entradas consecutivamente por filas desde la primera hasta la n -ésima entrada en cada fila. Todas las entradas que no han sido asignadas de la matriz $a_1(t)$ y el tensor cúbico $a_2(t)$ se igualarán a 0.

Usando la notación que se acaba de introducir, las ecuaciones de estado (5.1),(5.3) para el vector $z(t) = [x(t), \theta(t)] \in R^{n+p}$ pueden reescribirse como

$$dz(t) = (a_1(t)z(t) + a_2(t)z(t)z^T(t))dt + \text{diag}[b(t), \beta(t)]d[N_1^T(t), N_3^T(t)]^T, \quad (5.5)$$

$$z(t_0) = [x_0, \theta_0],$$

donde la matriz $a_1(t)$ y el tensor cúbico $a_2(t)$ ya han sido definidos. La ecuación (5.5) es bilineal con respecto al vector de estado extendido $z(t) = [x(t), \theta(t)]$.

5.2.2. Principio de Separación

Se puede observar que el principio de separación sigue siendo válido para un sistema estocástico lineal con parámetros desconocidos. Incluso, reemplacemos el estado bilineal no medible $z(t) = [x(t), \theta(t)]$, que satisface (5.1),(5.3), con su estimado óptimo $m(t)$ sobre las observaciones lineales $y(t)$ (5.2), el cual es obtenido utilizando el siguiente filtro óptimo para estados bilineales sobre observaciones lineales (ver [7] para el correspondiente problema de filtrado planteado y su solución)

$$dm(t) = (a_1(t)m(t) + a_2(t)m(t)m^T(t) + [B(t) | 0_{p \times l}]u(t)dt + a_2(t)P(t)dt + P(t)[A(t), 0_{m \times p}]^T(G(t)G^T(t))^{-1}[dy(t) - A(t)m(t)dt], \quad (5.6)$$

$$m(t_0) = [E(x(t_0) | F_t^Y), E(\theta(t_0) | F_t^Y)],$$

$$dP(t) = (a_1(t)P(t) + P(t)a_1^T(t) + 2a_2(t)m(t)P(t) + a_2(t)P(t) * 1 + (a_2(t)(2m(t)P(t) + P(t) * 1))^T + (\text{diag}[b(t), \beta_p])(\text{diag}[b(t), \beta_p]^T))dt - \quad (5.7)$$

$$P(t)[A(t), 0_{m \times p}]^T (G(t)G^T(t))^{-1} [A(t), 0_{m \times p}] P(t) dt,$$

$$P(t_0) = E((z(t_0) - m(t_0))(z(t_0) - m(t_0))^T | F_t^Y),$$

donde $(a_2(t)P(t) * 1)_{ih} = \sum_{j,k} a_{2\ ijk}(t) P_{kj}(t) 1_h$, $0_{m \times p}$ es la matriz cero, de dimensión $m \times p$; $P(t)$ es la varianza condicional del error de estimación $z(t) - m(t)$ con respecto a las observaciones $Y(t)$.

Recuerde que $\hat{z}(t) = m(t) = [\hat{x}(t), \hat{\theta}(t)]$ es el estimado óptimo para el vector de estado $z(t) = [x(t), \theta(t)]$, basado en el proceso de observación $Y(t) = \{y(s), t_0 \leq s \leq t\}$, que minimiza la norma Euclidiana

$$H = E[(z(t) - \hat{z}(t))^T (z(t) - \hat{z}(t)) | F_t^Y] \quad (5.8)$$

a cada momento del tiempo t . Aquí, $E[\xi(t) | F_t^Y]$ representa el valor esperado condicional de un proceso estocástico $\xi(t) = (z(t) - \hat{z}(t))^T (z(t) - \hat{z}(t))$ 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 [53], este estimado óptimo está dado por el valor esperado condicional

$$\hat{z}(t) = m(t) = E(z(t) | F_t^Y)$$

del estado del sistema $z(t)$ con respecto a la σ - álgebra F_t^Y generada por el proceso de observación $Y(t)$ en el intervalo $[t_0, t]$. La función matricial

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

es el estimado de la varianza del error.

Se puede verificar fácilmente (ver [44]) que el problema de control óptimo para el estado del sistema (5.1) y la función de costo (5.4) es equivalente al problema de control óptimo para el estimado (5.6) y la función de costo J representada como

$$\begin{aligned}
J = E\{ & \frac{1}{2}m^T(T)\Phi_1m(T) + \frac{1}{2}\int_{t_0}^T u^T(s)R(s)u(s)ds + \\
& \frac{1}{2}\int_{t_0}^T m^T(s)L_1(s)m(s)ds + \\
& \frac{1}{2}\int_{t_0}^T tr[P(s)L_1(s)]ds + \frac{1}{2}tr[P(T)\Phi_1]\},
\end{aligned} \tag{5.9}$$

donde $\Phi_1 = diag[\Phi, 0_{p \times p}]$, $L_1 = diag[L, 0_{p \times p}]$, y $tr[A]$ denotan la traza de una matriz A . Dado que los primeros tres términos de J son independientes del estimado del parámetro desconocido $\hat{\theta}(t)$, la función de costo (5.9) puede minimizarse mediante dos pasos consecutivos: Primero, los primeros tres términos de J se minimizan asumiendo que el parámetro $\theta(t)$ es una función conocida que depende del tiempo, i.e, la función de costo reducida

$$\begin{aligned}
M = E\{ & \frac{1}{2}\hat{x}^T(T)\Phi\hat{x}(T) + \frac{1}{2}\int_{t_0}^T u^T(s)R(s)u(s)ds + \\
& \frac{1}{2}\int_{t_0}^T \hat{x}^T(s)L(s)\hat{x}(s)ds \}
\end{aligned} \tag{5.10}$$

se emplea en el primer paso. Como resultado, el control óptimo es obtenido minimizando la función de costo reducida M (5.10) para el estimado del estado $m(t)$ (5.6) bajo la asunción de que el parámetro $\theta(t)$ es conocido. Segundo, los últimos dos términos de J son minimizados disminuyendo la norma de la varianza del error de estimación $P(t)$ en cada momento del tiempo t . Finalmente, el valor mínimo del criterio J se determina utilizando (5.9). Esta sección presenta el principio de separación para sistemas lineales con parámetros desconocidos.

5.2.3. Solución del Problema de Control Óptimo

Como primer paso, se debe notar que la ecuación del estimado óptimo del estado (5.6) es lineal en m bajo la suposición de que el parámetro $\theta(t)$ es una función conocida

que depende del tiempo. Incluso, en este caso, $a_2 = 0$ por construcción, y las ecuaciones (5.6),(5.7) se convierten en las ecuaciones del filtro lineal óptimo de Kalman-Bucy ([39]). Por lo tanto, la solución óptima está dada por el controlador óptimo LQG para sistemas lineales ([44]), el cual se aplica a la ecuación del estimado (5.6), asumiendo que el parámetro $\theta(t)$ es conocido y $a_2 = 0$, y la función de costo reducida(5.10). La correspondiente ley de control óptimo está dada por

$$u^*(t) = (R(t))^{-1}B^T(t)Q(t)m(t), \quad (5.11)$$

donde la función matricial $Q(t)$ es la solución de la siguiente ecuación de Riccati

$$\begin{aligned} \dot{Q}(t) = & -a^T(\theta, t)Q(t) - Q(t)a(\theta, t) + L(t) - \\ & Q(t)B(t)R^{-1}(t)B^T(t)Q(t), \end{aligned}$$

con la condición final $Q(T) = \Phi$.

Como segundo paso, recordando que el parámetro $\theta(t)$ es desconocido, el estimado para $\theta(t)$ debería ser asignado para minimizar la función de costo (5.9). De cualquier manera, el mejor estimado que minimiza la norma de la varianza del error de estimación $P(t)$, y en consecuencia la función de costo (5.9) para todo tiempo t , está dado por $\hat{\theta}(t)$, la segunda componente de $m(t) = \hat{z}(t)$, en vista de (5.8). Por lo tanto, la matriz de ganancia de la ecuación de Riccati toma la forma

$$\begin{aligned} \dot{Q}(t) = & -a^T(\hat{\theta}(t), t)Q(t) - Q(t)a(\hat{\theta}(t), t) + L(t) - \\ & Q(t)B(t)R^{-1}(t)B^T(t)Q(t), \end{aligned} \quad (5.12)$$

con la condición final $Q(T) = \Phi$.

Al sustituir el control óptimo (5.11) en la ecuación (5.6), se obtiene la siguiente ecuación del estimado controlado de manera óptima.

$$dm(t) = (a_1(t)m(t) + a_2(t)m(t)m^T(t) + a_2(t)P(t))dt + \quad (5.13)$$

$$[B(t) \mid 0_{p \times l}](R(t))^{-1}B^T(t)Q(t)m(t)dt + \\ P(t)[A(t), 0_{m \times p}]^T(G(t)G^T(t))^{-1}[dy(t) - A(t)m(t)dt],$$

con la condición inicial $m(t_0) = [E(x(t_0) \mid F_t^Y), E(\theta(t_0) \mid F_t^Y)]$.

Por lo tanto, la ecuación del estimado del estado controlado óptimamente (5.13), la ecuación de la matriz de ganancia (5.12), la ley de control óptimo (5.11), y la ecuación de la varianza (5.7) proporcionan la solución completa al problema del controlador óptimo para sistemas lineales con parámetros desconocidos.

5.3. Ejemplo

Esta sección presenta un ejemplo del diseño del controlador óptimo para el sistema bilineal (5.1) con un parámetro desconocido θ (5.3) sobre observaciones lineales (5.2), empleando el esquema (5.11)–(5.13), y comparándolo con el controlador para el sistema de referencia (5.1),(5.2) con exactamente un parámetro conocido θ .

Considere un sistema bilineal con la ecuación de estado dada por

$$\dot{x}_1(t) = x_2(t), x_1(0) = x_{10}, \quad (5.14)$$

$$\dot{x}_2(t) = \theta x_1(t) + u(t) + \psi_1(t), x_2(0) = x_{20},$$

y el proceso de observación dado por

$$y(t) = x_1(t) + \psi_2(t), \quad (5.15)$$

donde $\psi_1(t)$ y $\psi_2(t)$ son ruidos blancos de Poisson, los cuales son la derivada débil en promedio cuadrático de procesos estandar de Poisson (ver [53]), y x_0 es una variable aleatoria de Poisson. Las ecuaciones (5.14) y (5.15) presentan la forma convencional para las ecuaciones (5.1) y (5.2), las cuales son utilizadas en la práctica [2].

El problema del controlador consiste en encontrar el control $u(t)$, $t \in [0, T]$, $T = 1,6$, que minimiza el criterio

$$J = \frac{1}{2}E\left[\int_0^T u^2(t)dt + \int_0^T x^2(t)dt\right]. \quad (5.16)$$

En otras palabras, el problema de control consiste en minimizar la energía total del estado x utilizando la mínima energía del control u .

Primero se construirá el controlador donde la ley de control y las matrices $P(t)$ y $Q(t)$ se calculan en la misma manera que para el controlador lineal óptimo para un sistema bilineal con exactamente un parámetro conocido θ , que es $u^*(t) = (R(t))^{-1}B^T(t)Q(t)m(t)$ (ver [44]). Como $B(t) = 1$ en (5.12) y $R(t) = 1$ en (5.16), la ley de control es igual a

$$u(t) = Q(t)m(t); \quad (5.17)$$

donde $m(t)$ satisface la ecuación

$$\begin{aligned} \dot{m}(t) = & a(\theta, t)m(t) + B(t)u(t) + \\ & P(t)A^T(t)G(t)G^T(t))^{-1}(y(t) - (A_0(t) + A(t)m(t))), \end{aligned}$$

$$m(t_0) = m_0 = E(x_0 | F_{t_0}^Y);$$

$Q(t)$ satisface la ecuación de Riccati

$$\begin{aligned} \dot{Q}(t) &= -a^T(\theta, t)Q(t) - Q(t)a(\theta, t) + L(t) - \\ &Q(t)B(t)R^{-1}(t)B^T(t)Q(t), \end{aligned}$$

con la condición final $Q(T) = \Phi$; y $P(t)$ satisface la ecuación de Riccati

$$\begin{aligned} \dot{P}(t) &= P(t)a^T(\theta, t) + a(\theta, t)P(t) + b(t)b^T(t) - \\ &P(t)A^T(t)(G(t)G^T(t))^{-1}A(t)P(t), \end{aligned}$$

con la condición inicial $P(t_0) = E((x_0 - m_0)(x_0 - m_0)^T | y(t_0))$.

Como $t_0 = 0$, $a(t) = 1$, $B(t) = 1$, $b(t) = 0$ en (5.14), $A_0(t) = 0$, $A(t) = 1$, $G(t) = 1$ en (5.15), y $L = 1$ y $\Phi = 0$ en (5.16), las últimas ecuaciones se transforman en

$$\begin{aligned} \dot{m}_1(t) &= m_2(t) + P_{11}(t)(y(t) - m_1(t)), \\ \dot{m}_2(t) &= m_1(t) + P_{12}(t)(y(t) - m_1(t)) + u(t), \end{aligned} \tag{5.18}$$

con $m_1(0) = E(x_{10} | y(0)) = m_{10}$ y $m_2(0) = E(x_{20} | y(0)) = m_{20}$,

$$\begin{aligned} \dot{Q}_{11}(t) &= 1 - 2Q_{12}(t) - Q_{12}^2(t), \\ \dot{Q}_{12}(t) &= -Q_{11}(t) - Q_{22}(t) - Q_{12}(t)Q_{22}(t), \\ \dot{Q}_{22}(t) &= 1 - 2Q_{12}(t) - Q_{22}^2(t), \end{aligned} \tag{5.19}$$

con la condición final $Q(1,6) = 0$.

$$\begin{aligned} \dot{P}_{11}(t) &= 2P_{12}(t) - P_{11}^2(t), \\ \dot{P}_{12}(t) &= P_{11}(t) + P_{22}(t) - P_{11}(t)P_{12}(t), \\ \dot{P}_{22}(t) &= 1 + 2P_{12}(t) - P_{12}^2(t), \end{aligned} \tag{5.20}$$

con la condición inicial $P(0) = E((x_0, \theta_0) - m(0))(x_0, \theta_0) - m(0))^T | y(0)) = P_0$.

Al sustituir el control (5.17) en (5.18), la ecuación del estimado controlado toma la forma

$$\dot{m}_2(t) = m_1(t) + P_{12}(t)(y(t) - m_1(t)) + m_1(t)Q_{12} + m_2(t)Q_{22}. \quad (5.21)$$

Para la simulación numérica del sistema (5.14),(5.15) y el controlador (5.17)–(5.21), se asignaron los siguientes valores iniciales $x_1(0) = x_2(0) = 4,2$, $m_1(0) = m_2(0) = 1$, $P_{12}(0) = 10$ y $P_{11}(0) = P_{22}(0) = 100$ y el valor del parámetro conocido $\theta = 1$. Los disturbios de Poisson $\psi_1(t)$ y $\psi_2(t)$ en (5.14) y (5.15) respectivamente, se realizaron utilizando el diagrama en Simulink sugerido en [5]. Note que el controlador (5.17)–(5.21) es de hecho no realizable, pues el valor del parámetro real $\theta = 1$ es desconocido, y esta simulación se lleva a cabo sólo con propósitos de comparación.

Los resultados de aplicar el controlador (5.17)–(5.21) al sistema(5.14),(5.15) se muestran en la Figura 1, la cual presenta la gráfica del control $u(t)$ obtenido en (5.17), el error de estimación entre las variables de estado (5.14) $x_1(t)$ y $x_2(t)$ y los estimados controlados (5.21) $m_1(t)$ y $m_2(t)$ y la gráfica del criterio (5.16) $J(t)$ en el intervalo $[0, 1,6]$. El criterio (5.16) en el tiempo final $T = 1,6$ toma el valor de $J(1,6) = 70,85$.

Ahora se aplicará el controlador óptimo (5.11)–(5.13), (5.7) para sistemas con parámetros desconocidos para el sistema (5.14), (5.15), asumiendo $\beta(t) = 1$ en (5.3). La ley de control (5.11) toma la forma

$$u^*(t) = Q(t)\hat{x}(t), \quad (5.22)$$

donde

$$\dot{\hat{x}}_1(t) = \hat{x}_2(t) + P_{11}(t)(y(t) - \hat{x}_1(t)), x_1(0) = x_{10}, \quad (5.23)$$

$$\dot{\hat{x}}_2(t) = P_{13}(t) + \hat{x}_1(t)\hat{\theta}(t) + u^*(t) + P_{12}(t)(y(t) - \hat{x}_1(t)), x_2(0) = x_{20}$$

$$\dot{\hat{\theta}}(t) = P_{13}(t)(y(t) - \hat{x}_1(t)), \quad \hat{\theta}(0) = \hat{\theta}_0, \quad (5.24)$$

y

$$\dot{Q}_{11}(t) = 1 - 2\hat{\theta}(t)Q_{12}(t) - Q_{12}^2(t),$$

$$\dot{Q}_{12}(t) = -Q_{11}(t) - \hat{\theta}(t)Q_{22}(t) - Q_{12}(t)Q_{22}(t), \quad (5.25)$$

$$\dot{Q}_{22}(t) = 1 - 2Q_{12}(t) - Q_{22}(t),$$

con la condición final $Q(1,6) = 0$,

$$\dot{P}_{11}(t) = 2P_{12}(t) - P_{11}^2(t),$$

$$\dot{P}_{12}(t) = P_{13}(t) + P_{22}(t) - P_{11}(t)P_{12}(t) + 2P_{11}(t)\hat{\theta}(t), \quad (5.26)$$

$$\dot{P}_{13}(t) = P_{23}(t) - P_{11}(t)P_{13}(t), \quad \dot{P}_{22}(t) = 1 + 2P_{13}(t) + 4\hat{\theta}(t)P_{12}(t) - P_{12}^2(t),$$

$$\dot{P}_{23}(t) = P_{13}(t) - P_{12}(t)P_{13}(t) + 2P_{13}(t)\hat{\theta}(t), \quad \dot{P}_{33}(t) = 1 - P_{13}^2(t),$$

con la condición inicial $P(t_0) = E((z(t_0) - m(t_0))(z(t_0) - m(t_0))^T | F_t^Y)$, and $z(t) = [x_1(t), x_2(t), \theta(t)]$.

Al sustituir el control (5.22) en (5.23), la ecuación del estimado controlado toma la forma

$$\begin{aligned} \dot{\hat{x}}_2(t) = & P_{13}(t) + \hat{x}_1(t)\hat{\theta}(t) + \hat{x}_1(t)Q_{12}(t) + \hat{x}_2(t)Q_{22}(t) + \\ & P_{12}(t)(y(t) - \hat{x}_1(t)) \end{aligned} \quad (5.27)$$

Los estimados obtenidos al resolver las ecuaciones (5.23)–(5.26) también se comparan con los estimados que satisfacen las ecuaciones del controlador convencional polinomial Gaussiano para el estado (5.14) sobre observaciones lineales incompletas (5.14):

$$\begin{aligned}\dot{\hat{x}}_{G1}(t) &= \hat{x}_{G2}(t) + P_{G11}(t)(y(t) - \hat{x}_{G1}(t)), \\ \dot{\hat{x}}_{G2}(t) &= P_{G13}(t) + \hat{x}_{G1}(t)\hat{\theta}(t) + \hat{x}_{G1}(t)Q_{G12}(t) + \\ &\quad \hat{x}_{G2}(t)Q_{G22}(t) + P_{G12}(t)(y(t) - \hat{x}_{G1}(t)),\end{aligned}\tag{5.28}$$

$$\dot{\hat{\theta}}(t) = P_{G13}(t)(y(t) - \hat{x}_{G1}(t)), \quad \hat{\theta}(0) = \hat{\theta}_0,\tag{5.29}$$

y

$$\begin{aligned}\dot{Q}_{G11}(t) &= 1 - 2\hat{\theta}(t)Q_{G12}(t) - Q_{G12}^2(t), \\ \dot{Q}_{G12}(t) &= -Q_{G11}(t) - \hat{\theta}(t)Q_{G22}(t) - Q_{G12}(t)Q_{G22}(t),\end{aligned}\tag{5.30}$$

$$\dot{Q}_{G22}(t) = 1 - 2Q_{G12}(t) - Q_{G22}(t),$$

$$\dot{P}_{G11}(t) = 2P_{G12}(t) - P_{G11}^2(t),$$

$$\dot{P}_{G12}(t) = P_{G22}(t) - P_{G11}(t)P_{G12}(t) + 2P_{G11}(t)\hat{\theta}(t),\tag{5.31}$$

$$\dot{P}_{G13}(t) = P_{G23}(t) - P_{G11}(t)P_{G13}(t), \quad \dot{P}_{G22}(t) = 1 + 4\hat{\theta}(t)P_{G12}(t) - P_{G12}^2(t),$$

$$\dot{P}_{G23}(t) = -P_{G12}(t)P_{G13}(t) + 2P_{G13}(t)\hat{\theta}(t), \quad \dot{P}_{G33}(t) = 1 - P_{G13}^2(t),$$

Para cada uno de los controladores (5.23)–(5.26) y (5.28)–(5.31) y el sistema de referencia (5.14), (5.15) involucrados en la simulación, se asignaron los siguientes valores iniciales: $x_1(0) = x_2(0) = 4,2$, $\hat{x}_1(0) = \hat{x}_2(0) = 1$, $\hat{\theta}(0) = 0,535$, $P_{12}(0) = P_{13}(0) = P_{23}(0) = 10$ y $P_{11}(0) = P_{22}(0) = P_{33}(0) = 100$. El parámetro real se asigna como $\theta = 1$. Los ruidos

de Poisson $\psi_1(t)$ y $\psi_2(t)$ en (5.14) y (5.15) respectivamente, se generaron utilizando el diagrama en Simulink sugerido en [5].

Los resultados al aplicar el controlador (5.11)–(5.13),(5.7) al sistema (5.14), (5.15) usando el controlador diseñado (5.23)–(5.26) se muestran en la Figura 2, la cual presenta la gráfica del control $u(t)$ obtenido en (5.22), los errores de estimación entre las variables reales del estado (5.14) $x_1(t)$ y $x_2(t)$, y los estimados controlados (5.23) $\hat{x}_1(t)$ y $\hat{x}_2(t)$, la gráfica del estimado del parámetro $\hat{\theta}(t)$ (5.24), y la gráfica del criterio (5.16) $J(t)$ en el intervalo $[0, 1,6]$

Los valores del estimado del parámetro $\hat{\theta}(t)$, y el criterio (5.16) en el tiempo final $T = 1,6$ son $\hat{\theta}(1,6) = 0,99$ y $J(1,6) = 73,38$.

Por otro lado, los resultados de aplicar el controlador (5.11)–(5.13), (5.7) al sistema (5.14), (5.15) usando el controlador convencional diseñado para sistemas con ruidos Gaussianos (5.28)–(5.31) se muestran en la Figura 3, la cual presenta la gráfica del control $u(t)$ obtenido en (5.22), los errores de estimación entre las variables reales del estado (5.14) $x_1(t)$ y $x_2(t)$, y los estimados controlados (5.28) $\hat{x}_{G1}(t)$ y $\hat{x}_{G2}(t)$, la gráfica del estimado del parámetro $\hat{\theta}(t)$ (5.29), y la gráfica del criterio (5.16) $J(t)$ en el intervalo $[0, 1,6]$. Los valores del estimado del parámetro $\hat{\theta}(t)$, y el criterio (5.16) en el tiempo final $T = 1,6$ son $\hat{\theta}(1,6) = 1,01$ y $J(1,6) = 78,14$.

Aunque el valor final del estimado del parámetro en ambos casos es muy cercano al valor real $\theta = 1$, en el caso del controlador diseñado, el valor del criterio J tiene un error de sólo el 3,5% del valor real, en contra del controlador convencional usado para sistemas con ruidos blancos Gaussianos el cual proporciona un error del 10,29% del valor real del criterio para sistemas con un valor de parámetro conocido. Esto verifica exitosamente el rendimiento general para sistemas con parámetros desconocidos y ruidos de Poisson.

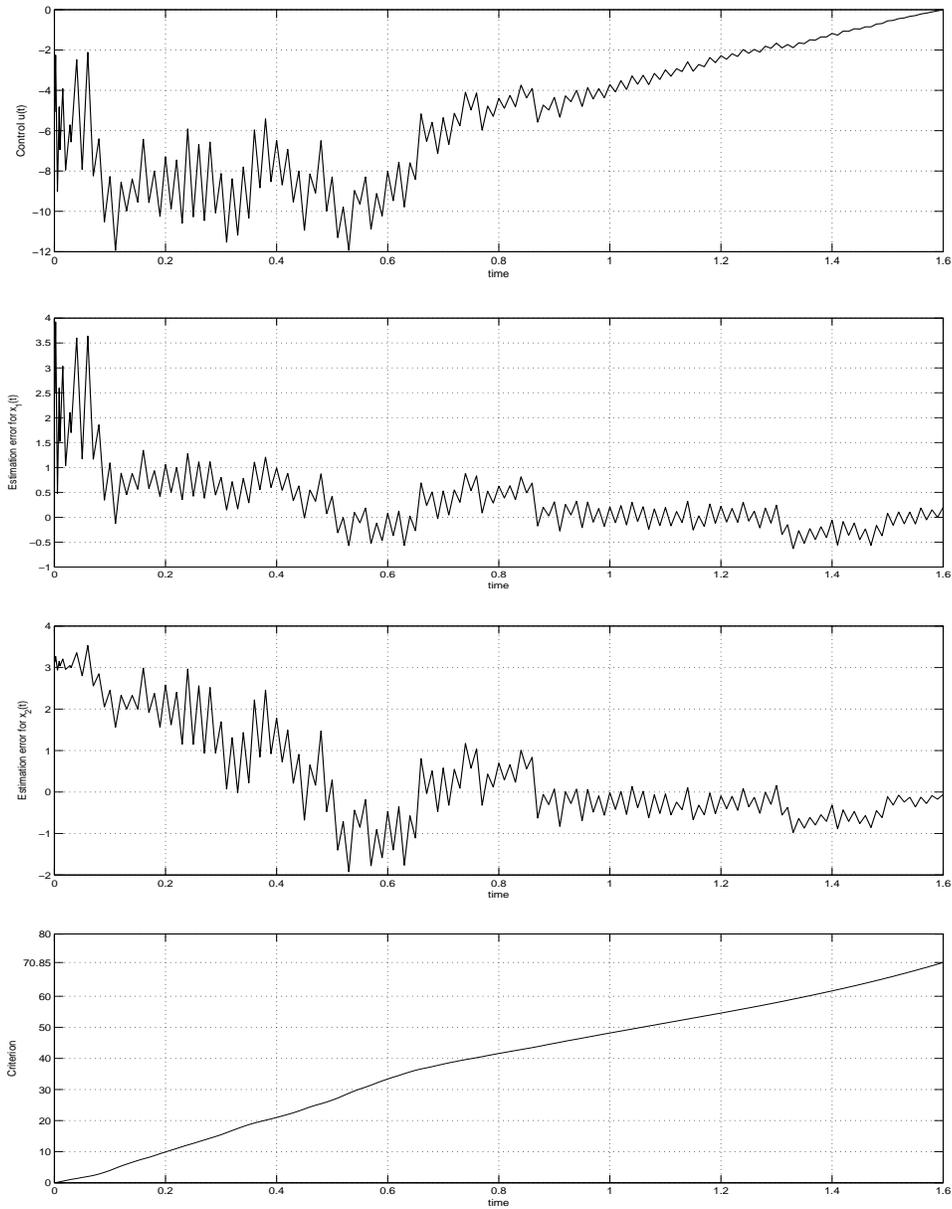


Figura 5.1: Gráfica del control $u(t)$ (5.17), el error entre las variables de estado reales $x_1(t)$ y $x_2(t)$, que satisfacen (5.14), y los estimados óptimos controlados $m_1(t)$ y $m_2(t)$ que satisfacen (5.18), y gráfica del criterio (5.16) $J(t)$ en el intervalo $[0, 1,6]$.

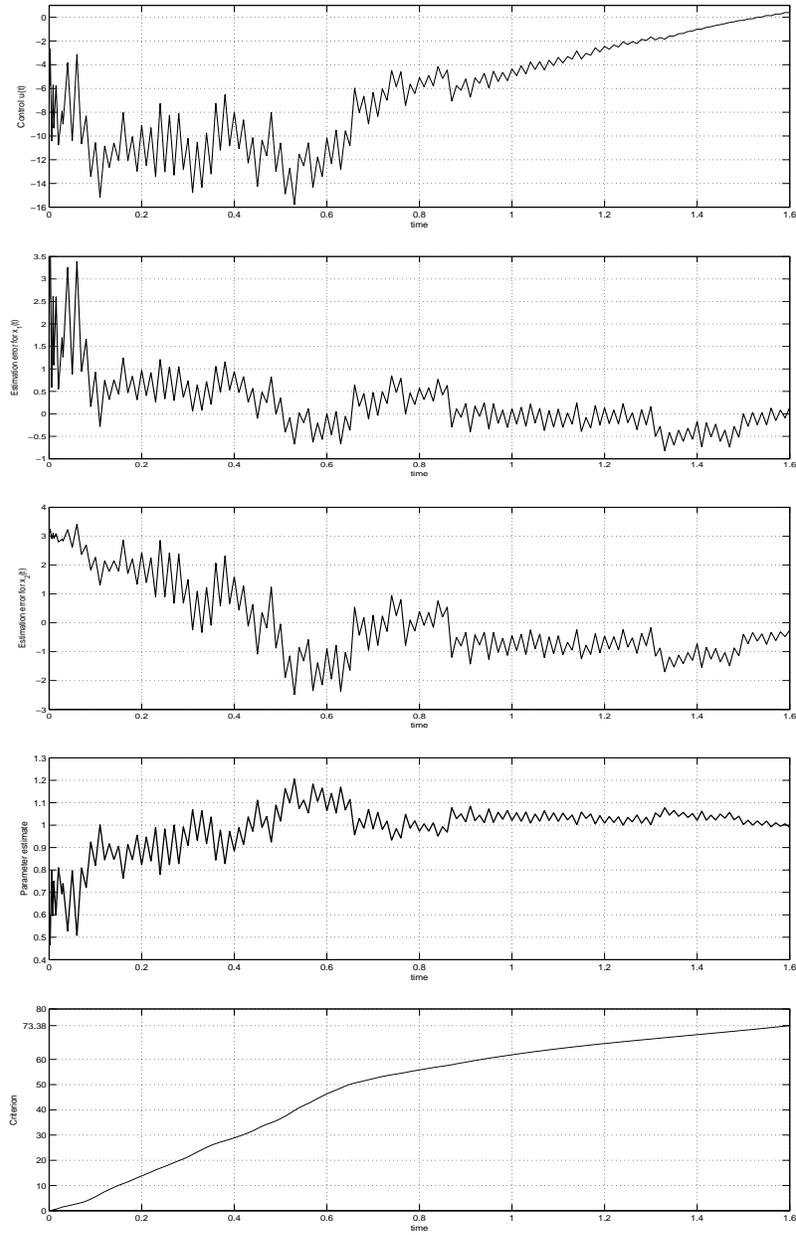


Figura 5.2: Gráfica del control $u(t)$ (5.22), el error entre las variables de estado reales $x_1(t)$ y $x_2(t)$, que satisfacen (5.14), y los estimados óptimos controlados $\hat{x}_1(t)$ y $\hat{x}_2(t)$ que satisfacen (5.23)–(5.26), la gráfica del estimado del parámetro $\hat{\theta}$ que satisface (5.24), y la gráfica del criterio (Ec. (5.16)) $J(t)$ en el intervalo $[0, 1,6]$.

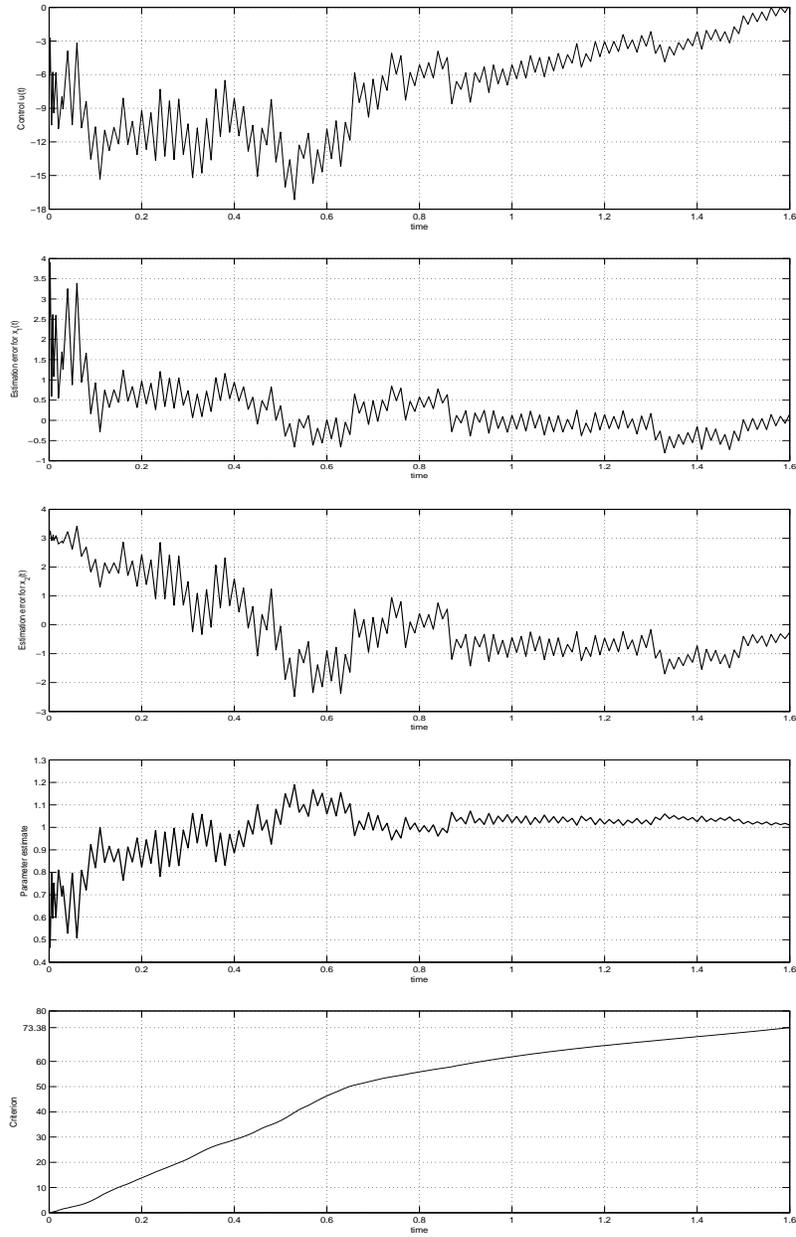


Figura 5.3: Gráfica del control $u(t)$ (5.22), el error entre las variables de estado reales $x_1(t)$ y $x_2(t)$, que satisfacen (14), y los estimados controlados $\hat{x}_{G1}(t)$ y $\hat{x}_{G2}(t)$ que satisfacen (5.28)–(5.31), la gráfica del estimado del parámetro $\hat{\theta}$ satisfying (5.29) , y la gráfica del criterio (Ec. (16)) $J(t)$ en el intervalo $[0, 1,6]$

Capítulo 6

Filtro Aproximado de Dimensión Finita para Estados Polinomiales con Ruido de Poisson sobre Observaciones Polinomiales

6.1. Planteamiento del Problema

Sea (Ω, F, P) un espacio completo de probabilidad con una familia de σ -álgebras $F_t, t \geq t_0$ creciente y continua por la derecha, y sean $(N_1(t), F_t, t \geq t_0)$ y $(N_2(t), F_t, t \geq t_0)$ dos procesos independientes de Poisson centralizados. El proceso aleatorio F_t -medible $(x(t), y(t))$ es descrito por una ecuación diferencial estocástica no lineal con un término drift polinomial para el estado del sistema

$$dx(t) = \rho(x, t)dt + \sigma(x, t)dN_1(t), \quad x(t_0) = x_0, \quad (6.1)$$

y el proceso de observación

$$dy(t) = h(x, t)dt + B(t)dN_2(t), \quad (6.2)$$

donde, $x(t) \in R^n$ es el vector de estado y $y(t) \in R^m$ es el vector de observación. La condición inicial $x_0 \in R^n$ es un vector de Poisson tal que x_0 , $N_1(t) \in R^p$, y $N_2(t) \in R^q$ son independientes. Se asume que $B(t)B^T(t)$ es una matriz definida positiva, por lo tanto $m \leq q$. Todos los coeficientes en (6.1)–(6.2) son funciones determinísticas de dimensiones apropiadas. Las soluciones de las ecuaciones diferenciales estocásticas de Ito (6.1) y (6.2) se consideran como soluciones débiles (ver, por ejemplo, [51] para definición). La misma definición se mantiene para las soluciones para las otras ecuaciones diferenciales estocásticas de Ito durante todo el capítulo.

Las funciones no lineales $\rho(x, t) \in R^n$, $\sigma(x, t) \in R^n$, y $h(x, t) \in R^m$ se consideran polinomios de n variables, donde las componentes del vector de estado $x(t) \in R^n$, son coeficientes que dependen del tiempo. Como $x(t) \in R^n$ es un vector, esto requiere una definición especial del polinomio para $n > 1$. De acuerdo con [4], un polinomio de grado p del vector $x(t) \in R^n$ se considera como una forma lineal p de n componentes de $x(t)$

$$\rho(x, t) = \alpha_0(t) + \alpha_1(t)x + \alpha_2(t)xx^T + \dots + \alpha_p(t)x \dots_p \text{ veces} \dots x, \quad (6.3)$$

donde $\alpha_0(t)$ es un vector de dimensión n , α_1 es una matriz de dimensión $n \times n$, α_2 es un tensor de dimensión $n \times n \times n$, α_p es un $(p+1)$ D tensor de dimensión $n \times \dots_{(p+1) \text{ veces}} \dots \times n$,

y $x \times \dots \times_p \text{ veces} \dots \times x$ es un p D tensor de dimensión $n \times \dots \times_p \text{ veces} \dots \times n$ obtenida de multiplicar el vector $x(t)$, p veces por sí mismo (ver definición [4]). Tal polinomio puede ser expresado en la siguiente forma de sumatoria

$$\begin{aligned} \rho_k(x, t) &= \alpha_0 \text{ }_k(t) + \sum_i \alpha_1 \text{ }_{ki}(t) x_i(t) + \sum_{ij} \alpha_2 \text{ }_{kij}(t) x_i(t) x_j(t) + \dots \\ &+ \sum_{i_1 \dots i_p} \alpha_p \text{ }_{ki_1 \dots i_p}(t) x_{i_1}(t) \dots x_{i_p}(t), \quad k, i, j, i_1, \dots, i_p = 1, \dots, n. \end{aligned}$$

El problema de estimación consiste en encontrar el estimado en promedio cuadrático $\hat{x}(t)$ del estado del sistema $x(t)$, basado en el proceso de observación $Y(t) = \{y(s), 0 \leq s \leq t\}$, que minimice el valor esperado condicional para la norma Euclidiana

$$J = E[(x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t)) \mid F_t^Y]$$

para cada momento del tiempo t . Donde, $E[\xi(t) \mid F_t^Y]$ representa el valor esperado condicional del proceso estocástico $\xi(t) = (x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t))$ 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 [53], el estimado en promedio cuadrático está dado por el valor esperado condicional

$$\hat{x}(t) = m_x(t) = E(x(t) \mid F_t^Y)$$

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

$$P(t) = E[(x(t) - m_x(t))(x(t) - m_x(t))^T \mid F_t^Y]$$

es la matriz de covarianza del error de estimación. De aquí en adelante, el problema de filtrado formulado se considerará en un intervalo de tiempo $[t_0, T_1]$, donde la solución de la ecuación de estado (6.1) existe y es casi acotada. Aparentemente, $T_1 < T^*$, donde T^*

es un tiempo de escape para el sistema (6.1). Note que dado que la condición inicial x_0 es una variable aleatoria de Poisson, los momentos condicionales del proceso $x(t)$ permanecen acotados para cualquier $t \leq T_1$ ([53]).

La solución propuesta para el problema de filtrado planteado se basa en las fórmulas para las diferenciales Ito del estimado en promedio cuadrático y la varianza del error de estimación (citado después de [53]), la cual se dará en la siguiente sección.

6.2. Diseño del Filtro

El problema de filtrado es resultado en el siguiente teorema.

Teorema 1. *El filtro en promedio cuadrático para el estado polinomial $x(t)$ (6.1) sobre observaciones polinomiales $y(t)$ (6.2) está dado por las siguientes ecuaciones para el estimado en promedio cuadrático $m(t) = [m_z(t), m_x(t)] = E([z(t), x(t)] | F_t^Y)$ y la matriz de covarianza del error de estimación $P(t) = E([(z(t), x(t)) - m(t))((z(t), x(t)) - m(t))^T | F_t^Y]$:*

$$dm(t) = E(\bar{f}(x, t) | F_t^Y)dt + P(t)[I, 0]^T (B(t)B^T(t))^{-1}(dy(t) - m_z(t)dt), \quad (6.4)$$

$$dP(t) = (E([(z(t), x(t)) - m(t)](\bar{f}(x, t))^T | F_t^Y] + E(\bar{f}(x, t)([z(t), x(t)] - m(t))^T | F_t^Y))dt + (6.5)$$

$$E(\bar{g}(x, t)\bar{g}^T(x, t) | F_t^Y) - P(t)[I, 0]^T (B(t)B^T(t))^{-1}[I, 0]P(t),$$

con las condiciones iniciales $m(t_0) = [m_z(t_0), m_x(t_0)] = E([z_0, x_0] | F_{t_0}^Y)$ and $P(t_0) = E([(z_0, x_0) - m(t_0)]([z_0, x_0] - m(t_0))^T | F_{t_0}^Y)$. Aquí, $\bar{f}(x, t) = [f(x, t), \rho(x, t)]$, $\bar{g}(x, t) =$

$[g(x, t), \sigma(x, t)]^T$,

$$f(x, t) = \frac{\partial h(x, t)}{\partial x} \rho(x, t) dt + \frac{\partial h(x, t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 h(x, t)}{\partial x^2} \sigma(x, t) \sigma^T(x, t) dt, \quad g(x, t) = \frac{\partial h(x, t)}{\partial x} \sigma(x, t),$$

y el estado polinomial adicional $z(t) = h(x, t)$ satisface la ecuación

$$dz(t) = \frac{\partial h(x, t)}{\partial x} \rho(x, t) dt + \frac{\partial h(x, t)}{\partial t} dt + \tag{6.6}$$

$$\frac{1}{2} \frac{\partial^2 h(x, t)}{\partial x^2} \sigma(x, t) \sigma^T(x, t) dt + \frac{\partial h(x, t)}{\partial x} \sigma(x, t) dN_1(t), \quad z(0) = z_0.$$

Si la condición inicial $[z_0, x_0]$ para el vector de estado extendido es condicionalmente de Poisson con respecto a las observaciones, el sistema de las ecuaciones de filtrado (6.4), (6.5) se convierte en un sistema de dimensión finita en forma cerrada después de expresar los momentos condicionales de orden superior del estado del sistema $x(t)$ con respecto a las observaciones $y(t)$ como funciones sólomente de los primeros dos momentos condicionales, $m(t)$ y $P(t)$.

Demostración. Se reformulará el problema, introduciendo el proceso estocástico $z(t) = h(x, t)$. Usando la fórmula Ito (ver [53]) para la diferencial estocástica de la función no lineal $h(x, t)$, donde $x(t)$ satisface la ecuación (6.1), la ecuación (6.6) es obtenida para $z(t)$

$$dz(t) = \frac{\partial h(x, t)}{\partial x} \rho(x, t) dt + \frac{\partial h(x, t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 h(x, t)}{\partial x^2} \sigma(x, t) \sigma^T(x, t) dt + \frac{\partial h(x, t)}{\partial x} \sigma(x, t) dN_1(t), \quad z(0) = z_0.$$

Note que la expresión $\frac{1}{2} \frac{\partial^2 h(x, t)}{\partial x^2} \sigma(x, t) \sigma^T(x, t)$ aparece en vista de la segunda derivada en x en la fórmula Ito.

Se asume que la condición inicial $[z_0, x_0]$ para el vector de estado extendido es un vector aleatorio condicionalmente de Poisson con respecto a las observaciones. Esta suposición es admisible en el área de filtrado, pues las distribuciones reales de $x(t)$ y $z(t)$ son desconocidas. Incluso, como se ve en [52], si solamente están disponibles los primeros dos momentos condicionales, la esperanza m_0 y la varianza P_0 , de un vector aleatorio $[z_0, x_0]$, la mejor aproximación para la distribución condicional desconocida de $[z_0, x_0]$ con respecto a las observaciones, es la distribución de Poisson con los mismos parámetros, m_0 y P_0 . Esto es también un corolario del teorema del límite central [59] en la teoría de probabilidad.

Un punto clave para derivaciones adicionales es que el lado derecho de la ecuación (6.6) es un polinomio en x . De hecho, como $h(x, t)$ es un polinomio en x , las funciones $\frac{\partial h(x, t)}{\partial x}$, $\frac{\partial h(x, t)}{\partial x}x(t)$, $\frac{\partial h(x, t)}{\partial t}$, y $\frac{\partial^2 h(x, t)}{\partial x^2}$ son también polinomios en x . Así, la ecuación (6.6) es una ecuación de estado polinomial con ruido multiplicativo polinomial. Puede escribirse en la forma compacta

$$dz(t) = f(x, t)dt + g(x, t)dN_1(t), \quad z(t_0) = z_0, \quad (6.7)$$

donde

$$f(x, t) = \frac{\partial h(x, t)}{\partial x} \rho(x, t)dt + \frac{\partial h(x, t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 h(x, t)}{\partial x^2} \sigma(x, t) \sigma^T(x, t) dt, \quad g(x, t) = \frac{\partial h(x, t)}{\partial x} \sigma(x, t).$$

En términos del proceso $z(t)$, la ecuación de observación (6.2) toma la forma

$$dy(t) = [I, 0][z(t), x(t)]dt + B(t)dN_2(t). \quad (6.8)$$

El problema de estimación reformulado consiste ahora en encontrar el estimado en promedio cuadrático $[m_z(t), m_x(t)]$ del estado del sistema $[z(t), x(t)]$, basado en el proceso de observación $Y(t) = \{y(s), 0 \leq s \leq t\}$. Este estimado en promedio cuadrático está dado por el valor esperado condicional

$$m(t) = [m_z(t), m_x(t)] = [E(z(t) | F_t^Y), E(x(t) | F_t^Y)]$$

del estado del sistema $[z(t), x(t)]$ con respecto a la σ - álgebra F_t^Y generada por el proceso de observación $Y(t)$ en el intervalo $[t_0, t]$. La función matricial

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

es la matriz de covarianza del error de estimación para este problema reformulado.

El sistema de filtrado obtenido incluye las dos ecuaciones, (6.6) (o (6.7)) y (6.1), para el estado parcialmente medible $[z(t), x(t)]$ y la ecuación (6.8) para las observaciones $y(t)$, donde $z(t)$ es un estado polinomial completamente medible con un ruido multiplicativo polinomial, $x(t)$ es un estado polinomial no medible, y $y(t)$ es un proceso de observación lineal directamente midiendo al estado $z(t)$. Aplicando el filtro en promedio cuadrático para estados polinomiales incompletamente medibles con un ruido multiplicativo polinomial sobre observaciones lineales (ver [7, 16]) para el sistema (6.7),(6.1),(6.8) se producen las ecuaciones de filtrado deseadas (6.4),(6.5). Finalmente, después de representar los momentos condicionales de orden superior del estado del sistema como funciones de el valor esperado $m(t)$ y la varianza del error $P(t)$ usando la propiedad de una variable aleatoria de Poisson $x(t) - m(t)$ de representar los momentos condicionales superiores del estado del sistema como funciones de la varianza $P(t)$, (ver [4, 7, 11] para detalles), se puede obtener un sistema de dimensión finita de las ecuaciones de filtrado, cerrado con respecto a $m(t)$ y $P(t)$, si la condición inicial $[z_0, x_0]$ para el vector de estado extendido es condicionalmente de Poisson. ■

Nota. Observe que algunos casos particulares del Teorema 1, como sistemas lineales o bilineales con ruidos dependientes, fueron previamente considerados en [4, 7, 16], donde fueron obtenidas las ecuaciones de filtrado en promedio cuadrático de dimensión finita explícitas. Por otro lado, el resultado general del Teorema 1, permite diseñar un filtro sub-óptimo en promedio cuadrático de dimensión finita para cualquier estado polinomial con ruido blanco de Poisson sobre observaciones polinomiales. Más aún, como cualquier función no lineal puede ser aproximada por un polinomio de cierto grado con cierta precisión, el resultado obtenido en el Teorema 1 es potencialmente útil para tratar el diseño de un filtro sub-óptimo en promedio cuadrático de dimensión finita para cualquier estado no lineal con ruido blanco de Poisson sobre observaciones con un término drift no lineal.

En la siguiente sección se presenta un ejemplo, donde se obtendrán las ecuaciones de filtrado en forma cerrada para un caso particular de funciones polinomiales escalares de segundo y tercer orden $\rho(x, t)$, $\sigma(x, t)$ y $h(x, t)$ en las ecuaciones (6.1) y (6.2). Así, la aplicación del mismo procedimiento resultaría en el diseño de un sistema cerrado de las ecuaciones de filtrado para cualesquier funciones polinomiales $\rho(x, t)$, $\sigma(x, t)$, y $h(x, t)$ en (6.1),(6.2).

6.3. Ejemplo: Problema de Filtrado para el Sensor de Tercer Grado para un Sistema Cuadrático

Esta sección presenta un ejemplo del diseño de un filtro de dimensión finita en forma cerrada para un estado cuadrático sobre observaciones polinomiales de tercer grado, donde la condición inicial para el vector de estado extendido se asume condicionalmente de Poisson con respecto a las observaciones.

Sean $x(t)$ un estado escalar no medible que satisface la ecuación cuadrática

$$dx(t) = x^2(t)dt + dn_1(t), \quad x(0) = x_0, \quad (6.9)$$

y $y(t)$ el proceso de observación dado por la ecuación escalar del sensor de tercer grado

$$dy(t) = x^3(t)dt + dn_2(t), \quad (6.10)$$

donde $n_1(t)$ y $n_2(t)$ son procesos de Poisson independientes entre ellos y de una variable aleatoria de Poisson x_0 que servirá como condición inicial en (6.9). El problema de filtrado consiste en encontrar el estimado en promedio cuadrático para el estado cuadrático(6.9), usando las observaciones del sensor de tercer grado (6.10).

El problema será reformulado, introduciendo el proceso estocástico $z(t) = h(x, t) = x^3(t)$. Utilizando la fórmula Ito (ver [53]) para la diferencial estocástica de la función cúbica $h(x, t) = x^3(t)$, donde $x(t)$ satisface la ecuación (6.9), se obtienen las siguientes ecuaciones para $z(t)$

$$dz(t) = (3x(t) + 3x^4(t))dt + 3x^2(t)dn_1(t), \quad z(0) = z_0.$$

Tomando en cuenta que $z(t) = x^3(t)$, la última ecuación toma la forma

$$dz(t) = (3x(t)(1 + z(t)))dt + 3x^2(t)dn_1(t), \quad z(0) = z_0. \quad (6.11)$$

Aquí, $\frac{\partial h(x,t)}{\partial x} = 3x^2(t)$, $\frac{1}{2}\frac{\partial h^2(x,t)}{\partial x^2} = 3x(t)$, y $\frac{\partial h(x,t)}{\partial t} = 0$, dado que $h(x, t)$ no depende de t , entonces, $f(x, t) = 3x(t) + 3x^4(t) = 3x(t)(1 + z(x))$ y $g(x, t) = 3x^2(t)$. La condición inicial $[z_0, x_0]$ se considera condicionalmente de Poisson con respecto a las observaciones (ver el párrafo anterior a (6.7) para detalles). En términos del proceso $z(t)$, la ecuación de observación (6.10) toma la forma

$$dy(t) = z(t)dt + dn_2(t). \quad (6.12)$$

El sistema de filtrado obtenido incluye dos ecuaciones, (6.11) y (6.9), para el estado parcialmente medible $[z(t), x(t)]$ y una ecuación (6.12) para las observaciones $y(t)$, donde $z(t)$ es un estado de segundo grado completamente medible con un ruido multiplicativo cuadrático, $x(t)$ es un estado cuadrático no medible, y $y(t)$ es un proceso de observación lineal que mide directamente al estado $z(t)$. Por lo tanto, el filtro diseñado en promedio cuadrático puede ser aplicado para resolver este problema. Las ecuaciones de filtrado(6.4),(6.5) toman la forma particular para el sistema (6.11),(6.9),(6.12)

$$dm_1(t) = (3m_2(t) + 3m_1(t)m_2(t) + 3P_{12}(t))dt + P_{11}(t)[dy(t) - m_1(t)dt], \quad (6.13)$$

$$dm_2(t) = (m_2^2(t) + P_{22}(t))dt + P_{12}(t)[dy(t) - m_1(t)dt], \quad (6.14)$$

con las condiciones iniciales $m_1(0) = E(z_0 | y(0)) = m_{10}$ y $m_2(0) = E(x_0 | y(0)) = m_{20}$,

$$\dot{P}_{11}(t) = 12P_{12}(t) + 9m_2^4(t) + 9P_{22}(t) + 36P_{22}(t)m_2(t) + 54P_{22}(t)m_2^2(t) + \quad (6.15)$$

$$27P_{22}^2(t) + 12P_{11}(t)m_2(t) - P_{11}^2(t),$$

$$\dot{P}_{12}(t) = 3P_{12}(t) + 7P_{22}(t) + 8P_{12}(t)m_2(t) + 3m_2^2(t) - P_{11}(t)P_{12}(t), \quad (6.16)$$

$$\dot{P}_{22}(t) = 1 + 2P_{22}(t) + 4P_{22}(t)m_2(t) - P_{12}^2(t), \quad (6.17)$$

con la condición inicial $P(0) = E((z_0, x_0)^T - m(0))(z_0, x_0)^T - m(0))^T | y(0)) = P_0$. Aquí, $m_1(t)$ es el estimado para el estado $z(t) = x^3(t)$ y $m_2(t)$ es el estimado para el estado $x(t)$.

Los estimados obtenidos al resolver las ecuaciones (6.13)–(6.17) se comparan con los estimados que satisfacen las ecuaciones de filtrado diseñadas para sistemas con ruido blanco Gaussiano para el estado cuadrático (6.11) sobre las observaciones polinomiales de tercer orden (6.10), los cuales son obtenidos al aplicar el Teorema 8.1 en [35]:

$$dm_{G1}(t) = (3m_{G2}(t) + 3m_{G1}(t)m_{G2}(t) + 3P_{G12}(t))dt + P_{G11}(t)[dy(t) - m_{G1}(t)dt], \quad (6.18)$$

$$dm_{G2}(t) = (m_{G2}^2(t) + P_{G22}(t))dt + P_{G12}(t)[dy(t) - m_{G1}(t)dt], \quad (6.19)$$

con las condiciones iniciales $m_{G1}(0) = E(z_0 | y(0)) = m_{G10}$ y $m_{G2}(0) = E(x_0 | y(0)) = m_{G20}$,

$$\dot{P}_{G11}(t) = 6P_{G12}(t) + 9m_{G2}^4(t) + 9P_{G22}(t) + 36P_{G22}(t)m_{G2}(t) + 54P_{G22}(t)m_{G2}^2(t) + (6.20)$$

$$27P_{G22}^2(t) + 12P_{G11}(t)m_{G2}(t) - P_{G11}^2(t),$$

$$\dot{P}_{G12}(t) = 6P_{G22}(t) + 8P_{G12}(t)m_{G2}(t) + 3m_{G2}^2(t) - P_{G11}(t)P_{G12}(t), \quad (6.21)$$

$$\dot{P}_{G22}(t) = 1 + 4P_{G22}(t)m_{G2}(t) - P_{G12}^2(t), \quad (6.22)$$

con la condición inicial $P_G(0) = E((z_0, x_0)^T - m_G(0))(z_0, x_0)^T - m_G(0))^T | y(0)) = P_{G0}$.

Los resultados de la simulación numérica se obtienen al resolver los sistemas de las ecuaciones de filtrado (6.13)–(6.17) y (6.18)–(6.22).

Los valores obtenidos de los estimados del estado $m_2(t)$, que satisfacen la ecuación (6.14), y $m_{G_2}(t)$, que satisface la ecuación (6.19), se comparan con los valores reales de la variable de estado $x(t)$ en (6.9).

Para los filtros (6.13)–(6.17), (6.18)–(6.22) y el sistema de referencia(6.11),(6.9),(6.12) envueltos en simulación, se asignan los siguientes valores iniciales: $x_0 = 0$, $z_0 = 18$, $m_{10} = m_{G_{10}} = 1000$, $m_{20} = m_{G_{20}} = 10$, $P_{11}(0) = P_{G_{11}}(0) = 15$, $P_{12}(0) = P_{G_{12}}(0) = 3$, $P_{22}(0) = P_{G_{22}}(0) = 1$. Los ruidos de Poisson $dn_1(t)$ y $dn_2(t)$ fueron generados utilizando el diagrama de simulación sugerido en [5]. Note que el poder del ruido puede ser cambiado variando los términos $\sigma(x, t)$ y $B(t)$ en (6.1),(6.2), lo que llevaría a cambiar los correspondiente términos de las ecuaciones de filtrado (6.4),(6.5). El intervalo de simulación se establece como $[0, 4,0575]$, pues el error de estimación dado por el filtro diseñado para sistemas con ruido blanco Gaussiano diverge a infinito en este tiempo.

La figura 1 muestra las gráficas de los errores entre el estado de referencia $x(t)$ (6.9) y su estimado $m_2(t)$ (6.14), y el estado de referencia $z(t) = x^3(t)$ (6.11) y su estimado $m_1(t)$ (6.13), en todo el intervalo de simulación de $t_0 = 0$ a $T = 4,0575$. Se puede observar que los errores de estimación convergen a los estados reales muy rápidamente y después mantienen el valor promedio de cero, a pesar de un error considerable en las condiciones iniciales, $m_{20} - x_0 = 10$, $m_{10} - z_0 = 982$. El error de estimación del estado $x(t)$ al tiempo final $T = 4,0575$ es igual a $m_2(4,0575) - x(4,0575) = 0,1646$. La Figura 2 muestra la gráfica de los errores entre el estado de referencia $x(t)$ (6.9) y el estimado $m_{G_2}(t)$ (6.19), y el estado de referencia $z(t) = x^3(t)$ (6.11) y su estimado $m_{G_1}(t)$ (6.18), en el intervalo de simulación de $t_0 = 0$ a $T = 4,0575$. Note que aunque las ecuaciones del sistema (6.13)–(6.14) coinciden con el sistema (6.18)–(6.19), el filtro diseñado muestra un completo rendimiento para cada momento del tiempo t , mientras que el error de estimación del estimado $m_{G_2}(t)$ del filtro diseñado para sistemas con ruido blanco Gaussiano diverge a infinito al tiempo $T = 4,0575$

Así, se puede concluir que el filtro obtenido (6.13)–(6.17) resuelve el problema de filtrado del sensor de tercer orden para el sistema (6.9),(6.10), donde el estado y las observaciones son polinomiales y ambas contienen ruido blanco de Poisson, y proporciona un estimado confiable del estado no medible.

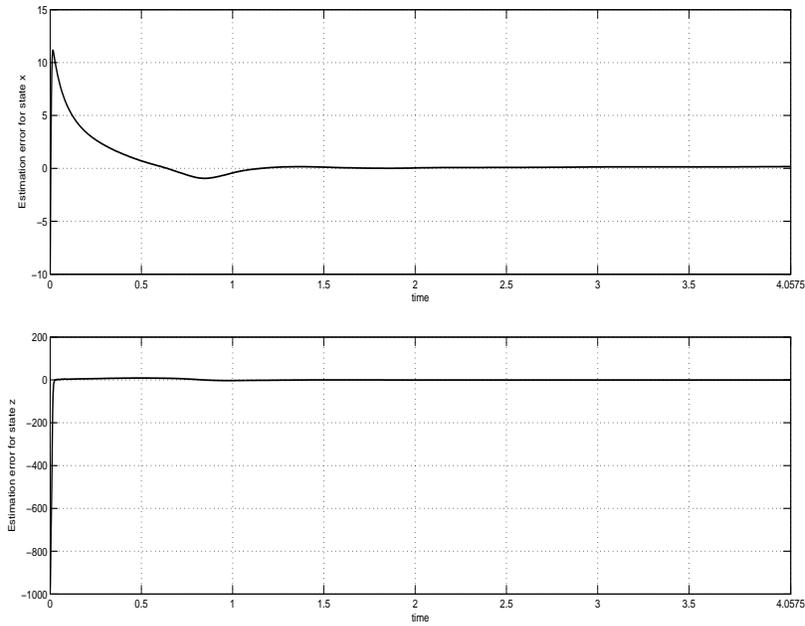


Figura 6.1: **Arriba.** Gráfica del error de estimación entre el estado de referencia $x(t)$ (15) y su estimado óptimo $m_2(t)$ (6.20) en el intervalo $[0, 4,0575]$. **Abajo.** Gráfica del error de estimación entre el estado de referencia $z(t)$ (6.17) y su estimado óptimo $m_1(t)$ (6.19) en el intervalo $[0, 4,0575]$.

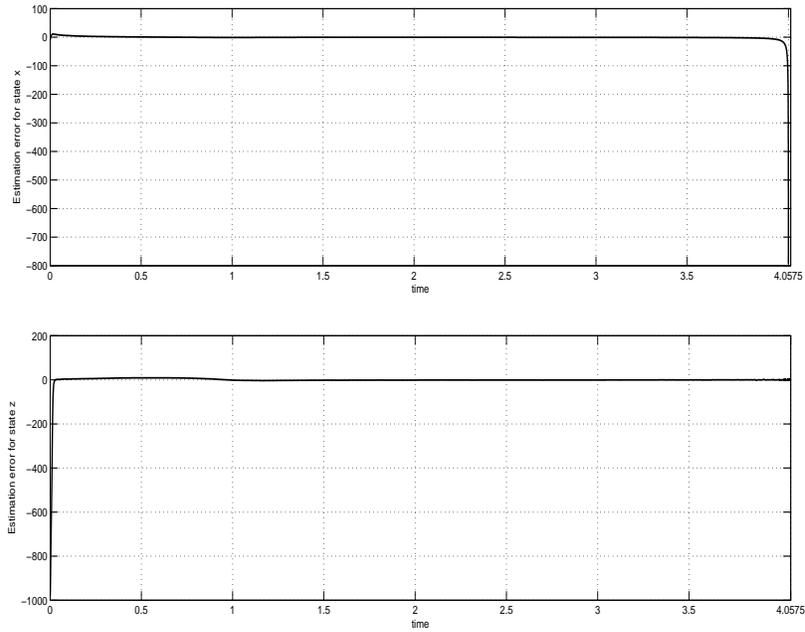


Figura 6.2: **Arriba.** Gráfica del error de estimación entre el estado de referencia $x(t)$ (6.15) y su estimado $m_{G_2}(t)$ (6.25) en el intervalo $[0, 4,0575]$. **Abajo.** Gráfica del error de estimación entre el estado de referencia $z(t)$ (6.17) y el estimado $m_{G_2}(t)$ (6.24) en el intervalo $[0, 4,0575]$.

Capítulo 7

Conclusiones y Trabajos Futuros

7.1. Conclusiones

En esta tesis se ha realizado un estudio comparativo entre dos algoritmos diseñados óptimamente, uno en contexto Gaussiano y el que se ha diseñado en este trabajo en contexto de Poisson. Este estudio se ha realizado sobre sistemas dinámicos con presencia de ruido blanco de Poisson tanto en el estado como en las observaciones.

Se han estudiado distintas técnicas, casos con estados no observables, parámetros desconocidos, sistemas estables e inestables. Los puntos concretos de interés en que se ha centrado el análisis han sido:

- o Estudiar qué algoritmos proporcionan mejores resultados en el caso de tener estados no observables en presencia de ruido de Poisson, medidos en relación a un criterio de costo definido.
- o Estudiar qué algoritmos proporcionan mejores resultados en el caso de tener variables de estado no observables y parámetros desconocidos.

- Estudiar qué algoritmos proporcionan un mejor controlador para sistemas lineales en presencia de ruido de Poisson, medidos en relación a un criterio de costo definido.
- Estudiar qué algoritmos proporcionan mejores resultados en el caso de tener observaciones polinomiales en presencia de ruido de Poisson, medidos en relación a un criterio de costo definido.

Los resultados y conclusiones que se obtuvieron tras la aplicación de los algoritmos diseñados en esta tesis pueden resumirse en los siguientes puntos:

1. Es necesaria la aplicación de un filtro óptimo para estados incompletamente medibles en presencia de ruidos de Poisson, teniendo en cuenta la naturaleza del ruido, pues los filtros convencionales no presentan un buen rendimiento.
2. Los filtros e identificadores Gaussianos para sistemas con ruidos de Poisson tienen un muy mal rendimiento, por lo que es necesario la aplicación de los filtros e identificadores en contexto de Poisson.
3. El controlador diseñado para sistemas lineales con parámetros desconocidos y ruido de Poisson mejoró el rendimiento del controlador convencional Gaussiano.
4. El filtro aproximado de dimensión finita diseñado para sistemas no lineales sobre observaciones polinomiales mostró una gran ventaja sobre los filtros convencionales, donde estos últimos no mostraron un buen desempeño en el caso planteado.

En general, el diseño de algoritmos de filtrado para sistemas en presencia de ruido de Poisson, merece ser estudiado consistentemente, pues a través de los resultados puede observarse que hay diferencias significativas entre tomar en cuenta las consideraciones de Poisson y hacer aproximaciones a una distribución Gaussiana.

7.2. Trabajos Futuros

De los resultados obtenidos se abre un conjunto amplio de posibles líneas de investigación. Se citan a continuación algunas de interés en relación con el presente trabajo:

- Aplicar los resultados obtenidos en este trabajo de investigación a áreas como la fotografía, centellografía, telecomunicaciones, etc.
- Ampliar el procedimiento de estimación, a través del diseño de filtros para sistemas polinomiales en presencia de ruido blanco de Poisson, utilizando técnicas de filtrado H-Infinito.
- Ampliar el procedimiento de identificación de parámetros para sistemas en presencia de ruido blanco de Poisson, utilizando técnicas de filtrado H-Infinito.
- Diseñar un filtro para sistemas polinomiales bajo observaciones lineales, donde el estado y la ecuación de observación son alterados por ruido aditivo de tipo Poisson y Gaussiano simultáneamente.

Además se sugiere utilizar plataformas distintas a MATLAB, puesto que no existen bloques para simular ruido de Poisson. Para simularlo, como se menciona en la tesis, se utilizó el diagrama en *Simulink* sugerido en [5]. Actualmente se cuenta con una herramienta llamada LabVIEW, que además tiene la capacidad de interactuar con MATLAB y que tiene incluida entre sus herramientas un bloque para simular ruido blanco de Poisson.

Apéndice A

Glosario

Definición: La observación de algún fenómeno bajo algunas condiciones y acciones, en un período de tiempo dado, es llamada una prueba.

Definición: Una característica cualitativa de una prueba consiste en registrar si los resultados de un experimento presentan algún efecto o no. Este efecto es llamado evento.

Definición: Una característica cuantitativa de una prueba consiste en determinar los valores de algunas variables obtenidas como un resultado de una prueba. Cada una de estas variables asume diferentes valores como resultado de una prueba, los cuales son imposibles de predecir. A estas variables se les llama variables aleatorias. Los valores específicos que toma una variable aleatoria son llamados valores simples o realizaciones de la variable aleatoria.

Definición: La proporción de el número de apariciones de un evento respecto al número total de pruebas es llamada la frecuencia del evento. Así si un evento aparece m veces en n pruebas, entonces la frecuencia en esta serie de pruebas es igual a m/n .

Definición: Dada la estabilidad de la frecuencia de un evento, y asumiendo que a todo evento le es asociado un número, a este número se le llama la probabilidad de este

evento. Esto es, el número en el cual la frecuencia tiende a ser estable. $P(A)$ denota la probabilidad del evento A .

Definición: Un evento elemental es aquel que no contiene algunos sub-eventos, excepto el evento imposible (ϕ) y a sí mismo.

Definición: El conjunto de todos los eventos elementales asociados con una prueba, es llamado el espacio de eventos elementales y usualmente es denotado por Ω .

Definición: Sea Ω un espacio de eventos elementales. Sea δ un conjunto de subconjuntos de Ω . δ es llamada una σ -álgebra si:

- Para toda $\delta_i \in \delta, \delta_i^c \in \delta$, donde $\delta_i^c = \{x \in \mathbb{R} / x \text{ no está en } \delta_i, \}$ es el evento complementario al evento δ_i .
- Si $\delta_1, \delta_2, \dots, \delta_n, \dots$ es una secuencia contable de elementos de δ , entonces $\bigcup \delta_n \in \delta$.
- $\phi \in \delta$.

Definición: Sea Ω un espacio de eventos elementales y A un evento de δ , la σ -álgebra definida en Ω . La función $P(A)$ es llamada probabilidad (o medida de probabilidad de A) si se cumplen las siguientes condiciones:

- $P(A) \geq 0$;
- $P(\Omega) = 1$;
- Si A_1, A_2, \dots es una secuencia finita o infinita de eventos mutuamente excluyentes $A_i \cap A_j = \phi$ para todas i, j tales que $i \neq j$ entonces: $P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$

(Ω, δ, P) forman un espacio de probabilidad.

Definición: Un espacio de eventos elementales Ω con una álgebra o σ -álgebra dada de conjuntos δ , y una probabilidad en δ definida como una medida no negativa $P(A)$, $A \in \delta$, es llamado un espacio de probabilidad y denotado por (Ω, δ, P) . Así, el espacio de probabilidad sirve como un modelo matemático de algún fenómeno aleatorio en teoría moderna de probabilidad.

Definición: La correspondencia entre algún conjunto de eventos y sus probabilidades es usualmente llamada distribución de probabilidad. Así, la probabilidad $P(A)$ como una función de un conjunto $A \in \delta$ define una distribución de probabilidad en δ .

Definición: Al conjunto de eventos, de los cuales es determinada su probabilidad, es llamado σ -álgebra de eventos y es denotado por δ .

Definición: Un conjunto contable de eventos A_k es llamado un conjunto completo de eventos si hasta el último de ellos aparece como resultado de una prueba. Es decir, los eventos $A_1, \dots, A_n, n < \infty$, forman un conjunto completo si $\bigcup A_k = \Omega$.

Definición: Si $P(B) \neq 0$, $B \in \mathcal{S}$ entonces la probabilidad condicional de algún evento $A \in \mathcal{S}$ relativo al evento B es determinada por la siguiente fórmula:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

El evento A es independiente de B si $P(A/B) = P(A)$ Y además

$$P(A \cap B) = P(A)P(B/A) = P(B)P(A/B)$$

Definición: Un proceso estocástico es una colección de variables aleatorias, $X_t, t \in T$, definidas sobre el mismo espacio de probabilidad, donde T es un conjunto indizado y X_t es una variable aleatoria para cada t . X_t también se denota por $X(t)$ y los valores que esta variable aleatoria asigna al evento elemental w se denotarán por $X(t, w)$.

Definición: Un proceso aleatorio $X(t)$ es llamado proceso con incrementos no correlacionados si para todos los intervalos disjuntos $[t_1, t_2), [t_3, t_4), t_1 < t_2 \leq t_3 < t_4$ los

incrementos correspondientes $X_{t_2} - X_{t_1}$, y $X_{t_4} - X_{t_3}$ del proceso $X(t)$ son no correlacionados.

$X(t)$ es un proceso con incrementos independientes si para cualquier \mathcal{N} , $t_0 < t_1 < \dots < t_{\mathcal{N}}$ las variables aleatorias $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_{\mathcal{N}}} - X_{t_{\mathcal{N}-1}}$, son independientes.

Dada la función aleatoria $X(t) = [X_1(t_1), \dots, X_n(t_n)]$, se define su función de densidad multidimensional $f(x_1, \dots, x_n; t_1, \dots, t_n)$ como la densidad conjunta del vector aleatorio $X(t) = [X_1(t), \dots, x_n(t)]$, la cual toma los valores de X_1, \dots, X_n , en los tiempos t_1, \dots, t_n .

Definición: Para la función aleatoria $X(t)$ con valores independientes, los valores de las variables X_{t_1}, \dots, X_{t_n} son independientes para algún $t_1, \dots, t_n \in T$ y algún número natural n . Para la función aleatoria $X(t)$ con valores independientes para $n = 1, 2, \dots$ es presentada la siguiente relación

$$f_n(x_1, \dots, x_n; t_1, \dots, t_n) = f_1(x_1, t_1)f_1(x_2, t_2)\dots f_1(x_n, t_n)$$

Así, todas las distribuciones multi-dimensionales de una función aleatoria con los valores independientes es determinada únicamente por sus distribuciones uni-dimensionales.

Definición: Dada la función aleatoria $X(t)$, $t \in T_1$ los conjuntos de funciones características uni-dimensional $g_1(\lambda; t)$ y multi-dimensional $g_n = g_n(\lambda_1, \lambda_2, \dots, \lambda_n; t_1, \dots, t_n)$ con los parámetros t y t_1, \dots, t_n son determinados respectivamente por:

$$\begin{aligned} g_1(\lambda, t) &= E(e^{i\lambda^T X(t)}), \\ g_n(\lambda_1, \lambda_2, \dots, \lambda_n; t_1, \dots, t_n) &= E(e^{i\lambda_1^T X(t_1) + \dots + i\lambda_n^T X(t_n)})(n = 2, 3, \dots) \end{aligned}$$

Definición: La probabilidad P_x es llamada la medida de probabilidad o la distribución de la variable aleatoria X , y está dada por:

$$P_x(A) = P(X \in A) = P(X^{-1}(A)), A \in \mathcal{A}$$

Bibliografía

- [1] Alcorta M.A., Basin M., Maldonado J.J. and Anguiano S.G., Sub-optimal risk-sensitive filtering for third degree polynomial stochastic systems, *ICIC Express Letters*, vol.2, pp. 371-377, 2008.
- [2] Åström K.J., *Introduction to Stochastic Control Theory*, Academic Press, 1970.
- [3] Bar-Shalom Y., Optimal simultaneous state estimation and parameter identification in linear discrete-time systems, *IEEE Trans. Automat. Contr.*, vol.17, pp. 308-319, 1972.
- [4] Basin M.V., *New Trends in Optimal Filtering and Control for Polynomial and Time-Delay Systems*, Springer, 2008.
- [5] Basin M.V., Alcorta-Garcia M.A. and Peña D., Simulation and optimal filtering for linear systems with Poisson white noises, *Dynamics of Continuous, Discrete and Impulsive Systems. Series B*, vol.14, pp. 221-231, 2007.
- [6] Basin M.V. and Calderon-Alvarez D., Alternative optimal filter for linear systems with multiple state and observation delays, *Int. J. of Innovative Computing, Information and Control*, vol.4, pp. 2889-2898, 2008.

- [7] Basin M.V. and Maldonado J.J., Mean-square filter design for nonlinear polynomial systems with Poisson noise, *Proc. 2011 American Control Conference*, pp. 612–617, 2011.
- [8] Basin M.V. and Maldonado J.J., Mean-square filtering for incompletely measured polynomial states confused with white Poisson noises, *Int. J. of Innovative Computing, Information and Control.*, vol.7, pp. 6085–6095, 2011.
- [9] Basin M.V. and Maldonado J.J., Mean-square filtering for Polynomial system states confused with Poisson noises over polynomial observations, *Modeling, Identification and Control*, vol.32, pp. 47-55, 2011.
- [10] Basin M.V. and Maldonado J.J., Optimal controller for stochastic linear systems with Poisson noises and unknown parameters, *IECON 2011*, pp. 3563-3568, 2011.
- [11] Basin M.V., Calderon-Alvarez D. and Skliar M. Optimal filtering for incompletely measured polynomial states over linear observations, *International J. Adaptive Control and Signal Processing*, vol.22, pp. 482-494, 2008.
- [12] Basin M.V. and Maldonado J.J., Optimal mean-square and parameter estimation for stochastic linear systems with Poisson noises, *Information Sciences*, vol.197, pp. 177-186, 2012.
- [13] Basin M.V., Perez J. and Calderon-Alvarez D., Optimal filtering for linear systems over polynomial observations, *Int. J. of Innovative Computing, Information and Control*, vol.4, pp. 313-320, 2008.
- [14] Basin M.V., Perez J. and Skliar M., Optimal filtering for polynomial system states with polynomial multiplicative noise, *Int. J. Robust and Nonlinear Control*, vol.16, pp. 287-298, 2006.

- [15] Basin M.V., Shi P. and Calderon-Alvarez D., Approximate finite-dimensional filtering for polynomial states over polynomial observations, *International Journal of Control*, vol.83, pp. 724-730, 2010.
- [16] Basin M.V., Shi P. and Calderon-Alvarez D., Optimal filtering for incompletely measured polynomial systems with multiplicative noises, *Circuits, Systems and Signal Processing*, vol.28, pp. 223-239, 2009.
- [17] Bellman R.E., Dynamic programming, *Princeton Univ. Press*, 1957.
- [18] Benes V.E., Exact finite-dimensional filters for certain diffusions with nonlinear drift, *Stochastics*, vol.5, pp. 65-92, 1981.
- [19] Ding Q. and Zhong M., On designing H_∞ fault detection filter for Markovian jump linear systems with polytopic uncertainties, *Int. J. of Innovative Computing, Information and Control*, vol.6, pp. 995-1004, 2010.
- [20] Dupé F.X. , Fadili M.J. and Starck J.-L., Image deconvolution under Poisson noise using sparse representations and proximal thresholding iteration, *Proc. IEEE ICASSP - Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 761-764, 2008.
- [21] Elliott L.J. and Krishnamurthy V., New finite-dimensional filters for estimation of continuous-time linear Gaussian systems, *SIAM. J. Contr. Optim.*, vol.35, pp. 1908-1923, 1997.
- [22] Elliott L.J. and Krishnamurthy V., New finite-dimensional filters for parameter estimation of discrete-time linear Gaussian models, *IEEE Trans. Automat. Contr.*, vol.44, pp. 938-951, 1999.

- [23] Fleming F.H. and Rishel R.W., *Deterministic and Stochastic Optimal Control*, Springer-Verlag, 1975.
- [24] Fleming W.H. and McEneaney, W.M., Robust limits of risk sensitive nonlinear filters, *Mathematics of Control, Signals and Systems*, vol.14, pp. 109-142, 2001.
- [25] Florentin J.J., Optimal control of continuous time, Markov, stochastic systems, *J. of Electronics and Control*, vol.10, pp. 473-488, 1961.
- [26] Florentin J.J., Partial observability and optimal control, *J. of Electronics and Control*, vol.13, pp. 263-279, 1962.
- [27] Fridman E. and Shaked U., On regional nonlinear H_∞ filtering, *Systems and Control Letters* vol.29, pp. 233-240, 1997.
- [28] Gao H. and Chen T, H_∞ estimation for uncertain systems with limited communication capacity, *IEEE Transactions on Automatic Control*, vol.52, pp. 2070-2084, 2007.
- [29] Gao H. and Wang C. A delay-dependent approach to robust H_∞ filtering for uncertain discrete-time state-delayed systems, *IEEE Transactions on Signal Processing*, vol.52, pp. 1631-1640, 2004.
- [30] Gao H., Lam J., Xie L. and Wang C., New approach to mixed H_2/H_∞ -filtering for polytopic discrete-time systems, *IEEE Transactions on Signal Processing*, vol.53, pp. 3183-3192, 2005.
- [31] Hameed I.A., Using the extended Kalman filter to improve the efficiency of greenhouse climate control, *Int. J. of Innovative Computing, Information and Control*, vol.6, pp. 2671-2680, 2010.

- [32] Hannequin P. and Mas J., Statistical and heuristic image noise extraction (SHINE): A new method for processing Poisson noise in scintigraphic images, *Phys. Med. Biol.*, vol. 47, pp. 4329-4344, 2002.
- [33] Hazewinkel M., Marcus S.I. and Sussmann H.J., Nonexistence of exact finite-dimensional filters for conditional statistics of the cubic sensor problem, *Systems and Control Letters*, vol.5, pp. 331-340, 1983.
- [34] Hussain A., Jaffar M.A. and Mirza M.A., Detail preserving fuzzy filter for impulsive noise removal, *Int. J. of Innovative Computing, Information and Control*, vol.5, pp. 3583-3591, 2009.
- [35] Jazwinski A.H., *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
- [36] Kalman R.E., A new approach to linear filtering and prediction problems, *J. Basic Eng.*, vol.82, pp. 34-35, 1960.
- [37] Kalman R.E., Contributions to the theory of optimal control, *Bol. Soc. Mat. Mex.*, vol.5, pp 102-119, 1960.
- [38] Kalman R.E., The theory of optimal control and the calculus of variations in mathematical optimization techniques, *Mathematical Optimizacion Techniques*, Univ. de Calif Press, pp.309-331, 1963.
- [39] Kalman R.E. and Bucy R.S., New results in linear filtering and prediction theory, *ASME Trans., Part D (J. of Basic Engineering)*, vol.83, pp. 95-108, 1961.
- [40] Kolmanovsky I. and Maizenberg T., Optimal containment control for a class of stochastic systems perturbed by Poisson and Wiener processes, *IEEE Trans. on Automatic Control*, vol. 47, pp. 1641-1645, 2002.

- [41] Kolmanovsky I. and Maizenberg T.L., Optimal containment control for a class of stochastic systems perturbed by Poisson and Wiener processes, *Proc. 2002 American Control Conf.*, pp. 322-327, 2002.
- [42] Kushner H.J., On differential equations satisfied by conditional probability densities of Markov processes, *SIAM J. Control*, vol.12. pp. 106-119, 1964.
- [43] Kushner H.J., Optimal stochastic control, *IRE Trans. on Aut. Control*, pp. 120-122, 1962.
- [44] Kwakernaak H. and Sivan R., *Linear Optimal Control Systems*, Wiley-Interscience, 1972.
- [45] Lewis F.L., Applied optimal control and estimation, *Prentice Hall*, 1992.
- [46] Liptser R.S. and Shiriyayev A.N., *Theory of Martingales*, Springer, 1989.
- [47] Lu H., Liang Z. and Chen D., A combined transformation of ordering SPECT sinograms for signal extraction from measurements with Poisson noise, *SPIE Proc.*, vol. 4322, pp. 1431-1438, 2001.
- [48] Mahmoud M. and Shi P., Robust Kalman filtering for continuous time-lag systems with Markovian jump parameters, *IEEE Transactions on Circuits and Systems*, vol.50, pp. 98-105, 2003.
- [49] Mara-Shabaik H.E., Mahmoud M.S. and Shi Y., H_2 and H_∞ filtering for linear systems with unknown inputs and polytopic uncertainty, *Int. J. of Innovative Computing, Information and Control*, vol.6, no. 5, pp. 2211-2220, 2010.
- [50] Nguang S.K. and Fu M.Y., Robust nonlinear H_∞ filtering, *Automatica* vol.32, pp. 1195-1199, 1996.

- [51] Oksendal B., *Stochastic Differential Equations*, Springer, 2006.
- [52] Pugachev V.S., *Probability Theory and Mathematical Statistics for Engineers*, Pergamon, 1984.
- [53] Pugachev V.S. and Sinitsyn I.N., *Stochastic Systems: Theory and Applications*, World Scientific, 2001.
- [54] Sawada Y. and Kondo J., Kalman filter based LEQG control of a parallel-structured single-link flexible arm mounted on moving base, *Int. J. of Innovative Computing, Information and Control*, vol.6, pp. 29-42, 2010.
- [55] Shen B., Wang Z., Shu H. and Wei G., H_∞ infinity filtering for nonlinear discrete-time stochastic systems with randomly varying sensor delays, *Automatica*, vol.45, pp. 1032-1037, 2009.
- [56] Sheng J., Chen T. and Shah S.L., Optimal filtering for multirate systems, *IEEE Trans. on Circuits and Systems*, vol.52, pp. 228-232, 2005.
- [57] Sheng J., Optimal filtering for multirate systems based on lifted models, *Proc. 2005 American Control Conf.*, pp. 3459-3461, 2005.
- [58] Shi P., Filtering on sampled-data systems with parametric uncertainty, *IEEE Transactions on Automatic Control*, vol.43, pp. 1022-1027, 1998.
- [59] Tucker H.G., *A Graduate Course in Probability*, Academic Press, 1967.
- [60] Wang Z., Ho D.W.C., Liu Y. and Liu X., Robust H_∞ infinity control for a class of nonlinear discrete time-delay stochastic systems with missing measurements, *Automatica*, vol.45, pp. 684-691, 2009.

- [61] Wang Z., Lam J. and Liu X., Nonlinear filtering for state delayed systems with Markovian switching, *IEEE Transactions on Signal Processing*, vol.51, pp. 2321-2328, 2003.
- [62] Wang Z., Liu Y. and Liu X., H_∞ infinity filtering for uncertain stochastic time-delay systems with sector-bounded nonlinearities, *Automatica*, 44, 1268-1277, 2008.
- [63] Wei G., Wang Z. and Shu H., Robust filtering with stochastic nonlinearities and multiple missing measurements, *Automatica*, vol.45, pp. 836-841, 2009.
- [64] Wonham W.M., Some applications of stochastic differential equations to nonlinear filtering, *SIAM J. Control*, vol.2, pp. 347-369, 1965.
- [65] Xie L.H., De Souza C.E. and Wang Y.Y., Robust filtering for a class of discrete-time uncertain nonlinear systems, *International Journal of Robust and Nonlinear Control*, vol.6, pp. 297-312, 1996.
- [66] Xu S.Y. and Chen T.W., Robust H_∞ filtering for uncertain impulsive stochastic systems under sampled measurements, *Automatica*, vol.39, pp. 509-516, 2003.
- [67] Xu S. and van Dooren P.V., Robust H_∞ -filtering for a class of nonlinear systems with state delay and parameter uncertainty, *Int. J. Control*, vol.75, pp. 766-774, 2002.
- [68] Xu S.Y., Lam J., Gao H.J. and Zhou Y., Robust H_∞ filtering for uncertain discrete stochastic systems with time delays, *Circuits, Systems and Signal Processing*, vol.24, pp. 753-770, 2005.
- [69] Yang X., Gao H., Shi P. and Duan G., Robust H_∞ control for a class of uncertain mechanical systems, *Int. J. Control*, vol.83, pp. 1303-1324, 2010.

- [70] Yau S.-T., 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.
- [71] Yaz E. and Yaz, Y., State estimation of uncertain nonlinear systems with general criteria, *Applied Mathematics Letters*, vol.14, pp. 605-610, 2001.
- [72] Zhang B., Fadili M.J. and Starck J.-L., Wavelets, ridgelets, and curvelets for Poisson noise removal, *IEEE Trans. on Image Processing*, vol.17, pp. 1093-1108, 2008.
- [73] Zhang B., Fadili M.J., Starck J.-L. and Dige S.W., Fast Poisson noise removal by biorthogonal haar domain hypothesis testing, *Statistical Methodology*, vol. 5, pp. 387-396, 2008.
- [74] Zhang W.H., Chen B.S. and Tseng C.S., Robust H_∞ filtering for nonlinear stochastic systems, *IEEE Transactions on Signal Processing*, vol.53, pp. 589-598, 2005.