

UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE INGENIERIA MECANICA Y ELECTRICA

DIVISION DE ESTUDIOS DE POSTGRADO



**ALGORITMOS DE FILTRADO DE KALMAN PARA
ESTADOS DE SISTEMAS ESTOCASTICOS SOBRE
OBSERVACIONES DISCRETAS Y CONTINUAS**

TESIS

**QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS
DE LA INGENIERIA ELECTRICA
CON ESPECIALIDAD EN CONTROL**

PRESENTA:

MARIO ARIEL VILLANUEVA LLANES

SAN NICOLAS DE LOS GARZA, N. L.

JULIO, 1999

TM
Z5853
.M2
FIME
1999
75

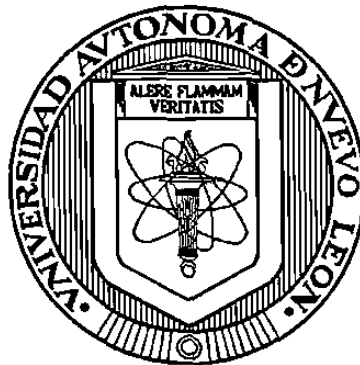
ALGORITMOS DE FILTRADO DE KALMAN PARA ESTADOS DE SISTEMAS
ESTOCASTICOS SOBRE OBSERVACIONES DISCRETAS Y CONTINUAS

M. A. V. L. L.



1020126386

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



**ALGORITMOS DE FILTRADO DE KALMAN PARA ESTADOS DE SISTEMAS
ESTOCÁSTICOS SOBRE OBSERVACIONES DISCRETAS Y CONTINUAS**

TESIS

**QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS DE LA INGENIERÍA ELÉCTRICA
CON ESPECIALIDAD EN CONTROL**

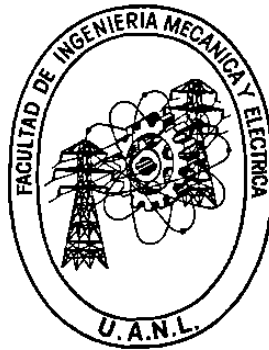
PRESENTA:

MARIO ARIEL VILLANUEVA LLANES

San Nicolás de los Garza, N. L.

Julio, 1999.

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



**ALGORITMOS DE FILTRADO DE KALMAN PARA ESTADOS DE SISTEMAS
ESTOCÁSTICOS SOBRE OBSERVACIONES DISCRETAS Y CONTINUAS**

TESIS

**QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS DE LA INGENIERÍA ELÉCTRICA
CON ESPECIALIDAD EN CONTROL**

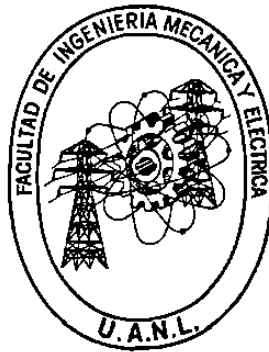
PRESENTA:

MARIO ARIEL VILLANUEVA LLANES

San Nicolás de los Garza, N. L.

Julio, 1999.

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



**ALGORITMOS DE FILTRADO DE KALMAN PARA ESTADOS DE SISTEMAS
ESTOCÁSTICOS SOBRE OBSERVACIONES DISCRETAS Y CONTINUAS**

TESIS

QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS DE LA INGENIERÍA ELÉCTRICA
CON ESPECIALIDAD EN CONTROL

PRESENTA:

MARIO ARIEL VILLANUEVA LLANES

San Nicolás de los Garza, N. L.

Julio, 1999.

0131-65360

TM
25 53
.M2
FIM-
1999
V5



FONDO
TESIS

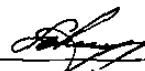


FONDO
TESIS

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

Los miembros del comité de tesis recomendamos que la tesis **ALGORITMOS DE FILTRADO DE KALMAN PARA ESTADOS DE SISTEMAS ESTOCÁSTICOS SOBRE OBSERVACIONES DISCRETAS Y CONTINUAS** realizada por el Ing. Mario Ariel Villanueva Llanes sea aceptada para su defensa como opción al Grado de Maestro en Ciencias de la Ingeniería Eléctrica con especialidad en Control.

El comité de tesis:



Asesor
Dr. Mikhail V. Basin



Coasesor
Dr. José A. de la O Serna



Coasesor
Dr. Salvador Acha Daza



Vb. Bo.
M. C. Roberto Villarreal Garza
División de Estudios de Postgrado

© Copyright 1999
por Mario Ariel Villanueva Llanes,
Reservados Todos los Derechos.

Dedicatoria

A mi padre Fidel Villanueva Madrid, por enseñarme a trabajar y a dar valor a las cosas.

A mi madre Beatriz Llanes Lara, por darme su amor y apoyarme en todas mis decisiones.

A mis Hermanos Iván Nazul, Carlos Norberto, Celia, Julian Alejandro y Diana Carolina, quienes siempre me han manifestado su cariño y apoyo.

A todos mis demás parientes, especialmente a mis Tíos Roger Encarnación y Mario Ernesto, por el apoyo moral que siempre me brindaron.

Agradecimientos

Al Dr. Mikhail V. Basin por la orientación que me brindó para obtener la formación y nivel académico que hicieron posible la realización de la presente tesis.

Al Dr. Salvador Acha Daza por sus valiosos comentarios que enriquecieron el contenido final de esta tesis, así como, por la amabilidad con la que me ayudó y me orientó en diferentes aspectos académicos.

Al Dr. José A. de la O Serna por contribuir con sus amplios conocimientos del tema para el mejoramiento de este trabajo.

A todos los profesores del Programa Doctoral por compartir abiertamente sus conocimientos durante la realización de mis estudios, especialmente al Dr. Joaquín Collado Moctezuma, al Dr. Jesús de León Morales, al Dr. Krishna K. Busawon y al Dr. Juan M. Ramírez Arredondo.

A todos los compañeros del Doctorado por la cordialidad y amabilidad con la que siempre me trataron, especialmente a César Elizondo González, Martha Aguilera Hernández, Guadalupe Acosta Villarreal, Irma R. Valadez Guzmán, Ma. de los Angeles Carrera González, Gustavo Pérez Hernández, David A. Díaz Romero, Oscar F. Huerta Guevara, Francisco H. Florenzano Lara, César Guerra Torres y Ramón Huereca Leyva.

Al CONACYT y a la Universidad Autónoma de Nuevo León por su apoyo económico.

Gracias.

Resumen

Algoritmos de Filtrado de Kalman para Estados de Sistemas Estocásticos sobre Observaciones Discretas y Continuas

Publicación No. _____
Mario Ariel Villanueva Llanes
Universidad Autónoma de Nuevo León
Facultad de Ingeniería Mecánica y Eléctrica
Profesor Asesor: Dr. Mikhail V. Basin
Julio, 1999.

El contenido de la presente tesis está orientado a resolver el problema de filtrado de sistemas con estados estocásticos del tipo Volterra, los cuales, están comprendidos dentro de la clase general de sistemas con retardo. El estudio se realiza en el área de filtrado en el dominio del tiempo haciendo una extensión del filtro de Kalman-Bucy [3], y del filtro condicional Gaussiano [6], para procesos de Ito-Volterra; apoyada en los resultados sobre filtrado de Y. V. Orlov y M. V. Basin expuestos en [8, 9].

La teoría desarrollada se presenta en el capítulo 6, y consiste en un teorema para el filtrado de procesos de Ito-Volterra sobre observaciones discontinuas, un teorema para el filtrado de procesos de Ito-Volterra sobre observaciones con retardo y un corolario para el filtrado de procesos de Ito-Volterra sobre observaciones con múltiples retardos. También se proporcionan ejemplos con el propósito de ilustrar la aplicabilidad de los resultados.

Los resultados expuestos, tienen su importancia científica como herramientas matemáticas en el diseño de filtros para sistemas estocásticos expresados en forma de ecuaciones integrales y que no pueden ser transformados a ecuaciones diferenciales. Lo anterior implica que al resolverse el problema de filtrado desde el punto de vista integral, implícitamente, se está resolviendo el mismo problema para toda aquella clase de sistemas diferenciales que pueden escribirse como dichos sistemas integrales. Con esto, se abarcan las aplicaciones usuales del filtro de Kalman y se tiene la oportunidad de explorar la solución de filtrado para sistemas más complejos que el caso diferencial no resuelve.

Notación

$a \vee b$	a ó b
$a \wedge b$	a y b
\cup	Unión de conjuntos
\cap	Intersección de conjuntos
\prod	Producto
\emptyset	Conjunto vacío
\in	Pertenece a
\notin	No pertenece a
\exists	Existe
\forall	Para toda
\Rightarrow	Implica que
\rightarrow	Tiende a
$:$	Tal que
$ $	Dado que ocurre o en términos de
$\ \cdot\ _l$	Norma l de \cdot
$ \cdot $	Valor absoluto de \cdot
\mathbb{R}	Conjunto de los números reales
\mathbb{R}^n	Espacio vectorial de dimensión n con componentes reales
$\mathbb{R}^{n \times m}$	Campo de las matrices de dimensión $n \times m$ con componentes reales
A	Matriz
A'	Matriz transpuesta de A
A^{-1}	Matriz inversa de A
A^c	Matriz complemento de A
$\det A$	Determinante de A
\inf	Infimo
\sup	Supremo
\dot{X}	Derivada completa de X con respecto al tiempo

Indice

1) Introducción

1.1) Motivación	1
1.2) El Problema del Filtrado	2
1.3) Antecedentes	3
1.4) Aportaciones Principales	4
1.4.1) Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discontinuas	4
1.4.2) Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discretas con Retardo	5
1.4.3) Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discretas con Múltiples Retardos	5
1.4.4) Ejemplos Matemáticos	5
1.4.5) Método de Simulación para Procesos de Ito-Volterra	6
1.5) Organización de la Tesis	6

2) Marco Teórico

2.1) Probabilidad y Estadística	8
2.1.1) Fundamentos Matemáticos de la Teoría de Probabilidad	8
2.1.2) Variable Aleatoria	9
2.1.3) Valor Esperado, Varianza y Covarianza	11
2.1.4) Independencia y Valor Esperado Condicional	14
2.1.5) Modos de Convergencia de las Variables Aleatorias	16

Indice (continuación)

2.2) Procesos Estocásticos	17
2.2.1) Procesos Estacionarios con Incrementos Independientes	17
2.2.2) Integral de Ito	20
2.2.3) Fórmula de Ito	21
2.3) Cálculo Avanzado	26
2.3.1) Integral de Lebesgue	26
2.3.2) Función de Variación Acotada	28
2.3.3) Integral de Ito-Volterra	29
3) Filtrado de Kalman	
3.1) Filtro de Wiener	33
3.1.1) Planteamiento del Problema	33
3.1.2) Solución al Problema de Filtrado de Wiener	34
3.2) Filtro de Kalman	35
3.2.1) Proyecciones Ortogonales	35
3.2.2) Planteamiento del Problema	37
3.2.3) Solución al Problema de Filtrado	38
3.3) Filtro de Kalman-Bucy	40
3.3.1) Planteamiento del Problema	40
3.3.2) Solución del Problema de Filtrado de Kalman-Bucy	41
3.4) Filtro Gaussiano Condicional	42
3.4.1) Propiedades de los Sistemas Gaussianos	42
3.4.2) Planteamiento del Problema	43
3.4.3) Solución del Problema	44

Índice (continuación)

4) Filtrado de Procesos de Ito-Volterra

4.1) Filtro Gaussiano Condicional para Procesos de Ito-Volterra	46
4.1.1) Planteamiento del Problema	46
4.1.2) Filtro Óptimo	47
4.2) Filtrado de Procesos de Ito-Volterra sobre Observaciones Continuas	49
4.2.1) Planteamiento del Problema	49
4.2.2) Filtro Óptimo	50
4.2.3) Ejemplo	53

5) Solución para Ecuaciones Diferenciales en Distribuciones con Funciones Regulares Discontinuas sobre el Lado Derecho

5.1) Funciones Generalizadas	56
5.1.1) Espacio Fundamental y Funciones Generalizadas	56
5.1.2) Propiedades de las Funciones Generalizadas	58
5.1.3) Ecuaciones Diferenciales en Distribuciones	59
5.2) Vibrosoluciones para Ecuaciones Diferenciales en Distribuciones con Funciones Regulares Discontinuas sobre el Lado Derecho	59
5.2.1) Vibrosoluciones	60
5.2.2) Ejemplo	68

6) Resultados

6.1) Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discontinuas	69
6.1.1) Declaración del Problema	69
6.1.2) Filtro Óptimo	70

Índice (continuación)

6.2) Forma Equivalente de las Ecuaciones de Filtrado	71
6.2.1) Ejemplo	72
6.3) Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discretas con Retardo	76
6.3.1) Planteamiento del Problema	76
6.3.2) Filtro Óptimo	76
6.3.3) Ejemplos	78
6.4) Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discretas con Múltiples Retardos	84
6.4.1) Descripción del Problema	84
6.4.2) Filtro Óptimo	84
6.4.3) Ejemplo	86
7) Conclusiones	
7.1) Aportaciones	89
7.2) Trabajos Futuros	90
Apéndice A	91
Apéndice B	98
Apéndice C	105
Referencias	107

Índice de Figuras

Figura 3.1: Modelo de Señal para el Filtro de Wiener	33
Figura 3.2: Estructura del Filtro de Wiener	34
Figura 3.3: Descomposición Ortogonal de un Vector Aleatorio	37
Figura 3.4: Modelo de un Sistema Lineal de Tiempo Discreto	38
Figura 3.5: Modelo del Filtro Óptimo de Kalman	39
Figura 3.6: Modelo de un Sistema Lineal	40
Figura 3.7: Modelo del Filtro Óptimo de Kalman-Bucy	41
Figura 4.1: Funciones: (a) de Correlación y (b)-(d) Características	54
Figura 4.2: Ecuaciones: (a) de Observación y (b) Estado/Estimado	55
Figura 6.1: Función de Correlación Discontinua	74
Figura 6.2: Ecuaciones: (a) de Observación Discontinua y (b) de Estado y Estimado	75

Índice de Tablas

Tabla 3.1 Suposiciones y Propiedades del Filtro de Wiener	35
Tabla 3.2 Diferencias entre el Filtro de Wiener y el Filtro de Kalman-Bucy	42
Tabla 6.1 Realización para el Estado y Ruido Gaussiano	80

Capítulo 1

Introducción

1.1 Motivación

En la vida moderna el concepto de *filtro* cobra cada vez más relevancia, esto es debido a que en la naturaleza todo tipo de proceso (ya sea biológico, químico, eléctrico, mecánico, etc...), no escapa a la presencia de ruidos que distorsionan alguna señal o referencia que se desea medir. Uno de los ejemplos más claros de cómo afecta el ruido a nuestras vidas cotidianas puede verse en el campo de las comunicaciones, y por citar algunos casos: las señales de radio o de televisión, las cuales, son comúnmente recibidas con ruido, de tal manera que en algunos casos es imposible una clara recepción; este problema ha motivado el uso de filtros para mejorar la calidad de operación de estos aparatos. Es precisamente la palabra *calidad*, la que encierra el propósito primordial de la existencia de la teoría del filtrado, pues basta simplemente con asociar la idea de *filtrar* con la de eliminar los componentes no deseados.

El contenido de la presente tesis está orientado a resolver el problema de filtrado de sistemas con estados estocásticos del tipo Volterra, los cuales, están comprendidos dentro de la clase general de sistemas con retardo. El estudio se realiza en el área de filtrado en el dominio del tiempo haciendo una extensión del filtro de Kalman-Bucy [3], y del filtro condicional Gaussiano [6], para procesos de Ito-Volterra; apoyada en los resultados sobre filtrado de Y. V. Orlov y M. V. Basin [8, 9].

Los resultados expuestos, tienen importancia científica como herramientas matemáticas en el diseño de filtros para sistemas estocásticos expresados en forma de ecuaciones integrales que no pueden ser transformados a ecuaciones diferenciales. Lo anterior implica que al resolverse el problema de filtrado desde el punto de vista integral, implícitamente, se está resolviendo el mismo problema para toda aquella clase de sistemas diferenciales que pueden escribirse como dichos sistemas integrales. Con esto se abarcan las aplicaciones usuales del filtro de Kalman y se tiene la oportunidad de explorar la solución de filtrado de sistemas más complejos que el caso diferencial no resuelve.

1.2 El Problema de Filtrado

El término *filtro*, es comúnmente usado para describir un componente físico o un modelo matemático programable computacionalmente, que está aplicado a un conjunto de eventos con ruido con la intención de extraer cierta información de interés. El ruido por su parte, puede ser generado de múltiples formas y físicamente podrá afectar a todo sistema. En cualquier evento, se usa un filtro para desempeñar tres operaciones básicas de procesamiento de la información:

- (1) *Filtrado*, el cual, consiste en la extracción de la información de interés al tiempo t , usando mediciones de datos desde un tiempo inicial t_0 hasta el tiempo t . En el proceso de señales en línea (por ejemplo el filtro de Kalman-Bucy), para formar el mejor estimado en t se emplea únicamente el valor de la observación en el instante t .
- (2) *Suavización*, en este proceso a diferencia del filtrado, la información de interés no necesita estar disponible al tiempo t , y las mediciones de los datos son hechas después del tiempo t . Lo anterior implica que existe un tiempo de retardo en producirse el resultado de interés.
- (3) *Predicción*, esta operación no es más que una previsión de la información en el tiempo $t + \tau$, para $\tau > 0$, cuando el proceso se encuentra en el tiempo t .

En el caso del filtrado, el problema consiste en encontrar el mejor estimado al tiempo t de los vectores reales o complejos, afectados por la presencia de ruidos aleatorios, al tiempo t y de los cuales se desea conocer su dinámica.

Para mejor comprensión del tipo de problemas de filtrado que este trabajo comprende, se requiere de la introducción de los siguientes conceptos considerando que $t \in \mathbb{R}$ y $t_0 \leq t$.

Definición 1.1 *Las observaciones continuas, son funciones tales que están disponibles en cada tiempo t y forman una función continua. Del mismo modo, las observaciones discretas, son funciones disponibles en un conjunto finito o contable de valores para los momentos $t_k, \forall k = 0, 1, 2, \dots$*

Definición 1.2 *Se dice que las observaciones discontinuas (con discontinuidades), son funciones disponibles en cada instante t y forman una función discontinua de variación acotada que incluye componentes continuo y discreto.*

Definición 1.3 Las observaciones discretas con retardos, son funciones disponibles en un conjunto finito o contable de valores para los momentos t_k , $\forall k = 0, 1, 2, \dots$, las cuales, pueden depender de los estados disponibles en los tiempos previos t_{ki} , $\forall i = 0, 1, 2, \dots$, tal que $t_{ki} \leq t_k$.

1.3 Antecedentes

La teoría de filtrado tuvo sus inicios en 1949 con los estudios del problema de filtrado de sistemas bajo la presencia de ruidos Gaussianos, este método ideado por *N. Wiener* [1], utilizó técnicas de interpolación y de extrapolación sobre series estacionarias en el tiempo y se le conoce como *filtrado de Wiener*. En 1958, como una aplicación de la teoría de la variable de estado, *Rudolf Emil Kalman (1930)*, elaboró una aproximación del modelo basado en el problema del filtrado de Wiener, resultado de gran trascendencia para la ciencia, y que meses después se le conocería como el célebre *filtro de Kalman* [2]. Más tarde, en 1961, Kalman y *R. S. Bucy* [3], presentaron una nueva aplicación a la teoría del filtrado, ésta se conoce como el *filtro de Kalman-Bucy*. Los trabajos en [2] y [3] desarrollaron originalmente el concepto de filtrado para sistemas lineales en tiempo discreto y en tiempo continuo respectivamente. Con base a lo anterior, en [20] *P. Frost* y *T. Kailath (1971)*, utilizando el mismo esquema del filtrado lineal, efectuaron la generalización de éste al caso no lineal. Por otra parte, el método universal para obtener las ecuaciones del filtrado de sistemas bajo la presencia de ruidos Gaussianos, es decir, la forma abstracta para funciones de densidades de probabilidad condicional que se establece en el teorema de correlación de procesos Gaussianos (teorema de correlación normal), fue originalmente planteada por *R. S. Liptser* y *A. N. Shirayev* en [11], en el año de 1977. En este sentido y para fines prácticos, la información más sobresaliente del área de la teoría del filtrado, se puede hallar en el trabajo [19], de *S. K. Mitter (1996)*.

Por otra parte y como producto del modelado matemático de la coexistencia entre dos poblaciones ecológicas, el Italiano *Vito Volterra (1860-1940)*, en los años 20, concluyó con una ecuación integral (la cual, puede ser o no, reducida al caso de una ecuación diferencial), para analizar las variaciones cíclicas observadas en las poblaciones de tiburones y los peces que le

sirven de alimento en el mar Adriático. Esta ecuación, combinada con la integral estocástica de *Ito*, dieron paso a la ecuación de *Ito-Volterra* y representa (cualitativamente), el modelo estocástico presa-depredador, otra cualidad aun más importante, es la de expresar matemáticamente todas aquellas situaciones dadas por los sistemas estocásticos escritos en forma de ecuación diferencial los cuales forman parte de la teoría del control lineal en variables de estado. Los primeros trabajos sobre la existencia y la unicidad de la ecuación de *Ito-Volterra* fueron realizados por *K. Ito* [4], en el año de 1979. Posteriormente, *A. V. Balakrishnan* [5], realizó en 1980 las primeras aplicaciones de este modelo a la teoría del control óptimo. En 1985, *M. L. Kleptsina* y *A. Yu. Veretennikov* [6], establecieron el procedimiento para obtener el filtro óptimo del proceso descrito sobre observaciones continuas de una ecuación reducible a una ecuación diferencial en el caso escalar. Un año más tarde *L. E. Shaiket* [7], publicaría el mismo resultado, pero para el caso vectorial. Finalmente *Y V Orlov* [8] y *M. V. Basin* [8, 9], en una serie de trabajos diversos, sentaron las bases necesarias para el análisis del filtrado de las ecuaciones de *Ito-Volterra* sobre observaciones discontinuas, de las cuales se obtienen las observaciones discretas con retardo presentadas por primera vez en este trabajo.

1.4 Aportaciones Principales

Las aportaciones de esta tesis son las siguientes:

1.4.1 Filtrado para Procesos de *Ito-Volterra* sobre Observaciones de *Ito-Volterra* Discontinuas

Este resultado es básicamente una extensión del filtro de Kalman-Bucy [3], y del filtro condicional Gaussiano para procesos de *Ito-Volterra* [6]. La ecuación que describe el comportamiento dinámico del sistema y la ecuación de observación, están dadas en la forma de una ecuación integral de *Ito-Volterra*, mismas que no pueden ser reducidas al caso de una ecuación diferencial ordinaria, lo cual dificulta el camino para hallar una ecuación de estimación óptima para ese sistema. Además, la ecuación de observación está caracterizada por una función de variación acotada, la cual puede ser representada por una función generalizada de orden cero como la función escalón. En este

sentido, se establecen las ecuaciones de filtrado óptimo, la función de correlación y la función característica para procesos de Ito-Volterra sobre observaciones de Ito-Volterra en la forma de ecuaciones integrales para los casos continuo y discontinuo respectivamente.

Esta aportación fue originalmente ideada por *M. V. Basin* y concluida en [24], publicación aceptada para su presentación en la Conferencia Americana de Control (ACC). Se incluye en el apéndice.

1.4.2 Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discretas con Retardo

Esta aportación es obtenida de la derivación del filtrado para procesos de Ito-Volterra sobre observaciones de Ito-Volterra discontinuas al incluir en la ecuación de observación una función Heaviside que la transforma en una ecuación discreta (de diferencias). Se establece mediante un teorema las condiciones para obtener el estimado óptimo, la función del error de estimación y la función de la mejor desviación del estimado del estado para sistemas de este tipo.

1.4.3 Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discretas con Múltiples Retardos

Esta aportación [32], es obtenida de la derivación del filtrado de procesos de Ito-Volterra sobre observaciones de Ito-Volterra con retardo, al substituir en el caso de observaciones discontinuas una función Heaviside con múltiples tiempos de escalonamiento. Se presentan las ecuaciones del mejor estimado, la función de correlación y la función característica.

Las dos contribuciones anteriores están puestas a consideración para participar en la Conferencia de Decisión y Control de 1999 (CDC99). Se incluye en el apéndice.

1.4.4 Ejemplos Matemáticos

En los respectivos casos de filtrado para procesos de Ito-Volterra sobre observaciones continuas, discontinuas, discretas con retardo y discretas con múltiples retardos; se presentan ejemplos matemáticos (sin aplicación práctica), con el fin de ilustrar la funcionalidad y las posibles aplicaciones de los resultados obtenidos. sin embargo, en el caso de observaciones con retardo, se sugieren aplicaciones prácticas (sin conocer las ecuaciones reales de los modelos), para el problema

de transmisión de fotografías satelitales a la correspondiente estación de control y análisis, así como, para el problema del mejoramiento de imágenes de placas radiográficas digitales [33].

1.4.5 Método de Simulación para Procesos de Ito-Volterra

Dado que en la actualidad no se cuenta con los programas necesarios para la simulación de sistemas expresados en forma de ecuaciones integrales, se presenta para los correspondientes ejemplos, un método aproximado de simular las ecuaciones de Ito-Volterra (basado en la ley de *Leibnitz*), el cual pese a no ser preciso, resulta satisfactorio para el propósito fundamental de dichos problemas. Para todos los casos, las simulaciones en computadora de estos sistemas se realizó con la ayuda del programa *Simnon 1.0*.

1.5 Organización de la Tesis

En el capítulo dos, se presenta una recopilación de la teoría de probabilidad, necesaria para la definición de los espacios de probabilidad, procesos estocásticos, la integral de Ito y su fórmula diferencial. También se revisan las definiciones de la integral de Lebesgue, la ecuación de Volterra y la integral de Ito-Volterra. En este capítulo, se pretende dar al lector una idea de los conceptos fundamentales para la comprensión de los resultados expuestos, así como para plasmar la serie de referencias en las que sustenta la base teórica de los mismos.

En el capítulo tres se introducen los resultados más importantes de la teoría clásica del filtrado, es decir, el filtro de Wiener, el filtro de Kalman y el filtro de Kalman-Bucy; haciendo una comparación de sus propiedades y soluciones, con el fin de establecer las ventajas y desventajas de cada uno de estos métodos y la manera en que a ido evolucionado la teoría general del filtrado de los sistemas Gaussianos.

Para tener un acercamiento a los resultados que se exponen en el capítulo seis, el capítulo cuatro encierra en su contenido la descripción del filtro condicional Gaussiano para procesos de Ito-Volterra (en los casos escalar y vectorial), descritos en [6] y [7]; y del problema de filtrado de Ito-Volterra sobre observaciones discontinuas escalares resuelto en [8].

Por otra parte, en el capítulo cinco se presenta un método para la solución de ecuaciones

diferenciales en distribuciones, el cual es necesario para la correspondiente solución del problema de filtrado para procesos de Ito-Volterra sobre observaciones discontinuas. Aquí, se propone el uso de una ecuación de medida equivalente a la solución de las ecuaciones de este tipo conocida como Vibrosolución. Este capítulo es fundamental en la descripción teórica de los problemas de filtrado que son investigados posteriormente.

El capítulo seis encierra los resultado que se ponen a consideración en esta tesis, descritos en la sección anterior de aportaciones, y comprenden dos teoremas y un corolario.

Finalmente en el capítulo siete, se presentan las conclusiones de los resultados obtenidos, así como sugerencias para una serie de trabajos a realizar en un futuro en esta misma línea de investigación.

Capítulo 2

Marco Teórico

2.1 Probabilidad y Estadística

2.1.1 Fundamentos Matemáticos de la Teoría de Probabilidad

Considere el problema de construir un modelo probabilístico para un experimento con un número *infinito* de resultados independientes.

Definición 2.1 *A todo resultado o elemento ω_i , $i = 0, 1, \dots$, ocurriendo en un experimento, se le conoce como punto muestra o evento elemental.*

En lo sucesivo (a menos que se indique lo contrario), consideraremos que en un mismo experimento, ω_i es independiente de ω_j , $\forall i \neq j$, $i, j = 0, 1, \dots$

Definición 2.2 *Al conjunto infinito de estos eventos elementales $\Omega = \{\omega_1, \dots\}$, se le conoce como espacio muestra.*

Definición 2.3 *Un evento, está definido como cualquier subconjunto $A \subset \Omega$, tal que, $\forall \omega_i \in \Omega$, y si $A, B \subset \Omega$, entonces:*

a) La *unión* de A y B está dada por:

$$A \cup B = \{\omega \in \Omega : \omega \in A \vee \omega \in B\},$$

donde, $A \cup B$, es el evento consistente de realizaciones de A o de B .

b) La *intersección* de A y B :

$$A \cap B = \{\omega \in \Omega : \omega \in A \wedge \omega \in B\},$$

donde, $A \cap B$, consiste de las realizaciones simultáneas de A y B .

c) El *complemento* de A :

$$A^c = \{\omega \in \Omega : \omega \notin A\}.$$

Definición 2.4 *Se dice que \emptyset , es el conjunto vacío o evento imposible. Por el contrario a Ω , se le conoce como el evento cierto.*

Definición 2.5 Sea Ω un conjunto infinito de puntos elementales ω . Una colección \mathcal{A} de subconjuntos de Ω , es llamada un álgebra si:

- a) $\Omega \in \mathcal{A}$,
- b) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}, A \cap B \in \mathcal{A}$,
- c) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

Definición 2.6 Un sistema \mathcal{F} de subconjuntos de Ω , es un σ -álgebra si:

- a) es un álgebra
- b) y además satisface que

$$\forall \{A_n\} \subseteq \mathcal{F} \Rightarrow \bigcup A_n \in \mathcal{F}.$$

Definición 2.7 El par (Ω, \mathcal{F}) es un espacio medible.

Definición 2.8 (Fundamental) Sea P , una función de probabilidad sobre \mathcal{F} , entonces, la triple ordenada (Ω, \mathcal{F}, P) , es llamada un modelo probabilístico o un espacio de probabilidad.

Nota 2.1 En la construcción de un modelo probabilístico para espacios no contables Ω (si Ω tiene potencia de continua [10]), se deberá considerar la función de probabilidad P , asignada a los eventos de Ω , y no a sus eventos elementales.

2.1.2 Variable Aleatoria

Definición 2.9 Sean $(\Omega_1, \mathcal{F}_1)$ y $(\Omega_2, \mathcal{F}_2)$, espacios medibles. Si X es un mapeo de Ω_1 en Ω_2 , denotado por $X : \Omega_1 \rightarrow \Omega_2$, entonces X es medible si:

$$\forall A \in \mathcal{F}_2, \{\omega : X(\omega) \in A\} = X^{-1}(A) \in \mathcal{F}_1.$$

Definición 2.10 Sean $(\Omega_1, \mathcal{F}_1)$ y $(\Omega_2, \mathcal{F}_2)$, espacios medibles. Sea $X : \Omega_1 \rightarrow \Omega_2$ medible. Si $\Omega_2 = \mathbb{R}, \mathbb{R}^n$ o $\mathbb{R}^{n \times m}$, entonces, se dice que \mathcal{F}_2 es una σ -álgebra de borel de subconjuntos de Ω_2 , de dimensión apropiada, denotado por $\mathcal{B}, \mathcal{B}^n$ o $\mathcal{B}^{n \times m}$ respectivamente.

Definición 2.11 Una función real $X = X(\omega)$, definida sobre (Ω, \mathcal{F}) , es una función \mathcal{F} -medible o una variable aleatoria, si:

$$\{\omega : X(\omega) \in B\} \in \mathcal{F}, \forall B \in \mathcal{B}.$$

Definición 2.12 Sea $(\Omega_1, \mathcal{F}_1, P)$, un espacio de probabilidad y $(\Omega_2, \mathcal{F}_2)$ un espacio medible. Si $X : \Omega_1 \rightarrow \Omega_2$ es medible, la función P_X definida por:

$$P_X(A) = P\{\omega : X(\omega) \in A\} = P(X \in A), \forall A \in \mathcal{F}_2,$$

es llamada la distribución de X . Si $X = (x_1, \dots, x_n)'$ es un vector aleatorio, entonces su función de distribución $F_X : \mathbb{R}^n \rightarrow [0, 1]$, está definida por:

$$F_X(x) = P_X \left(\prod_{i=1}^n (-\infty, x_i] \right) = P_X(X_1 \leq x_1, \dots, X_n \leq x_n),$$

donde $x = (x_1, \dots, x_n)'$; F_X es llamada la función de distribución acumulada de las variables aleatorias X_1, \dots, X_n .

Estas funciones se definen sobre un espacio Euclideano [10], y tienen las siguientes propiedades:

- a) F_X es no decreciente y continua por la derecha $\forall x_i, 0 \leq i \leq n$,
 - b) Si $x_i < y_i \Rightarrow F_X(x_1, \dots, x_i, \dots, x_n) \leq F_X(x_1, \dots, y_i, \dots, x_n)$,
 - c) $\lim_{x_1 \rightarrow -\infty} F_X(x) = 0, \lim_{x_1 \rightarrow \infty} F_X(x) = 1$.
- $$\begin{array}{ccc} \vdots & & \vdots \\ x_n \rightarrow -\infty & & x_n \rightarrow \infty \end{array}$$

Cualquier función $F : \mathbb{R}^n \rightarrow [0, 1]$, satisfaciendo las propiedades anteriores, es la distribución de un vector aleatorio. La *distribución marginal* $F_{X_{n_1}, \dots, X_{n_k}}$, puede ser obtenida de F_X para cualquier subconjunto X_{n_1}, \dots, X_{n_k} de la siguiente forma:

$$F_{X_{n_1}, \dots, X_{n_k}}(x_1, \dots, x_n) = \lim_{y_j \rightarrow \infty} F_X(y_1, \dots, x_1, \dots, x_k, \dots, y_n),$$

$$j \notin \{n_1, \dots, n_k\}$$

Definición 2.13 Se dice que X tiene una función de densidad de probabilidad, si existe una función $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$, tal que

$$F_X(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(y_1, \dots, y_n) dy_n, \dots, dy_1. \quad (2.1)$$

Definición 2.14 Se dice que X es absolutamente continua, si X tiene una función densidad de probabilidad (2.1). Entonces, para cada punto de continuidad de F

$$\frac{\partial^n F_X}{\partial x_1, \dots, \partial x_n} \Big|_{x=x_1, \dots, x_n} = f(x_1, \dots, x_n).$$

Además, si X es absolutamente continua, las distribuciones marginales están dadas por

$$F_X(x_1, \dots, x_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{\infty} f(y_1, \dots, y_n) dy_n, \dots, dy_1.$$

2.1.3 Valor Esperado, Varianza y Covarianza

Definición 2.15 Sea X una variable aleatoria definida sobre un espacio de probabilidad (Ω, \mathcal{F}, P) . El valor esperado de X es:

$$E(X) = \int_{\Omega} X dP, \quad (2.2)$$

si este existe, es el promedio de X sobre el espacio de probabilidad completo. Como tal, el valor esperado, media, esperanza matemática o integral de Lebesgue de X con medida de probabilidad P , proporciona información probabilística burda concerniente a X .

Nota 2.2 El valor esperado existe (es finito), si y sólo si $E(|X|) < \infty$, una situación descrita por la notación: $X \in \mathbb{L}^1(\Omega, \mathcal{F}, P)$, o simplemente $X \in \mathbb{L}^1$. En este caso,

$$|E(X)| \leq E(|X|).$$

Cuando X y Y son variables aleatorias, algunas propiedades del valor esperado son:

a) Si a, b son constantes,

$$E(aX + bY) = aE(X) + bE(Y),$$

b) Se dice que $X \leq Y$ si la variable aleatoria $Y - X$ sólo toma valores no negativos, entonces si $X \leq Y$,

$$E(X) \leq E(Y).$$

c) Desigualdad de *Jensen*. Si $g(x)$ es una función convexa, tal que para $0 < \lambda < 1$,

$$g(\lambda x + (1 - \lambda)y) \leq \lambda g(x) + (1 - \lambda)g(y)$$

entonces

$$g(E(X)) \leq E(g(X)).$$

d) Si $X \geq 0$, y $a > 0$,

$$P(X \geq a) \leq \frac{E(X)}{a}. \quad (2.3)$$

A la relación (2.3), se le conoce como la desigualdad de Markov.

Un cambio de las variables en (2.2), de tal modo que X sea un vector aleatorio n -dimensional y Y sea un vector aleatorio m -dimensional, dado por $Y = g(X)$, donde $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sea medible, entonces

$$E(Y) = \int_{\Omega} g(x)dP = \int_{\mathbb{R}^n} g(x)dF_X(x), \quad (2.4)$$

En particular, tomando a g como la función identidad, tal que $g(x) = x$, entonces

$$E(X) = \int_{\mathbb{R}^n} x dF_X(x);$$

y si X tiene una función de densidad de probabilidad f , entonces, de (2.1) y (2.4) se obtiene que

$$E(X) = \int_{\mathbb{R}^n} x f(x) dx.$$

Suponga que $g(x) = |x|^p$, $p \geq 1$. Entonces, la correspondiente colección de variables aleatorias X , para las cuales, $Y = g(X)$ tiene un valor esperado finito, forman un espacio lineal normado completo (espacio de *Banach*, [10, 14]), denotado por $\mathbb{L}^p(\Omega, \mathcal{F}, P)$. La norma en este espacio es:

$$\|X\|_p = \{E(|X|^p)\}^{\frac{1}{p}},$$

donde la desigualdad del triángulo

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p \quad (2.5)$$

es llamada desigualdad de *Minkowski*. Si $p \geq q$, entonces, $\mathbb{L}^p \subseteq \mathbb{L}^q$. Sean X y Y vectores aleatorios n -dimensionales en \mathbb{L}^p y \mathbb{L}^q respectivamente, donde $p \geq 1$ y $\frac{1}{p} + \frac{1}{q} = 1$, entonces,

$$E(X'Y) = \sum_{i=1}^n E(X_i Y_i),$$

existe, es finito y satisface la desigualdad de Holder:

$$|E(X'Y)| \leq \|X\|_p \|Y\|_q. \quad (2.6)$$

En el caso de que $p = 2$, $E(X'Y)$ define un producto interior y (2.6), es conocida como la desigualdad de *Cauchy-Schwarz*, y el espacio \mathbb{L}^2 es un *espacio de Hilbert*. En el contexto probabilístico, la norma \mathbb{L}^2 , es referida como *media cuadrática* o *norma media cuadrática*.

Definición 2.16 Sean $X, Y \in \mathbb{L}^2$ variables aleatorias, entonces la covarianza de X y Y (o función de *crosscorrelación*), es una matriz $n \times n$, definida por

$$\text{Cov}(X, Y) = E([X - E(X)][Y - E(Y)]') \quad (2.7)$$

para el caso escalar

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

si $\text{Cov}(X, Y) = 0$, se dice que X y Y son no correlacionadas. La $\text{Cov}(X, X)$ se denota como $\text{Cov}(X)$, y para el caso $X \in \mathbb{R}^n$ se conoce como la *varianza* $V(X)$ (o función de *autocorrelación*), y está dada por

$$V(X) = E[X - E(X)]^2 = E(X)^2 - [E(X)]^2, \quad (2.8)$$

la raíz cuadrada, $\sigma = \sqrt{V(X)}$, es llamada *desviación estándar* de X . Suponga $m = E(X)$ y $\sigma^2 = V(X)$ ambas finitas. Usando (2.3) para la variable no negativa $(X - \mu)$, se obtiene la *desigualdad de Chebyshev*:

$$\text{si } a > 0, P((X - m)^2 \geq a) \leq \frac{\sigma^2}{a}.$$

Si $n = 1$, $k \in \mathbb{N}$ y $X \in \mathbb{L}^k$, $k = 0, 1, \dots$, entonces, $E(X^k)$ y $E[(X - m)^k]$ definen el k -ésimo *momento* y el k -ésimo *momento central* respectivamente.

Definición 2.17 La función característica de un vector aleatorio $X \in \mathbb{R}^n$ con distribución finita F_X , está dada por

$$\varphi_X(u) = E(e^{ju'X}) = \int_{\mathbb{R}^n} e^{ju'X} dF_X, \quad u \in \mathbb{R}^n, \quad (2.9)$$

y la distribución es determinada únicamente por φ .

Ejemplo 2.1 Considere el vector aleatorio $X = (X_1, \dots, X_n)$, con distribución normal o Gaussiana (denotado por $\mathcal{N}(m, Q)$), es decir, X para $m \in \mathbb{R}^n$ y $Q \in \mathbb{R}^{n \times n}$, tal que $Q = Q' > 0$; tiene una función de densidad de probabilidad

$$f(x) = [(2\pi)^n \det Q]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x - m)'Q^{-1}(x - m)\right\}. \quad (2.10)$$

Si X tiene una matriz de covarianza $\text{Cov}(X) = Q$ y función característica $\varphi_X(u) = \exp\{j(u'm - \frac{1}{2}u'Qu)\}$, entonces, para $n = 1$, X con distribución $\mathcal{N}(m, \sigma^2)$ tiene momentos centrales

$$E[(X - m)^k] = \begin{cases} 0, & k \text{ impar} \\ 1 \cdot 3 \cdot 5 \cdots (k - 1)\sigma^k, & k \text{ par} \end{cases} \quad (2.11)$$

2.1.4 Independencia y Valor Esperado Condicional

Definición 2.18 Se dice que los eventos A_1, A_2, \dots, A_n , en el espacio de probabilidad (Ω, \mathcal{F}, P) son independientes, si para cada subconjunto $\{i_1, i_2, \dots, i_k\}$ del conjunto de enteros $\{1, 2, \dots, n\}$

$$P\left(A_{i_1} \cap \cdots \cap A_{i_k}\right) = P(A_{i_1}) \cdots P(A_{i_k}) \quad (2.12)$$

Si los eventos A_1, A_2, \dots, A_n , son secuenciales en el tiempo, (2.12) implica que la $P(A_i)$ no depende de la ocurrencia de los eventos elementales anteriores o consecutivos. Se dice que las sub- σ -álgebras $\mathcal{F}_1, \dots, \mathcal{F}_n$ son independientes si (2.12) se satisface para todo evento $A_i \in \mathcal{F}_i$, $i = 1, 2, \dots, n$. Los vectores aleatorios X_1, \dots, X_n , definidos sobre (Ω, \mathcal{F}, P) , son independientes si las correspondientes σ -álgebras $\mathcal{F}_1(X_1), \dots, \mathcal{F}_n(X_n)$, generadas por los vectores aleatorios son independientes.

Definición 2.19 Sea X un vector aleatorio en $\mathbb{L}^1(\Omega, \mathcal{F}, P)$, y sea \mathcal{G} , una sub- σ -álgebra de \mathcal{F} . Note que X no es necesariamente un vector aleatorio sobre el espacio de probabilidad (Ω, \mathcal{G}, P) , cuando éste no necesariamente es \mathcal{G} -medible. Ahora bien, denotando por $E(X | \mathcal{G})$ al vector aleatorio Y sobre el espacio (Ω, \mathcal{G}, P) , tal que satisface

$$\int_B Y dP = \int_B X dP, B \in \mathcal{G}. \quad (2.13)$$

Este vector aleatorio $Y = E(X | \mathcal{G})$, es llamado el valor esperado condicional o la esperanza condicional de X dado \mathcal{G} .

La existencia de Y , está garantizada por el teorema de Radon-Nikodym, mismo que puede ser consultado en las referencias [10, 12]. Como X representa un porción de toda la información de los eventos en \mathcal{F} , $E(X | \mathcal{G})$ representa la porción de la información llevada por X , la cual, está relacionada a las sub- σ -álgebras \mathcal{G} . El qué tan burda sea esta información depende de qué tanto X y \mathcal{G} estén relacionados. En un extremo, si X es \mathcal{G} -medible, entonces,

$$E(X | \mathcal{G}) = X; \quad (2.14)$$

y por otro lado si X y \mathcal{G} son independientes

$$E(X | \mathcal{G}) = E(X). \quad (2.15)$$

Generalmente para sub- σ -álgebras $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$,

$$E(E(X | \mathcal{G}_1) | \mathcal{G}_2) = E(E(X | \mathcal{G}_2) | \mathcal{G}_1) = E(X | \mathcal{G}_1). \quad (2.16)$$

De lo anterior se puede concluir que

$$E(E(X | \mathcal{G})) = E(X). \quad (2.17)$$

Algunas de la propiedades más importantes de (2.13), son:

- a) Para a y b constantes, $E(aX + bY | \mathcal{G}) = aE(X | \mathcal{G}) + bE(Y | \mathcal{G})$.
- b) De acuerdo con la nota 2.2 (b), si $X \leq Y$, entonces $E(X | \mathcal{G}) \leq E(Y | \mathcal{G})$.
- c) Desigualdad de Jensen. Si g es convexa, entonces, $g(E(X | \mathcal{G})) \leq E(g(X) | \mathcal{G})$; en particular, $|E(X | \mathcal{G})| \leq E(|X| | \mathcal{G})$.

Las propiedades (a), (b) y (c) anteriores, se satisfacen *convergiendo con probabilidad 1*, concepto

que se define en la siguiente sección.

2.1.5 Modos de Convergencia de las Variables Aleatorias

El conjunto de soluciones de una ecuación diferencial estocástica, puede considerarse como un conjunto de variables aleatorias independientes, donde cada uno de estos elementos o parámetros, está indexado con un correspondiente intervalo continuo que forma parte de esa solución particular. En virtud de lo anterior, surge la pregunta acerca del comportamiento asintótico de esas variables aleatorias, en relación a si éstas convergen a un valor específico o si tienden al infinito. Para conocer tal comportamiento asintótico se hace necesario establecer los modos de convergencia para dicho conjunto de soluciones.

Existen tres modos de convergencia:

1° Suponga que $\{X_n\}$ es una secuencia de variables aleatorias definidas en un espacio común de probabilidad (Ω, \mathcal{F}, P) . *convergencia con probabilidad 1* (c.p.1.), o al menos convergencia segura de X_n a una variable aleatoria X definida sobre (Ω, \mathcal{F}, P) , implica que:

$$P(\omega : |X_n(\omega) - X(\omega)| \rightarrow 0, \text{ cuando } n \rightarrow \infty) = 1. \quad (2.18)$$

2° L^2 o *convergencia media cuadrática* de X_n a X , implica que:

$$\|X_n - X\|_2 \equiv [E |X_n - X|^2]^{\frac{1}{2}} \rightarrow 0, \text{ cuando } n \rightarrow \infty, \quad (2.19)$$

y finalmente,

3° X_n converge a una variable X en Probabilidad o X_n *converge estocásticamente* a X si, para cualquier $\epsilon > 0$,

$$P(\omega : |X_n(\omega) - X(\omega)| \geq \epsilon) \rightarrow 0, \text{ cuando } n \rightarrow \infty. \quad (2.20)$$

La expresión en (2.20), es equivalente a:

$$E \left(\frac{|X_n - X|}{1 + |X_n - X|} \right) \rightarrow 0, \text{ cuando } n \rightarrow \infty. \quad (2.21)$$

y define una métrica o una función de la distancia entre las variables aleatorias. La convergencia estocástica es el más débil de los tres modos de convergencia:

$$(2.18) \Rightarrow (2.20) \Leftarrow (2.19).$$

Cuando $|X_n| \leq Y$, para algún $Y \in L^2(\Omega, \mathcal{F}, P)$, (2.18) resulta ser el más fuerte de los modos de

convergencia y los dos restantes resultan ser equivalentes:

$$(2.18) \Rightarrow (2.20) \Leftrightarrow (2.19) \Leftarrow (2.18),$$

la implicación $(2.18) \Rightarrow (2.19)$, bajo las condiciones dadas, se conoce como el *teorema de convergencia dominada*.

2.2 Procesos Estocásticos

Definición 2.20 Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y T un conjunto arbitrario. Una familia de variables aleatorias

$$X(t, \omega) : \omega \in \Omega, t \in T, \quad (2.22)$$

es llamada un proceso estocástico con índice T . Note que $\forall t \in T$ fija,

$$X(t) = X(t, \cdot) \quad (2.23)$$

define una variable aleatoria sobre el espacio de probabilidad dado y que $\forall \omega$ fija implica que

$$X(\cdot, \omega), \quad (2.24)$$

sea una función valuada realmente definida sobre T . A esta función también se le conoce como *trayectoria, función muestra o realización del proceso*.

Si $\Omega = \mathbb{R}, \mathbb{R}^n$ o $\mathbb{R}^{n \times m}$ y $T \subset \mathbb{R}$, entonces, $X(t, \omega) \in \mathbb{R}, \mathbb{R}^n$ o $\mathbb{R}^{n \times m}$ respectivamente, es un proceso estocástico con dominio de tiempo T . En forma particular si:

- $T = \{1, 2, \dots\}$, llamaremos a $X(\xi_1, \xi_2, \dots)$, un proceso estocástico de tiempo discreto.
- $T = [0, 1], (-\infty, \infty), [0, \infty), \dots$, llamaremos a $X(\xi_t), t \in T$, un proceso estocástico de tiempo continuo.

2.2.1 Procesos Estacionarios con Incrementos Independientes

Definición 2.21 Se dice que un proceso estocástico $X(t)$ es estrictamente estacionario si sus distribuciones de dimensión finitas son invariantes bajo desplazamiento del tiempo, es decir,

$\forall t \in \mathbb{R}$, tal que $t_j, t_j + t \in T$, y $j \in \mathbb{N}^+$:

$$F_{t_1+t, \dots, t_n+t}(x_1, \dots, x_n) = F_{t_1, \dots, t_n}(x_1, \dots, x_n) \quad (2.25)$$

Si además, $\forall t \in \mathbb{R}$, $X(t) \in \mathbb{L}^2$, entonces $\forall s, t$

$$E(X(t)) = \mu \quad (2.26)$$

y

$$\text{Cov}(X(t), X(s)) = C(t - s) \quad (2.27)$$

donde μ es una constante y C es una función que depende únicamente de la diferencia $t - s$.

Definición 2.22 Al proceso \mathbb{L}^2 que satisface (2.26) y (2.27), se le llama proceso ampliamente estacionario.

Definición 2.23 Un Proceso con incrementos independientes se define $\forall \{t_j\} \subseteq T$, con $t_j < t_{j+1}$, como un proceso con diferencias $X(t_{j+1}) - X(t_j)$ independientes.

Definición 2.24 Si $X(t)$, es un proceso aditivo \mathbb{L}^2 continuo con incrementos independientes, entonces, $\forall s < t \in T$, $X(t) - X(s)$, son variables aleatorias Gaussianas o normales y si $X(t_0)$, para algún $t_0 \in T$, tiene distribución normal, entonces $X(t)$, es un proceso Gaussiano.

Definición 2.25 Sea $T = [0, \infty]$. Un proceso escalar \mathbb{L}^2 continuo $X(t)$ con índice T , que tiene incrementos independientes estacionarios y que satisface:

- i) $P(X(0) = 0) = 1$,
- ii) $E(X(t)) = 0, \forall t \in T$,
- iii) $V(X(t)) = t, \forall t \in T$,

es llamado un movimiento Browniano (estándar), o proceso de Wiener. Se denotará a este como $W(t)$.

Nota 2.3 Si $X(t)$, es un proceso estocástico continuo c.p.1, $X(0) = 0, \forall t \geq 0$, satisfaciendo:

$$\mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}, \text{ si } t_1 < t_2,$$

y si además $\forall t > s \geq 0$,

- i) $X(t)$ es \mathcal{F}_t -medible,
- ii) $E([X(t) - X(s)] | \mathcal{F}_s) = 0$, c.p.1,
- iii) $E([X(t) - X(s)]^2 | \mathcal{F}_s) = t - s$, c.p.1.

Entonces, $X(t)$, es un proceso de Wiener

Definición 2.26 Se le llama ruido blanco Gaussiano, a todo proceso Gaussiano estacionario escalar $\mathcal{N}(t)$ para $-\infty < t < \infty$ con $E(\mathcal{N}(t)) = 0$, y una función de densidad espectral $f(\lambda)$ constante sobre el conjunto de los reales, esto es, si $C(t) = E(\mathcal{N}(s)\mathcal{N}(t+s))$, entonces,

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j\lambda t} C(t) dt = \frac{K}{2\pi}, \quad \forall \lambda \in \mathbb{R}, \quad (2.28)$$

para alguna constante K .

El ruido blanco Gaussiano, es también conocido simplemente como ruido blanco por analogía con la luz blanca, la cual está compuesta de diferentes frecuencias con las mismas intensidades.

Nótese que (2.28), implica que la función de covarianza $Cov(t) = \delta(t)$, donde $\delta(t)$ es la función delta Dirac; por lo que el proceso no está correlacionado para tiempos distintos de cero, por lo tanto, es independiente en cada tiempo puesto que es Gaussiano. En particular

$$Cov(0) = \int_{-\infty}^{\infty} f(\lambda) d\lambda = \infty.$$

La naturaleza de la covarianza $Cov(t)$, indica que este proceso no puede ser realizado, y que, el ruido blanco no es un proceso estocástico en el sentido usual. La relación entre el ruido blanco Gaussiano $\mathcal{N}(t)$ y un proceso estándar de Wiener $W(t)$, puede entenderse formalmente considerando la covarianza de este último

$$Cov(W(t, s)) = E(W(t)W(s)) = \min(t, s).$$

Cuando existe la covarianza del proceso derivativo satisfaciendo

$$Cov\left(\frac{dW(t, s)}{dt}\right) = \frac{\partial^2}{\partial s \partial t} Cov(W(t, s)),$$

entonces, en este caso

$$Cov\left(\frac{dW(t, s)}{dt}\right) = \frac{\partial^2}{\partial s \partial t} \min(t, s) = \delta(t - s); \quad (2.29)$$

es decir, la covarianza de la derivada de un proceso de Wiener es la covarianza del ruido blanco.

Definición 2.27 Sea $X(t)$ un proceso estocástico definido sobre un espacio de estado \mathbb{R}^n con índice $J = [t_0, T] \subseteq [0, \infty)$. Para $\tau_1, \tau_2 \in J$, tal que $\tau_1 \leq \tau_2$, defina

$$\mathcal{F}([\tau_1, \tau_2]) = \mathcal{F}(X(\tau), \tau_1 \leq \tau \leq \tau_2)$$

donde $\mathcal{F}([\tau_1, \tau_2])$ es generada por todos los conjuntos de la forma

$$\{\omega : X(s_1, \omega) \in B_1, \dots, X(s_m, \omega) \in B_m\} = \{X(s_1) \in B_1, \dots, X(s_m) \in B_m\},$$

$\tau_1 \leq s_1 < s_2 < \dots < s_m \leq \tau_2$, $B_i \in \mathcal{B}^n$, con m un número entero positivo. $X(t)$ es llamado un proceso de Markov, si satisface que: para $n \geq 1$, $t_0 \leq t_1 < \dots < t_n < t \leq T$ y $B \in \mathcal{B}^n$,

$$P(X(t) \in B \mid X(t_1), \dots, X(t_n)) = P(X(t) \in B \mid X(t_n)).$$

2.2.2 Integral de Ito

Para fines de definición considere a $W(t)$ un proceso de Wiener estándar definido sobre un espacio de probabilidad completo (Ω, \mathcal{F}, P) , y sea $\{\mathcal{F}(t) : t \in [a, b]\}$, una familia de sub- σ -álgebras de \mathcal{F} satisfaciendo las condiciones dadas en subsección 2.1.2.

Sea \mathcal{L}^2 , la clase de todas las funciones aleatorias $f(t, \omega)$, satisfaciendo:

- A) f es medible sobre $[a, b] \times \Omega$.
- B) $f(t, \cdot)$, es $\mathcal{F}(t)$ -medible, $\forall t \in [a, b]$.
- C) $\int_a^b E\{f^2(t, \omega)\} dt < \infty$.

Sea \mathcal{E} , la subclase de \mathcal{L}^2 consistente de todas la funciones aleatorias que además satisfacen:

- D) existe una partición $a = t_0 < t_1 < \dots < t_n = b$, tal que $f(t, \cdot) = f(t_i, \cdot)$, $t_i \leq t < t_{i+1}$.

El espacio \mathcal{L}^2 equipado con la norma:

$$\|f\| = \left[\int_a^b E\{f^2(t, \omega)\} dt \right]^{\frac{1}{2}},$$

forma un espacio real de *Hilbert*. Así, considere a $\mathbb{L}^2(\Omega)$, el espacio usual de Hilbert de funciones aleatorias F sobre (Ω, \mathcal{F}, P) , con:

$$\|F\|_2 = \{E|F|^2\}^{\frac{1}{2}} < \infty.$$

Definición 2.28 Sea el mapeo lineal $\mathcal{I} : \mathcal{L}^2 \rightarrow \mathbb{L}^2$ sobre \mathcal{E} dado por:

$$I(f) = \sum_{k=1}^m f(t_{k-1}) \Delta W_k,$$

donde $f \in \mathcal{E}$, tal que, $f(t) = f(t_{k-1})$, $t_{k-1} \leq t < t_k$, y $\Delta W_k = [W(t_k) - W(t_{k-1})]$. A $\mathcal{I}(f)$ se le conoce como la integral de Ito y se denota por:

$$\mathcal{I}(f) = \int_a^b f(t) dW(t). \quad (2.30)$$

Nota 2.4 En la definición anterior $f \in \mathcal{E}$, pero para el caso de una función medible $f \in \mathcal{L}^2$, se requiere elegir $\{f_n\} \subseteq \mathcal{E}$ tal que $f_n \rightarrow f$, c.p.1. Como $\{f_n\} \in \mathcal{L}^2$ es una secuencia de Cauchy, entonces, por linealidad y continuidad de \mathcal{I} , la secuencia $\{\mathcal{I}(f_n)\}$ es una secuencia de Cauchy en el espacio completo \mathbb{L}^2 , por lo tanto, la definición de la integral de Ito se puede extender para el espacio completo \mathcal{L}^2 donde resulta que:

$$\mathcal{I}(f) = \lim_{n \rightarrow \infty} \mathcal{I}(f_n) = \lim_{n \rightarrow \infty} \sum_{k=1}^m f_n(t_{k-1}) \Delta W_k,$$

esto implica que,

$$\mathcal{I}(f) = \lim_{n \rightarrow \infty} \int_a^b f_n(t) dW(t), \quad (2.31)$$

la cual, se denotará igual que (2.30).

Propiedades de la integral de Ito:

- Ia)** $E(\mathcal{I}(f)) = 0$.
- Ib)** $\|\mathcal{I}(f)\|_2 = \|f\|$.

2.2.3 Fórmula de Ito

Sea \mathcal{L}_s la clase de funciones aleatorias f las cuales satisfacen las condiciones (A) y (B) para las clases \mathcal{L}^2 y que además satisface:

- E)** $\int_a^b f^2(t, \omega) dt < \infty$, c. p.1.

Si f y g son funciones aleatorias con $g \in \mathcal{L}_s$ y f satisface (A), (B) y (E), entonces para $a \leq t \leq b$,

la ecuación:

$$X(t) = X(a) + \int_a^t f(s)ds + \int_a^t g(s)dW(s), \quad (2.32)$$

define un proceso estocástico con trayectoria continua y probabilidad 1 (ver [13]), donde la primera integral del lado derecho de (2.32), es una integral ordinaria en el sentido de Lebesgue [10], (la cual, se define posteriormente en la Sección 2.3), de límite superior que satisface (E), y la segunda integral es la integral de Ito.

Definición 2.29 *Sea la ecuación (2.32), entonces, $X(t)$ tiene diferencial estocástico:*

$$dX(t) = f(t)dt + g(t)dW(t). \quad (2.33)$$

La ecuación (2.33), caracteriza cualquier proceso cuyos incrementos $X(t) - X(s)$, $t > s$, están dados por:

$$X(t) - X(s) = \int_s^t f(r)dr + \int_s^t g(r)dW(r).$$

Si f y g no son aleatorias, los intervalos disjuntos son independientes y Gaussianos con:

$$E[X(t) - X(s)] = \int_s^t f(r)dr,$$

y

$$V[X(t) - X(s)] = \int_s^t g^2(r)dr.$$

Teorema 2.1 *Suponga que $X(t)$ tiene diferencial estocástico (2.33). Si $F(t, x)$ es una función determinística valuada realmente definida para toda $t \in [a, b]$, $x \in \mathbb{R}$, con derivadas parciales continuas $\partial F/\partial t$, $\partial F/\partial x$, y $\partial^2 F/\partial x^2$, entonces el proceso $F(t, X(t))$ tiene un diferencial estocástico sobre $[a, b]$ dado por:*

$$dF(t, X(t)) = \tilde{f}(t, X(t))dt + \tilde{g}(t, X(t))dW(t), \quad (2.34)$$

donde

$$\tilde{f}(t, x)dt = \left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x}f + \frac{1}{2} \frac{\partial^2 F}{\partial x^2}g^2 \right] (t, x),$$

y

$$\tilde{g}(t, x) = \left[\frac{\partial F}{\partial x} g \right] (t, x).$$

Demostración: Sea $Y(t) = F(t, X(t))$. Suponga que (2.34) se cumple, entonces se demostrará que:

$$\begin{aligned} Y(t) - Y(s) &= \int_s^t \left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x} f + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} g^2 \right] (u, X(u)) du \\ &\quad + \int_s^t \left[\frac{\partial F}{\partial x} g \right] (u, X(u)) dW(u). \end{aligned} \quad (2.35)$$

El resultado se prueba primero para el caso de f y g constantes. Sea $s = t_0 < \dots < t_n = t$, y denote el incremento $h(t_{k+1}) - h(t_k)$, para cualquier función h , por Δh_k . Entonces se puede escribir

$$Y(t) - Y(s) = F(t, X(t)) - F(s, X(s)) = \sum_{k=0}^{n-1} \Delta F_k.$$

Del teorema de Taylor (ver pag. 212 de [14]), $\exists \alpha_k, \beta_k \in [0, 1]$, tal que

$$\begin{aligned} \Delta F_k &= \frac{\partial F}{\partial t}(t_k + \alpha_k \Delta t_k, X(t_k)) \Delta t_k + \frac{\partial F}{\partial x}(t_k, X(t_k)) \Delta X_k \\ &\quad + \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(t_k, X(t_k) + \beta_k \Delta X_k) (\Delta X_k)^2. \end{aligned} \quad (2.36)$$

Por la continuidad de $\partial F/\partial t$ y $\partial^2 F/\partial x^2$, y la continuidad de $X(t)$, convergiendo c.p.1, se obtiene que el $\max \Delta X_k \rightarrow 0$, y que

$$\frac{\partial F}{\partial t}(t_k + \alpha_k \Delta t_k, X(t_k)) \rightarrow \frac{\partial F}{\partial t}(t_k, X(t_k)), \text{ c.p.1,}$$

del mismo modo

$$\frac{\partial^2 F}{\partial x^2}(t_k, X(t_k) + \beta_k \Delta X_k) \rightarrow \frac{\partial^2 F}{\partial x^2}(t_k, X(t_k)), \text{ c.p.1.}$$

Con f y g constantes, $\Delta X_k = f \Delta t_k + g \Delta W_k$, por lo tanto

$$\sum_{k=0}^{n-1} (\Delta X_k)^2 - (g \Delta W_k)^2 = f^2 \sum_{k=0}^{n-1} (\Delta t_k)^2 + 2fg \sum_{k=0}^{n-1} \Delta W_k \Delta t_k. \quad (2.37)$$

Ambas sumas del lado derecho de (2.37) tienden a cero cuando $\max \Delta t_k \rightarrow 0$. De las ecuaciones (2.36) y (2.37) se obtiene que:

$$\begin{aligned}
Y(t) - Y(s) &= \lim_{\max \Delta t_k \rightarrow 0} \sum_{k=0}^{n-1} \Delta F_k \quad (2.38) \\
&= \lim_{\max \Delta t_k \rightarrow 0} \sum_{k=0}^{n-1} \left\{ \left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x} f + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} g^2 \right] (t_k, X(t_k)) \right\} \Delta t_k \\
&\quad + \lim_{\max \Delta W_k \rightarrow 0} \sum_{k=0}^{n-1} \left\{ \left[\frac{\partial F}{\partial x} g \right] (t_k, X(t_k)) \right\} \Delta W_k \\
&\quad + \lim_{\substack{\max \Delta t_k \rightarrow 0 \\ \max \Delta W_k \rightarrow 0}} \sum_{k=0}^{n-1} \left\{ \left[\frac{1}{2} \frac{\partial^2 F}{\partial x^2} g^2 \right] \right\} \{(\Delta W_k)^2 - \Delta t_k\}
\end{aligned}$$

Los primeros dos límites del lado derecho de (2.37) son los términos sobre el lado derecho de (2.35).

Sea $\gamma_k = (\Delta W_k)^2 - \Delta t_k$ y sea $I_k^{(N)}$ la función indicador (Pag. 4, [13]), del evento

$$\{|X(t_i)|\} \leq N, \forall i \leq k.$$

Ahora, dado que γ_k son independientes, $E(\gamma_k) = 0$, y $E(\gamma_k^2) = 2(\Delta t_k)^2$, se obtiene que

$$\begin{aligned}
E \left(\sum_{k=0}^{n-1} \frac{\partial^2 F}{\partial x^2} (t_k, X(t_k)) I_k^{(N)} \gamma_k \right)^2 &= \sum_{k=0}^{n-1} E \left(\frac{\partial^2 F}{\partial x^2} (t_k, X(t_k)) I_k^{(N)} \gamma_k \right)^2 \\
&\leq \sup_{\substack{s < u < t \\ |x| \leq N}} \left(\frac{\partial^2 F}{\partial x^2} (u, x) \right)^2 \sum_{k=0}^{n-1} 2(\Delta t_k)^2 \rightarrow 0,
\end{aligned}$$

cuando $\max \Delta t_k \rightarrow 0$, esto establece que el límite en (2.37) tiende a cero y

$$P(|X(t_i)| \leq N, \forall i \leq k) \rightarrow 1 \text{ cuando } N \rightarrow \infty. \quad (2.39)$$

Lo anterior se puede probar si definimos

$$Q_N = \left\{ \sup_{s \leq u \leq t} |X(u)| > N \right\};$$

entonces, $P(Q_N) \rightarrow 0$ cuando $N \rightarrow \infty$. Por lo tanto, (2.39) se satisface puesto que

$$\left\{ I_k^{(N)} = 0, \text{ para algún } k \right\} = \bigcup_k \{|X(t_i)| > N, \text{ para algún } i < k\} \subseteq Q_N;$$

esto completa la prueba para el caso en que f y g sean funciones constantes (escalón).

Para el caso general, suponga que $\{f_n(u)\}$ y $\{g_n(u)\}$ son secuencias de funciones escalón tales que, en probabilidad

$$\int_s^t |f_n(u) - f(u)| du \rightarrow 0 \text{ y } \int_s^t |g_n(u) - g(u)|^2 du \rightarrow 0.$$

Entonces la secuencia

$$X_n(u) = X(s) + \int_s^u f_n(r) dr + \int_s^u g_n(r) dW(r)$$

converge en probabilidad a $X(u)$ cuando $n \rightarrow \infty$. Sin pérdida de generalidad, considerando que en cada subsecuencia la convergencia en probabilidad puede ser reemplazada por convergencia con probabilidad 1, y más aún que $X_n \rightarrow X$ uniformemente sobre $[s, t]$. La formula de Ito satisfaciendo funciones escalón (constantes), como se mostró anteriormente, establece que $\forall n$,

$$\begin{aligned} Y_n(t) - Y_n(s) &= F(t, X_n(t)) - F(s, X_n(s)) \\ &= \int_s^t \left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x} f_n + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} g_n^2 \right] (u, X_n(u)) du \\ &\quad + \int_s^t \left[\frac{\partial F}{\partial x} g_n \right] (u, X_n(u)) dW(u). \end{aligned} \quad (2.40)$$

El lado derecho de (2.39) converge con probabilidad 1 para $Y(t) - Y(s)$. Puesto que $X_n \rightarrow X$, $f_n \rightarrow f$, y $g_n \rightarrow g$ en probabilidad, la desigualdad del triángulo para cada argumento en el sentido de convergencia estocástica produce:

$$\left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x} f_n + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} g_n^2 \right] (u, X_n(u)) \rightarrow \left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x} f + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} g^2 \right] (u, X(u)) \quad (2.41)$$

y

$$\left[\frac{\partial F}{\partial x} g_n \right] (u, X_n(u)) \rightarrow \left[\frac{\partial F}{\partial x} g \right] (u, X(u)). \quad (2.42)$$

De hecho, dada la posibilidad de considerar subsecuencias, (2.41) se puede considerar que converge con probabilidad 1. Puesto que X es continua con probabilidad 1, esta es acotada con probabilidad 1, además, debido a que todas las funciones en (2.41) son subsecuentemente acotadas, el teorema de convergencia dominada, puede ser aplicado sobre cada trayectoria continua para obtener que la

primera integral en (2.39) tiende a

$$\int_s^t \left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x} f + \frac{1}{2} \frac{\partial^2 F}{\partial x^2} g^2 \right] (u, X(u)) du,$$

con probabilidad 1. Similarmente, de (2.42)

$$\int_s^t \left[\frac{\partial F}{\partial x} g_n \right] (u, X_n(u)) du \rightarrow \int_s^t \left[\frac{\partial F}{\partial x} g \right] (u, X(u)) du$$

con probabilidad 1, lo cual implica convergencia en el sentido \mathcal{L}_s . Consecuentemente, la segunda integral en (2.39) tiende a

$$\int_s^t \left[\frac{\partial F}{\partial x} g \right] (u, X(u)) dW(u).$$

en probabilidad, y una vez más es posible considerar subsecuencias, las que convergen con probabilidad 1. Así, el lado derecho de (2.39) tiende a (2.35). ■

Para el caso en que $X(t)$ sea m -dimensional, su diferencial estocástico estará dado por:

$$dX(t) = f(t)dt + G(t)dW(t)$$

donde $W(t) = (W_1(t), \dots, W_m(t))'$ es un proceso de Wiener m -dimensional con todos sus componentes independientes, $f(t) = (f_1(t), \dots, f_m(t))'$ es una función vectorial n -dimensional con componentes en \mathbb{L}^1 y c.p.1, y $G(t) = \{g_{ij}(t)\}$ es una función matricial $n \times m$ -dimensional con sus componentes g_{ij} en \mathcal{L}_s . Para la demostración vea la referencia [13].

2.3 Cálculo Avanzado

2.3.1 Integral de Lebesgue

Definición 2.30 *Sea f una función simple, es decir, una función contable μ -medible [10] tomando valores distintos*

$$y_1, y_2, \dots, y_n, \dots$$

La integral de Lebesgue de f sobre un conjunto A , denotada por:

$$\int_A f(x) d\mu, \tag{2.43}$$

se define como:

$$\sum_n y_n \mu(A_n), \quad (2.44)$$

donde:

$$A_n : \{x : x \in A, f(x) = y_n\}$$

es tal que (2.44) es absolutamente convergente, es decir,

$$\sum_n |y_n \mu(A_n)| < \infty.$$

Si la integral de Lebesgue de f existe, entonces f es una función integrable o sumable (con respecto a la medida μ), sobre el conjunto A .

Definición 2.31 Para una función medible f , no necesariamente simple, se dice que f es integrable o sumable, sobre un conjunto A si existe una secuencia $\{f_n\}$ de funciones integrables simples convergiendo uniformemente a f sobre A . El límite

$$\lim_{n \rightarrow \infty} \int_A f_n(x) d\mu, \quad (2.45)$$

es llamado la integral de Lebesgue de f sobre A , denotado de igual forma que en (2.43).

Sea f una función sumable definida sobre un espacio X , equipado con una medida σ -aditiva μ (ver la ref: [10]). Entonces, (2.43) existe $\forall A \subset X$ medible. Si $X \subset \mathbb{R}$, equipado con una medida ordinaria de Lebesgue μ , y si $A = [a, b]$, entonces es posible expresar a (2.43), como:

$$\int_a^b f(x) dx,$$

o equivalentemente

$$\int_a^b f(t) dt, \quad (2.46)$$

en términos de una nueva variable de integración limitada t . Suponga ahora que fijamos el límite inferior a , y que el límite superior b es variable haciendo notar esto por la sustitución de b por la de x . Entonces, (2.46) se reduce a:

$$\int_a^x f(t) dt, \quad (2.47)$$

expresión a la que se le conoce como la *integral indefinida de Lebesgue*.

propiedades de la integral de Lebesgue indefinida, si f es una función sumable sobre $[a, b]$:

- a) La $\frac{d}{dx} \int_a^x f(t)dt$, existe y es finita, para toda x ,
- b) $\frac{d}{dx} \int_a^x f(t)dt = f(x)$, se satisface, excepto para conjuntos de medida cero (ver [10]).

2.3.2 Función de Variación Acotada

Definición 2.32 Una función f definida sobre un intervalo $[a, b]$, se dice es de variación acotada, si existe una constante $C > 0$ tal que:

$$\sum_{k=1}^n |f(x_k) - f(x_{k-1})| \leq C \quad (2.48)$$

para toda partición

$$a = x_0 < x_1 < \dots < x_n = b \quad (2.49)$$

del intervalo $[a, b]$ por puntos de subdivisión x_0, x_1, \dots, x_n .

Definición 2.33 Sea f una función de variación acotada. Entonces la variación total de f sobre $[a, b]$, está dada por

$$V_a^b(f) \equiv \sup \sum_{k=1}^n |f(x_k) - f(x_{k-1})| \quad (2.50)$$

donde la mínima cota superior es tomada sobre todas las posibles particiones (finitas) (2.49) del intervalo $[a, b]$.

Algunas propiedades de las funciones de variación acotadas son:

- a) Si $a < b < c$, entonces, $V_a^c(f) = V_a^b(f) + V_b^c(f)$.
- b) La función $v(x) = V_a^x(f)$, es no decreciente.
- c) Si f es de variación acotada sobre $[a, b] \Rightarrow \exists v, g$ funciones no decrecientes, tales que: $f = v(x) - g(x)$.
- d) Si f es sumable sobre $[a, b]$, entonces la integral indefinida $\Phi(x) = \int_a^x f(t)dt$, es una función de variación acotada sobre $[a, b]$.

2.3.3 Ecuación de Ito-Volterra

En la sección anterior se habló de los inicios de la ecuación integral de Volterra. También, se menciono que ésta puede ser o no reducida al caso de una ecuación diferencial. Cuando es posible expresar el problema en forma de ecuaciones diferenciales, la solución del mismo puede ser calculada por los métodos convencionales ya conocidos, por ahora, se presentan las características de aquellas ecuaciones de Volterra que no pueden ser llevadas a la forma de una ecuación diferencial.

Definición 2.34 Se conoce como ecuación de Volterra a la ecuación integral

$$y(x) = f(x) + \lambda \int_a^x K(x, \xi) y(\xi) d\xi, \quad (2.51)$$

donde el límite superior de integración es variable (similar a (2.47)), y corresponde a un problema de valor inicial, x y ξ son variables independientes, λ es una constante, $K(x, \xi)$ es una función dada y es llamada el Kernel o tasa de crecimiento, $f(x)$ es un término no homogéneo. En general $f(x)$, $K(x, \xi)$ y la solución $y(x)$ son funciones continuas en el intervalo asociado.

Definición 2.35 Se conoce como ecuación de Fredholm, a la ecuación integral

$$y(x) = f(x) + \lambda \int_a^b K(x, \xi) y(\xi) d\xi, \quad (2.52)$$

donde la integral está definida sobre el intervalo $[a, b]$, y corresponde a un problema de valor en la frontera. En esta ecuación los términos se definen en la misma forma que los términos de (2.51).

Cada ecuación ((2.51) o (2.52)), puede ser clasificada dependiendo de que si aparece o no el término $y(x)$ en el lado izquierdo de la ecuación, es decir,

$$0 = f(x) + \lambda \int_a^b K(x, \xi) y(\xi) d\xi,$$

es una ecuación de Fredholm de *primera clase*. En este caso, la función desconocida aparece únicamente bajo el signo de la integral. Si la función desconocida está presente afuera de la integral, tal y como en (2.52), se dice que la ecuación es de *segunda clase*.

La solución general de la ecuación de Fredholm se establece en [21], y es a partir de ésta que se puede conocer un método de solución para la ecuación de Volterra. Otro método para conocer la solución de una ecuación de Volterra se presenta a continuación.

Considere la ecuación de Volterra de segunda clase (2.51), con kernel integrable cuadrado, tal que

$$\int_0^{\infty} \int_0^{\infty} K^2(x, y) dx dy \leq N$$

existe y es acotada. Asuma que existe algún tipo de solución para la serie

$$y(x) = \sum_{n=0}^{\infty} \lambda^n \phi_n(x). \quad (2.53)$$

Substituyendo (2.53) dentro de (2.51), se obtiene que

$$\sum_{n=0}^{\infty} \lambda^n \phi_n(x) = f(x) + \sum_{n=0}^{\infty} \lambda^{n+1} \int_0^x K(x, \xi) \phi_n(\xi) d\xi. \quad (2.54)$$

Asociando términos del mismo orden, resulta que:

$$\begin{aligned} \lambda^0, \phi_0(x) &= f(x), \\ \lambda^1, \phi_1(x) &= \int_0^x K(x, \xi) \phi_0(\xi) d\xi \\ &= \int_0^x K_1(x, \xi) f(\xi) d\xi, \\ \lambda^2, \phi_2(x) &= \int_0^x K(x, \xi_1) \phi_1(\xi_1) d\xi_1 \\ &= \int_0^x K_1(x, \xi_1) d\xi_1 \int_0^{\xi_1} K_1(\xi_1, \xi) f(\xi) d\xi \\ &= \int_{\xi_1=0}^x \int_{\xi=0}^{\xi_1} K_1(x, \xi_1) K_1(\xi_1, \xi) f(\xi) d\xi_1 d\xi \\ &= \int_0^x K_2(x, \xi) f(\xi) d\xi, \\ &\vdots \\ \lambda^n, \phi_n(x) &= \int_0^x K_n(x, \xi) f(\xi) d\xi, \end{aligned}$$

donde

$$\begin{aligned} K_2(x, \xi) &= \int_{\xi_1=\xi}^x K_1(x, \xi_1) K_1(\xi_1, \xi) d\xi_1, \\ &\vdots \\ K_n(x, \xi) &= \int_{\xi_1=\xi}^x K_1(x, \xi_1) K_{n-1}(\xi_1, \xi) d\xi_1, \end{aligned}$$

por lo tanto,

$$y(x) = f(x) + \sum_{n=1}^{\infty} \lambda^n \int_0^x K_n(x, \xi) f(\xi) d\xi.$$

Con

$$K_1(x, \xi) = K(x, \xi),$$

se puede definir un nuevo kernel $\Gamma(x, \xi; \lambda)$, llamado el Kernel Revolvente:

$$\Gamma(x, \xi; \lambda) = \sum_{n=1}^{\infty} \lambda^{n-1} K_n(x, \xi).$$

En términos del kernel revolvente, la solución se puede escribir como:

$$y(x) = f(x) + \lambda \int_0^x \Gamma(x, \xi; \lambda) f(\xi) d\xi.$$

A esta solución se le conoce como la solución de series de *Neumann* para la ecuación de Volterra de segunda clase.

Para probar si es posible que la solución de Neumann válida para (2.51), también lo es para el caso de una ecuación de Volterra de primera clase, considere la ecuación

$$f(x) = \int_0^x K(x, \xi) y(\xi) d\xi, \quad (2.55)$$

la cual es una ecuación de Volterra de primera clase. Diferenciando (2.55) con respecto a x :

$$\dot{f}(x) = K(x, x)y(x) + \int_0^x \frac{\partial K(x, \xi)}{\partial x} y(\xi) d\xi.$$

Sea $K(x, x) \neq 0$, en el intervalo $(0, x)$, tal que se puede definir:

$$F(x) = \frac{\dot{f}(x)}{K(x, x)}; \quad \bar{K}(x, \xi) = -\frac{1}{K(x, x)} \frac{\partial K(x, \xi)}{\partial x}.$$

Entonces, (2.55), se puede escribir como:

$$y(x) = F(x) + \int_0^x \bar{K}(x, \xi) y(\xi) d\xi,$$

la cual, es una ecuación de Volterra de segunda clase y puede ser resuelta por el método mencionado anteriormente. Otros tópicos sobre este tema, se pueden consultar en [15].

Ejemplo 2.2 Considere la ecuación integral

$$y(x) = x - \int_0^x (x - \xi)y(\xi)d\xi,$$

donde $\lambda = -1$; $K(x, \xi) = x - \xi$ y $f(x) = x$. Del procedimiento anterior se obtiene que

$$\begin{aligned}\phi_0(x) &= x, \\ \phi_1(x) &= \int_0^x (x - \xi)\xi d\xi = \left(\frac{x\xi^2}{2} - \frac{\xi^3}{3} \right) \Big|_0^x = \frac{x^3}{6}, \\ \phi_2(x) &= \int_0^x (x - \xi)\frac{\xi^3}{6} d\xi = \frac{1}{6} \left(\frac{x\xi^4}{4} - \frac{\xi^5}{5} \right) \Big|_0^x = \frac{x^5}{120},\end{aligned}$$

con lo que la solución

$$y(x) = x - \frac{x^3}{6} + \frac{x^5}{120} + \dots,$$

resulta ser, en este caso particular, una expansión de $\sin x$. ■

Definición 2.36 Sean las ecuaciones (2.30), (2.31) y (2.51). Sea la combinación de (2.30) y (2.51), o de (2.31) y (2.51); representada por:

$$y(t) = x(t) + \int_a^t K(t, s)y(s)dW(s), \quad (2.56)$$

entonces, a (2.56), se le llama ecuación de Ito-Volterra. Nótese que (2.56), preserva todas las propiedades mencionadas anteriormente para los procesos estocásticos, integrales de Lebesgue indefinidas y funciones de variación acotada.

Capítulo 3

Filtrado de Kalman

3.1 Filtro de Wiener

El filtrado de Wiener está ligado principalmente a dos ideas importantes: sistemas dinámicos y estimación óptima en presencia de ruido. La cantidad a estimar es un objetivo móvil que evoluciona con el sistema dinámico.

3.1.1 Planteamiento del Problema

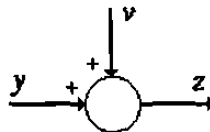


Figura 3.1: Modelo de Señal para el Filtro de Wiener

Considere el arreglo de la figura 3.1, en el cual, se describe una señal $y(\cdot)$, un ruido contaminante $v(\cdot)$ y una medición $z(\cdot)$. El problema se puede plantear para el caso en tiempo continuo o discreto, ya sea escalar o vectorial, por lo que sin pérdida de generalidad, considere el caso de señales continuas escalares definidas sobre el intervalo de tiempo $(-\infty, \infty)$. Asuma que $y(\cdot)$ y $v(\cdot)$ son funciones muestra (definición (2.24)), de procesos aleatorios estacionarios, independientes y con media cero. Asuma también que se conocen las funciones de correlación de las señales aleatorias o sus transformadas de Fourier llamadas espectros $\phi_{yy}(j\omega)$, $\phi_{vv}(j\omega)$, para toda $\omega \in \mathbb{R}$.

La tarea del filtro de Wiener es usar las mediciones $z(\cdot)$ para estimar $y(\cdot)$. Para tal fin, se requiere que la estimación:

- Sea *causal*, es decir, que $y(t)$ se estime usando $z(s)$ para toda $s < t$;
- $z(t)$, esté *disponible* para todo momento de estimación t ;
- $\hat{y}(t)$ sea *óptima*, es decir, deberá ofrecer el mínimo error medio cuadrático, tal que $E\{[y(t) - \hat{y}(t)]^2\}$ sea minimizado.

En el caso de que $y(\cdot)$ y $z(\cdot)$ sean Gaussianos, esto implica que $\hat{y}(t)$ es una media condicional estimada, dada por, $\hat{y}(t) = E[y(t) | z(s)], \forall s < t$.

3.1.2 Solución al Problema de Filtrado de Wiener

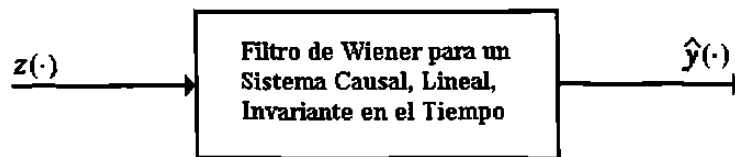


Figura 3.2: Estructura del Filtro de Wiener

La solución para este problema está ilustrado en la figura 3.2. Asuma que el bloque etiquetado como filtro de Wiener es un sistema lineal, causal, invariante en el tiempo y estable; descrito por una respuesta al impulso o por la función de transferencia $h(\cdot)$, de este modo, el mejor estimado estará dado por:

$$\hat{y}(t) = \int_{-\infty}^t h(t-s)z(s)ds, \quad (3.1)$$

por lo que la solución al problema se resume en aplicar un método para determinar $h(\cdot)$ a partir de los espectros conocidos. Uno de los procedimientos que llevan a la determinación de $h(\cdot)$, es conocido como *factorización espectral* [1]. Para tal caso, la condición necesaria y suficiente para que $\hat{y}(t)$ sea el estimado óptimo de $y(t)$, es que exista una función $A(t, s)$ que satisfaga la relación:

$$Cov[y(t)z'(s)] - \int_{t_0}^t A(t, s)Cov[z(t)z'(s)]ds = 0, \quad (3.2)$$

conocida como la ecuación de *Wiener-Hopf*, la cual, es una relación que involucra dos funciones:

- la función de *autocorrelación* de la medida $z(\cdot)$, es decir, $Cov[z(t)z'(s)]$, y,
- la función de *crosscorrelación* entre la señal del sistema $y(t)$ y la medida $z(\cdot)$, esto es, $Cov[y(t)z'(s)]$.

El procedimiento de factorización espectral no resulta ser de interés del presente trabajo, por lo que se sugiere consultar [1], y [22], o para el caso vectorial, se sugiere ver el capítulo 7 de [23].

La siguiente tabla relaciona algunas de las características importantes del problema de filtrado de Wiener y su solución.

Tabla 3.1: Suposiciones y Propiedades del filtro de Wiener

Tiempo inicial t_0	$t_0 = -\infty$.
Procesos Aleatorios	Necesariamente Estacionarios.
Señal $y(\cdot)$	<i>i</i>) Espectro $\phi_{yy}(\cdot)$, suficiente para resolver el problema
	<i>ii</i>) $y(\cdot)$ estacionaria.
	<i>iii</i>) $\phi_{yy}(\cdot)$ no necesariamente racional.
Ruido de Medición $v(\cdot)$	<i>i</i>) Usualmente independiente de $y(t)$.
	<i>ii</i>) Estacionario.
	<i>iii</i>) No necesariamente Blanco
	<i>iv</i>) $\phi_{vv}(\cdot)$ no necesariamente es racional.
Filtro de Wiener	Lineal invariante en el tiempo y estable pero no necesariamente con función de transferencia racional.
Dificultad en el cálculo	Factorización espectral.
Cantidad Estimada	$y(t)$.

3.2 Filtro de Kalman

Uno de los más celebres resultados en la teoría del filtrado es sin duda el filtro de Kalman [2]. A partir de la ecuación de Wiener-Hopf, Kalman mostró que la solución del problema de filtrado de Wiener se puede obtener desde el punto de vista de variables de estado, por lo que a través de esta sección, se tratará con sistemas dinámicos lineales discretos para los que en las referencias [17, 26], se dan las condiciones de estabilidad, controlabilidad y observabilidad entre otras propiedades para esta clase de sistemas. Otra de las herramientas utilizadas por Kalman, fue el concepto de ortogonalidad [2], el cual fue necesario para la conclusión del estimado óptimo por medio de mínimos cuadrados. Así, a continuación se describe brevemente la idea de las proyecciones ortogonales.

3.2.1 Proyecciones Ortogonales

El cálculo explícito del estimado óptimo como una función de variables observadas es en general imposible. Sin embargo existe una excepción importante cuando los procesos son Gaussianos.

Por otra parte, si intentamos obtener el estimado óptimo bajo la restricción de una función de pérdida [2] tal que $L(\epsilon) = \epsilon^2$, con el requerimiento adicional de que el estimado sea una función lineal de las variables observadas; se obtiene un estimado idéntico al obtenido en el caso Gaussiano sin la suposición de linealidad o de la función de pérdida cuadrática.

En estos casos, la solución explícita del problema de estimación se puede entender de modo más sencillo con la ayuda de la figura 3.3. Cabe mencionar que el principio de ortogonalidad para variables aleatorias aquí expuesto, tiene su importancia en el hecho de que se aplica a todos los problemas de filtrado tratados de aquí en adelante.

Considere el conjunto de variables valuadas realmente $y(t_0), \dots, y(t)$. Entonces el conjunto de todas las combinaciones lineales de tales variables aleatorias con coeficientes $\alpha_i \in \mathbb{R}$, $i = 0, 1, \dots$, con

$$\sum_{i=t_0}^t \alpha_i y(i), \quad (3.3)$$

forma un espacio vectorial o *variedad lineal* denotada por $\mathcal{Y}(t)$. Dado que $t \in \mathbb{R}$ no es fija, $\mathcal{Y}(t)$ debe considerarse como un subespacio de dimensión finita contenido en el espacio de todas las posibles observaciones, es decir, $\mathcal{Y}(t)$ es de potencia continua.

Dados dos vectores arbitrarios $u, v \in \mathcal{Y}(t)$, se dice que u es ortogonal a v (denotado por $u \perp v$), si $E(uv) = 0$. Seleccionando una base ortonormal [10], de vectores $e_{t_0}, e_{t_1}, \dots, e_t$ en $\mathcal{Y}(t)$, tal que, cualquier vector en $\mathcal{Y}(t)$ se puede representar como una combinación lineal única de $e_{t_0}, e_{t_1}, \dots, e_t$, y

$$E(e_i e_j) = \delta_{ij} = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases} \quad \forall i, j = t_0, \dots, t. \quad (3.4)$$

Así, cualquier vector $\bar{x} \in \mathcal{Y}(t)$ está dado por

$$\sum_{i=t_0}^t \alpha_i e_i,$$

entonces, los coeficientes α_i , se determinan por medio de la ecuación (3.4), esto es

$$E(\bar{x} e_j) = E\left(\sum_{i=t_0}^t \alpha_i e_i\right) e_j = \sum_{i=t_0}^t \alpha_i E(e_i e_j) = \alpha_j, \text{ si } i = j. \quad (3.5)$$

La ecuación (3.5) implica que cualquier variable aleatoria x (no necesariamente en $\mathcal{Y}(t)$), tiene una

única descomposición en:

- una parte $\bar{x} \in \mathcal{Y}(t)$, y,
- un componente $\tilde{x} \perp \mathcal{Y}(t)$.

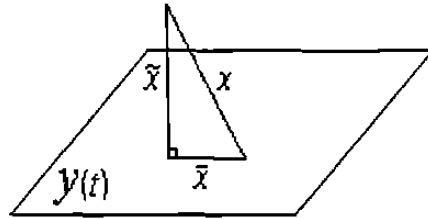


Figura 3.3: Descomposición Ortogonal de un Vector Aleatorio.

La figura 3.3, proporciona una clara idea de la descomposición del vector aleatorio x . Por lo que,

$$x = \bar{x} + \tilde{x} = \sum_{i=t_0}^t (E(xe_i))e_i + \tilde{x}. \quad (3.6)$$

Así, \bar{x} puede determinarse de un único modo por la ecuación (3.6) y obviamente es un vector en $\mathcal{Y}(t)$. Por lo tanto, \tilde{x} puede también ser determinada únicamente, con lo que solo resta probar que \tilde{x} es ortogonal a $\mathcal{Y}(t)$:

$$E(\tilde{x}e_i) = E(x - \bar{x})e_i = E(xe_i) - E(\bar{x}e_i),$$

note que las coordenadas de \bar{x} con respecto a la base $e_{t_0}, e_{t_1}, \dots, e_t$, están dadas en la forma $E(\bar{x}e_i)$ como en la ecuación (3.5), o en la forma $E(xe_i)$ como en (3.6). Puesto que las coordenadas son únicas, $E(xe_i) = E(\bar{x}e_i)$, $i = t_0, \dots, t$; por lo que $E(\tilde{x}e_i) = 0$ y \tilde{x} es ortogonal a cada vector base e_i y por lo tanto a $\mathcal{Y}(t)$.

Nota 3.1 Dada la ecuación (3.6), cuando $E(\tilde{x}\tilde{x}'|\mathcal{Y}(t))$ es minimizado, entonces, $\bar{x} = E(x|\mathcal{Y}(t))$ es el estimado óptimo de x con respecto a la base generada en $\mathcal{Y}(t)$.

3.2.2 Planteamiento del Problema

Considere el modelo del sistema de la figura 3.4, tal que, el comportamiento dinámico está descrito por:

$$x(k+1) = F(k)x(k) + G(k)u(k), \quad x(0) = x_0, \quad (3.7)$$

$$y(k) = H(k)x(k) + v(k), \quad k = 0, 1, \dots, \quad (3.8)$$

donde para $m \leq n$, $u(k) \in \mathbb{R}^n$, $v(k) \in \mathbb{R}^m$ son procesos Gaussianos independientes con media cero; $x(k) \in \mathbb{R}^n$, $y(k) \in \mathbb{R}^m$; y $F(k), G(k) \in \mathbb{R}^{n \times n}$, $H(k) \in \mathbb{R}^{p \times n}$ son matrices cuyos componentes son funciones del tiempo no aleatorias.

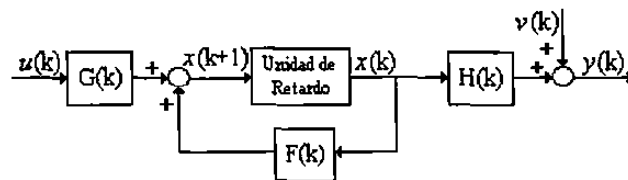


Figura 3.4: Modelo de un Sistema Lineal de Tiempo Discreto

De acuerdo a la nota 3.1, el problema de filtrado, consiste en encontrar el mejor estimado $\hat{x}(k+1 | k)$ de $x(k+1)$ dado el conjunto de valores observados $y(k_0), \dots, y(k)$, para $k_0 \leq l \leq k$, tal que $E((x(l) - \hat{x}(l))(x(l) - \hat{x}(l))' | k)$, sea minimizado, es decir, que el error $\tilde{x}(l | k) = (x(l) - \hat{x}(l) | k)$ sea ortogonal a la variedad lineal generada por $y(k_0), \dots, y(k)$.

3.2.3 Solución al Problema de Filtrado

La solución al problema de filtrado de Kalman puede resumirse en el siguiente:

Teorema 3.1 *Considere el sistema descrito en (3.7) y (3.8). El estimado óptimo $\hat{x}(k+1 | k)$ de $x(k+1)$ dadas las observaciones $y(k_0), \dots, y(k)$, es generado por:*

$$\hat{x}(k+1 | k) = [F(k) - \Delta^*(k)H(k)]\hat{x}(k | k-1) + \Delta^*(k)y(k), \quad (3.9)$$

y que corresponde al comportamiento del sistema lineal dinámico de la figura 3.5, con $\Delta^(k) \in \mathbb{R}^{n \times p}$ conocida como la ganancia óptima del estimador. El error de estimación está dado por:*

$$\tilde{x}(k+1 | k) = [F(k) - \Delta^*(k)H(k)]\tilde{x}(k | k-1) + u(k). \quad (3.10)$$

La matriz de covarianza del error de estimación (función de correlación), está dada por

$$\text{Cov}[\tilde{x}(k | k-1)] = E[\tilde{x}(k | k-1)\tilde{x}'(k | k-1)] = P(k), \quad (3.11)$$

donde se satisface que $\forall k \geq k_0$,

$$\Delta^*(k) = F(k)P(k)H'(k)[H(k)P(k)H'(k) + R(k)]^{-1} \quad (3.12)$$

y

$$P(k+1) = [F(k) - \Delta^*(k)H(k)]P(k)F'(k) + Q(k), \quad P(k_0) = P_0, \quad (3.13)$$

con valores medios cuadráticos:

$$E \begin{bmatrix} u(k) \\ v(k) \end{bmatrix} \begin{bmatrix} u'(k) & v'(k) \end{bmatrix} = \begin{bmatrix} Q(k) & 0 \\ 0 & R(k) \end{bmatrix},$$

$$P_0 = E[x_0 x_0'].$$

La demostración de este teorema se encuentra en [2].

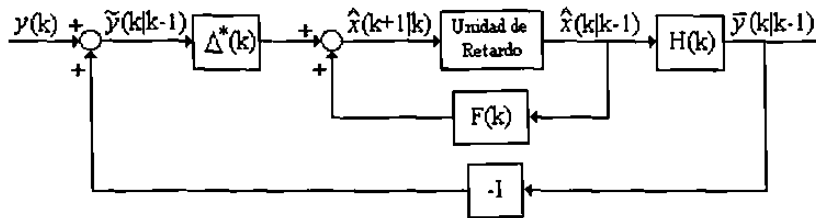


Figura 3.5: Modelo del Filtro Óptimo de Kalman

Nota 3.2 *Substituyendo (3.12) en la ecuación (3.13) se obtiene la ecuación de diferencias no lineal:*

$$P(k+1) = F(k)P(k)F'(k) - F(k)P(k)H'(k)[H(k)P(k)H'(k) + R(k)]^{-1} \\ H(k)P(k)F'(k) + Q(k).$$

la cual, análogamente a (3.2), juega el mismo papel en la teoría convencional del filtrado al depender de los valores $Q(k)$, $R(k)$ y P_0 , pero con una solución mucho más simple.

3.3 Filtro de Kalman-Bucy

En esta sección se establece una generalización del filtro de Kalman al caso de tiempo continuo conocido como filtro de Kalman-Bucy.

3.3.1 Planteamiento del Problema

Considere el modelo de señal de la figura 3.6, tal que,

$$\frac{dx(t)}{dt} = F(t)x(t) + G(t)w(t), \quad (3.14)$$

$$z(t) = H(t)x(t) + v(t), \quad (3.15)$$

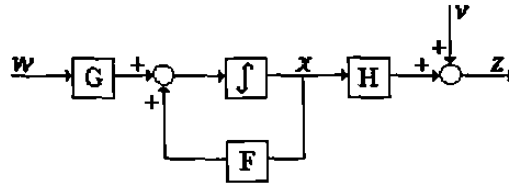


Figura 3.6: Modelo de un Sistema Lineal.

donde, F, G, H son matrices de dimensiones $n \times n$, $n \times m$, y $p \times n$ respectivamente. Los procesos $w(\cdot)$ y $v(\cdot)$ son ruidos blancos con media cero, tal que,

$$E \begin{bmatrix} w(t) \\ v(t) \end{bmatrix} \begin{bmatrix} w'(s) & v'(s) \end{bmatrix} = \begin{bmatrix} Q(t) & S(t) \\ S'(t) & R(t) \end{bmatrix} \delta(t-s), \quad (3.16)$$

con $R(t) = R'(t) > 0$, $\forall t$. Suponga que $w(\cdot)$ y $v(\cdot)$ son independientes, entonces, $S(t) \equiv 0$, por lo tanto, $Q(t) = Q'(t) \geq 0$.

En primera instancia, es necesario asumir un tiempo inicial finito t_0 . De este modo, se puede considerar que $x(t)$ es una variable aleatoria Gaussiana, de media x_0 y covarianza P_0 . La ecuación (3.14) define un modelo Gauss-Markov, puesto que $x(\cdot)$ es un proceso de Markov, el cual, es Gaussiano (de hecho $[x(\cdot)z'(\cdot)]'$ también es Gaussiano y un proceso de Markov).

La tarea de estimación es usar las mediciones $z(s)$, para $s \leq t$, con el propósito de estimar $x(t)$;

el estimado, denotado por $\hat{x}(t)$, está disponible para todo momento de estimación t y minimiza $E([x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]')$, donde a $\tilde{x}(t) = x(t) - \hat{x}(t)$, se le conoce como el error de estimación.

3.3.2 Solución del Problema de Filtrado de Kalman-Bucy

Defina $P(t) = P'(t) \geq 0$, como la solución de

$$\dot{P} = PF' + FP - PH'R^{-1}HP + GQG', \quad P(t_0) = P_0, \quad (3.17)$$

a (3.17) se le conoce como la ecuación diferencial de Riccati (ver [18, 25]). Considere la figura 3.7, y sea $\hat{x}(t)$ la solución de

$$\frac{d\hat{x}(t)}{dt} = F(t)\hat{x}(t) + P(t)H'(t)R^{-1}(t)[z(t) - H\hat{x}(t)], \quad (3.18)$$

entonces,

$$E([x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]') = P(t), \quad (3.19)$$

se satisface. Es decir, el desempeño del estimador óptimo medido por la covarianza del error de estimación está dado por la solución de (3.17).

En la tabla 3.2, se resumen las diferencias entre el filtro de Wiener y el filtro de Kalman-Bucy, considerando las características descritas en la tabla 3.1.

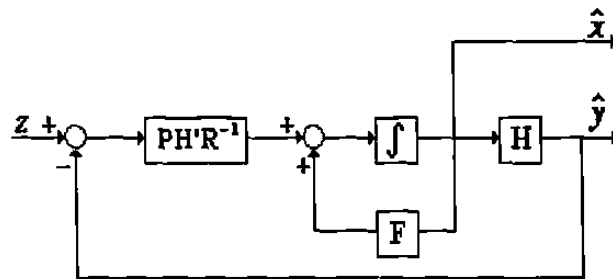


Figura 3.7: Modelo del Filtro Óptimo de Kalman-Bucy

Tabla 3.2: Diferencias entre el filtro de Wiener y el filtro de Kalman-Bucy

Wiener	Kalman-Bucy
$t_0 = -\infty$	$t_0 \geq -\infty$
Estacionario	No requiere ser estacionario
Dimensión infinita	Dimensión finita
Mediciones de ruidos Gaussianos	Mediciones de ruido blanco Gaussiano
Factorización espectral	Solución de la ecuación de Riccati
Estimación de la señal del sistema	Estimación de los Estados del sistema.

3.4 Filtro Gaussiano Condicional

En esta sección se revisará uno de los resultados más importantes de la teoría del filtrado sobre procesos estocásticos con distribuciones Gaussianas.

3.4.1 Propiedades de los Sistemas Gaussianos

Los sistemas dinámicos estudiados por Wiener, Kalman y Bucy incluyen ruidos Gaussianos afectando al comportamiento de los mismos, por lo que de aquí en adelante nos referiremos a esta clase de modelos únicamente como sistemas Gaussianos. Una generalización para el filtrado de sistemas Gaussianos está dada por el *teorema de correlación normal* descrito en esta sección, lo que hace necesario exponer las siguientes propiedades de los vectores Gaussianos mediante el siguiente teorema.

Teorema 3.2 Sea $X \in \mathbb{R}^n$ un vector Gaussiano, entonces,

- i) Los componentes de un vector Gaussiano son no correlacionados si y sólo si ellos son independientes.
- ii) Un vector $X = (X_1, \dots, X_n)$ es Gaussiano si y sólo si, para todo vector $\lambda = (\lambda_1, \dots, \lambda_n)$, $\lambda_k \in \mathbb{R}$, la variable aleatoria $(X, \lambda) = \lambda_1 X_1 + \dots + \lambda_n X_n$ tiene distribución Gaussiana.

Demostración:

- i) Si los componentes de $X = (X_1, \dots, X_n)$ no están correlacionados, esto implica que la forma de la función característica $\varphi_X(t)$ (definida en (2.9)), es el producto de funciones características, por

lo tanto, (del teorema 4, capítulo 2, sección 12 de [12]); los componentes son independientes. El inverso es evidente, puesto que la independencia siempre implica la ausencia de correlación.

ii) Si $X = (X_1, \dots, X_n)$ es Gaussiano, de (2.10) se obtiene que $\forall t \in \mathbb{R}$,

$$E[\exp\{jt(\lambda_1 X_1 + \dots + \lambda_n X_n)\}] = \exp\left\{jt\left(\sum \lambda_k m_k\right) - \frac{t^2}{2}\left(\sum r_{kl}\lambda_k\lambda_l\right)\right\},$$

con lo que consecuentemente (X, λ) tiene una distribución $\mathcal{N}(\sum \lambda_k m_k, \sum r_{kl}\lambda_k\lambda_l)$.

Por el contrario, si (X, λ) tiene una distribución Gaussiana implica que:

$$\begin{aligned} E[\exp\{j(X, \lambda)\}] &= \exp\left\{jE(X, \lambda) - \frac{V(X, \lambda)}{2}\right\} \\ &= \exp\left\{j\sum \lambda_k E(X_k) - \frac{1}{2}\sum \lambda_k\lambda_l \text{Cov}(X_k, X_l)\right\}, \end{aligned}$$

puesto que los valores $\lambda_1, \dots, \lambda_n$ son arbitrarios, entonces, por definición $X = (X_1, \dots, X_n)$ es Gaussiano. ■

Nota 3.3 Sea (X, Y) un vector Gaussiano con $X = (X_1, \dots, X_n)$ y $Y = (Y_1, \dots, Y_m)$. Si X y Y son no correlacionados, esto es $\text{Cov}(X_i, Y_j) = 0$, $i, j = 1, \dots, n, m$; entonces X y Y son independientes.

3.4.2 Planteamiento del Problema

Considere que el comportamiento de un sistema dinámico lineal está descrito por un proceso o vector de estados Gaussiano $X \in \mathbb{R}^n$, con una ecuación de observación $Y \in \mathbb{R}^m$, tal que el vector aleatorio (X, Y) es Gaussiano. Considere que únicamente Y es observable y que X sólo puede estimarse como producto de la observación de Y .

Sea $m = m(x)$ una función medible. El problema consiste en determinar la variable aleatoria $m(X)$, la cual, es un estimador de X en términos de Y ; y la función del error $E([X - m(X)][X - m(X)]')$ del estimador, la cual se entiende es en el sentido de mínimos cuadrados. El estimador $m(X)$ será el óptimo cuando

$$P \triangleq \inf_m E([X - m(X)][X - m(X)]'),$$

donde el ínfimo es tomado sobre todas las funciones medibles $m = m(x)$.

Por el principio de ortogonalidad que se describe (3.6), el estimado óptimo de X en términos de Y

es $E(X|Y(t))$, donde $Y(t)$ es el proceso de observación.

3.4.3 Solución del Problema

Asuma que

$$m_X = E(X), \quad m_Y = E(Y),$$

son vectores columna, y que

$$V_{XX} = Cov(X, X) = \|Cov(X_i, X_j)\|, \quad 1 \leq i, j \leq n,$$

$$V_{XY} = Cov(X, Y) = \|Cov(X_i, Y_j)\|, \quad 1 \leq i \leq n, 1 \leq j \leq m,$$

$$V_{YY} = Cov(Y, Y) = \|Cov(Y_i, Y_j)\|, \quad 1 \leq i, j \leq m,$$

son matrices. Asumiendo que $Cov(Y, Y)$ tiene inversa, la solución al problema de filtrado del sistema Gaussiano (X, Y) se puede resumir en el siguiente:

Teorema 3.3 *Para un vector Gaussiano (X, Y) , el estimador óptimo $E(X|Y)$ de X en términos de Y , y la matriz de error*

$$P = E([X - E(X|Y)][X - E(X|Y)]'),$$

están dados por las fórmulas

$$E(X|Y) = m_X + V_{XY}V_{YY}^{-1}(Y - m_Y), \quad (3.20)$$

$$P = V_{XX} - V_{XY}V_{YY}^{-1}V_{XY}'. \quad (3.21)$$

Demostración: Formando el vector

$$\eta = (X - m_X) - V_{XY}V_{YY}^{-1}(Y - m_Y).$$

puede verse que $E[\eta(Y - m_Y)'] = 0$, esto implica que η no está correlacionado con $(Y - m_Y)$. Pero puesto que (X, Y) es Gaussiano, entonces el vector (η, Y) también es Gaussiano. Por lo que de la nota 3.3, η y $(Y - m_Y)$ son independientes. Entonces, η y Y son independientes, y

consecuentemente $E(\eta|Y) = E(\eta) = 0$. Por lo tanto,

$$E[X - m_X|Y] - V_{XY}V_{YY}^{-1}(Y - m_Y) = 0$$

con lo cual, se establece (3.20).

Para probar (3.21), se debe considerar la covarianza condicional

$$Cov(X, X|Y) \triangleq E([X - E(X|Y)][X - E(X|Y)]'|Y).$$

Puesto que $X - E(X|Y) = \eta$, y que η y Y son independientes, se obtiene que

$$\begin{aligned} Cov(X, X|Y) &= E[\eta\eta'|Y] = E[\eta\eta'] \\ &= V_{XX} + V_{XY}^{-1}V_{YY}V_{YY}^{-1}V'_{XY} - 2V_{XY}V_{YY}^{-1}V_{YY}V_{YY}^{-1}V'_{XY} \\ &= V_{XX} - V_{XY}V_{YY}^{-1}V'_{XY}, \end{aligned}$$

con lo cual, se obtiene (3.21). ■

Capítulo 4

Filtrado de Procesos de Ito-Volterra

Todo modelo escrito en forma de ecuación diferencial puede ser llevado a la forma de ecuación integral (el inverso no necesariamente se cumple), por lo que los resultados del capítulo tres únicamente son válidos para el caso de ecuaciones diferenciales. Por lo tanto, resolver el problema de filtrado para ecuaciones integrales implica resolver el caso de las ecuaciones diferenciales correspondientes. Como una extensión de la teoría expuesta anteriormente, en este capítulo se presentan los resultados concernientes al filtrado de sistemas Gaussianos cuyo comportamiento dinámico está dado por el modelo matemático de la ecuación (2.56). En particular, los modelos con la estructura de (2.56), se incluyen dentro de la clase de procesos estocásticos con retardo, de aquí se sigue la importancia que estos sistemas tienen en las áreas de comunicaciones y el control.

4.1 Filtro Gaussiano Condicional para Procesos de Ito-Volterra

Básicamente en esta sección se presenta un resumen del trabajo de Kleptsina y Veretennikov [6], del problema de filtrado para procesos de Ito-Volterra sobre observaciones de una ecuación diferencial en el caso escalar, el cual, nos brinda la metodología fundamental a seguir en el diseño del filtro para sistemas donde las ecuaciones de estado y de observación son procesos de Ito-Volterra que no pueden llevarse al caso diferencial.

4.1.1 Planteamiento del Problema

Por simplicidad, se denotará al vector de estados $x(t) = x_t$ y a todas las demás variables continuas en el tiempo de la misma manera. Considere el espacio de probabilidad completo (Ω, \mathcal{F}, P) con \mathcal{F}_t una familia de σ -álgebras (definición 2.6), continuas por la derecha y sean (W_t^1, \mathcal{F}_t) y (W_t^2, \mathcal{F}_t) procesos independientes de Wiener. Asuma que (x_t, y_t) es la solución continua del sistema:

$$x_t = \int_0^t b(t, s) dW_s^2 + \int_0^t (a_0(t, s) + a_1(t, s)x_s) ds, \quad (4.1)$$

$$y_t = \int_0^t B(s)dW_s^1 + \int_0^t (A_0(s) + A_1(s)x_s)ds, \quad (4.2)$$

para $0 \leq s \leq t$. Las funciones $B(s)$, $b(t, s)$, $A_i(s)$ y $a_i(t, s)$; $i = 1, 2$; son acotadas, \mathcal{F}_s -(no anticipatorias), suaves en t uniformemente en s y continuas en s . Sea la ecuación

$$x_s^t = \int_0^s b(s, r)dW_r^2 + \int_0^s (a_0(s, r) + a_1(s, r)x_r)dr, \quad (4.3)$$

para $0 \leq r \leq s \leq t$, donde (4.3) puede expresarse en forma de ecuación diferencial. Sea \mathcal{F}_t^y el σ -álgebra generada por las observaciones $y_t \in C_{[0, T]}$, $\forall t \in [0, T]$, entonces,

$$\begin{aligned} m_t &= E(x_t | \mathcal{F}_t^y), \\ P_t &= E[(x_t - m_t)^2 | \mathcal{F}_t^y], \\ m_s^t &= E(x_s^t | \mathcal{F}_t^y), \\ \delta_s &= x_s - m_s, \\ \delta_s^t &= x_s^t - m_s^t, \\ f(t, s) &= E(\delta_s^t \delta_s | \mathcal{F}_t^y), \end{aligned}$$

y sea

$$d\bar{W}_s = \frac{dy_s - (A_0(s) + A_1(s)E(y_s))ds}{B(s)}.$$

El problema es obtener el mejor estimado m_t de x_t basado en el σ -álgebra \mathcal{F}_t^y , $0 \leq s \leq t$, tal que el error o función de correlación P_t sea minimizado. Para tal efecto, se hace necesaria la determinación de la mejor desviación caracterizada por la función $f(t, s)$, de la cual se obtiene P_t al tomar $t = s$.

4.1.2 Filtro Óptimo

El mejor estimado m_t , la función de correlación P_t y la función característica $f(t, s)$ del sistema (4.1), están dados por el siguiente:

Teorema 4.1 *Las siguientes condiciones se satisfacen:*

$$m_t = \int_0^t (a_0(t, s) + a_1(t, s)m_s)ds + \int_0^t \frac{A_1(s)}{B(s)} f(t, s)d\bar{W}_s, \quad (4.4)$$

$$P_t = \int_0^t \left(2a_1(t, s)f(t, s) + b^2(t, s) - \frac{A_1^2(s)}{B^2(s)}f^2(t, s) \right) ds, \quad (4.5)$$

$$\begin{aligned} f(t, s) = & \int_0^s b(t, r)b(s, r)dr + \int_0^s (a_1(t, r)f(s, r) + a_1(s, r)f(t, r))dr \\ & - \int_0^s \frac{A_1^2(s)}{B^2(s)}f(t, r)f(s, r)dr. \end{aligned} \quad (4.6)$$

Este sistema tiene una solución única con $P_t \geq 0$, continua en t .

Demostración: En virtud del teorema 3.3, se obtiene que

$$m_s^t = \int_0^s (a_0(t, r) + a_1(t, r)m_r)dr + \int_0^s \frac{A_1(r)}{B(r)} E[(x_r^t x_r - m_r^t m_r) | \mathcal{F}_t^y] d\bar{W}_r.$$

Asumiendo que $s = t$, se obtiene (4.4).

Como se mencionó anteriormente, es necesario calcular en primer orden la ecuación (4.6). Por definición $f(t, s) = E(\delta_s^t \delta_s | \mathcal{F}_s^y)$. Usando la fórmula de Ito dada en (2.34),

$$\begin{aligned} d_s \delta_s^u &= \left(a_1(u, s) - \frac{A_1^2(s)}{B^2(s)} f(u, s) \right) \delta_s ds + b(u, s) dW_s^2 - \frac{A_1(s)}{B(s)} f(u, s) dW_s^1, \\ d_s \delta_s^u \delta_s^t &= \delta_s^t \left(a_1(u, s) - \frac{A_1^2(s)}{B^2(s)} f(u, s) \right) \delta_s ds - \delta_s^t \left(\frac{A_1(s)}{B(s)} f(u, s) dW_s^1 - b(u, s) dW_s^2 \right) \\ &+ \delta_s^u \left(a_1(t, s) - \frac{A_1^2(s)}{B^2(s)} f(t, s) \right) \delta_s ds - \delta_s^u \left(\frac{A_1(s)}{B(s)} f(t, s) dW_s^1 - b(t, s) dW_s^2 \right) \\ &+ \left(b(u, s)b(t, s) + \frac{A_1^2(s)}{B^2(s)} f(u, s)f(t, s) \right) ds. \end{aligned}$$

Del teorema 3.3, se obtiene:

$$\begin{aligned} E(\delta_s^u \delta_s^t | \mathcal{F}_s^y) |_{u=s} &= \int_0^s [a_1(s, r) E(\delta_r^t \delta_r | \mathcal{F}_r^y) + a_1(t, r) E(\delta_r^s \delta_r | \mathcal{F}_r^y) \\ &+ b(u, r)b(t, r) - \frac{A_1^2(r)}{B^2(r)} f(t, r)f(s, r)] dr \\ &+ \int_0^s \frac{A_1(r)}{B(r)} E(\delta_r^s \delta_r^t \delta_r | \mathcal{F}_r^y) d\bar{W}_r \end{aligned}$$

Pero del ejemplo 2.1, $E(\delta_r^s \delta_r^t \delta_r | \mathcal{F}_r^y) = 0$ y dado que $E(\delta_r^t \delta_r | \mathcal{F}_r^y) = f(t, r)$, se obtiene (4.6), con la cual, se determina (4.5) a al tomar $s = t$. ■

En [7], se presenta la extensión del resultado del teorema 4.1 al caso vectorial.

4.2 Filtrado de Procesos de Ito-Volterra sobre Observaciones Continuas

A continuación se describirá y se mostrará la solución del problema de filtrado para procesos de Ito-Volterra sobre observaciones de tipo Ito- Volterra en caso vectorial de tiempo continuo resuelto en [24]. Cabe hacer la aclaración de que en [6, 7], las observaciones del vector de estados, son realizadas por una ecuación integral transformable en una ecuación diferencial; en este nuevo caso, la ecuación de observación del sistema también es una ecuación integral pero no puede ser llevada a la forma de una ecuación diferencial. Sin embargo, la metodología en el diseño del filtro es similar al caso de la sección anterior.

4.2.1 Planteamiento del Problema

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad completo con una familia de σ -álgebras \mathcal{F}_t , $t \geq 0$, crecientes y continuas por la derecha; y sea $(W_t^1, \mathcal{F}_t, t \geq 0)$ y $(W_t^2, \mathcal{F}_t, t \geq 0)$ procesos de Wiener independientes. Sea (x_t, y_t) un proceso aleatorio (función \mathcal{F}_t -medible), tal que la ecuación de estado:

$$x_t = \int_0^t (a_0(t, s) + a(t, s)x_s)ds + \int_0^t b(t, s)dW_s^2, \quad (4.7)$$

y la ecuación de observación:

$$y_t = \int_0^t (A_0(t, s) + A(t, s)x_s)ds + \int_0^t B(t, s)dW_s^1, \quad (4.8)$$

donde $x_t \in \mathbb{R}^n$; $y_t \in \mathbb{R}^m$; $a_0(t, s)$, $a(t, s)$, $b(t, s)$ son suaves en t uniformemente en s y continuas en s ; y $A_0(t, s)$, $A(t, s)$, y $B(t, s)$ son continuas en t y s . Sean $A(t, s)$ y $B(t, s)$ matrices no nulas, tal que $B(t, s)B'(t, s) > 0$.

El problema de filtrado consiste en encontrar el mejor estimado m_t de x_t al tiempo t basado en el proceso de observación $Y(t) = \{y_t, 0 \leq s \leq t\}$, es decir, se busca la esperanza condicional

$$m_t = E(x_t | \mathcal{F}_t^Y).$$

Sea la función de correlación:

$$P_t = E((x_t - m_t)(x_t - m_t)' | \mathcal{F}_t^Y)$$

y la función característica:

$$f(t, s) = E((x_s^t - m_s^t)(x_s - m_s)' | \mathcal{F}_{t,s}^Y),$$

donde

$$x_s^t = \int_0^s (a_0(t, r) + a(t, r)x_r)dr + \int_0^s b(t, r)dW_r^2, \quad (4.9)$$

$\mathcal{F}_{t,s}^Y$ es el σ -álgebra generada por el proceso estocástico:

$$y_s^t = \int_0^s (A_0(t, s) + A(t, s)x_s)ds + \int_0^s B(t, s)dW_s^1, \quad (4.10)$$

y

$$m_s^t = E(x_s^t | \mathcal{F}_{t,s}^Y). \quad (4.11)$$

donde (4.9), (4.10) y (4.11) son ecuaciones que pueden ser escritas en forma diferencial debido a que t es simplemente un parámetro.

4.2.2 Filtro Óptimo

Teorema 4.2 *El mejor estimado m_t del sistema (4.7) sobre las observaciones de (4.8), la función de correlación P_t y la función característica $f(t, s)$ satisfacen las siguientes ecuaciones:*

$$m_t = \int_0^t (a_0(t, s) + a(t, s)m_s)ds + \int_0^t f(t, s)A'(t, s)(B(t, s)B'(t, s))^{-1} [dy_s - (A_0(t, s) + A(t, s)m_s)ds], \quad (4.12)$$

$$P_t = \int_0^t [a(t, s)f'(t, s) + f(t, s)a'(t, s) + b(t, s)b'(t, s)]ds - \int_0^t f(t, s)A'(t, s)(B(t, s)B'(t, s))^{-1}A(t, s)f'(t, s)ds, \quad (4.13)$$

$$\begin{aligned}
f(t, s) &= \int_0^s [a(t, r)f'(s, r) + f(t, r)a'(s, r) + \frac{1}{2}(b(t, r)b'(s, r) + b(s, r)b'(t, r))]dr \\
&\quad - \int_0^s [f(t, r)A'(s, r)(B(s, r)B'(s, r))^{-1}A(s, r)f'(s, r) + f(t, r)A'(t, r) \\
&\quad (B(t, r)B'(t, r))^{-1}A(t, r)f'(s, r) - \frac{1}{2}f(t, r)A'(t, r)(B(t, r)B'(s, r))^{-1} \\
&\quad A(s, r)f'(s, r) - \frac{1}{2}f(s, r)A'(s, r)(B(s, r)B'(t, r))^{-1}A(t, r)f'(t, r)]dr, \quad (4.14)
\end{aligned}$$

Demstración: Considerando el problema de filtrado para la ecuación del estado x_s^t sobre las observaciones del proceso y_s^t . La ecuación (4.9) para x_s^t y la ecuación (4.10) para y_s^t son de hecho ecuaciones diferenciales con respecto a s , donde t es un parámetro. Así, el teorema 3.3 es aplicable, por lo tanto, se obtiene que

$$\begin{aligned}
m_s^t &= \int_0^s (a_0(t, r) + a(t, r)m_r)dr + \int_0^s E(x_r^t x_r' - m_r^t m_r' | \mathcal{F}_{t,r}^Y)A'(t, r) \\
&\quad (B(t, r)B'(t, r))^{-1}[dy_r^t - (A_0(t, r) + A(t, r)m_r)dr].
\end{aligned}$$

Del mismo modo que en el teorema 4.1, igualando $s = t$, se obtiene la ecuación (4.12) para m_r .

Análogamente al teorema 4.1, se demostrará primero (4.14) para $f(t, s)$. De la fórmula (2.34) de Ito resulta que

$$\begin{aligned}
d(x_r^t - m_r^t) &= [a(t, r) - f(t, r)A'(t, r)(B(t, r)B'(t, r))^{-1}A(t, r)](x(r) - m_r)dr \\
&\quad + b(t, r)dW_r^2 - f(t, r)A'(t, r)(B(t, r)B'(t, r))^{-1}B(t, r)dW_r^1,
\end{aligned}$$

$$d((x_r^t - m_r^t)(x_r^u - m_r^u)') = (x_r^t - m_r^t)(x(r) - m_r)'[a(u, r) - f(u, r)A'(u, r)$$

$$(B(u, r)B'(u, r))^{-1}A(u, r)]'dr + [a(t, r) - f(t, r)A'(t, r)(B(t, r)B'(t, r))^{-1}$$

$$A(t, r)](x(r) - m_r)(x_r^u - m_r^u)'dr + (1/2)(b(t, r)b'(u, r) + b(u, r)$$

$$b'(t, r))dr + (1/2)f(t, r)A'(t, r)(B(t, r)B'(u, r))^{-1}A(u, r)f'(u, r)dr$$

$$\begin{aligned}
& + (1/2)f(u, r)A'(u, r)(B(u, r)B'(t, r))^{-1}A(t, r)f'(t, r)]dr \\
& - (x_r^t - m_r^t)[f(u, r)A'(u, r)(B(u, r)B'(u, r))^{-1}B(u, r)dW_r^1 \\
& - b(u, r)dW_r^2]' - [f(t, r)A'(t, r)(B(t, r)B'(t, r))^{-1}B(t, r)dW_r^1 \\
& - b(t, r)dW_r^2](x_r^u - m_r^u)'.
\end{aligned}$$

Integrando con respecto a r , para $0 \leq r \leq s$, igualando $u = s$, y usando el teorema 3.3 resulta que:

$$\begin{aligned}
E((x_s^t - m_s^t)(x_s - m_s)' | \mathcal{F}_{t,s}^Y) &= \int_0^s [E((x_r^t - m_r^t)(x_r - m_r)' | \mathcal{F}_{t,r}^Y)a'(s, r) \\
& + a(t, r)E((x_r - m_r)(x_r^s - m_r^s)' | \mathcal{F}_{s,r}^Y) + (1/2)(b(t, r)b'(s, r) \\
& + b(s, r)b'(t, r))dr + (1/2)f(t, r)A'(t, r)(B(t, r)B'(s, r))^{-1}A(s, r) \\
& f'(s, r)dr + (1/2)f(s, r)A'(s, r)(B(s, r)B'(t, r))^{-1}A(t, r)f'(t, r) \\
& - E((x_r^t - m_r^t)(x_r - m_r)' | \mathcal{F}_{t,r}^Y)(f(s, r)A'(s, r)(B(s, r)B'(s, r))^{-1} \\
& A(s, r))' - f(t, r)A'(t, r)(B(t, r)B'(t, r))^{-1}A(t, r)E((x_r - m_r)(x_r^s \\
& - m_r^s)' | \mathcal{F}_{s,r}^Y)]dr + \int_0^s [E((x_r^t - m_r^t)(x_r^s - m_r^s)'(x_r - m_r)' | \mathcal{F}_{t,r}^Y)] \\
& A'(t, r)(B(t, r)B'(t, r))^{-1}[dy_r^t - (A_0(t, r) + A(t, r)m_r)dr]. \tag{4.15}
\end{aligned}$$

Note que el último término de la ecuación (4.15), es igual a cero; de acuerdo al ejemplo 2.1. De este modo, la ecuación (4.15) produce (4.14). Finalmente, la ecuación (4.13) es obtenida de (4.14), debido a que $P_t = f(t, t)$. ■

4.2.3 Ejemplo

Ejemplo 4.1 Considere el sistema:

$$x(t) = \int_0^t [e^{-3t+s} + e^{-t-s}x(s)]ds + \int_0^t e^{t-2s}dW^1(s), \quad (4.16)$$

$$y(t) = \int_0^t [e^{4(t-s)} + e^{-2t-s}x(s)]ds + \int_0^t e^{3(t-s)}dW^2(s), \quad (4.17)$$

donde $W^1(s)$ y $W^2(s)$ son procesos independientes de Wiener

El mejor estimado, la función de correlación y la función característica satisfacen:

$$m_t = \int_0^t [e^{-3t+s} + e^{-t-s}m_s]ds + \int_0^t e^{-8t+5s}f(t,s)[dy_s - (e^{4(t-s)} + e^{-2t-s}m_s)ds], \quad (4.18)$$

$$P_t = \int_0^t [e^{2t-4s} + 2e^{-t-s}f(t,s)]ds - \int_0^t e^{-10t+4s}f^2(t,s)ds, \quad (4.19)$$

$$f(t,s) = \int_0^s [e^{-t-r}f(s,r) + e^{-s-r}f(t,r) + e^{t-4r+s} - (e^{-10s+4r} + e^{-10t+4r} - e^{-5t+4r-5s})f(t,r)f(s,r)]dr. \quad (4.20)$$

Note que,

$$f(s,s) = P_s = \int_0^s [e^{2s-4r} + 2e^{-s-r}f(s,r)]dr - \int_0^s e^{-10s+4r}f^2(s,r)dr, \quad (4.21)$$

y que diferenciando (4.21) con respecto a s :

$$\dot{P}_s = e^{-2s} + 2e^{-2s}P_s - e^{-6s}P_s^2 + \int_0^s \frac{\partial}{\partial s} [e^{2s-4r} + 2e^{-s-r}f(s,r) - e^{-10s+4r}f^2(s,r)]dr. \quad (4.22)$$

La ecuación (4.22) se puede aproximar como:

$$\dot{P}_s = e^{-2s} + 2e^{-2s}P_s - e^{-6s}P_s^2. \quad (4.23)$$

Así, se obtiene la curva descrita en la figura 4.1 (a), la cual, converge al valor fijo de 0.84777, con lo que el criterio de mínimos cuadrados se satisface.

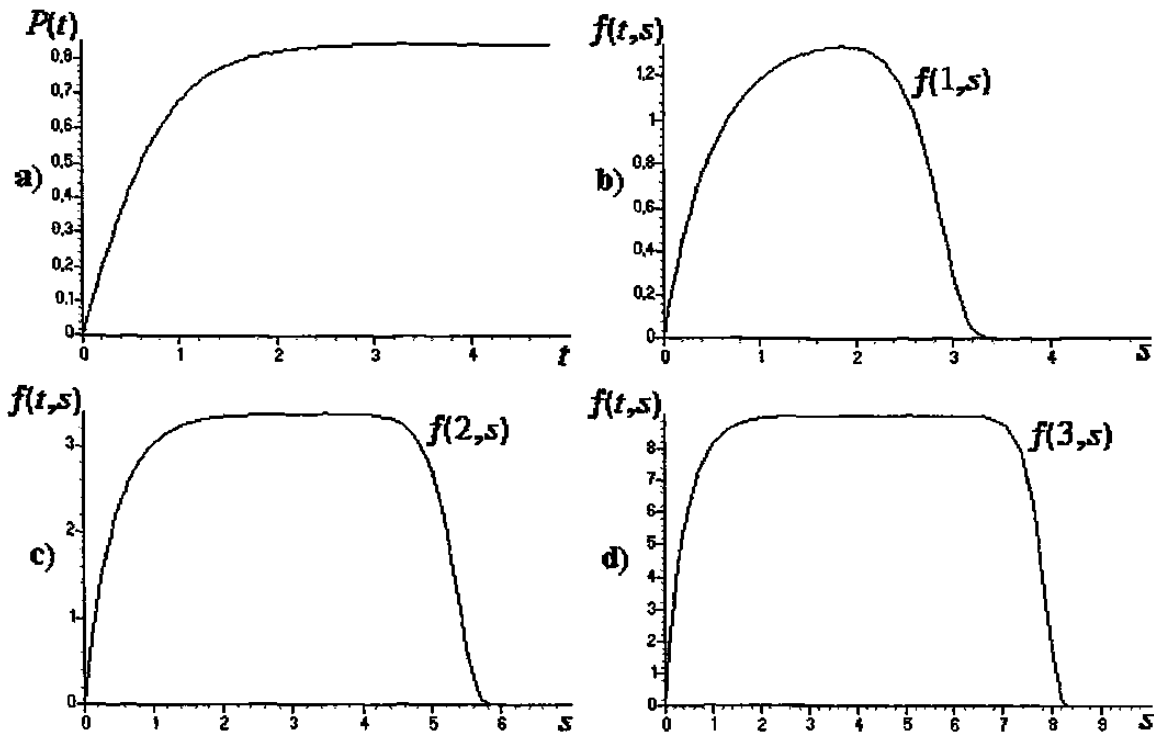


Figura 4.1: Funciones (a) de Correlación y (b)-(d) Características

Para saber si el comportamiento de $f(t, s)$ satisface el criterio de mínimos cuadrados, se toma la derivada de (4.20) con respecto a s y fijando $t = 1$:

$$\begin{aligned} \frac{d}{ds} f(1, s) &= e^{-1-s} f(s, s) + e^{-2s} f(1, s) + e^{-1-3s} - (e^{-6s} + e^{-10+4s} - e^{-5-s}) f(1, s) f(s, s) \\ &+ \int_0^s \frac{\partial}{\partial s} [e^{-1-r} f(s, r) + e^{-s-r} f(1, r) + e^{1-4r+s} - (e^{-10s+4r} \\ &+ e^{-10+4r} - e^{-5+4r-5s}) f(1, r) f(s, r)] dr, \end{aligned} \quad (4.24)$$

de igual modo que en el caso de $P(t)$, (4.24) puede expresarse como:

$$\dot{f}(1, s) = e^{1-3s} + e^{-1-s} P_s + e^{-2s} f(1, s) - (e^{-6s} + e^{-10+4s} - e^{-5-s}) f(1, s) P_s, \quad (4.25)$$

para $t = 2$, resulta que (4.20) :

$$\dot{f}(2, s) = e^{2-3s} + e^{-2-s} P_s + e^{-2s} f(2, s) - (e^{-6s} e^{-20+4s} - e^{-10-s}) f(2, r) P_s, \quad (4.26)$$

y para $t = 3$:

$$\dot{f}(3, s) = e^{3-3s} + e^{-3-s}P_s + e^{-2s}f(3, s) - (e^{-6s} + e^{-30+4s} - e^{-15-s})f(3, r)P_s. \quad (4.27)$$

Los resultados gráficos de las ecuaciones (4.25)-(4.27) se aprecian en la figura 4.1 (b)-(d) respectivamente, de donde se puede concluir que existe un comportamiento generalizado para la función $f(\cdot, s)$ que converge al origen, por lo tanto, el cálculo del mejor estimado del estado está bien definido y para efectos de la simulación de (4.18), considere las aproximaciones de las ecuaciones (4.16) y (4.17), tal que:

$$\dot{x}(t) = e^{-2t}(1 + x(t)) + e^{-t}\psi(t), \quad (4.28)$$

$$\dot{y}(t) = 1 + e^{-3t}x(t) + \psi(t), \quad (4.29)$$

$$\dot{m}_t = e^{-2t}(1 + m_t) + [e^{-6t}(x(t) - m_t) + e^{-3t}\psi(t)]P_t. \quad (4.30)$$

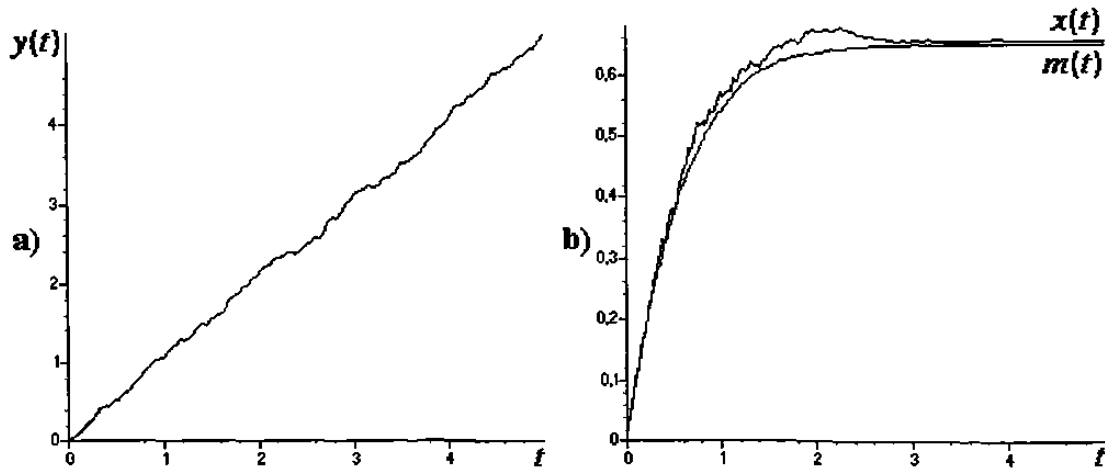


Figura 4.2: Ecuaciones (a) de Observación y (b) Estado/Estimado

La figura 4.2, representa una realización para el sistema (4.16)-(4.17). De la figura 4.2 (b) se puede apreciar que el estimado $m(t)$ aproxima en su trayectoria al estado $x(t)$. ■

Note que en la figura 4.1, las curvas (b)-(d) describen únicamente el comportamiento cualitativo de la función $f(t, s)$, y no así sus valores reales, los cuales, no deben exceder los valores de la función de correlación $P(t)$, pero que dada la aproximación no se pueden obtener.

Capítulo 5

Solución para Ecuaciones Diferenciales en Distribuciones con Funciones Regulares Discontinuas sobre el Lado Derecho

Con la intención de obtener las ecuaciones del filtrado para procesos de Ito-Volterra sobre observaciones discontinuas del tipo Integral, se presenta en este capítulo un método para resolver las ecuaciones diferenciales que contienen funciones "generalizadas" con discontinuidades sobre el lado derecho. Este algoritmo presentado en [8, 9], evita las complicaciones que encierran las técnicas de solución de ecuaciones diferenciales en distribuciones caracterizadas por funciones de variación acotada, y que se conoce como Vibrosolución.

5.1 Funciones Generalizadas

El grado de generalidad ligada a la noción de *función* varía dependiendo del problema en cuestión. Determinados problemas requieren que las funciones sean continuas, en otros casos requieren funciones diferenciables una o varias veces; sin embargo, existen situaciones en que el manejo de las funciones se vuelve inadecuado, aun cuando se entiendan en el sentido más general. Los problemas de esta naturaleza se resuelven por medio de funciones generalizadas y abarcan casos donde las funciones no tienen derivadas.

5.1.1 Espacio Fundamental y Funciones Generalizadas

Sea \mathbb{K} el conjunto de todas las funciones finitas φ definidas sobre $(-\infty, \infty)$ con derivadas continuas de todos los ordenes, donde cada función $\varphi \in \mathbb{K}$ finita se desvanece al exterior de algún intervalo. Si \mathbb{K} está equipado con las operaciones adición y multiplicación escalar, entonces, \mathbb{K} forma un espacio lineal y la convergencia en este espacio está dada por la siguiente definición.

Definición 5.1 Una sucesión $\{\varphi_n\}$ de funciones en \mathbb{K} se dice que converge a una función $\varphi \in \mathbb{K}$ (denotado por $\varphi_n \rightarrow \varphi$), si:

- a) Existe un intervalo fuera del cual, todas las funciones φ_n se desvanecen.
- b) La secuencia $\{\varphi_n^{(k)}\}$ de derivadas de orden k convergen uniformemente a $\varphi^{(k)}$ sobre su intervalo para toda $k = 0, 1, 2, \dots$

Al espacio lineal \mathbb{K} equipado con esta clase de convergencia se le conoce como el *espacio fundamental*, y a las funciones $\varphi \in \mathbb{K}$ se les llama *funciones fundamentales*.

Definición 5.2 Toda funcional lineal continua $T(\varphi)$ definida sobre el espacio fundamental \mathbb{K} es llamada una función generalizada sobre $(-\infty, \infty)$, donde la continuidad de $T(\varphi)$ establece que cuando $\varphi_n \rightarrow \varphi$ en \mathbb{K} implica que $T(\varphi_n) \rightarrow T(\varphi)$.

Sea $f(x)$ una función integrable sobre cada intervalo finito ($f(x)$ es *localmente integrable*). Entonces, $f(x)$ genera una función generalizada a través de expresión:

$$T_f(\varphi) = \int_{-\infty}^{\infty} f(x)\varphi(x)dx. \quad (5.1)$$

Las funciones generalizadas de la forma de (5.1) son clasificadas como *regulares*, y aquellas de estructura diferente como *singulares*.

Ejemplo 5.1 La "función delta"

$$T(\varphi) = \varphi(0),$$

es una funcional lineal continua definida sobre \mathbb{K} con estructura singular. No obstante, se puede representar como

$$T(\varphi) = \int_{-\infty}^{\infty} \delta(x)\varphi(x)dx,$$

donde $\delta(x)$ es una función ficticia nula excepto cuando $x = 0$, tal que,

$$\int_{-\infty}^{\infty} \delta(x)dx = 1,$$

por lo tanto, se puede expresar formalmente que

$$T(\varphi) = \int_{-\infty}^{\infty} \delta(x)\varphi(x)dx = \varphi(0) \int_{-\infty}^{\infty} \delta(x)dx = \varphi(0). \blacksquare$$

5.1.2 Propiedades de las Funciones Generalizadas

Sean

$$T_f(\varphi) = \int_{-\infty}^{\infty} f(x)\varphi(x)dx, \quad T_g(\varphi) = \int_{-\infty}^{\infty} g(x)\varphi(x)dx,$$

donde f y g son localmente integrables y $\varphi \in \mathbb{K}$, $\alpha \in \mathbb{R}$, entonces,

$$(T_f + T_g)(\varphi) = T_f(\varphi) + T_g(\varphi) = T_{f+g}(\varphi)$$

y

$$(\alpha T_f)(\varphi) = \alpha T_f(\varphi) = T_{\alpha f}(\varphi).$$

Definición 5.3 Una sucesión de funciones generalizadas $\{T_n\}$ se dice que converge a una función generalizada T si $T_n(\varphi) \rightarrow T(\varphi) \forall \varphi \in \mathbb{K}$. El espacio de funciones generalizadas con esta noción de convergencia se denota por \mathbb{K}^* . Este modo de convergencia es la convergencia estrella débil [10] de las funcionales lineales continuas sobre \mathbb{K} .

Asuma que la derivada de f existe y que es localmente integrable. Entonces, la derivada de una función regular generalizada está dada por

$$\frac{dT}{dx}(\varphi) = \int_{-\infty}^{\infty} \dot{f}(x)\varphi(x)dx. \quad (5.2)$$

Integrando (5.2) por partes y usando el hecho de que cada función φ se desvanece al exterior de algún intervalo infinito, se obtiene que

$$\frac{dT}{dx}(\varphi) = - \int_{-\infty}^{\infty} f(x) \dot{\varphi}(x)dx, \quad (5.3)$$

la cual, es una expresión para $\frac{dT}{dx}$ que no involucra la derivada de f .

Extendiendo (5.3) al caso singular, resulta que derivada $\frac{dT}{dx}$ es la funcional definida por la fórmula:

$$\frac{dT}{dx}(\varphi) = -T(\dot{\varphi}) \quad (5.4)$$

donde el lado derecho de (5.4) es una funcional lineal continua y es por si misma una función generalizada.

Ejemplo 5.2 Considere la función Heaviside o función escalón:

$$f(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 & \text{si } x > 0, \end{cases} \quad (5.5)$$

definiendo la funcional lineal

$$T(\varphi) = \int_{-\infty}^{\infty} f(x)\varphi(x)dx = \int_0^{\infty} \varphi(x)dx.$$

de la definición de (5.4),

$$\frac{dT}{dx}(\varphi) = - \int_0^{\infty} \dot{\varphi}(x)dx = \varphi(0),$$

por lo tanto, la derivada de (5.5) es la función delta.

5.1.3 Ecuaciones Diferenciales en Distribuciones

Considerando que φ es una función fundamental en \mathbb{K} , entonces, la ecuación

$$\dot{x}(t) = f(x, u, t) + b(x, u, t) \dot{u}(t), \quad x(t_0) = x_0, \quad (5.6)$$

se satisface en el estricto sentido de que $\forall \varphi \in \mathcal{C}_{(-\infty, \infty)}$,

$$\int_{-\infty}^{\infty} \varphi(t)dx(t) = \int_{-\infty}^{\infty} f(x, u, t)\varphi(t)dt + \int_{-\infty}^{\infty} b(x, u, t)\varphi(t)du(t), \quad (5.7)$$

donde $f(x, u, t)$ es una función continua con respecto a todos sus argumentos, $b(x, u, t)$ es una función continua con respecto a x y discontinua con respecto a u, t ; y $u(t)$ es una función de variación acotada. A la ecuación de la clase de (5.6) se le conoce como *ecuación diferencial en distribuciones* y fueron estudiadas por primera vez en [34].

5.2 Vibrosoluciones para Ecuaciones Diferenciales en Distribuciones con Funciones Regulares Discontinuas sobre el Lado Derecho

Considere la ecuación diferencial en distribuciones de dimensión n ,

$$\dot{x}(t) = f(x, u, t) + b(x, u, t) \dot{u}(t), \quad x(t_0) = x_0, \quad (5.8)$$

donde $f(\xi, v, s)$ y $b(\xi, v, s)$ son funciones continuas por partes en las variables abstractas (ξ, v, s) ; y

$u(t) \in \mathbb{R}^m$ es una función de variación acotada con componentes $u_i(t)$ no decrecientes y continuos sobre la izquierda,

$$\dot{u}_i(t) \geq 0, \quad |u_i(t)| < \infty, \quad u_i(t) = u_i(t-), \quad i = 1, \dots, m,$$

tal que, para $t_2 \geq t_1$,

$$u(t_2) \geq u(t_1) \text{ si } u_i(t_2) \geq u_i(t_1).$$

De acuerdo con la referencia [29], se dice que las funciones $f(\xi, v, s)$ y $b(\xi, v, s)$ son continuas por partes en un dominio acotado $G \subset \mathbb{R}^{n+m+1}$ si G es la unión finita de aquellos dominios en los cuales $f(\xi, v, s)$ y $b(\xi, v, s)$ son continuas hasta su correspondiente límite, con un conjunto de medida cero consistente de los límites de esos dominios. Por lo que, $f(\xi, v, s)$ y $b(\xi, v, s)$ son continuas por partes en \mathbb{R}^{n+m+1} si son continuas por partes en cualquier dominio acotado $G \subset \mathbb{R}^{n+m+1}$.

Sea $u(t) \in \mathbb{R}^m$ una función absolutamente continua. Entonces, la solución de (5.8) se entiende es una solución absolutamente continua para la ecuación diferencial generalizada $\dot{x}(t) \in F(x, t)$, donde $F(x, t)$, es el mínimo conjunto convexo que contiene la secuencia de valores x_n , cuyo límite $x_n \rightarrow x$ y t constante de la función

$$f(x_n, u(t), t) + b(x_n, u(t), t) \dot{u}(t)$$

a menos que (x_n, t) pertenezca a su conjunto de discontinuidades.

En el caso de que $u(t) \in \mathbb{R}^m$ sea una función no decreciente, se define una solución para (5.8) siguiendo aproximadamente el concepto de vibrosolución utilizado en [8, 9, 24].

5.2.1 Vibrosoluciones

Un criterio útil para definir el concepto de convergencia estrella débil en el espacio topológico de las funciones de variación acotada, está dado por el *principio de selección de Helly* [10], de acuerdo con éste, una secuencia de funciones de variación acotada $u_k(t)$ tiene *convergencia estrella débil*, denotado por $* - \lim u_k(t)$; si satisface las condiciones siguientes:

- (1) $\lim u_k(t_0) = u(t_0)$, cuando $k \rightarrow \infty$;
- (2) $\lim u_k(t) = u(t)$, cuando $k \rightarrow \infty$, $t_0 \leq t$, para todos los puntos de continuidad de la función $u(t)$;

$$(3) \sup_k V_{t_0}^t u_k(t) \leq \infty, \forall t_0 \leq t.$$

Definición 5.4 Una vibrosolución para (5.8) con una función no decreciente $u(t)$, es una función $x(t)$ que satisface las siguientes condiciones:

- a) $x(t)$ es continua sobre la izquierda.
- b) Sea $u_k(t)$ una secuencia de funciones no decrecientes absolutamente continuas que convergen a $u(t)$ en la topología estrella débil del espacio de funciones de variación acotada sobre el lado derecho, es decir, si

$$* - \lim u_k(t) = u(t), \quad k \rightarrow \infty, \quad t \geq t_0,$$

entonces, la secuencia $x_k(t)$ de soluciones para la ecuación

$$\dot{x}_k(t) = f(x_k, u_k, t) + b(x_k, u_k, t) \dot{u}_k(t), \quad x_k(t_0) = x_0,$$

converge a $x(t)$ en la misma topología

$$* - \lim x_k(t) = x(t), \quad k \rightarrow \infty, \quad t \geq t_0.$$

Del mismo modo que en el caso de las ecuaciones diferenciales en distribuciones con funciones regulares sobre el lado derecho, la existencia de una vibrosolución para la ecuación (5.8) está estrechamente relacionada a la existencia de la solución de un sistema asociado de ecuaciones diferenciales exactas.

Teorema 5.1 Las siguientes condiciones se satisfacen:

- i) las funciones $f(x, u, t)$ y $b(x, u, t)$ son acumuladamente continuas por partes en (x, u, t) y sus dominios (continuos), son localmente conectados;
- ii) las funciones $f(x, u, t)$ y $b(x, u, t)$ satisfacen (excepto para sus conjuntos de límites con medida cero), un lado de la condición de Lipschitz

$$(x - y, f(x, u, t) - f(y, u, t)) \leq L_1(t, u)(x - y)^2,$$

$$(x - y, b(x, u, t) - b(y, u, t)) \leq L_2(t, u)(x - y)^2,$$

donde $L_1(t, u)$ y $L_2(t, u)$ son integrables con respecto a las variables (t, u) . Si (5.8) tiene una vibrosolución, entonces, el sistema de ecuaciones diferenciales exactas

$$\frac{d\xi}{du} = b(\xi, u, s), \quad \xi(w) = z, \quad (5.9)$$

es solucionable en el cono $K = \{u \in \mathbb{R}^m : u_i \geq w_i, i = 1, \dots, m\}$ de direcciones positivas para una condición inicial arbitraria $w \in \mathbb{R}^m, z \in \mathbb{R}^n$ y $s \geq t_0$.

Demostración: Considere una vibrosolución $x(t)$ para (5.8) con la condición inicial $x(s) = z$ y

con $u(t) = w + (v - w)\mathcal{X}(t - s)$, donde $\mathcal{X}(t)$ es la función Heaviside, $w, v \in \mathbb{R}^m$, $v \geq w$ y $s \geq t_0$. Las condiciones del teorema 5.1 aseguran la existencia de esta vibrosolución. Se mostrará que bajo las condiciones del teorema, la identidad

$$x(s+) = y(1) \quad (5.10)$$

se satisface, donde $x(s+) = \lim_{t \rightarrow s+} x(t)$, cuando $t \rightarrow s+$, y $y(\tau)$ es una solución de la ecuación

$$\frac{dy}{d\tau} = b(y, w + (v - w)\tau, s)(v - w), \quad 0 \leq \tau \leq 1, \quad y(0) = z, \quad (5.11)$$

Note que el teorema de existencia y unicidad para ecuaciones diferenciales con discontinuidades sobre el lado derecho [29], y las condiciones del teorema 5.1 tienen una única solución absolutamente continua $y(\tau)$ dado que $v \geq w$.

Ahora, se mostrará que bajo la suposiciones del teorema 5.1 las funciones $y_k(\tau) = x_k(s + \tau/k)$, $0 \leq \tau \leq 1$, $k = 1, \dots$; donde x_k es la vibrosolución de (5.8) con la condición inicial $x_k(s) = z$ y con la función absolutamente continua sobre el lado derecho

$$u_k(t) = \begin{cases} w & \text{para } t \leq s, \\ w + k(v - w)(t - s) & \text{para } s \leq t \leq s + 1/k, \\ v & \text{para } t \geq s + 1/k; \end{cases}$$

satisface las ecuaciones

$$\frac{dy_k}{d\tau} = \frac{1}{k} f \left(y_k, w + (v - w)\tau, s + \frac{\tau}{k} \right) + b \left(y_k, w + (v - w)\tau, s + \frac{\tau}{k} \right) (v - w), \quad y_k(0) = z,$$

y la condición

$$x_k \left(s + \frac{1}{k} \right) = y_k(1). \quad (5.12)$$

Dado el teorema de la dependencia continua de la solución para una ecuación diferencial con discontinuidades sobre el lado derecho [29], la secuencia de funciones absolutamente continuas $y_k(\tau)$ converge a la solución absolutamente continua $y(\tau)$ en la norma Euclideana, esto es,

$$\lim_{k \rightarrow \infty} \|y_k(\tau) - y(\tau)\| = 0, \quad \forall \tau \in [0, 1].$$

De lo anterior se tiene que $\lim y_k(1) = y(1)$ cuando $k \rightarrow \infty$, y por (5.12),

$$\lim_{t \rightarrow s+k \rightarrow \infty} \lim_{k \rightarrow \infty} x_k(t) = y(1).$$

Puesto que

$$\lim_{t \rightarrow s+k \rightarrow \infty} \lim_{k \rightarrow \infty} x_k(t) = x(s+)$$

y $x(t)$ es una vibrosolución de (5.8), de donde se llega a (5.10).

Considere la función

$$\xi(z, w, v, s) = y(1) = z + \int_0^1 b(y(\tau), w + (v - w)\tau, s)(v - w)d\tau, \quad (5.13)$$

donde $y(\tau)$ es la solución de (5.11). Bajo las condiciones del teorema 5.1, la solución $y(\tau)$ existe y es única para $v \geq w$, entonces, la solución de $\xi(z, w, v, s)$ es única y definida sobre el cono de direcciones positivas $K = \{u \in \mathbb{R}^m : u_i \geq w_i, i = 1, \dots, m\}$.

La prueba del hecho de que

$$\frac{d\xi(z, w, v, s)}{dv} = b(\xi(z, w, v, s), v, s)$$

reproduce la última parte de la prueba del teorema 1 de [30], por lo que se omite. Así, la función $\xi(z, w, v, s)$ definida por (5.13) es la única solución del sistema (5.9) sobre K . ■

Una vez establecidas las condiciones de existencia y unicidad de las vibrosoluciones, se presenta en el siguiente teorema, algunas condiciones adicionales sobre la función $b(x, u, t)$, para las cuales, las condiciones necesarias y suficientes de la existencia de una vibrosolución coinciden.

Teorema 5.2 Sean $w \in \mathbb{R}^m$ y $z \in \mathbb{R}^n$ condiciones iniciales arbitrarias, y sea $s \geq t_0$, tal que se satisfacen:

- i) Las funciones $f(x, u, t)$, $b(x, u, t)$, $\partial b(x, u, t)/\partial x$ y $\partial b(x, u, t)/\partial t$ son acumuladamente continuas por partes en (x, u, t) y sus dominios continuamente localmente conectados;
- ii) las funciones $f(x, u, t)$ y $b(x, u, t)$ satisfacen (excepto para sus conjuntos de límites con medida cero), un lado de la condición de Lipschitz.

Si el sistema (5.9) de ecuaciones diferenciales exactas es solucionable sobre el cono:

$$K = \{u \in \mathbb{R}^m : u_i \geq w_i, i = 1, \dots, m\},$$

entonces, la ecuación (5.8) tiene una vibrosolución.

Demostración: Suponga que el teorema 5.2 se satisface, entonces, el sistema (5.9) tiene una solución $\xi(z, w, u, t)$ sobre K para $t \geq t_0$. Se busca una solución para (5.8) correspondiente a la función no decreciente $u(t)$ en la forma

$$x(t) = \xi(z(t), u_0, u(t), t) \quad (5.14)$$

donde $u_0 = u(t_0)$ y $u(t) \geq u_0$.

Acorde a la definición del sistema (5.9), la expresión (5.14) implica la existencia de

$$x(t) = z(t) + \int_{u_0}^{u(t)} b(\xi(v), v, t) dv,$$

o de

$$x(t) = z(t) + \int_0^T b(z(t) + y(r), u_0 + w(r), t) \dot{w}(r) dr, \quad (5.15)$$

donde $T = t - t_0$ es el tiempo en el que la trayectoria del sistema (5.9) alcanza el punto $x(t)$, y $y(r)$ es la solución de (5.9) correspondiente a la función no decreciente $w(r) = u(r) - u_0$.

De la resolución del sistema (5.9) sobre K [31], se tiene que

$$\int_0^{T'} b(z(t) + y(r), u_0 + w(r), t) \dot{w}(r) dr = 0 \quad (5.16)$$

para toda función suave $w(r)$ continua por partes en el intervalo $[0, T']$ del ortante positivo en \mathbb{R}^m ; es decir, $\forall w(r) \in \mathbb{R}^m$, tal que $w_i(r) \geq 0$, $i = 0, 1, \dots$, y $w(0) = w(T') = 0$; aquí T' es el tiempo de revolución sobre el ciclo $w(r) \in \mathbb{R}^m$.

Sea $w(r) \geq 0$ un ciclo suave por partes en \mathbb{R}^m tal que la correspondiente solución del sistema (5.9) con condición inicial $z(t)$ alcance el punto $x(t)$ en el tiempo $T = t - t_0$. Entonces, la ecuación (5.16) se puede escribir en la forma

$$\int_0^T b(z(t) + y(r), u_0 + w(r), t) \dot{w}(r) dr + \int_T^{T'} b(z(t) + y(r), u_0 + w(r), t) \dot{w}(r) dr =$$

$$\int_0^T b(z(t) + y(r), u_0 + w(r), t) \dot{w}(r) dr + \int_0^{T'-T} b(x(t) + y(r), u_0 + w(r), t) \dot{w}(r) dr = 0. \quad (5.17)$$

Por substitución de (5.15) en (5.17), se obtiene

$$x(t) - z(t) + \int_0^{T'-T} b(x(t) + y(r), u_0 + w(r), t) \dot{w}(r) dr = 0,$$

esto es,

$$z(t) = x(t) + \int_0^{T'-T} b(x(t) + y(r), u_0 + w(r), t) \dot{w}(r) dr,$$

o

$$z(t) = x(t) + \int_{u(t)}^{u_0} b(\xi(v), u(t), t) \dot{w}(v) dv. \quad (5.18)$$

De la representación (5.18), tal como se muestra en [30], la fórmula de inversión

$$z(t) = \xi(x(t), u(t), u_0, t) \quad (5.19)$$

es válida si $u(t) \geq u_0$. En particular, de (5.19), $z(t_0) = x(t_0) = x_0$.

Acorde a la definición de la solución del sistema (5.9), la fórmula de Leibnitz

$$\xi(z, w, u, s) = z + \int_w^u b_1(\xi(z, w, u, s), v, s) dv, \quad u \geq w,$$

es válida excepto cuando $b_1(\xi, v, s)$ sea igual a $b(\xi, v, s)$ o pertenezca al lado derecho de la correspondiente ecuación diferencial generalizada. Las derivadas $\partial\xi/\partial z$, $\partial\xi/\partial w$ y $\partial\xi/\partial s$, únicamente para los lados derechos continuos [27], satisfacen (excepto para sus conjuntos de límites con medida cero), ciertos sistemas variacionales de ecuaciones diferenciales lineales [29]. En el caso de esta prueba, los sistemas variacionales son sistemas lineales de ecuaciones diferenciales exactas con discontinuidades en el lado derecho. Bajo las condiciones del teorema 5.2, estos sistemas son solucionables sobre K [29]. Consecuentemente, las derivadas $\partial\xi/\partial z$, $\partial\xi/\partial w$ y $\partial\xi/\partial s$ existen (excepto para las condiciones establecidas anteriormente), y son continuas por partes. La existencia de la derivada $\partial\xi/\partial u$ sobre un conjunto de medida completa [10] está dada por la continuidad absoluta de la solución $\xi(z, w, u, s)$ para el sistema (5.9).

Una vez establecido que las derivadas existen, se mostrará que $z(t)$ satisface la siguiente ecuación con discontinuidades sobre el lado derecho:

$$\dot{z} = \varphi(z(t), u_0, u(t), t), \quad z(t_0) = x_0, \quad (5.20)$$

donde,

$$\varphi(z, u_0, u, t) = \frac{\partial \xi(\xi(z, u_0, u, t), u, u_0, t)}{\partial z} f(\xi(z, u_0, u, t), u, t) + \frac{\partial \xi(\xi(z, u_0, u, t), u, u_0, t)}{\partial s}.$$

En vista de que $\partial \xi / \partial z$, f y $\partial \xi / \partial u$ son continuas por partes, la función $\varphi(z, u_0, u, t)$ también es continua por partes; consecuentemente (5.20) tiene solución.

Sea $u_k(t)$, $k = 1, 2, \dots$; $* - \lim u_k(t) = u(t)$, cuando $k \rightarrow \infty$, $t \geq t_0$; una secuencia de funciones absolutamente continuas convergiendo a $u(t)$ en el sentido topológico estrella débil del espacio de las funciones de variación acotada. La ecuación (5.8) con funciones $u_k(t)$ sobre el lado derecho es una ecuación diferencial ordinaria con discontinuidades en el lado derecho:

$$\dot{x}_k(t) = f(x_k, u_k, t) + b(x_k, u_k, t) \dot{u}_k(t), \quad x_k(t_0) = x_0. \quad (5.21)$$

Por el teorema de la existencia y unicidad para ecuaciones diferenciales ordinarias con discontinuidades sobre el lado derecho [29], y de las condiciones *i*) y *ii*) de 5.2 resulta que (5.21) tiene una única solución. Además, de (5.19) se tiene que

$$\xi(x_k(t), u_k(t), u_0, t) = z_k(t)$$

es la única solución para la ecuación

$$\dot{z}_k(t) = \varphi(z_k(t), u_0, u_k(t), t), \quad z_k(t_0) = x_0. \quad (5.22)$$

Sea $* - \lim u_k(t) = u(t)$, cuando $k \rightarrow \infty$, $t \geq t_0$. Entonces, del teorema de la dependencia continua de la solución sobre el lado derecho [29], el $* - \lim z_k(t) = z^*(t)$, cuando $k \rightarrow \infty$, $t \geq t_0$, donde $z^*(t)$ es la solución de (5.20). También, esta solución es única, puesto que las soluciones $z_k(t)$ para (5.22) son únicas para las funciones $u_k(t)$. De este modo, $z^*(t)$ es una vibrosolución para la ecuación (5.20). Debido a que el mapeo (5.14) es continuo y uno a uno, entonces, $x^*(t) \triangleq \xi(z^*(t), u_0, u(t), t)$ es la vibrosolución deseada para (5.8). Aun más, $\sup_k V_{t_0}^t x_k(t) < \infty$, para $t \geq t_0$, ya que las variaciones de las funciones $z_k(t)$ y $u_k(t)$ son uniformemente acotadas en virtud de la convergencia

$$* - \lim_{k \rightarrow \infty} u_k(t) = u(t), \quad * - \lim_{k \rightarrow \infty} z_k(t) = z^*(t), \quad t \geq t_0,$$

en la topología estrella débil del espacio de las funciones de variación acotada. ■

Nota 5.1 Las vibrosoluciones pueden ser definidas, y las condiciones necesarias y suficientes para su existencia pueden ser establecidas para funciones $u(t)$ no monotónicas, las cuales, cumplan con un lado de las condiciones de Lipschitz para la función $b(x, u, t) \text{sign}(\dot{u}(t))$ y para las funciones $b(x_k, u_k, t) \dot{u}_k(t)$ para toda $k = 1, 2, \dots$

Tal y como se establece en [28, 30], sólo las funciones $x_k(t)$ que corresponden a las funciones $u_k(t)$ absolutamente continuas son las soluciones clásicas para las ecuaciones diferenciales en distribuciones de la forma (5.8). Además, no es claro cómo pueden interpretarse los saltos de la vibrosolución en los puntos de discontinuidad de la función no decreciente $u(t)$, dado que la integral

$$\int_{t_0}^t b(x(s), u(s), s) du(s)$$

no puede definirse en la forma ordinaria de la integral de Lebesgue-Stieltjes [10]. Por lo tanto, parece ser apropiado obtener una ecuación con medida equivalente a la vibrosolución de (5.8) como una solución ordinaria que permita el cálculo explícito de los saltos en los puntos de discontinuidad de $u(t)$.

Teorema 5.3 *Asuma que las condiciones de teorema 5.2 se satisfacen. Entonces, la solución de la ecuación diferencial con medida*

$$dx(t) = f(x, u, t)dt + b(x, u, t)du_c(t) + \sum_{t_i} G(x(t_i-), u(t_i-), \Delta u(t_i), t_i)d\mathcal{X}(t - t_i), \quad (5.23)$$

$x(t_0) = x_0, i = 0, 1, \dots$, coincide con la vibrosolución de (5.8), tal que,

$$G(z, w, u, s) = \xi(z, w, w + u, s) - z,$$

donde la función $\xi(z, w, u, s)$, $u \geq w$, $s \geq t_0$, es la solución del sistema (5.9) de ecuaciones diferenciales exactas; $u_c(t)$ es el componente continuo de la función no decreciente $u(t)$; y $\Delta u(t_i) = u(t_i+) - u(t_i-)$ son los saltos de $u(t)$ en sus puntos de discontinuidad t_i . $\mathcal{X}(t - t_i)$ es la función Heaviside.

Demostración: la demostración es similar a la prueba del teorema 3 de la referencia [30] bajo la suposición de que las funciones $f(x, u, t)$, $b(x, u, t)$, $\partial b(x, u, t)/\partial x$ y $\partial b(x, u, t)/\partial t$ son continuas. ■

En conclusión, se puede decir que el teorema 5.3 permite resolver numéricamente (5.8) entre los puntos de discontinuidad de $u(t)$ y conocer explícitamente sus saltos mediante el uso de (5.9) y de (5.13). ■

5.2.2 Ejemplo

Ejemplo 5.3 *Considere el sistema*

$$\dot{x}(t) = x(t) u(t), \quad x(0-) = x_0, \quad (5.24)$$

donde $u(t)$ es la función Heaviside y

$$\Delta x = G(x_0, 0, 1, 0).$$

De (5.9) se tiene que

$$\frac{dK}{du} = b(K, u, s), \quad K(v) = z, \quad (5.25)$$

por lo que,

$$b(x, u, s) = x \Rightarrow b(K, u, s) = K,$$

esto implica que (5.25) se pueda expresar como:

$$\frac{dK}{du}(u) = K(u), \quad (5.26)$$

resolviendo (5.26), resulta que

$$K(u) = ze^{u-v}.$$

En virtud del teorema 5.3, se puede hallar una ecuación de la forma (5.23), tal que,

$$G(z, v, v + u, s) = ze^{v+u-v} - z = z(e^u - 1),$$

con lo que $\Delta x = x_0(e^1 - 1)$, por lo tanto, la vibrosolución del sistema (5.24) tiene una ecuación con medida equivalente a

$$\begin{aligned} x(0+) &= x(0-) + \Delta x \\ &= x_0 + x_0(e^1 - 1) = 2.718x_0. \quad \blacksquare \end{aligned}$$

Capítulo 6

Resultados

En este capítulo se presenta como un caso particular del problema de filtrado para procesos de Ito-Volterra sobre observaciones de integrales continuas, solución a los problemas de filtrado cuando dicha ecuación de observación es discontinua y cuando es discreta ya sea con uno o múltiples retardos. También, se presentan ejemplos matemáticos para cada uno de estos casos.

Las aplicaciones prácticas para el filtrado de procesos de esta naturaleza, se intuyen en el campo del reconocimiento satelital, procesos químicos industriales, procesamiento digital de imágenes y el ya mencionado caso ecológico presa-depredador entre otros. Durante el desarrollo de los resultados, se emplearon los programas computacionales *Maple* y *Simmon* para transformar y verificar las extensas fórmulas que aparecen.

6.1 Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discontinuas

6.1.1 Declaración del Problema

Sea el par $(x(t), y(t))$ un proceso estocástico \mathcal{F}_t -medible tal que:

$$x(t) = \int_0^t (a_0(t, s) + a(t, s)x(s))ds + \int_0^t b(t, s)dW^2(s), \quad (6.1)$$

$$y_i(t) = \int_0^t [A_{0i}(t, s) + (A_i(t, s), x(s))]du_i(s) + \int_0^t B_i(t, s)dW^1(u_i(s)), \quad (6.2)$$

donde para toda $i = 1, \dots, m$, $A(t, s) = (A_1(t, s), \dots, A_m(t, s))'$, $A_i(t, s) \in R^n, \forall i$; (a, b) es el producto escalar en R^n . Considere una función vectorial de variación acotada $u(t) = (u_1(t), \dots, u_m(t)) \in R^m$, la cuál, es no decreciente en el siguiente sentido: $u(t_2) \geq u(t_1)$ cuando $t_2 \geq t_1$ si $u_i(t_2) \geq u_i(t_1)$.

El resto de la notación y el problema de estimación es formulado exactamente como en el problema descrito en la sección 4.2.

6.1.2 Filtro Óptimo

En [9] se sugiere un procedimiento para obtener las ecuaciones del filtro óptimo de (6.1) sobre observaciones discontinuas a partir de las ecuaciones de filtrado sobre observaciones continuas.

Para poder aplicar tal método, es necesario considerar las siguientes acciones:

- Asumiendo que $u(t) \in R^m$ en (6.2) es absolutamente continua, se escriben las ecuaciones de Ito-Volterra sobre observaciones continuas obtenidas en la sección 4.2.
- Así, se asume dentro de las ecuaciones obtenidas, una función vectorial de variación acotada $u(t) \in R^m$ no decreciente, con la idea de que la derivada $\dot{u}(t)$ puede ser una función generalizada.

De lo anterior se obtiene el siguiente sistema de ecuaciones de Ito-Volterra sobre observaciones discontinuas:

$$m(t) = \int_0^t (a_0(t, s) + a(t, s)m(s))ds + \int_0^t f(t, s)A'(t, s)(B(t, s)B'(t, s))^{-1} [dy(s) - (A_0(t, s) + A(t, s)m(s))du(s)], \quad (6.3)$$

$$P(t) = \int_0^t [a(t, s)f'(t, s) + f(t, s)a'(t, s) + b(t, s)b'(t, s)]ds - \int_0^t f(t, s)A'(t, s) (B(t, s)B'(t, s))^{-1} A(t, s)f'(t, s)du(s), \quad (6.4)$$

$$\begin{aligned} f(t, s) = & \int_0^s [a(t, r)f'(s, r) + f(t, r)a'(s, r) + \frac{1}{2}(b(t, r)b'(s, r) + b(s, r)b'(t, r))]dr \\ & - \int_0^s [f(t, r)A'(s, r)(B(s, r)B'(s, r))^{-1} A(s, r)f'(s, r) \\ & + f(t, r)A'(t, r)(B(t, r)B'(t, r))^{-1} A(t, r)f'(s, r) \\ & - \frac{1}{2}f(t, r)A'(t, r)(B(t, r)B'(s, r))^{-1} A(s, r)f'(s, r) \\ & - \frac{1}{2}f(s, r)A'(s, r)(B(s, r)B'(t, r))^{-1} A(t, r)f'(t, r)]du(r). \end{aligned} \quad (6.5)$$

Las ecuaciones obtenidas (6.3)-(6.5), son ecuaciones integrales con respecto a un vector de medida discontinua generada por una función de variación acotada no decreciente $u(t)$. Continuando con el esquema de [9], para poder distinguir los saltos de las ecuaciones obtenidas en sus correspondientes

puntos de discontinuidad, en la siguiente sección se utiliza el teorema 5.3.

6.2 Forma Equivalente de las Ecuaciones de Filtrado

Dado que las ecuaciones (6.3)-(6.5) pueden ser escritas de la forma de (5.8), y puesto que satisfacen las condiciones de los teoremas 5.1 y 5.2, entonces, por el teorema 5.3 se obtiene que las ecuaciones con medida equivalente para las ecuaciones de filtrado (6.3)-(6.5) toman la forma:

$$m(t) = \int_0^t (a_0(t, s) + a(t, s)m(s))ds + \int_0^t f(t-, s-)[I + A'(t, s)(B(t, s)B'(t, s))^{-1} \\ A(t, s)f(t-, s-)\Delta u(s)]^{-1} A'(t, s)(B(t, s)B'(t, s))^{-1} \\ [dy(s) - (A_0(t, s) + A(t, s)m(s-))du(s)], \quad (6.6)$$

$$P(t) = \int_0^t [a(t, s)f'(t, s) + f(t, s)a'(t, s) + b(t, s)b'(t, s)]ds - \int_0^t f(t-, s- \\ [I + A'(t, s)(B(t, s)B'(t, s))^{-1} A(t, s)f(t-, s-)\Delta u(s)]^{-1} \\ A'(t, s)(B(t, s)B'(t, s))^{-1} A(t, s)f'(t-, s-)du(s), \quad (6.7)$$

$$f(t, s) = \int_0^s [a(t, r)f'(s, r) + f(t, r)a'(s, r) + \frac{1}{2}(b(t, r)b'(s, r) + b(s, r)b'(t, r))]dr \\ - \int_0^s [f(t, r-)[I + (A'(s, r)(B(s, r)B'(s, r))^{-1} A(s, r)f(s, r-) + A'(t, r) \\ (B(t, r)B'(t, r))^{-1} A(t, r)f(t, r-) - \frac{1}{2}A'(s, r)(B(s, r)B'(t, r))^{-1} A(t, r)f(t, r-) \\ - \frac{1}{2}A'(t, r)(B(t, r)B'(s, r))^{-1} A(s, r)f(s, r-))\Delta u(r)]^{-1} A'(s, r)(B(s, r)B'(s, r))^{-1} \\ A(s, r)f'(s, r-) + f(s, r-)[I + (A'(s, r)(B(s, r)B'(s, r))^{-1} A(s, r)f(s, r-) + A'(t, r) \\ (B(t, r)B'(t, r))^{-1} A(t, r)f(t, r-) - \frac{1}{2}A'(s, r)(B(s, r)B'(t, r))^{-1} A(t, r)f(t, r-) \\ - \frac{1}{2}A'(t, r)(B(t, r)B'(s, r))^{-1} A(s, r)f(s, r-))\Delta u(r)]^{-1} A'(t, r)(B(t, r)B'(t, r))^{-1}$$

$$\begin{aligned}
& A(t, r)f'(t, r-) - \frac{1}{2}f(s, r-)[I + (A'(s, r)(B(s, r)B'(s, r))^{-1}A(s, r)f(s, r-) + A'(t, r) \\
& \quad (B(t, r)B'(t, r))^{-1}A(t, r)f(t, r-) - \frac{1}{2}A'(s, r)(B(s, r)B'(t, r))^{-1}A(t, r)f(t, r-) \\
& \quad - \frac{1}{2}A'(t, r)(B(t, r)B'(s, r))^{-1}A(s, r)f(s, r-)]\Delta u(r)]^{-1}A'(s, r)(B(s, r)B'(t, r))^{-1} \\
& A(t, r)f'(t, r-) - \frac{1}{2}f(t, r-)[I + (A'(s, r)(B(s, r)B'(s, r))^{-1}A(s, r)f(s, r-) + A'(t, r) \\
& \quad (B(t, r)B'(t, r))^{-1}A(t, r)f(t, r-) - \frac{1}{2}A'(s, r)(B(s, r)B'(t, r))^{-1}A(t, r)f(t, r-) \\
& \quad - \frac{1}{2}A'(t, r)(B(t, r)B'(s, r))^{-1}A(s, r)f(s, r-)]\Delta u(r)]^{-1}A'(t, r) \\
& \quad (B(t, r)B'(s, r))^{-1}A(s, r)f'(s, r-)]du(r), \tag{6.8}
\end{aligned}$$

donde la función $f(t, s)$ es continua en t .

Teorema 6.1 *Las ecuaciones (6.6)-(6.8) son el estimado óptimo, la función de correlación y la función característica respectivamente, en el problema de filtrado (6.1), (6.2).*

Demostración: la demostración está dada por la construcción de la ecuación con medida equivalente a la vibrosolución de cada una de las respectivas ecuaciones (6.3)-(6.5), tal y como se describió en el capítulo 5. ■

6.2.1 Ejemplo

Ejemplo 6.1 *Considere la ecuación de estado (4.16) del sistema descrito en el ejemplo 4.1, y sea la ecuación de observación:*

$$y(t) = \int_0^t [e^{4(t-s)} + e^{-2t-s}x(s)]du(s) + \int_0^t e^{3(t-s)}dW^2(u(s)), \tag{6.9}$$

donde $u(t) = \mathcal{X}(t - t_1) + \mathcal{X}(t - t_2)$ para $0 < t_1 < t_2$, es decir

$$u(t) = \begin{cases} 0 & \text{si } t \leq t_1, \\ 1 & \text{si } t_1 < t \leq t_2, \\ 2 & \text{si } t_2 < t. \end{cases}$$

tal que $\Delta u(t_i) = u(t_i+) - u(t_i-)$.

Con la ayuda de la computadora, se substituyen los respectivos coeficientes en las ecuaciones (6.6)-(6.8), con lo que se obtienen el mejor estimado, la función de correlación y la función característica

$$m(t) = \int_0^t (e^{-3t+s} + e^{-t-s}m(s))ds + \int_0^t e^{-8t+5s}[1 + e^{-10t+4s}f(t-, s-)\Delta u(s)]^{-1} f(t-, s-)[dy(s) - (e^{4(t-s)} + e^{-2t-s}m(s-))du(s)], \quad (6.10)$$

$$P(t) = \int_0^t [2e^{-t-s}f(t, s) + e^{2t-4s}]ds - \int_0^t e^{-10t+4s}f^2(t-, s-)[1 + e^{-10t+4s}f(t-, s-)\Delta u(s)]^{-1} du(s), \quad (6.11)$$

$$f(t, s) = \int_0^s [e^{-t-r}f(s, r) + e^{-s-r}f(t, r) + e^{s-4r+t}]dr - \int_0^s \{ [e^{-10s+4r} + e^{-10t+4r} - e^{-5s+4r-5t}] [1 + ((e^{-10s+4r} - \frac{1}{2}e^{-5s+4r-5t})f(s, r-) + (e^{-10t+4r} - \frac{1}{2}e^{-5s+4r-5t})f(t, r-)) \Delta u(r)]^{-1} f(t, r-)f(s, r-) \} du(r). \quad (6.12)$$

Igualando $t = s$, $f(s, s) = P(s)$, y diferenciando (6.12) con respecto a s :

$$\dot{P}(s) = e^{-2s} + 2e^{-2s}P(s) - e^{-6s}[1 + e^{-6s}P(s)\Delta u(s)]^{-1}P^2(s). \quad (6.13)$$

Si se asume que $u(t)$ es absolutamente continua, entonces, la dinámica de la función de correlación es la misma que en el caso de observaciones continuas que se muestra en la figura 4.1 (a). Por otra parte, de acuerdo con el planteamiento $u(t)$ es discontinua, esto implica que los saltos de la varianza están dados por la expresión:

$$\Delta P(t_i) = P(t_i+) - P(t_i-),$$

donde,

$$P(t+) = \int_0^{t+} [2e^{-t-s} f(t, s) + e^{2t-4s}] ds - \int_0^{t+} e^{-10t+4s} f^2(t-, s-) [1 + e^{-10t+4s} f(t-, s-) \Delta u(s)]^{-1} du(s), \quad (6.14)$$

$$P(t-) = \int_0^{t-} [2e^{-t-s} f(t, s) + e^{2t-4s}] ds - \int_0^{t-} e^{-10t+4s} f^2(t-, s-) [1 + e^{-10t+4s} f(t-, s-) \Delta u(s)]^{-1} du(s), \quad (6.15)$$

por lo tanto,

$$\Delta P(t) = -e^{-10t+4s} f^2(t-, s-) [1 + e^{-10t+4s} f(t-, s-) \Delta u(s)]^{-1} \Delta u(s), \quad (6.16)$$

con $t_1 = 3/2$ y $t_2 = 3$, los saltos están dados por $\Delta u(3/2) = \Delta u(3) = 1$, con lo que se obtiene

$$\begin{aligned} \Delta P(3/2) &= -e^{-9} P^2(3/2) [1 + e^{-9} P(3/2)]^{-1} \\ &= -e^{-9} (0.7823)^2 [1 + e^{-9} (0.7823)]^{-1} \\ &= -7.5518 \times 10^{-5}, \end{aligned}$$

y

$$\begin{aligned} \Delta P(3) &= -e^{-18} (0.8444)^2 [1 + e^{-18} (0.8444)]^{-1} \\ &= -1.0859 \times 10^{-8}, \end{aligned}$$

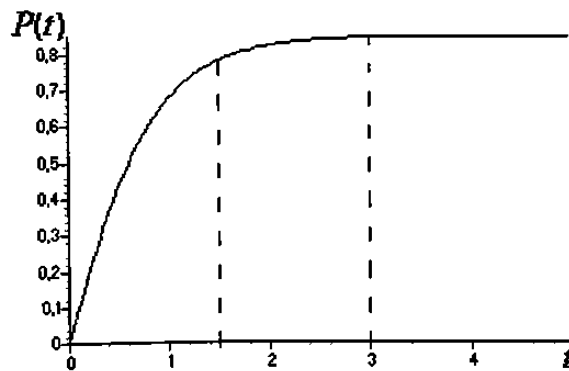


Figura 6.1: Función de Correlación Discontinua

donde $P(3/2) = 0.7823$ y $P(3) = 0.8444$. Tal y como se puede ver en la figura 6.1, los saltos de la

función de correlación son pequeños, por lo que se obtiene para la función de error una curva similar a la de la figura 4.1 (a) (pero con los correspondientes puntos de discontinuidad), que converge al valor de 0.84769.

Del mismo modo, los saltos del mejor estimado están dados por:

$$\Delta m(t) = e^{-3t} [1 + e^{-6t} P(t) \Delta u(t)]^{-1} P(t-) [\Delta y(t) - (1 + e^{-3t} m(t-)) \Delta u(t)], \quad (6.17)$$

entonces,

$$\begin{aligned} \Delta m(3/2) &= e^{-9/2} [1 + e^{-9} (0.7823)]^{-1} (0.7823) [2.0065 - 1 - e^{-9/2} (0.603)] \\ &= 8.688 \times 10^{-3} \end{aligned}$$

y

$$\begin{aligned} \Delta m(3) &= e^{-9} [1 + e^{-18} (0.8444)]^{-1} (0.8444) [2.0001 - 1 - e^{-9} (0.632)] \\ &= 1.0421 \times 10^{-4}. \end{aligned}$$

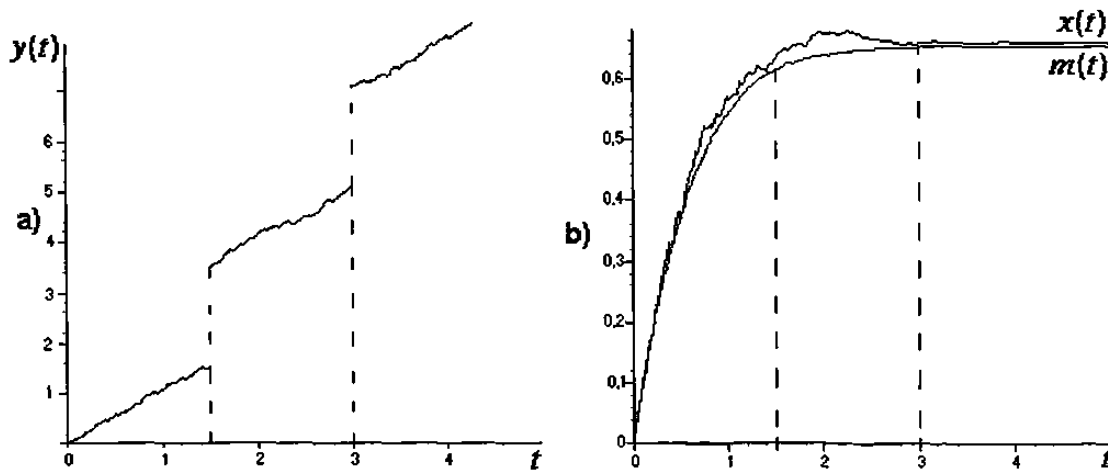


Figura 6.2: Ecuaciones (a) de Observación Discontinua y (b) de Estado y Estimado

Los saltos de la ecuación (6.17) están representados en la figura 6.2 (b) como simples puntos de discontinuidad dada la magnitud de cada uno de ellos. Para lograr lo anterior, es necesario calcular

$$\Delta y(t) = [1 + e^{-3t} x(t)] \Delta u(t) + \Delta u(t) = e^{-3t} x(t) + 2, \quad (6.18)$$

entonces,

$$\Delta y(3/2) = e^{-9/2}x(3/2) + 2 = e^{-9/2}(0.584) + 2 = 2.0065,$$

$$\Delta y(3) = e^{-9}x(3) + 2 = e^{-9}0.649 + 2 = 2.0001;$$

los saltos que describen el comportamiento de la ecuación (6.18) están representados en la figura 6.2 (a). Los valores fijos de $x(3/2) = 0.584$, $x(3) = 0.649$, $m(3/2) = 0.603$ y $m(3) = 0.632$ se obtienen del caso de observaciones continuas resuelto en el ejemplo 4.1. ■

6.3 Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discretas con Retardo

6.3.1 Planteamiento del Problema

Considere ahora el problema de filtrado para un proceso de Ito-Volterra sobre observaciones discretas. Sea la ecuación de estado (6.1), y la ecuación de observación:

$$y(t_j) = A_0(t_j, t_i) + A(t_j, t_i)x(t_i) + B(t_j, t_i)\psi(t_i) \quad (6.19)$$

donde $y(t_j) \in R^m$, $j = 0, 1, \dots$; $x(t_i)$ es el valor del estado en el momento t_i disponible para la observación en el momento t_j , $A(t_j, t_i) \in R^{m \times n}$ son matrices de transición, y $\psi(t_i)$ son ruidos Gaussianos. El modelo de observación tiene un único tiempo de retardo $t_j - t_i$. Note que (6.19) se puede obtener de (6.2) mediante la substitución de $u(t) = \mathcal{X}(t - t_i)$, donde $\mathcal{X}(t - t_i)$ es la función Heaviside. Sea $\Delta u(t_k) = u(t_k+) - u(t_k-)$, los saltos de la función $u(t_k)$ al tiempo t_k , para toda $k = 1, 2, \dots$

El problema del filtrado es definido de forma generalizada para el caso de observaciones discretas, en el mismo sentido descrito en la Sección anterior.

6.3.2 Filtro Óptimo

Teorema 6.2 *Dadas las ecuaciones (6.6)-(6.8), entonces, el mejor estimado $m(t)$, la función de correlación $P(t)$ y la función característica $f(t, s)$ del sistema descrito por (6.1) y (6.19), satisfacen*

las siguientes ecuaciones entre los consecutivos momentos de observación t_j :

$$m(t) = m(t_{j+}) + \int_{t_{j+}}^t (a_0(t, s) + a(t, s)m(s))ds, \quad (6.20)$$

$$P(t) = P(t_{j+}) + \int_{t_{j+}}^t [a(t, s)f'(t, s) + f(t, s)a'(t, s) + b(t, s)b'(t, s)]ds, \quad (6.21)$$

$$f(t, s) = f(t, t_{j+}) + \int_{t_{j+}}^s [a(t, r)f'(s, r) + f(t, r)a'(s, r) + \frac{1}{2}(b(t, r)b'(s, r) + b(s, r)b'(t, r))]dr, \quad (6.22)$$

con saltos:

$$\begin{aligned} \Delta m(t_j) &= f(t_{j-}, t_{j-})[I + A'(t_j, t_j)(B(t_j, t_j)B'(t_j, t_j))^{-1}A(t_j, t_j)f(t_{j-}, t_{j-})]^{-1}A'(t_j, t_j) \\ &\quad (B(t_j, t_j)B'(t_j, t_j))^{-1}y(t_j) - f(t_{j-}, t_{i-})[I + A'(t_j, t_i)(B(t_j, t_i)B'(t_j, t_i))^{-1}A(t_j, t_i) \\ &\quad f(t_{j-}, t_{i-})]^{-1}A'(t_j, t_i)(B(t_j, t_i)B'(t_j, t_i))^{-1}(A_0(t_j, t_i) + A(t_j, t_i)m(t_{i-})), \end{aligned} \quad (6.23)$$

$$\begin{aligned} \Delta P(t_j) &= -f(t_{j-}, t_{i-})[I + A'(t_j, t_i)(B(t_j, t_i)B'(t_j, t_i))^{-1}A(t_j, t_i)f(t_{j-}, t_{i-})]^{-1}A'(t_j, t_i) \\ &\quad (B(t_j, t_i)B'(t_j, t_i))^{-1}A(t_j, t_i)f'(t_{j-}, t_{i-}), \end{aligned} \quad (6.24)$$

$$\begin{aligned} \Delta f(t, t_j) &= -f(t, t_{i-})[I + A'(t_j, t_i)(B(t_j, t_i)B'(t_j, t_i))^{-1}A(t_j, t_i)f(t_{j-}, t_{i-}) + A'(t, t_i) \\ &\quad (B(t, t_i)B'(t, t_i))^{-1}A(t, t_i)f(t, t_{i-}) - \frac{1}{2}A'(t_j, t_i)(B(t_j, t_i)B'(t, t_i))^{-1} \\ &\quad A(t, t_i)f(t, t_{i-}) - \frac{1}{2}A'(t, t_i)(B(t, t_i)B'(t_j, t_i))^{-1}A(t_j, t_i)f(t_{j-}, t_{i-})]^{-1} \\ &\quad A'(t_j, t_i)(B(t_j, t_i)B'(t_j, t_i))^{-1}A(t_j, t_i)f'(t_{j-}, t_{i-}) \\ &\quad -f(t_{j-}, t_{i-})[I + A'(t_j, t_i)(B(t_j, t_i)B'(t_j, t_i))^{-1}A(t_j, t_i)f(t_{j-}, t_{i-}) + A'(t, t_i) \\ &\quad (B(t, t_i)B'(t, t_i))^{-1}A(t, t_i)f(t, t_{i-}) - \frac{1}{2}A'(t_j, t_i)(B(t_j, t_i)B'(t, t_i))^{-1} \end{aligned}$$

$$\begin{aligned}
& A(t, t_i) f(t, t_i-) - \frac{1}{2} A'(t, t_i) (B(t, t_i) B'(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-)]^{-1} \\
& \quad A'(t, t_i) (B(t, t_i) B'(t, t_i))^{-1} A(t, t_i) f'(t, t_i-) \\
& + \frac{1}{2} f(t_j, t_i-) [I + A'(t_j, t_i) (B(t_j, t_i) B'(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-) + A'(t, t_i) \\
& \quad (B(t, t_i) B'(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \frac{1}{2} A'(t_j, t_i) (B(t_j, t_i) B'(t, t_i))^{-1} \\
& \quad A(t, t_i) f(t, t_i-) - \frac{1}{2} A'(t, t_i) (B(t, t_i) B'(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-)]^{-1} \\
& \quad A'(t_j, t_i) (B(t_j, t_i) B'(t, t_i))^{-1} A(t, t_i) f'(t, t_i-) \\
& + \frac{1}{2} f(t, t_i-) [I + A'(t_j, t_i) (B(t_j, t_i) B'(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-) + A'(t, t_i) \\
& \quad (B(t, t_i) B'(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \frac{1}{2} A'(t_j, t_i) (B(t_j, t_i) B'(t, t_i))^{-1} \\
& \quad A(t, t_i) f(t, t_i-) - \frac{1}{2} A'(t, t_i) (B(t, t_i) B'(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-)]^{-1} \\
& \quad A'(t, t_i) (B(t, t_i) B'(t_j, t_i))^{-1} A(t_j, t_i) f'(t_j, t_i-), \tag{6.25}
\end{aligned}$$

Demostración: Entre cada momento de observación discreta, la ecuación de estados tiene un comportamiento continuo, por lo que, tomando de (6.6)-(6.8) sus correspondientes intervalos continuos sujetos a una condición inicial sobre el lado derecho de las discontinuidades se escriben los componentes del estimado, correlación y característica, que en este caso son las integrales (6.20)-(6.4). Por otro lado y con la ayuda de la computadora, se substituye $u(t) = \mathcal{X}(t - t_i)$ en las ecuaciones (6.6)-(6.8) y se obtienen (6.20)-(6.25). ■

6.3.3 Ejemplos

Ejemplo 6.2 Considere la ecuación de estado (4.16) del sistema descrito en el ejemplo 4.1, y sea la ecuación de observación discreta:

$$y(t_j) = e^{t_j - t_i} x(t_i) - e^{-\frac{1}{2} t_j - t_i} \psi(t_i), \tag{6.26}$$

donde $t_i = 0, 1, \dots$

Se tiene que, entre los consecutivos momentos de observación t_j :

$$m(t) = m(t_{j+}) + \int_{t_{j+}}^t [e^{-3t+s} + e^{-t-s}m(s)]ds, \quad (6.27)$$

$$P(t) = P(t_{j+}) + \int_{t_{j+}}^t [e^{2t-4s} + 2e^{-t-s}f(t, s)]ds, \quad (6.28)$$

$$f(t, s) = f(t, t_{j+}) + \int_{t_{j+}}^s [e^{t-4r+s} + e^{-t-r}f(s, r) + e^{-s-r}f(t, r)]dr. \quad (6.29)$$

Los saltos de estas funciones en los tiempos de observación t_j , están caracterizados por:

$$\begin{aligned} \Delta m(t_j) &= e^{3t_j} f(t_{j-}, t_{j-}) [1 + e^{3t_j} f(t_{j-}, t_{j-})]^{-1} y(t_j) - e^{2t_j+t_i} f(t_{j-}, t_{i-}) \\ &\quad [1 + e^{3t_j} f(t_{j-}, t_{i-})]^{-1} e^{t_j-t_i} m(t_{i-}), \end{aligned} \quad (6.30)$$

$$\Delta P(t_j) = -e^{3t_j} [1 + e^{3t_j} f(t_{j-}, t_{i-})]^{-1} f^2(t_{j-}, t_{i-}), \quad (6.31)$$

$$\begin{aligned} \Delta f(t, t_j) &= -(e^{3t_j} + e^{3t} - e^{\frac{3}{2}t_j + \frac{3}{2}t}) [1 + e^{3t_j} f(t_j, t_{i-}) + e^{3t} f(t, t_{i-}) - \frac{1}{2} e^{\frac{3}{2}t_j + \frac{3}{2}t} f(t, t_{i-}) \\ &\quad - \frac{1}{2} e^{\frac{3}{2}t_j + \frac{3}{2}t} f(t_j, t_{i-})]^{-1} f(t, t_{i-}) f(t_j, t_{i-}). \end{aligned} \quad (6.32)$$

Suponga que los momentos de observación están dados para $t_j = 0, 3/2, 3, 9/2, 6, \dots$; y que el estado se puede observar en los momentos $t_i = 0, 1/2, 1, 3/2, 2, \dots$; entonces, la ecuación de observación (6.26) al tiempo $3/2$ con retardo en $1/2$, está dada por:

$$\begin{aligned} y(3/2) &= e^{\frac{3}{2}-1/2} x(1/2) - e^{-\frac{3}{4}-1/2} \psi(1/2) \\ &= e^1 (0.391) - e^{-\frac{5}{4}} (0.177) \\ &= 1.0628 - 5.0711 \times 10^{-2} \\ &= 1.0121 \end{aligned} \quad (6.33)$$

donde los valores aleatorios de una de las realizaciones de $x(t)$ y de $\psi(t)$ se toman de la tabla 6.1.

Tabla 6.1: Realización para el Estado y Ruido Gaussiano

$x(0) = 0$	$\psi(0) = 0$
$x(1/2) = 0.391$	$\psi(1/2) = 0.177$
$x(1) = 0.555$	$\psi(1) = 0.545$
$x(3/2) = 0.610$	$\psi(3/2) = 0.401$
$x(2) = 0.643$	$\psi(2) = 0.631$
$x(5/2) = 0.615$	$\psi(5/2) = 0.958$
$x(3) = 0.649$	$\psi(3) = 0.672$

El cálculo del estimado $m(t)$, se obtiene (aproximadamente), de la derivada completa de la ecuación (6.27), es decir,

$$\dot{m}(t) = e^{-2t} + e^{-2t}m(t), \quad m(0+) = 0, \quad (6.34)$$

de la ecuación (6.34), se puede obtener de modo inmediato $m(1/2) = 0.372$, $m(1) = 0.541$ y $m(3/2) = 0.608$, tal que para el siguiente tiempo de observación discreta es necesario tomar la condición inicial como $m(3/2+) = m(3/2) + \Delta m(3/2)$. Entonces,

$$\begin{aligned} \Delta m(3/2) &= e^{\frac{7}{2}} f(3/2, 3/2) [1 + e^{9/2} f(3/2, 3/2)]^{-1} y(3/2) \\ &\quad - e^{\frac{7}{2}} f(3/2, 1/2) [1 + e^{9/2} f(3/2, 1/2)]^{-1} e^1 m(1/2) \\ &= e^{\frac{7}{2}} (2.049) [1 + e^{9/2} (2.049)]^{-1} (1.0121) \\ &\quad - e^{\frac{7}{2}} (1.4027) [1 + e^{9/2} (1.4027)]^{-1} e^1 (0.372) \\ &= 0.37032 - 0.36907 \\ &= 1.2455 \times 10^{-3} \end{aligned}$$

Del mismo modo los saltos de la varianza al tiempo de observación t_j :

$$\begin{aligned} \Delta P(3/2) &= -e^{9/2} [1 + e^{9/2} f(3/2, 1/2)]^{-1} f^2(3/2, 1/2) \\ &= -e^{9/2} [1 + e^{9/2} (1.4027)]^{-1} (1.4027)^2 \\ &= -1.3916 \end{aligned}$$

donde los valores $f(3/2, 1/2) = 1.4027$, $f(3/2, 1) = 1.8834$ y de $f(3/2, 3/2) = 2.0493$ se

obtienen de la función $f(3/2, s)$:

$$\dot{f}(3/2, s) = e^{\frac{3}{2}-3s} + e^{-\frac{3}{2}-s}P(s) + e^{-2s}f(3/2, s),$$

con

$$\dot{P}(t) = e^{-2t} + 2e^{-2t}P(t).$$

Con esto, la trayectoria de la ecuación (6.27) en el siguiente intervalo satisface:

$$\begin{aligned} m(2) &= 0.634 + 1.2455 \times 10^{-3} \\ &= 0.635, \end{aligned}$$

$$\begin{aligned} m(5/2) &= 0.643 + 1.2455 \times 10^{-3} \\ &= 0.644, \end{aligned}$$

y

$$\begin{aligned} m(3) &= 0.647 + 1.2455 \times 10^{-3} \\ &= 0.648. \end{aligned}$$

Los siguientes puntos de observación t_j con retardo y los saltos posteriores, se calculan de la misma manera, por lo que se concluye con esta parte y se analizarán los cambios implicados en el estimado cuando se toma un retardo menor al estudiado anteriormente, es decir, cuando la ecuación de observación al tiempo $3/2$ tiene un retardo en 1. Para este nuevo caso, la nueva ecuación de observación está dada por:

$$\begin{aligned} y(3/2) &= e^{\frac{1}{2}}x(1) - e^{-\frac{7}{4}}\psi(1) \\ &= 0.91504 - 9.4707 \times 10^{-2} \\ &= 0.82033, \end{aligned} \tag{6.35}$$

y el salto del estimado y de la varianza por:

$$\Delta m(3/2) = e^4 f(3/2, 3/2)[1 + e^{9/2} f(3/2, 3/2)]^{-1} y(3/2)$$

$$\begin{aligned}
& -e^4 f(3/2, 1)[1 + e^{9/2} f(3/2, 1)]^{-1} e^{\frac{1}{2}} m(1) \\
= & e^4 (2.049)[1 + e^{9/2} (2.049)]^{-1} (0.82033) \\
& -e^4 (1.8834)[1 + e^{9/2} (1.8834)]^{-1} e^{\frac{1}{2}} (0.541) \\
= & -4.2949 \times 10^{-2}
\end{aligned}$$

$$\begin{aligned}
\Delta P(3/2) &= -e^{9/2} [1 + e^{9/2} f(3/2, 1)]^{-1} f^2(3/2, 1) \\
&= -e^{9/2} [1 + e^{9/2} (1.8834)]^{-1} (1.8834)^2 \\
&= -1.8723.
\end{aligned}$$

Considere ahora que el filtrado se efectúa sin retardos, entonces, la ecuación de observación, el salto del estimado y de la función de correlación en el tiempo 3/2, estará dada respectivamente por:

$$\begin{aligned}
y(3/2) &= x(3/2) - e^{-\frac{9}{4}} \psi(3/2) \\
&= 0.610 - e^{-\frac{9}{4}} (0.401) \\
&= 0.56773,
\end{aligned}$$

$$\begin{aligned}
\Delta m(3/2) &= e^{\frac{9}{2}} f(3/2, 3/2)[1 + e^{9/2} f(3/2, 3/2)]^{-1} y(3/2) \\
&\quad - e^{\frac{9}{2}} f(3/2, 3/2)[1 + e^{9/2} f(3/2, 3/2)]^{-1} m(3/2) \\
&= e^{\frac{9}{2}} (2.0493)[1 + e^{9/2} (2.0493)]^{-1} (0.56773) \\
&\quad - e^{\frac{9}{2}} (2.0493)[1 + e^{9/2} (2.0493)]^{-1} (0.608) \\
&= 0.56467 - 0.60473 \\
&= -0.04006
\end{aligned}$$

y

$$\begin{aligned}
\Delta P(3/2) &= -e^{\frac{9}{2}} [1 + e^{\frac{9}{2}} f(3/2, 3/2)]^{-1} f^2(3/2, 3/2) \\
&= -e^{\frac{9}{2}} [1 + e^{\frac{9}{2}} (2.0493)]^{-1} (2.0493)^2 \\
&= -2.0383. \blacksquare
\end{aligned}$$

Note que $\Delta P(1/2)$, $\Delta P(1)$ y $\Delta P(3/2)$ tienen el comportamiento cualitativo esperado, pero no así

la magnitud apropiada, y esto es, debido a que los valores de la función $f(t, s)$ son aproximados, situación que se explicó con anterioridad al final del ejemplo 4.1.

Como conclusión se puede señalar que la magnitud del ruido Gaussiano obtenido en la realización de la ecuación de observación con el correspondiente retardo, determina de manera significativa el tamaño y sentido del salto del estimado. También, se aprecia que a medida que el retardo es menor con respecto al tiempo de observación, la varianza tiene un salto negativo de mayor tamaño por lo que la estimación es mejorada a medida que el retardo es menor.

Ejemplo 6.3 *Considere el ejemplo 6.2. Asuma que la ecuación de observación (6.26) es un proceso acumulado de cada una de las observaciones con retardo en $1/2$, 1 y $3/2$, es decir, $y(3/2)$ contiene múltiples retardos valuados independientemente.*

Entonces,

$$y(3/2) = 1.0121 + 0.82033 + 0.56773 = 2.4002,$$

por lo que el estimado está dado por:

$$\begin{aligned} \Delta m(3/2) &= 1.2455 \times 10^{-3} - 4.2949 \times 10^{-2} - 4.006 \times 10^{-2} \\ &= -8.1764 \times 10^{-2} \end{aligned}$$

con lo que después del correspondiente salto, el estimado:

$$\begin{aligned} m(2) &= 0.634 - 8.1764 \times 10^{-2} = 0.552, \\ m(5/2) &= 0.643 - 8.1764 \times 10^{-2} = 0.561, \\ m(3) &= 0.647 - 8.1764 \times 10^{-2} = 0.565. \end{aligned}$$

El procedimiento que aquí se termina, continua con el cálculo de $\Delta m(3)$, y así sucesivamente. ■

Note que la información retardada es independiente entre sí, esto implica que al momento de se dispone de un conjunto de datos separados pero que se conocen al mismo tiempo, este resultado es útil cuando se desea conocer un conjunto de eventos independientes que ocurren en el pasado y de los que solo se tiene posibilidad de saber de ellos en un tiempo fijo después y se le conoce como *procesado por tandas*. Una aplicación práctica de este resultado puede darse en el área espacial, específicamente, en la transmisión de fotografías satelitales a la correspondiente estación de control

y análisis.

Otro problema importante y que se resuelve en la siguiente sección, es cuando se requiere que los datos referentes al conjunto de eventos pasados estén fusionados o relacionados entre sí y disponibles para su observación en un tiempo fijo t_j . En este sentido, se pueden hallar aplicaciones prácticas en el mejoramiento de imágenes, particularmente el mejoramiento de las imágenes de placas radiográficas digitales como las empleadas por *R. Alvarado* y *J. A. de la O* en [33].

6.4 Filtrado para Procesos de Ito-Volterra sobre Observaciones de Ito-Volterra Discretas con Múltiples Retardos

6.4.1 Descripción del Problema

Sea la ecuación de estado (6.1), y la ecuación de observación:

$$y(t_j) = \sum_{i=0} [A_0(t_j, t_{ji}) + A(t_j, t_{ji})x(t_{ji}) + B(t_j, t_{ji})\psi(t_{ji})], \quad (6.36)$$

donde $y(t_j) \in R^m$, $j = 0, 1, \dots$; $x(t_{ji})$ son los valores del sistema de estado en los momentos t_{ji} disponibles para la observación en los momentos t_j , $A(t_j, t_{ji}) \in R^{m \times n}$ son matrices de transición, y $\psi(t_{ji})$ son ruidos Gaussianos. Al igual que en el caso de observaciones discretas con un tiempo de retardo, la ecuación (6.36) se puede definir a partir de substituir $u(t) = \sum_{i=0} \mathcal{X}(t - t_{ji})$ en la ecuación de observación (6.2) del caso discontinuo, por lo tanto, el modelo de observación tiene n tiempos de retardo $t_j - t_{ji}$. Deberá considerarse en el planteamiento que el conjunto retardos utilizados es de medida cero.

Una vez más, el problema se define en el sentido mencionado con anterioridad.

6.4.2 Filtro Óptimo

Corolario 6.3 Sean las ecuaciones (6.6)-(6.8). Entonces el mejor estimado $m(t)$, la función de error $P(t)$, y la función característica $f(t, s)$ para el sistema (6.1)-(6.36) entre los consecutivos momentos de observación t_j , satisfacen las ecuaciones (6.20)-(6.22). Además los saltos de estas

funciones en los tiempos de observación t_j , están caracterizados por:

$$\begin{aligned} \Delta m(t_j) &= f(t_{j-}, t_{j-}) [I + A'(t_j, t_j)(B(t_j, t_j)B'(t_j, t_j))^{-1}A(t_j, t_j)f(t_{j-}, t_{j-})]^{-1} \\ &\quad A'(t_j, t_j)(B(t_j, t_j)B'(t_j, t_j))^{-1}y(t_j) - \sum_{i=0}^{j-1} \{f(t_{j-}, t_{ji-}) \\ &\quad [I + A'(t_j, t_{ji})(B(t_j, t_{ji})B'(t_j, t_{ji}))^{-1}A(t_j, t_{ji})f(t_{j-}, t_{ji-})]^{-1} \\ &\quad A'(t_j, t_{ji})(B(t_j, t_{ji})B'(t_j, t_{ji}))^{-1}(A_0(t_j, t_{ji}) + A(t_j, t_{ji})m(t_{ji-}))\}, \end{aligned} \quad (6.37)$$

$$\begin{aligned} \Delta P(t_j) &= - \sum_{i=0}^{j-1} \{f(t_{j-}, t_{ji-}) [I + A'(t_j, t_{ji})(B(t_j, t_{ji})B'(t_j, t_{ji}))^{-1}A(t_j, t_{ji})f(t_{j-}, t_{ji-})]^{-1} \\ &\quad A'(t_j, t_{ji})(B(t_j, t_{ji})B'(t_j, t_{ji}))^{-1}A(t_j, t_{ji})f'(t_{j-}, t_{ji-})\}, \end{aligned} \quad (6.38)$$

$$\begin{aligned} \Delta f(t, t_j) &= - \sum_{i=0}^{j-1} \{f(t, t_{ji-}) [I + A'(t_j, t_{ji})(B(t_j, t_{ji})B'(t_j, t_{ji}))^{-1}A(t_j, t_{ji})f(t_j, t_{ji-}) \\ &\quad + A'(t, t_{ji})(B(t, t_{ji})B'(t, t_{ji}))^{-1}A(t, t_{ji})f(t, t_{ji-}) - \frac{1}{2}A'(t_j, t_{ji}) \\ &\quad (B(t_j, t_{ji})B'(t, t_{ji}))^{-1}A(t, t_{ji})f(t, t_{ji-}) - \frac{1}{2}A'(t, t_{ji})(B(t, t_{ji})B'(t_j, t_{ji}))^{-1} \\ &\quad A(t_j, t_{ji})f(t_j, t_{ji-})]^{-1}A'(t_j, t_{ji})(B(t_j, t_{ji})B'(t_j, t_{ji}))^{-1}A(t_j, t_{ji})f'(t_j, t_{ji-}) \\ &\quad + f(t_j, t_{ji-}) [I + A'(t_j, t_{ji})(B(t_j, t_{ji})B'(t_j, t_{ji}))^{-1}A(t_j, t_{ji})f(t_j, t_{ji-}) \\ &\quad + A'(t, t_{ji})(B(t, t_{ji})B'(t, t_{ji}))^{-1}A(t, t_{ji})f(t, t_{ji-}) - \frac{1}{2}A'(t_j, t_{ji}) \\ &\quad (B(t_j, t_{ji})B'(t, t_{ji}))^{-1}A(t, t_{ji})f(t, t_{ji-}) - \frac{1}{2}A'(t, t_{ji})(B(t, t_{ji})B'(t_j, t_{ji}))^{-1} \\ &\quad A(t_j, t_{ji})f(t_j, t_{ji-})]^{-1}A'(t, t_{ji})(B(t, t_{ji})B'(t, t_{ji}))^{-1}A(t, t_{ji})f'(t, t_{ji-}) \\ &\quad - \frac{1}{2}f(t_j, t_{ji-}) [I + A'(t_j, t_{ji})(B(t_j, t_{ji})B'(t_j, t_{ji}))^{-1}A(t_j, t_{ji})f(t_j, t_{ji-}) \\ &\quad + A'(t, t_{ji})(B(t, t_{ji})B'(t, t_{ji}))^{-1}A(t, t_{ji})f(t, t_{ji-}) - \frac{1}{2}A'(t_j, t_{ji}) \end{aligned}$$

$$\begin{aligned}
& (B(t_j, t_{j_i})B'(t, t_{j_i}))^{-1}A(t, t_{j_i})f(t, t_{j_i-}) - \frac{1}{2}A'(t, t_{j_i})(B(t, t_{j_i})B'(t_j, t_{j_i}))^{-1} \\
& A(t_j, t_{j_i})f(t_j, t_{j_i-})]^{-1}A'(t_j, t_{j_i})(B(t_j, t_{j_i})B'(t, t_{j_i}))^{-1}A(t, t_{j_i})f'(t, t_{j_i-}) \\
& - \frac{1}{2}f(t, t_{j_i-})[I + A'(t_j, t_{j_i})(B(t_j, t_{j_i})B'(t_j, t_{j_i}))^{-1}A(t_j, t_{j_i})f(t_j, t_{j_i-}) \\
& + A'(t, t_{j_i})(B(t, t_{j_i})B'(t, t_{j_i}))^{-1}A(t, t_{j_i})f(t, t_{j_i-}) - \frac{1}{2}A'(t_j, t_{j_i}) \\
& (B(t_j, t_{j_i})B'(t, t_{j_i}))^{-1}A(t, t_{j_i})f(t, t_{j_i-}) - \frac{1}{2}A'(t, t_{j_i})(B(t, t_{j_i})B'(t_j, t_{j_i}))^{-1} \\
& A(t_j, t_{j_i})f(t_j, t_{j_i-})]^{-1}A'(t, t_{j_i})(B(t, t_{j_i})B'(t_j, t_{j_i}))^{-1}A(t_j, t_{j_i})f'(t_j, t_{j_i-})\}. \quad (6.39)
\end{aligned}$$

Demostración: De manera similar al caso anterior, substituyendo $u(t) = \sum_{i=0}^{\infty} \mathcal{X}(t - t_{j_i})$ en las ecuaciones (6.6)-(6.8) obtenidas en el caso del filtrado sobre observaciones discontinuas, se obtienen de modo directo, las ecuaciones (6.20)-(6.22) y (6.37)-(6.39). ■

6.4.3 Ejemplo

Ejemplo 6.4 Considere el ejemplo 6.2. Sea la ecuación de observación discreta:

$$y(t_j) = \sum_{i=0}^{\infty} [e^{t_j-t_{j_i}}x(t_{j_i}) - e^{-\frac{1}{2}t_j-t_{j_i}}\psi(t_{j_i})], \quad (6.40)$$

donde $t_i = 0, 1, \dots$

Entonces, el mejor estimado $m(t)$, la función de correlación $P(t)$ y la función característica $f(t, s)$ del sistema descrito por (4.16) y (6.40), satisfacen las ecuaciones (6.27), (6.28) y (6.29) entre los consecutivos momentos de observación t_j .

Suponga los momentos de observación dados anteriormente en el ejemplo 6.2; y que el estado se puede observar en los momentos $t_{j_i} = 0, 1/2, 1, 3/2, 2, \dots$; entonces, los saltos de estas funciones en los tiempos de observación t_j , están caracterizados por:

$$\begin{aligned}
\Delta m(t_j) = e^{3t_j} f(t_j-, t_{j-}) [1 + e^{3t_j} f(t_j-, t_{j-})]^{-1} y(t_j) - \sum_{i=0}^{\infty} \{ e^{2t_j+t_{j_i}} f(t_j-, t_{j_i-}) \\
[1 + e^{3t_j} f(t_j-, t_{j_i-})]^{-1} e^{t_j-t_{j_i}} m(t_{j_i-}) \}, \quad (6.41)
\end{aligned}$$

$$\Delta P(t_j) = - \sum_{i=0} \{e^{3t_j} [1 + e^{3t_j} f(t_{j-}, t_{j_i-})]^{-1} f^2(t_{j-}, t_{j_i-})\}, \quad (6.42)$$

$$\begin{aligned} \Delta f(t, t_j) = & - \sum_{i=0} \{ (e^{3t_j} + e^{3t} - e^{\frac{3}{2}t_j + \frac{3}{2}t}) [1 + e^{3t_j} f(t_j, t_{j_i-}) + e^{3t} f(t, t_{j_i-}) - \frac{1}{2} e^{\frac{3}{2}t_j + \frac{3}{2}t} f(t, t_{j_i-}) \\ & - \frac{1}{2} e^{\frac{3}{2}t_j + \frac{3}{2}t} f(t_j, t_{j_i-})]^{-1} f(t, t_{j_i-}) f(t_j, t_{j_i-}) \}. \end{aligned} \quad (6.43)$$

La metodología por emplear es la misma a la utilizada en los ejemplos 6.2 y 6.3, por lo que se consideran los siguientes casos:

Primero: Asuma que la ecuación de observación en 3/2 tiene retardos en 3/2 y 1 a la vez, esto implica que:

$$\begin{aligned} y(3/2) &= e^{\frac{1}{2}} x(1) - e^{-\frac{7}{4}} \psi(1) + x(3/2) - e^{-\frac{9}{4}} \psi(3/2) \\ &= 0.91504 - 9.4707 \times 10^{-2} + 0.610 - e^{-\frac{9}{4}} (0.401) \\ &= 0.56773 + 0.82033 \\ &= 1.3881, \end{aligned}$$

con lo que el estimado tiene un salto:

$$\begin{aligned} \Delta m(3/2) &= e^{\frac{9}{2}} f(3/2, 3/2) [1 + e^{9/2} f(3/2, 3/2)]^{-1} y(3/2) \\ &\quad - e^4 f(3/2, 1) [1 + e^{9/2} f(3/2, 1)]^{-1} e^{\frac{1}{2}} m(1) \\ &\quad - e^{\frac{9}{2}} f(3/2, 3/2) [1 + e^{9/2} f(3/2, 3/2)]^{-1} m(3/2) \\ &= e^{\frac{9}{2}} (2.0493) [1 + e^{9/2} (2.0493)]^{-1} (1.3881) \\ &\quad - e^4 (1.8834) [1 + e^{9/2} (1.8834)]^{-1} e^{\frac{1}{2}} (0.541) \\ &\quad - e^{\frac{9}{2}} (2.0493) [1 + e^{9/2} (2.0493)]^{-1} (0.608) \\ &= 1.3807 - 0.60473 - 0.53782 \\ &= 0.23815. \end{aligned}$$

Segundo: Tomando retardos en 1/2, 1 y 3/2 para la ecuación de observación en 3/2 :

$$y(3/2) = \sum_{i=0} [e^{\frac{3}{2}-t_i} x(t_{j_i}) - e^{-\frac{3}{4}-t_i} \psi(t_{j_i})]$$

$$\begin{aligned}
&= e^{\frac{3}{2}}x(0) + e^1(0.391) + e^{\frac{1}{2}}(0.555) + (0.610) \\
&\quad - e^{-\frac{3}{4}}\psi(0) - e^{-\frac{5}{4}}(0.177) - e^{-\frac{7}{4}}(0.545) - e^{-\frac{9}{4}}(0.401) \\
&= 2.5879 - 0.18768 \\
&= 2.4002.
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
\Delta m(3/2) &= e^{9/2}f(3/2, 3/2)[1 + e^{9/2}f(3/2, 3/2)]^{-1}y(3/2) \\
&\quad - \sum_{i=0} \{e^{3+t_{ji}}f(3/2, t_{ji-})[1 + e^{9/2}f(3/2, t_{ji-})]^{-1}e^{\frac{3}{2}-t_{ji}}m(t_{ji-})\} \\
&= e^{9/2}(2.0493)[1 + e^{9/2}(2.0493)]^{-1}(2.4002) \\
&\quad - e^{\frac{7}{2}}(1.4027)[1 + e^{9/2}(1.4027)]^{-1}e^1(0.372) \\
&\quad - e^4(1.8834)[1 + e^{\frac{9}{2}}(1.8834)]^{-1}e^{\frac{1}{2}}(0.541) \\
&\quad - e^{\frac{9}{2}}(2.0493)[1 + e^{\frac{9}{2}}(2.0493)]^{-1}(0.608) \\
&= 2.3873 - 0.36907 - 0.53782 - 0.60473 \\
&= 0.87568. \blacksquare
\end{aligned}$$

En este último caso puede apreciarse que el estimado es peor con respecto al del caso mostrado en el ejemplo 6.3. También, note que mientras el tiempo de los retardos incluidos al momento de la observación sea mayor, el salto del estimado es de mayor magnitud y por consiguiente es menos efectivo. Más aún, tanto el estimado como la varianza de este caso difieren en gran parte de sus valores reales dado que en los cálculos de los valores fijos del estimado, función de correlación y función característica se han utilizado aproximaciones burdas que no brindan valores satisfactorios, pero que pese a esto en los casos anteriores fueron aceptables.

Independientemente de los problemas que las aproximaciones ocasionan en el cálculo de los estimados de esta clase, hay que considerar que en este ejemplo toda la información obtenida de una ecuación con múltiples retardos está fusionada, implicando que el ruido total de la fusión también está acumulado, por lo cual, la perturbación resulta ser de considerable tamaño y afecta en la misma medida al estimado tal como se mencionó en el ejemplo 6.2.

Capítulo 7

Conclusiones

7.1 Aportaciones

Básicamente en este trabajo se realizó el diseño de métodos para filtrado de procesos integrales con estados estocásticos comprendidos dentro la clase de sistemas con retardo.

En el capítulo seis se pueden ver los resultados propuestos, los cuales, forman parte de una extensión de la teoría general del filtrado de Kalman a ecuaciones integrales. Las aportaciones se pueden resumir como:

- (1) Diseño de un algoritmo de solución para problemas de filtrado de procesos de Ito-Volterra, (ecuaciones integrales y/o ecuaciones diferenciales, bajo la presencia de ruidos Gaussianos), sobre observaciones de Ito-Volterra discontinuas .
- (2) Desarrollo de un procedimiento de solución para problemas de filtrado para procesos de Ito-Volterra sobre observaciones de Ito-Volterra discretas con un tiempo de retardo.
- (3) Solución del problema de filtrado de las ecuaciones integrales estocásticas sobre observaciones discretas con múltiples tiempos de retardo.
- (4) Diseño de ejemplos matemáticos para esta clase de filtros sobre observaciones continuas, discontinuas y discretas con retardos.
- (5) Desarrollo de un método aproximado de simulación computacional para sistemas integrales estocásticos con la ayuda del programa Simnon.

Con esto, un determinado problema de filtrado con estructura diferencial se puede resolver mediante las ecuaciones obtenidas igual que en el caso de problema resuelto por Kalman-Bucy, lo que puede servir de prueba rápida para cualquiera de estos resultados.

De este trabajo se puede concluir que es posible la implementación de un método capaz de resolver (de modo eficiente), el problema del filtrado cuando las ecuaciones del modelo estén representadas por formas complejas (como la de una ecuación integral estocástica), mismas que en la teoría convencional su resolución requiere de una mayor serie de restricciones y un cálculo complicado. Así es como con los resultados obtenidos, se pueden tener los correspondientes filtros y resolver otros casos particulares, los que posteriormente pueden ser llevados al tema del control.

7.2 Trabajos Futuros

Existen algunos casos particulares del filtrado para procesos de Ito-Volterra que no son publicados aún y que pueden resolverse con un algoritmo similar al de los problemas aquí presentados. Tales casos por resolver son:

- filtrado de procesos de Ito-Volterra sobre observaciones discontinuas o discretas con múltiples tiempos de retardos, cuando la ecuación del estado está caracterizada por una función de variación acotada, es decir, cuando el vector de estados es discontinuo.
- filtrado de procesos de Ito-Volterra discretos sobre observaciones discretas con múltiples retardos.
- resolver problemas físicos que tienen la estructura de un proceso de Ito-Volterra, tales como el de la posición global de satélites y el procesamiento de materiales químicos industriales entre otros.

Esto en conjunto con los estudios anteriores del área, permitirá completar la teoría general del filtrado para sistemas con estados estocásticos. También, en forma conjunta es posible abordar los temas de suavización y predicción para esta clase de sistemas, con el objetivo de abarcar la solución para un mayor número de problemas relacionados con el concepto del filtro.

Otra área por completar es la teoría general del control de sistemas con estados estocásticos, la cual, resulta ser por dualidad, del interés y competencia del material aquí presentado.

Apéndice A

"ON FILTERING PROBLEMS OVER ITO-VOLTERRA OBSERVATIONS"

REFERENCIA

[24]

Artículo aceptado por la Conferencia Americana del Control (ACC), San Diego California, 2-4 de junio de 1999.



Conference Editorial Board

Control Systems Society



From: Professor J. Jim Zhu
 3216-E CEBA Building
 Division of Engineering Services
 Louisiana State University
 Baton Rouge, LA 70803, USA

Contact:
 Email: css.ceb@rsip.lsu.edu
 Fax: +1 225 388 4893
 Voice: +1 225 388 6028
 Internet: <http://mail.eng.lsu.edu/~css.ceb>

January 29, 1999

To: Michael V. Basin
 Dept. of Elect. & Mech. Eng.
 Autonomous Univ. of Nuevo Leon
 Apdo postal 30-F, Cd. Universitaria
 San Nicolas de los Garza, Nuevo Leon CP 66450,
 MEXICO

phone +52.83294020x5773
 fax 52.83764514
 email mbsin@gama.fime.uanl.mx

Re: Paper ID **ACC99-IEEE1166 / FA16-5**
 Paper Title On filtering problems over Ito-Volterra observations

Paper type Regular

Authors	(PIN)	(surname)	(first name)	(affiliation)
	16649	Basin	Michael V.	Autonomous Univ. of Nuevo Leon
	31949	Villanueva L.	Mario A.	Universidad Autonoma de Nuevo Leon

Dear Dr. Basin:

This letter concerns your manuscript ACC99-IEEE1166 submitted to the 1999 American Control Conference through IEEE Society Review Chair. I am pleased to inform you that your paper was accepted as a REGULAR paper for presentation in Session FA16-5 at ACC99. Please see the notification letter mailed separately from the Program Chair and Publication Chair regarding manuscript preparation, advance program, and other important matters.

I am enclosing useful comments on your paper, if any, from the reviewers. I urge you to make every effort to address these comments in your final revision. If you need to change the title or authorship from the data listed above, do it on the Paper Certification Form found in the Author Kit. Also note that camera-ready manuscripts are due MARCH 1.

For your convenience, we have set up a searchable Early-Edition Technical Program at the web URL: <http://130.39.192.90/cebopenweb/conferences/acc99/>. Thank you for your contribution to the 1999 ACC.

Sincerely,



J. Jim Zhu,
 IEEE Society Review Chair

Reviewer/AE comments enclosed No reviewer/AE comments enclosed

ON FILTERING PROBLEMS OVER ITO-VOLTERRA OBSERVATIONS

Michael V. Basin and Mario A. Villanueva Llanes

Department of Electrical and Mechanical Engineering
Autonomous University of Nuevo Leon
Apdo postal 30-F, C.P. 66450, San Nicolas de los Garza
Nuevo Leon, Mexico
Phone: (52 8) 329 4020, ext. 5773
Fax: (52 8) 376 4514
E-mail: mbasin@gama.fime.uanl.mx

Abstract. In this paper, the Kalman-Bucy filter is designed for an Ito-Volterra process over Ito-Volterra observations that cannot be reduced to the case of a differential observation equation. The Kalman-Bucy filter is then designed for an Ito-Volterra process over discontinuous Ito-Volterra observations.

1. Introduction

Ito-Volterra processes and their applications to the optimal control theory have been studying from [1,2]. The first optimal filter for an Ito-Volterra process over scalar observations given by a differential equation was designed in [3], and then the optimal filter was obtained [4] over vector ones.

Continuing the research initiated in [3,4], this paper develops the optimal filter for an Ito-Volterra process over Ito-Volterra observations that cannot be reduced, unlike [3,4], to the case of a differential observation equation. Based on the obtained filtering equations over continuous observations, the Kalman-Bucy filter is then designed for an Ito-Volterra process over discontinuous Ito-Volterra observations. In this sense, this paper follows the series of papers ([5–8]) devoted to filtering over discontinuous observations. Some remarks concerning significance and applicability of this type of observations can be found there.

The paper is divided into two parts describing continuous and discontinuous observations, respectively. Proofs of theorems are given in Appendix.

2. Filtering over Continuous Observations

2.1. Problem Statement

Let (Ω, F, P) be a complete probability space with an increasing right-continuous family of σ -algebras $F_t, t \geq 0$, and let $(W^1(t), F_t, t \geq 0)$ and $(W^2(t), F_t, t \geq 0)$ be independent Wiener processes. The partly observed F_t -measurable random process

$(x(t), y(t))$ satisfies the Ito-Volterra equations

$$x(t) = \int_0^t (a_0(t, s) + a(t, s)x(s))ds + \int_0^t b(t, s)dW^2(s), \quad (1)$$

$$y(t) = \int_0^t (A_0(t, s) + A(t, s)x(s))ds + \int_0^t B(t, s)dW^1(s). \quad (2)$$

Here $x(t) \in R^n$ is a nonobserved component, and $y(t) \in R^m$ is an observed one for the process $(x(t), y(t))$. Functions $a_0(t, s)$, $a(t, s)$, $b(t, s)$ are smooth in t uniformly in s and continuous in s , and functions $A_0(t, s)$, $A(t, s)$, and $B(t, s)$ are continuous in t, s . Let $A(t, s)$ be a nonzero matrix, and $B(t, s)$ be a nonzero matrix, such that, $B(t, s)B^T(t, s) > 0$.

The estimation problem is to find the best estimate for the Ito-Volterra process $x(t)$ at time t based on the observation process $Y(t) = \{y(s), 0 \leq s \leq t\}$, that is the conditional expectation $m(t) = E(x(t) | F_t^Y)$. Let $P(t) = E((x(t) - m(t))(x(t) - m(t))^T | F_t^Y)$ be the correlation function, where the symbol a^T means transposition of a vector (matrix) a .

The above statement generalizes the problem statement given in [3,4] to the Ito-Volterra observation equation (2), which cannot be reduced to a differential equation. As in [3,4], it is impossible to obtain a closed system of filtering equations for variables $m(t)$ and $P(t)$ due to the Volterra nature of the equations (1) and (2). Designing a closed filter requires introducing the additional function $f(t)$ characterizing a deviation of the best estimate $m(t)$ from the real state $x(t)$:

$$f(t, s) = E((x_s^t - m_s^t)(x(s) - m(s))^T | F_{t,s}^Y), \quad (3)$$

where

$$x_s^t = \int_0^s (a_0(t, r) + a(t, r)x(r))dr + \int_0^s b(t, r)dW^2(r), \quad (4)$$

$F_{t,s}^Y$ is the σ -algebra generated by the stochastic process y_s^t

$$y_s^t = \int_0^s (A_0(t, s) + A(t, s)x(s))ds + \int_0^s B(t, s)dW^1(s), \quad (5)$$

$$\text{and } m_s^t = E(x_s^t | F_{t,s}^Y).$$

2.2. Optimal Filter

The optimal filter over continuous observations is given by the following theorem.

Theorem 1. The best estimate $m(t)$ of the system state (1) over the observations (2), its correlation function $P(t)$, and the function $f(t)$ (see (3)) satisfy the following filtering equations

$$m(t) = \int_0^t (a_0(t, s) + a(t, s)m(s))ds + \quad (6)$$

$$\int_0^t f(t, s)A^T(t, s)(B(t, s)B^T(t, s))^{-1} \times \\ [dy(t) - (A_0(t, s) + A(t, s)m(s))ds],$$

$$P(t) = \int_0^t [a(t, s)f^T(t, s) + f(t, s)a^T(t, s) + \quad (7)$$

$$b(t, s)b^T(t, s)]ds - \int_0^t f(t, s)A^T(t, s) \times$$

$$(B(t, s)B^T(t, s))^{-1}A(t, s)f^T(t, s)ds,$$

$$f(t, s) = \int_0^s [a(t, r)f^T(s, r) + f(t, r)a^T(s, r) + \quad (8)$$

$$(1/2)(b(t, r)b^T(s, r) + b(s, r)b^T(t, r))]dr -$$

$$\int_0^s [f(t, r)A^T(s, r)(B(s, r)B^T(s, r))^{-1} \times$$

$$A(s, r)f^T(s, r) +$$

$$f(t, r)A^T(t, r)(B(t, r)B^T(t, r))^{-1} \times$$

$$A(t, r)f^T(s, r) -$$

$$(1/2)f(t, r)A^T(t, r)(B(t, r)B^T(s, r))^{-1} \times$$

$$A(s, r)f^T(s, r) -$$

$$(1/2)f(s, r)A^T(s, r)(B(s, r)B^T(t, r))^{-1} \times$$

$$A(t, r)f^T(t, r)]dr.$$

3. Filtering over Discontinuous Observations

3.1. Problem Statement

Consider generalization of the filtering problem examined in Section 2 to the case of discontinuous observations. Let the partly observed F_t -measurable random process $(x(t), y(t))$ be given by the following Ito-Volterra equations

$$x(t) = \int_0^t (a_0(t, s) + a(t, s)x(s))ds + \int_0^t b(t, s)dW^2(s), \quad (9)$$

$$y_i(t) = \int_0^t (A_{0i}(t, s) + (A_i(t, s), x(s)))du_i(s) +$$

$$+ \int_0^t B_i(t, s)dW^1(u_i(s)), \quad i = 1, \dots, m, \quad (10)$$

where $A(t, s) = (A_1(t, s), \dots, A_m(t, s))^T$, $A_i(t, s) \in R^n$, $i = 1, \dots, m$, (a, b) is the scalar product in R^n ,

and the rest of the notation is the same as in Section 2.

The observation process is characterized by a vector bounded variation function $u(t) = (u_1(t), \dots, u_m(t)) \in R^m$, which is nondecreasing in the following sense: $u(t_2) \geq u(t_1)$ as $t_2 \geq t_1$ if $u_i(t_2) \geq u_i(t_1)$ for $i = 1, \dots, m$. This model of observations enables one to consider continuous and discrete observations in the common form: continuous observations correspond to the continuous component of a bounded variation function $u(t)$, and discrete observations correspond to its function of jumps.

The estimation problem is formulated exactly as in Section 2. All the remarks of Section 2 concerning possibility of obtaining a closed system of filtering equations remain valid in this case. We also keep the notation of Section 2 for functions $f(t, s)$, x_s^t , m_s^t , and y_s^t .

3.2. Optimal Filter

In [7], the filtering procedure is suggested to obtain filtering equations over discontinuous observations proceeding from the known filtering equations over continuous ones. To apply the filtering procedure to the examined problem is to complete the following actions:

- assuming a vector function $u(t) \in R^m$ in an observation equation (10) to be absolutely continuous, write the Ito-Volterra filtering equations over continuous observations obtained in Section 2 (see (6)–(8));
- in thus obtained equations assume a vector bounded variation function $u(t) \in R^m$ to be an arbitrary nondecreasing one again, keeping in mind that a derivative $\dot{u}(t)$ can be a generalized function of zero singularity order (for example, δ -function).

As a result the following system of Ito-Volterra equations over discontinuous observations (10) is obtained

$$m(t) = \int_0^t (a_0(t, s) + a(t, s)m(s))ds + \quad (11)$$

$$\int_0^t f(t, s)A^T(t, s)(B(t, s)B^T(t, s))^{-1} \times \\ [dy(t) - (A_0(t, s) + A(t, s)m(s))du(s)],$$

$$P(t) = \int_0^t [a(t, s)f^T(t, s) +$$

$$f(t, s)a^T(t, s) + b(t, s)b^T(t, s)]ds - \quad (12)$$

$$\int_0^t f(t, s)A^T(t, s)(B(t, s)B^T(t, s))^{-1} \times \\ A(t, s)f^T(t, s)du(s),$$

$$f(t, s) = \int_0^s [a(t, r)f^T(s, r) + f(t, r)a^T(s, r) + \quad (13)$$

$$(1/2)(b(t, r)b^T(s, r) + b(s, r)b^T(t, r))]dr -$$

$$\int_0^s [f(t, r)A^T(s, r)(B(s, r)B^T(s, r))^{-1} \times$$

$$A(s, r)f^T(s, r) +$$

$$f(t, r)A^T(t, r)(B(t, r)B^T(t, r))^{-1} \times$$

$$A(t, r)f^T(s, r) -$$

$$(1/2)f(t, r)A^T(t, r)(B(t, r)B^T(s, r))^{-1} \times$$

$$A(s, r)f^T(s, r) -$$

$$(1/2)f(s, r)A^T(s, r)(B(s, r)B^T(t, r))^{-1} \times$$

$$A(t, r)f^T(t, r)]du(r).$$

Here, multiplication by an m -dimensional measure $du(t)$ should be understood in the componentwise sense, as in the observation equation (10).

The obtained equations (11)–(13) are integral equations with integration w.r.t. a vector discontinuous measure generated by a nondecreasing bounded variation function $u(t)$. The further investigation will follow the standard scheme suggested in [7]. It will be specified how to understand the solution of these equations and how to compute jumps of the solution at the discontinuity points of $u(t)$. Actually, these jumps reflect reaction of the filtering variables (the estimate $m(t)$ and its characteristics $P(t)$ and $f(t, s)$) to appearance of discrete measurements in the environment of a continuous signal. The final step is to prove that the introduced solution of the system (11)–(13) really yields the optimal estimate and its correlation characteristics.

3.3. Solution and Jumps: Theoretical Background

To avoid unnecessary complication of formulas, let us study an integral equation with integration w.r.t. a vector discontinuous measure generated by a nondecreasing bounded variation function in the general form and develop theoretical constructions for it. Namely, consider an integral equation in the form

$$x(t) = x(t_0) + \int_{t_0}^t f(x, u, t, s)ds + b(x, u, t, s)du(s), \quad (14)$$

where functions $f(x, u, t, s)$ and $b(x, u, t, s)$ are continuous in x, u, t, s ; $u(t) = (u_1(t), \dots, u_m(t)) \in R^m$ is a nondecreasing bounded variation function. The set of discontinuity points of $u(t)$ is considered a countable set of isolated points.

The solution of the equation (14) is introduced as follows (cf. [7]).

Definition. The left-continuous function $x(t)$ is said to be a vibrosolution of the equation (14), if the $*$ -weak convergence (see [10]) of an arbitrary sequence of absolutely continuous nondecreasing functions $u^k(t) \in R^m$ to a non-decreasing function $u(t) \in R^m$ in the bounded variation functions space

$$* - \lim u^k(t) = u(t)$$

implies the analogous convergence

$$* - \lim x^k(t) = x(t)$$

of corresponding solutions $x^k(t)$ of the equation

$$x^k(t) = x(t_0) + \int_{t_0}^t f(x^k, u^k, t, s)ds + b(x^k, u^k, t, s)du^k(s),$$

and the unique limit $x(t)$ is regardless of a choice of an approximating sequence $\{u^k(t)\}, k = 1, 2, \dots$

The existence and uniqueness conditions for the vibrosolution of the equation (14) are given in the next theorem. Let us note that a vibrosolution is expected to be a function discontinuous at discontinuity points of $u(t)$.

Theorem 2. Let

1) functions $f(x, u, t, s), b(x, u, t, s), \partial b(x, u, t, s)/\partial x, \partial b(x, u, t, s)/\partial t, \partial b(x, u, t, s)/\partial s$ be continuous in x, u, t, s ;

2) functions $f(x, u, t), b(x, u, t)$ satisfy the Lipschitz condition in x ;

3) the $n \times m$ -dimensional system of differential equations in differentials

$$\frac{d\xi}{du} = b(\xi, u, t, t), \quad \xi(\omega) = z, \quad (15)$$

is solvable on the cone of positive directions

$K = \{u \in R^m : u_i \geq \omega_i, i = 1, \dots, m\}$ with arbitrary initial values $\omega \in R^m, \omega \geq u(t_0)$, and $z \in R^n$.

Then there exists the unique vibrosolution of the equation (14).

Jumps of the vibrosolution of the equation (14) at the discontinuity points of the function $u(t)$ can be computed using the following equivalent equation with a measure.

Theorem 3. Let the conditions of Theorem 2 hold. Then the integral equation (14) and the equivalent equation with a measure

$$x(t) = x(t_0) + \int_{t_0}^t f(x, u, t, s)ds + b(x, u, t, s)du^c(s) + \sum_{t_i} G(x(t_i-), u(t_i-), \Delta u(t_i), t_i)d\chi(t - t_i),$$

have the same unique solution regarded for the equation (14) as a vibrosolution. Here $G(z, \omega, u, t) = \xi(z, \omega, \omega + u, t) - z$ and $\xi(z, \omega, u, t)$ is the solution of the system in differentials (15); $x(t_i-)$ and $x(t_i+)$ are values of the function $x(t)$ at a discontinuity point t_i from the left and right, respectively; $u^c(t)$ is the continuous component of a nondecreasing function $u(t)$, $\Delta u(t_i) = u(t_i+) - u(t_i-)$ is the jump of a function $u(t)$ at t_i , t_i are the discontinuity points of a function $u(t)$, $\chi(t - t_i)$ is a Heaviside function.

3.4. Equivalent Form of Filtering Equations

Thus, Theorem 2 yields existence and uniqueness of the vibrosolution $\{m(t), P(t), f(t, s)\}$ to the system of filtering equations (11)–(13), and Theorem 3 brings out the method for computing jumps of the vibrosolution $\{m(t), P(t), f(t, s)\}$ at the discontinuity points of the function $u(t)$ (i.e., at the discontinuity points of observations). Indeed, in view of Theorem 3, the

equivalent equations with a measure for the filtering equations (11)–(13) take the form

$$m(t) = \int_0^t (a_0(t, s) + a(t, s)m(s))ds + \quad (16)$$

$$\int_0^t f(t-, s-)[I + A^T(t, s)(B(t, s)B^T(t, s))^{-1} \times \\ A(t, s)f(t-, s-)\Delta u(s)]^{-1}A^T(t, s)(B(t, s)B^T(t, s))^{-1} \times \\ [dy(t) - (A_0(t, s) + A(t, s)m(s-))du(s)],$$

$$P(t) = \int_0^t [a(t, s)f^T(t, s) + f(t, s)a^T(t, s) + \quad (17)$$

$$b(t, s)b^T(t, s)]ds - \int_0^t f(t-, s-)[I + A^T(t, s) \times$$

$$(B(t, s)B^T(t, s))^{-1}A(t, s)f(t-, s-)\Delta u(s)]^{-1} \times \\ A^T(t, s)(B(t, s)B^T(t, s))^{-1}A(t, s)f^T(t-, s-)\Delta u(s),$$

$$f(t, s) = \int_0^s [a(t, r)f^T(s, r) + f(t, r)a^T(s, r) + \quad (18)$$

$$(1/2)(b(t, r)b^T(s, r) + b(s, r)b^T(t, r))]dr -$$

$$\int_0^s [f(t, r-)[I + (A^T(s, r)(B(s, r)B^T(s, r))^{-1}A(s, r) \times$$

$$f(s, r-)+A^T(t, r)(B(t, r)B^T(t, r))^{-1}A(t, r)f(t, r-)-$$

$$(1/2)A^T(s, r)(B(s, r)B^T(s, r))^{-1}A(t, r)f(t, r-)-$$

$$(1/2)A^T(t, r)(B(t, r)B^T(s, r))^{-1} \times$$

$$A(s, r)f(s, r-)\Delta u(r)]^{-1} \times$$

$$A^T(s, r)(B(s, r)B^T(s, r))^{-1}A(s, r)f^T(s, r-)+$$

$$f(s, r-)[I + (A^T(s, r)(B(s, r)B^T(s, r))^{-1}A(s, r) \times$$

$$f(s, r-)+A^T(t, r)(B(t, r)B^T(t, r))^{-1}A(t, r)f(t, r-)-$$

$$(1/2)A^T(s, r)(B(s, r)B^T(t, r))^{-1}A(t, r)f(t, r-)-$$

$$(1/2)A^T(t, r)(B(t, r)B^T(s, r))^{-1} \times$$

$$A(s, r)f(s, r-)\Delta u(r)]^{-1} \times$$

$$A^T(t, r)(B(t, r)B^T(t, r))^{-1}A(t, r)f^T(t, r-)-$$

$$(1/2)f(s, r-)[I + (A^T(s, r)(B(s, r)B^T(s, r))^{-1}A(s, r) \times$$

$$f(s, r-)+A^T(t, r)(B(t, r)B^T(t, r))^{-1}A(t, r)f(t, r-)-$$

$$(1/2)A^T(s, r)(B(s, r)B^T(t, r))^{-1}A(t, r)f(t, r-)-$$

$$(1/2)A^T(t, r)(B(t, r)B^T(s, r))^{-1} \times$$

$$A(s, r)f(s, r-)\Delta u(r)]^{-1} \times$$

$$A^T(s, r)(B(s, r)B^T(t, r))^{-1}A(t, r)f^T(t, r-)-$$

$$(1/2)f(t, r-)[I + (A^T(s, r)(B(s, r)B^T(s, r))^{-1}A(s, r) \times$$

$$f(s, r-)+A^T(t, r)(B(t, r)B^T(t, r))^{-1}A(t, r)f(t, r-)-$$

$$(1/2)A^T(s, r)(B(s, r)B^T(t, r))^{-1}A(t, r)f(t, r-)-$$

$$(1/2)A^T(t, r)(B(t, r)B^T(s, r))^{-1} \times$$

$$A(s, r)f(s, r-)\Delta u(r)]^{-1} \times$$

$$A^T(t, r)(B(t, r)B^T(s, r))^{-1}A(s, r)f^T(s, r-)]du(r),$$

where I is the $n \times n$ -dimensional identity matrix. The function $f(t, s)$ is continuous in t .

It is readily verified that jumps of the filtering variables $m(t)$, $P(t)$, and $f(t, s)$ at a discontinuity point t_i of $u(t)$ are equal to the expressions under the integral signs in the right-hand sides of (16)–(18), upon substituting the jumps $\Delta u(t_i)$ and $\Delta y(t_i)$ for differentials $du(t)$ and $dy(t)$, respectively.

The equations with a measure (16)–(18) completely determine the behavior of the filtering variables $m(t)$, $P(t)$, and $f(t, s)$, i.e., the complete reaction of the filtering variables to a composition of continuous and discrete measurements. The next optimality theorem is the final step in solution of the filtering problem for an Ito-Volterra process over Ito-Volterra discontinuous observations.

Theorem 4. The solutions $m(t)$, $P(t)$, and $f(t, s)$ of the equations (16)–(18) are the optimal estimate in the filtering problem (9), (10), its correlation function, and its correlation characteristic (3), respectively.

4. Appendix

Proof of Theorem 1. Let us consider the filtering problem for the state x_s^t over the observation process y_s^t . The equation (4) for x_s^t and the equation (5) for y_s^t are actually differential equations with respect to s , where t is a parameter. Thus, the principal filtering theorem (the correlation theorem for conditionally Gaussian processes, see Theorem 8.6 in [9]) is applicable, and we obtain

$$m_s^t = \int_0^s (a_0(t, r) + a(t, r)m(r))dr +$$

$$\int_0^s E(x_r^t x^T(r) - m_r^t m^T(r) | F_{t, r}^Y) A^T(t, r)(B(t, r) \times$$

$$B^T(t, r))^{-1} \{dy_r^t - (A_0(t, r) + A(t, r)m(r))dr\}.$$

Equating $s = t$ yields, in view of (3), the equation (6) for $m(t)$.

Let us now prove the equation (8) for $f(t, s)$. The Ito's formula yields

$$d(x_r^t - m_r^t) = [a(t, r) -$$

$$f(t, r)A^T(t, r)(B(t, r)B^T(t, r))^{-1} \times$$

$$A(t, r)](x(r) - m(r))dr + b(t, r)dW^2(r) -$$

$$f(t, r)A^T(t, r)(B(t, r)B^T(t, r))^{-1}B(t, r)dW^1(r),$$

and

$$d((x_r^t - m_r^t)(x_r^u - m_r^u)^T) = (x_r^t - m_r^t)(x(r) - m(r))^T \times$$

$$[a(u, r) - f(u, r)A^T(u, r)(B(u, r)B^T(u, r))^{-1} \times$$

$$A(u, r)]^T dr + [a(t, r) - f(t, r)A^T(t, r)(B(t, r) \times$$

$$B^T(t, r))^{-1}A(t, r)](x(r) - m(r))(x_r^u - m_r^u)^T dr +$$

$$(1/2)(b(t, r)b^T(u, r) + b(u, r)b^T(t, r))dr +$$

$$(1/2)f(t, r)A^T(t, r)(B(t, r)B^T(u, r))^{-1} \times$$

$$A(u, r)f^T(u, r)dr + (1/2)f(u, r)A^T(u, r) \times$$

$$\begin{aligned}
& (B(u, r)B^T(t, r))^{-1}A(t, r)f^T(t, r)]dr - \\
& (x_r^t - m_r^t)[f(u, r)A^T(u, r)(B(u, r)B^T(u, r))^{-1} \times \\
& B(u, r)dW^1(r) - b(u, r)dW^2(r)]^T - \\
& [f(t, r)A^T(t, r)(B(t, r)B^T(t, r))^{-1}B(t, r)dW^1(r) - \\
& b(t, r)dW^2(r)](x_r^u - m_r^u)^T.
\end{aligned}$$

Integrating with r from 0 to s , equating $u = s$, and using Theorem 8.6 in [9], we obtain

$$\begin{aligned}
& E((x_s^t - m_s^t)(x(s) - m(s))^T | F_{t,s}^Y) = \\
& \int_0^s [E((x_r^t - m_r^t)(x(r) - m(r))^T | F_{t,r}^Y)a^T(s, r) + \\
& a(t, r)E((x(r) - m(r))(x_r^s - m_r^s)^T | F_{s,r}^Y) + \\
& (1/2)(b(t, r)b^T(s, r) + b(s, r)b^T(t, r))dr + \quad (19) \\
& (1/2)f(t, r)A^T(t, r)(B(t, r)B^T(s, r))^{-1} \times \\
& A(s, r)f^T(s, r)dr + (1/2)f(s, r)A^T(s, r) \times \\
& (B(s, r)B^T(t, r))^{-1}A(t, r)f^T(t, r) - \\
& E((x_r^t - m_r^t)(x(r) - m(r))^T | F_{t,r}^Y) \times \\
& (f(s, r)A^T(s, r)(B(s, r)B^T(s, r))^{-1}A(s, r))^T - \\
& f(t, r)A^T(t, r)(B(t, r)B^T(t, r))^{-1}A(t, r) \times \\
& E((x(r) - m(r))(x_r^s - m_r^s)^T | F_{s,r}^Y)]dr + \\
& \int_0^s [E((x_r^t - m_r^t)(x_r^s - m_r^s)^T (x(r) - m(r))^T | F_{t,r}^Y) \times \\
& A^T(t, r)(B(t, r)B^T(t, r))^{-1} \times \\
& [dy_r^t - (A_0(t, r) + A(t, r)m(r))dr].
\end{aligned}$$

Let us note that the latter addendum is equal to zero, because all the processes under the expectation sign are zero mean conditionally Gaussian. Thus, in view of (3), the equation (19) yields the equation (8). Finally, the equation (7) follows from (8), because $P(t) = f(t, t)$.

Proof of Theorem 2. By virtue of the theorem conditions, the equation (14) has the unique absolutely continuous solution on the continuity intervals of the function $u(t)$ (see [1]). Moreover, this absolutely continuous solution is a vibrosolution, in view of the Lebesgue's bounded convergence theorem (see [10]) which yields existence of the limit required in the definition of a vibrosolution on the continuity intervals of $u(t)$. Thus, it remains to prove existence of the vibrosolution of (14) only in neighborhoods of the isolated discontinuity points t_i , $i = 1, 2, \dots$, of the function $u(t)$.

In accordance with the theorem conditions, the system (15) has the solution $\xi(z, \omega, u, t_i)$ on the cone K , where t_i is an isolated discontinuity point of $u(t)$. Let us seek the solution of (14) corresponding to a non-decreasing function $u(t)$ in the form

$$x(t) = \xi(z(t), u_i, u(t), t_i),$$

where $u_i = u(t_i)$. Then, it can be readily shown, following the proof of Theorem 1 in [11], that the inversion formula

$$z(t) = \xi(x(t), u(t), u_i, t_i)$$

is valid for $u(t) \geq u_i$ in a neighborhood of the isolated discontinuity point t_i . The rest of the proof also follows the proof of Theorem 2 in [11], taking into account that behavior of the vibrosolution of (14) is investigated in a neighborhood of t_i .

Proof of Theorem 3 is similar to that of Theorem 3 in [11].

Proof of Theorem 4 is similar to that of Theorem 4 in [6].

5. References

1. K. Ito, "On the existence and uniqueness of solutions of stochastic integral equations of the Volterra type," *Kodai Math. Journal*, no. 2, pp. 158-170, 1979.
2. A.V. Balakrishnan, "On stochastic bang-bang control," *Lecture Notes in Control and Information Sciences*, vol. 25, New York: Springer-Verlag, pp. 221-238, 1980.
3. M.L. Kleptsina and A.Yu. Veretennikov, "On filtering and properties of conditional laws of Ito-Volterra processes," *Statistics and Control of Stochastic Processes. Steklov Seminar. 1984*, New York: Optimization Software Inc., Publication Division, pp. 179-196, 1985.
4. L.E. Shaikhet, "On an optimal control problem of partly observable stochastic Volterra processes," *Problems of Control and Inform. Theory*, vol. 16, no. 6, pp. 439-448, 1987.
5. M.V. Basin, "On filtering of Ito-Volterra processes over discrete-continuous observations," *Automation and Remote Control*, vol. 53, no. 8, Part 1, pp. 1183-1192, 1992.
6. M.V. Basin, "On ellipsoidal filtering over discrete-continuous observations with a vector measure," *Automation and Remote Control*, vol. 56, no. 7, Part 1, pp. 973-981, 1995.
7. Yu.V. Orlov and M.V. Basin, "On minmax filtering over discrete-continuous observations," *IEEE Trans. Automat. Contr.*, vol. AC-40, no. 9, pp. 1623-1626, 1995.
8. M.V. Basin, "Ellipsoidal filtering for a state of an infinite-dimensional system," *Automation and Remote Control*, vol. 58, nos. 7 and 8, pp. 1110-1118 and 1323-1329, 1997.
9. R.S. Liptser and A.N. Shirayev, *Statistics of Random Processes. I*, New York: Springer-Verlag, 1977.
10. A.N. Kolmogorov and S.V. Fomin, *Introductory Real Analysis*, New York: Dover, 1975.
11. M.V. Basin, "Vibrosolution of a differential equation in distributions with discontinuous regular functions in right-hand side," *Math. Notes*, vol. 58, no. 1, pp. 685-691, 1995.

Apéndice B

"ON FILTERING PROBLEMS OVER OBSERVATIONS WITH DELAYS"

REFERENCIA
[32]

Artículo candidato a participar en la Conferencia de Decisión y Control (CDC), en la Ciudad de Phoenix Arizona, el mes de diciembre de 1999.

From: Dr. Michael V. Basin
Dept. of Elect. & Mech. Eng.
Autonomous Univ. of Nuevo Leon
Apdo postal 30-F, Cd. Universitaria
San Nicolas de los Garza Nuevo Leon CP 66450 MEXICO



Please copy and stick the above barcode to the BACK of your mail envelope for submitting your manuscript for timely submission updates.

To: Prof. J. Jim Zhu - CDC99
College of Engineering
Louisiana State University
3216 CEBA Bldg.
Baton Rouge LA 70803-5901 USA
FAX: 225/388-4893

Date: Tue Feb 23 13:10:55 1999 (U.S. Central Time)

Re: CDC99-REG0090

Dear Prof. Zhu - CDC99:

This is to confirm our intention to submit a paper titled "*On filtering problems over observations with delays*" which has been assigned the Paper ID: CDC99-REG0090. The paper is written by the following author(s):

<i>Basin, Michael V.</i>	[16649]	<i>Autonomous Univ. of Nuevo Leon</i>	<i>Corresponding Author</i>
<i>Villanueva L., Mario A.</i>	[31949]	<i>Universidad Autonoma de Nuevo Leon</i>	
<i>Valadez Guzman, Irma Rosario</i>	[32727]	<i>Univ. of Nuevo Leon</i>	

The following keywords best describe the subject matter and main contribution of our paper, with 3 indicating the most relevant ones:

Significance 3: *Stochastic TimeDelay Estimation KalmanBucyFiltering*

Significance 2: *Linear DiscreteTime ContinuousTime TargetTracking*

Significance 1:

The manuscript will be delivered *electronically* to arrive on or before March 1, 1999. We understand that at this time electronic delivery is experimental in nature and is on a voluntary basis. We will follow the instructions for electronic delivery closely, and will have printed copies of our manuscript express mailed to you upon your request should it become necessary.

By signing this letter I certify that this paper has not been published elsewhere, nor has it been submitted to other conferences or submitted to CDC99 through other categories. I understand that multiple submissions of this paper to this or other conferences will result in automatic rejection of this paper by CDC99. However, I may submit this manuscript for simultaneous review to IEEE CSS and SIAM journals or other IEEE journals according to their guidelines.

Sincerely,

[signature required. Please print, sign, and fax or mail.]

Dr. Michael V. Basin

ON FILTERING PROBLEMS OVER OBSERVATIONS WITH DELAYS

Michael V. Basin, Mario A. Villanueva Llanes, and Irma R. Valadez Guzman

Department of Electrical and Mechanical Engineering
Autonomous University of Nuevo Leon
Apdo postal 30-F, C.P. 66450, San Nicolas de los Garza
Nuevo Leon, Mexico
E-mail: mbasin@gama.fime.uanl.mx

Abstract. In this paper, the optimal filtering equations over observations with delays are obtained for an Ito-Volterra process, proceeding from discontinuous observations of the Ito-Volterra type. It is shown that the designed filter can be significantly simplified for a dynamic system state governed by a differential equation. The filtering problems are considered over observations with single time delay, multiple delays, and a set of delays of the continuum power.

Keywords. Kalman-Bucy filtering, observations with delays, Ito-Volterra process

1. Introduction

A number of papers devoted to the Kalman filtering problem over observations with delays have been recently published (see [1–3] and bibliography there). In those papers, the authors suggested various direct or recurrent algorithms based on the classical Kalman model of observations [4] or the equation for the conditional probability density of a system state function with respect to observations [5]. However, analytic closed-form solution of the filtering problem over delayed observations readily follows, as shown in this paper, from the filtering equations over discontinuous observations of the Ito-Volterra type, which have been obtained in [6].

Following [6], the filtering problem for an Ito-Volterra process over observations with delays is first considered. As in the previous works [6–8] devoted to the Ito-Volterra filtering, it is impossible to obtain a closed system of filtering equations with respect to only two variables, the optimal estimate and its variance, and is necessary to introduce an additional cross-correlation function. Using these three variables, the filtering equations are obtained over observations with single time delay, as well as with multiple time delays allowing data fusion (see, for example, [9,10]). Then, it is shown that one can obtain more simplified filtering equations over delayed observations for a dynamic system state, if a system is governed by a differential equation. Using only two variables, the optimal estimate and its variance, the filtering equations for a dynamic system state are ob-

tained over observations with multiple delays and a set of delays of the continuum power. This simplification enables one to use conventional numerical methods applicable to differential equations for solving the obtained filtering ones.

The paper is organized as follows. The optimal filter over observations with delays for an Ito-Volterra process is designed in Section 2. Section 3 presents the optimal filter for a dynamic system state governed by a differential equation. Subsections correspond to various configurations of the set of observation delays: single delay, multiple delays, and a set of delays of the continuum power.

2. Filtering for Ito-Volterra process

2.1. Problem Statement

The following statement of the Kalman-Bucy filtering problem over discrete observations with delays is considered. Let (Ω, F, P) be a complete probability space with an increasing right-continuous family of σ -algebras $F_t, t \geq 0$, and let $(W^1(t), F_t, t \geq 0)$ be a Wiener process. The nonobserved F_t -measurable random process $x(t)$ satisfies the Ito-Volterra equation

$$x(t) = \int_0^t (a_0(t, s) + a(t, s)x(s))ds + \int_0^t b(t, s)dW^1(s), \quad (1)$$

and the F_t -measurable discrete observations with delays are given by

$$y(t_j) = A_0(t_j, t_i) + A(t_j, t_i)x(t_i) + B(t_j, t_i)\psi(t_i). \quad (2)$$

Here, $y(t_j) \in R^m$ are discrete observations at time moments $t_j, j = 0, 1, \dots$, and $x(t_i)$ is a value of the system state at a moment t_i available for measurement at the observation moment $t_j, A(t_j, t_i) \in R^{m \times n}$ are observation matrices, $\psi(t_i)$ are Gaussian noises independent of $W^1(t)$ and acting at the moments t_i . Let $A(t_j, t_i)$ be nonzero matrices, and $B(t_j, t_i)B^T(t_j, t_i)$ be positive definite ones.

The estimation problem is to find the best estimate for the Ito-Volterra process $x(t)$ at time t based on the observations $Y(t) = \{y(t_j), 0 \leq t_j \leq t\}$, that is the conditional expectation $m(t) = E(x(t) | F_t^Y)$.

Let $P(t) = E((x(t) - m(t))(x(t) - m(t))^T | F_t^Y)$ be the correlation function, where the symbol a^T means transposition of a vector (matrix) a , and F_t^Y be the σ -algebra generated by the observations $Y(t)$.

The solution of this filtering problem over discrete observations with delays is based on the solution of the filtering problem over Ito-Volterra discontinuous observations obtained in [6]. The model of Ito-Volterra discontinuous observations in [6] was defined as

$$y_i(t) = \int_0^t (A_{0i}(t, s) + (A_i(t, s), x(s))) du_i(s) + \int_0^t B_i(t, s) dW_i^2(u_i(s)), \quad i = 1, \dots, m, \quad (3)$$

where $A(t, s) = (A_1(t, s), \dots, A_m(t, s))^T$, $A_i(t, s) \in R^n$, $i = 1, \dots, m$; $B(t, s) = (B_1(t, s), \dots, B_m(t, s))^T$, $B_i(t, s) \in R^k$ is the i th row of the matrix $B(t, s)$; $A(t, s)$ is a nonzero matrix and $B(t, s)B^T(t, s)$ is a positive definite one; $(W^2(t), F_t, t \geq 0)$ is a Wiener process independent of $W^1(t)$; (a, b) is the scalar product in R^n .

The observation process $y(t)$ in (3) is characterized by a vector bounded variation function $u(t) = (u_1(t), \dots, u_m(t)) \in R^m$, which is nondecreasing in the following sense: $u(t_2) \geq u(t_1)$ as $t_2 \geq t_1$ if $u_i(t_2) \geq u_i(t_1)$ for $i = 1, \dots, m$. This model of observations enables one to consider continuous and discrete observations in the common form: continuous observations correspond to the continuous component of a bounded variation function $u(t)$, and discrete observations correspond to its function of jumps.

Let us note that if the observation distribution function $u(t)$ is assumed to be piecewise constant with unit jump at the time moment t_i when information on the system state $x(t)$ is available, then the model of discontinuous observations (3) turns to the discrete observations with delays (2), considering $y(t_j) \in R^m$ as a vector of m components: $y(t_j) = (y_1(t_j), \dots, y_m(t_j))^T$. Thus, the desired filtering equations over observations with delays should follow from the results obtained in [6].

As well as in [6] and the previous works [7,8], in this problem it is impossible to obtain a closed system of filtering equations for variables $m(t)$ and $P(t)$ due to the Volterra nature of the observation equation (3), which cannot be reduced to a differential equation. Designing a closed filter requires introducing the additional function $f(t)$ characterizing a deviation of the best estimate $m(t)$ from the real state $x(t)$:

$$f(t, s) = E((x_s^t - m_s^t)(x(s) - m(s))^T | F_{t,s}^Y), \quad (4)$$

where

$$x_s^t = \int_0^s (a_0(t, r) + a(t, r)x(r)) dr + \int_0^s b(t, r) dW^1(r), \quad (5)$$

$F_{t,s}^Y$ is the σ -algebra generated by the random variables $Y(t, s) = \{y(t_j, s_i), 0 \leq t_j \leq t, 0 \leq s_i \leq s\}$

$$y(t_j, s_i) = A_0(t_j, s_i) + A(t_j, s_i)x(s_i) + B(t_j, s_i)\psi(s_i), \quad (6)$$

$$\text{and } m_s^t = E(x_s^t | F_{t,s}^Y).$$

Let us note that a closed system of filtering equations can be obtained for only two variables, $m(t)$ and $P(t)$, if a system state (1) is governed by a differential equation, in other words, in the case of a dynamic system. This important particular case is considered in Section 3.

2.2. Optimal Filter

As it was mentioned, the filtering equations for an Ito-Volterra process over discrete observations with delays follow from the filtering equations over Ito-Volterra discontinuous observations obtained in [6], if the observation distribution function $u(t)$ in (3) is the Heaviside function $\chi(t - t_i)$ with unit jump at the time moment t_i when information on the system state $x(t)$ is available (see the observation equation (2)). This yields the following result.

Proposition. The optimal estimate $m(t)$ for the state vector (1) over the discrete observations with delays (2), its correlation function $P(t)$, and its correlation characteristic (4) $f(t, s)$ satisfy the following equations between the observation moments t_j

$$m(t) = m(t_j+) + \int_{t_j+}^t (a_0(t, s) + a(t, s)m(s)) ds, \quad (7)$$

$$P(t) = P(t_j+) + \int_{t_j+}^t [a(t, s)f^T(t, s) + f(t, s)a^T(t, s) + b(t, s)b^T(t, s)] ds, \quad (8)$$

$$f(t, s) = f(t, t_j+) + \int_{t_j+}^s [a(t, r)f^T(s, r) + f(t, r)a^T(s, r) + (1/2)(b(t, r)b^T(s, r) + b(s, r)b^T(t, r))] dr, \quad (9)$$

and their jumps at the moments t_j of discrete observations are equal to

$$\Delta m(t_j) = f(t_j, t_i-) [I + \quad (10)$$

$$A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1} \times A(t_j, t_i)f(t_j, t_i-)]^{-1} A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1} \times [y(t_j) - (A_0(t_j, t_i) + A(t_j, t_i)m(t_i-))],$$

$$\Delta P(t_j) = -f(t_j, t_i-) [I + A^T(t_j, t_i) \times \quad (11)$$

$$(B(t_j, t_i)B^T(t_j, t_i))^{-1} A(t_j, t_i)f(t_j, t_i-)]^{-1} \times A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1} A(t_j, t_i)f^T(t_j, t_i-),$$

$$\Delta f(t, t_j) = -[f(t, t_i-) [I + \quad (12)$$

$$(A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1} A(t_j, t_i)f(t_j, t_i-) + A^T(t, t_i)(B(t, t_i)B^T(t, t_i))^{-1} A(t, t_i)f(t, t_i-) -$$

$$(1/2)A^T(t_j, t_i)(B(t_j, t_i)B^T(t, t_i))^{-1} A(t, t_i)f(t, t_i-) - (1/2)A^T(t, t_i)(B(t, t_i)B^T(t_j, t_i))^{-1} \times A(t_j, t_i)f(t_j, t_i-)]^{-1} \times A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1} A(t_j, t_i)f^T(t_j, t_i-) + f(t_j, t_i-) [I + (A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1}$$

$$\begin{aligned}
& A(t_j, t_i) f(t_j, t_i-) + \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
(1/2) & A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
& (1/2) A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} \times \\
& A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} \times \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f^T(t, t_i-) - \\
& (1/2) f(t_j, t_i-) [I + (A^T(t_j, t_i) \times \\
& (B(t_j, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-) + \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
(1/2) & A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
& (1/2) A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} \times \\
& A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} \times \\
& A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f^T(t, t_i-) - \\
& (1/2) f(t, t_i-) [I + (A^T(t_j, t_i) \times \\
& (B(t_j, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-) + \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
(1/2) & A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
& (1/2) A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} \times \\
& A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} \times \\
& A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f^T(t_j, t_i-),
\end{aligned}$$

where $m(s-)$, $P(s-)$ and $f(t, s-)$ are values from the left of $m(s)$, $P(s)$ and $f(t, s)$, considered as functions of s , at points s and (t, s) , respectively. Note that the function $f(t, s)$ is continuous in its first argument t . A number of numerical and approximate analytical methods for solving Volterra equations (7)–(9) can be found in [11].

2.3. Observations with Multiple Delays

The filtering equations over Ito-Volterra observations obtained in [6] enable us to consider multiple observation delays in the equation (2), thus allowing data fusion [9,10]. The corresponding observation process is given by

$$\begin{aligned}
y(t_j) &= \sum_{t_i} A_0(t_j, t_i) + \\
& A(t_j, t_i) x(t_i) + B(t_j, t_i) \psi(t_i). \quad (13)
\end{aligned}$$

Here, $y(t_j) \in R^m$ are discrete observations at time moments t_j , $j = 0, 1, \dots$, and $x(t_i)$, $t_i = t_1, t_2, \dots$, is a set of values of the system state at moments t_i available for measurement at the observation moment t_j , $A(t_j, t_i) \in R^{m \times n}$ are observation matrices, $\psi(t_i)$ are Gaussian noises independent of $W^1(t)$ and acting at the moments t_i . Let $A(t_j, t_i)$ be nonzero matrices, and $B(t_j, t_i) B^T(t_j, t_i)$ be positive definite ones.

The key point (based on the results of [6]) to obtain the optimal filter in this case is to assume that the observation distribution function $u(t)$ in (3) is a linear combination of the Heaviside functions $\sum_{t_i} \chi(t - t_i)$

with unit jumps at the time moments t_i when information on system states $x(t)$ is available (see the observation equation (13)). This implies that the filtering equations (7)–(9) between the observation moments t_j remain the same and the equations (10)–(12) for jumps of the filtering variables at t_j take the form

$$\Delta m(t_j) = \sum_{t_i} \{f(t_j, t_i-) [I + \quad (14)$$

$$\begin{aligned}
& A^T(t_j, t_i) (B(t_j, t_i) B^T(t_j, t_i))^{-1} \times \\
& A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} A^T(t_j, t_i) (B(t_j, t_i) B^T(t_j, t_i))^{-1} \times \\
& [y(t_j) - (A_0(t_j, t_i) + A(t_j, t_i) m(t_i-))],
\end{aligned}$$

$$\Delta P(t_j) = - \sum_{t_i} \{f(t_j, t_i-) [I + A^T(t_j, t_i) \times \quad (15)$$

$$\begin{aligned}
& (B(t_j, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} \times \\
& A^T(t_j, t_i) (B(t_j, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f^T(t_j, t_i-), \\
& \Delta f(t, t_j) = - \sum_{t_i} \{f(t, t_i-) [I + \quad (16)
\end{aligned}$$

$$\begin{aligned}
& (A^T(t_j, t_i) (B(t_j, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-) + \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
(1/2) & A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
& (1/2) A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} \times
\end{aligned}$$

$$\begin{aligned}
& A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} \times \\
& A^T(t_j, t_i) (B(t_j, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f^T(t_j, t_i-) + \\
& f(t_j, t_i-) [I + (A^T(t_j, t_i) (B(t_j, t_i) B^T(t_j, t_i))^{-1}
\end{aligned}$$

$$\begin{aligned}
& A(t_j, t_i) f(t_j, t_i-) + \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
(1/2) & A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
& (1/2) A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} \times
\end{aligned}$$

$$\begin{aligned}
& A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} \times \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f^T(t, t_i-) - \\
& (1/2) f(t_j, t_i-) [I + (A^T(t_j, t_i) \times
\end{aligned}$$

$$\begin{aligned}
& (B(t_j, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-) + \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
(1/2) & A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
& (1/2) A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} \times
\end{aligned}$$

$$\begin{aligned}
& A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} \times \\
& A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f^T(t, t_i-) - \\
& (1/2) f(t, t_i-) [I + (A^T(t_j, t_i) \times
\end{aligned}$$

$$\begin{aligned}
& (B(t_j, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f(t_j, t_i-) + \\
& A^T(t, t_i) (B(t, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
(1/2) & A^T(t_j, t_i) (B(t_j, t_i) B^T(t, t_i))^{-1} A(t, t_i) f(t, t_i-) - \\
& (1/2) A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} \times
\end{aligned}$$

$$\begin{aligned}
& A(t_j, t_i) f(t_j, t_i-) \Big]^{-1} \times \\
& A^T(t, t_i) (B(t, t_i) B^T(t_j, t_i))^{-1} A(t_j, t_i) f^T(t_j, t_i-).
\end{aligned}$$

Note that it is also possible to consider a set of observation delays (data fusion) of the continuum power, i.e., the system states are available for observation at every previous time moment. However, this problem completely coincides the filtering problem over Volterra observations, which has been considered in the general form in [6].

The next section presents the results for observations with delays in the important particular case of a dynamic system, if the state equation is a differential one.

3. Filtering for Dynamic System State

3.1. Observations with Multiple Delays

Let the general assumptions of Section 2 for a probability space and a Wiener process be satisfied. Let the nonobserved F_t -measurable random process $x(t)$ satisfy the state differential equation

$$dx(t) = (a_0(t) + a(t)x(t))dt + b(t)dW^1(t), \quad (17)$$

and the F_t -measurable discrete observations with multiple delays be given by the observation equation (13).

This simplification of the state equation (in comparison to the Ito-Volterra equation (1)) yields the following result based on the filtering equations (7)–(9) and (14)–(16). Let us note that the equality $x_s^t = x(s)$ holds due to coincidence of the equations (17) and (5) in this case. The σ -algebra $F_{t,s}^Y$ coincides with the σ -algebra F_s^Y , because, in view of (6), the σ -algebra $F_{t,s}^Y$ is actually specified by the random variables $x(t_i)$, $t_i = t_1, t_2, \dots \leq s$, up to the moment s for any t and, therefore, $F_{t,s}^Y = F_{s,s}^Y = F_s^Y$. Thus, we can conclude that

$$\begin{aligned} f(t, s) &= E((x_s^t - m_s^t)(x(s) - m(s))^T | F_{t,s}^Y) = \\ E((x(s) - m(s))(x(s) - m(s))^T | F_s^Y) &= P(s) \quad (18) \\ \text{and } m_s^t &= E(x_s^t | F_{t,s}^Y) = E(x(s) | F_s^Y) = m(s). \quad (19) \end{aligned}$$

Two last equalities imply that it is possible to obtain a closed system of filtering equations with respect to only two variables, the optimal estimate $m(t)$ and its variance $P(t)$, as it was done in the standard Kalman-Bucy problem, although we still assume that observations are given with delays. Namely, the optimal estimate $m(t)$ for the state vector (17) over the discrete observations with multiple delays (13) and its correlation function $P(t)$ satisfy the following equations between the observation moments t_j

$$m(t) = m(t_j+) + \int_{t_j+}^t (a_0(s) + a(s)m(s))ds, \quad (20)$$

$$\begin{aligned} P(t) &= P(t_j+) + \int_{t_j+}^t [a(s)P(s) + \\ &P(s)a^T(s) + b(s)b^T(s)]ds, \quad (21) \end{aligned}$$

and their jumps at the moments t_j of discrete observations are equal to

$$\Delta m(t_j) = \sum_{t_i} \{P(t_i-)[I + \quad (22)$$

$$\begin{aligned} &A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1} \times \\ &A(t_j, t_i)P(t_i-)]^{-1}A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1} \times \\ &[y(t_j) - (A_0(t_j, t_i) + A(t_j, t_i)m(t_i-))], \\ \Delta P(t_j) &= - \sum_{t_i} \{P(t_i-)[I + A^T(t_j, t_i) \times \quad (23) \\ &(B(t_j, t_i)B^T(t_j, t_i))^{-1}A(t_j, t_i)P(t_i-)]^{-1} \times \\ &A^T(t_j, t_i)(B(t_j, t_i)B^T(t_j, t_i))^{-1}A(t_j, t_i)P(t_i-)\}. \end{aligned}$$

Thus, the filtering equations (20)–(23) over discrete observations with delays have the same structure, up to necessary generalizations, as the standard Kalman-Bucy filtering equations for a continuous system state over discrete observations. Indeed, the systems (20)–(21) and (22)–(23) are closed with respect to only two filtering variables $m(t)$ and $P(t)$ and do not contain $f(t, s)$, unlike the equations (7)–(9) and (14)–(16). This significantly simplifies solution of these equations: it is possible to use conventional numerical methods for solving the equations (20)–(21) and to analytically compute jumps of the filtering variables at observation points, using the direct formulas (22)–(23).

The next subsection presents the optimal filter for the dynamic system state (17) over observations allowing a set of delays of the continuum power, i.e., the observation equation is assumed to be an integral equation of the Ito-Volterra type in the general form.

3.2. Observations with Continuum Delays

The following problem statement is considered. Let (Ω, F, P) be a complete probability space with an increasing right-continuous family of σ -algebras F_t , $t \geq 0$, and let $(W^1(t), F_t, t \geq 0)$ and $(W^2(t), F_t, t \geq 0)$ be independent Wiener processes. The partly observed F_t -measurable random process $(x(t), y(t))$ satisfies the equations

$$dx(t) = (a_0(t) + a(t)x(t))dt + b(t)dW^1(t), \quad (24)$$

$$\begin{aligned} y_i(t) &= \int_0^t (A_{0i}(t, s) + (A_i(t, s), x(s)))du_i(s) + \\ &+ \int_0^t B_i(t, s)dW_i^2(u_i(s)), \quad i = 1, \dots, m. \quad (25) \end{aligned}$$

Here, $x(t) \in R^n$ is a nonobserved component, and $y(t) \in R^m$ is an observed one for the process $(x(t), y(t))$. Functions $a_0(t)$, $a(t)$, $b(t)$ and functions $A_0(t, s)$, $A(t, s)$, $B(t, s)$ are continuous in t and s . Let $A(t, s)$ be a nonzero matrix and $B(t, s)B^T(t, s)$ be a positive definite matrix. The function $u(t) = (u_1(t), \dots, u_m(t)) \in R^m$ is a vector nondecreasing function of bounded variation. The rest of the notation is the same as for the equation (3).

The estimation problem is to find the best estimate for the system state (24) $x(t)$ at time t based on the observation process $Y(t) = \{y(s), 0 \leq s \leq t\}$, that is the conditional expectation $m(t) = E(x(t) | F_t^Y)$. Let $P(t) = E((x(t) - m(t))(x(t) - m(t))^T | F_t^Y)$ be the correlation function and F_t^Y be the σ -algebra generated by the observations $Y(t)$.

Let us note that the observation equation (25) actually gives us observations with a set of delays of the continuum power, because the observation process $y(t)$ at a point t depend on all values of the state vector $x(s)$: from $x(0)$ to $x(t)$. Thus, the allowed set of delays includes all time shifts from 0 to t , i.e., coincides with all the interval $[0, t]$, which has the power of continuum.

Since all the previous subsection propositions concerning validity of the equalities (18)–(19) are also satisfied for the continuous Ito-Volterra observations (25), the desired filtering equations readily follow from the filtering equations for an Ito-Volterra process over Ito-Volterra observations obtained in [6]. As a result, the following system of filtering equations for the dynamic system state (24) over the Ito-Volterra observations (25) is obtained.

$$m(t) = \int_0^t (a_0(s) + a(s)m(s))ds + \quad (26)$$

$$\int_0^t P(s-) [I + A^T(t, s)(B(t, s)B^T(t, s))^{-1} \times \\ A(t, s)P(s-)\Delta u(s)]^{-1} A^T(t, s)(B(t, s)B^T(t, s))^{-1} \times \\ [dy(t) - (A_0(t, s) + A(t, s)m(s-))du(s)], \\ P(t) = \int_0^t [a(s)P(s) + P(s)a^T(s) + \quad (27) \\ b(s)b^T(s)]ds - \int_0^t P(s-) [I + A^T(t, s) \times \\ (B(t, s)B^T(t, s))^{-1} A(t, s)P(s-)\Delta u(s)]^{-1} \times \\ A^T(t, s)(B(t, s)B^T(t, s))^{-1} A(t, s)P(s-)du(s),$$

where $\Delta u(s)$ is a jump of the function $u(s)$ at s , $P(s-)$ and $m(s-)$ are values from the left of the functions $P(s)$ and $m(s)$ at s , I is the $n \times n$ -dimensional identity matrix. Note that multiplication by an m -dimensional measure $du(t)$ should be understood in the componentwise sense, as in the observation equation (25).

If, in the equation (25), $u(t) = t$ and the measure $du(t)$ is equal to the Lebesgue measure dt , we obtain the case of pure continuous observations with continuum delays. Therefore, the filtering equations (26)–(27) are valid with $\Delta u(s) = 0$, $du(t) = t$, and the continuous functions $m(s-) = m(s)$ and $P(s-) = P(s)$.

If the bounded variation function $du(t)$ and, therefore, the observation process $y(t)$ have both continuous and discrete components, we examine the case of discontinuous observations with continuum delays. The corresponding jumps of the filtering variables $m(t)$ and $P(t)$ at a discontinuity point t_i of $u(t)$ (and $y(t)$) are equal to the expressions under the integral

signs in the right-hand sides of (26)–(27), upon substituting the jumps $\Delta u(t_i)$ and $\Delta y(t_i)$ for differentials $du(t)$ and $dy(t)$, respectively.

Since the system of filtering equations (26)–(27) is also closed with respect to only two filtering variables $m(t)$ and $P(t)$ and do not contain $f(t, s)$, unlike the filtering equations for an Ito-Volterra process obtained in [6], all the remarks concerning simplification of numerical solution of these equations remain valid as well as in the previous subsection.

4. References

1. T.D. Larsen, N.A. Andersen, O. Ravn, and N.K. Poulsen, "Incorporation of the time-delayed measurements in a discrete-time Kalman filter," *Proc. 37th Conference on Decision and Control 1998, (Tampa, FL, December 1998)*, pp. 3972–3977, 1998.
2. F.H. Hsiao and S.T. Pan, "Robust Kalman filter synthesis for uncertain multiple time-delay stochastic systems," *J. of Dynamic Systems, Measurement, and Control*, 1996.
3. H.L. Alexander, "State estimation for distributed systems with sensing delay," *SPIE. Data Structures and Target Classification*, vol. 1470, 1991.
4. R.E. Kalman, "A new approach to linear filtering and prediction problems," *ASME Trans., Part D (J. of Basic Engineering)*, vol. 82, pp. 35–45, 1960.
5. P. Frost and T. Kailath, "An innovations approach to least squares estimation — Part III: Nonlinear estimation in white Gaussian noise," *IEEE Trans. Automat. Contr.*, vol. AC-16, no. 6, pp. 217–226, 1971.
6. M.V. Basin and M.A. Villanueva Llanes, "On filtering problems over Ito-Volterra observations," *Proc. American Control Conference 1999 (San Diego, CA, June 1999)*, Paper no. ACC99-IEEE1166, Track no. FA16–5.
7. M.L. Kleptsina and A.Yu. Veretennikov, "On filtering and properties of conditional laws of Ito-Volterra processes," *Statistics and Control of Stochastic Processes. Steklov Seminar. 1984*, New York: Optimization Software Inc., Publication Division, pp. 179–196, 1985.
8. L.E. Shaikhet, "On an optimal control problem of partly observable stochastic Volterra processes," *Problems of Control and Inform. Theory*, vol. 16, no 6, pp. 439–448, 1987.
9. Y. Bar-Shalom and X.R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, Artech House, 1993.
10. C.K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, Berlin et al.: Springer-Verlag, 1991.
11. A.S. Cakmak, J.F. Botha, and W.G. Gray, *Computational and Applied Mathematics for Engineering Analysis*, Southampton: Springer-Verlag, Computational Mechanics Publications, 1987.

Apéndice C

**”ALGORITMOS DE FILTRADO DE KALMAN EN SISTEMAS DINÁMICOS Y DE
VOLTERRA”**

SEMINARIO

Seminario impartido dentro de la 5a. semana de ingeniería electrónica del Instituto Tecnológico de Nuevo Laredo, Tamaulipas, el 26 de noviembre de 1998.

SEP

INSTITUTO TECNOLÓGICO de nuevo laredo



Otorga el presente Reconocimiento

Al C. Ing. Mario Ariel Villanueva Llanes

Por su brillante conferencia:

"Algoritmos de filtrados de Kalman en sistemas dinámicos y de Volterra"
Dentro de la:

1969-1998

5a. Semana

INGENIERIA ELECTRONICA
UN PASO... A LA EXCEDENCIA

23 AL 27 DE NOVIEMBRE DE 1998

Nuevo Laredo, Tam., 26 de noviembre de 1998



SECRETARÍA DE
EDUCACIÓN PÚBLICA
INSTITUTO TECNOLÓGICO
DE NUEVO LAREDO
DIRECCIÓN

Ing. Fidel Aguilón Hernández

Director

Referencias

- [1] N. Wiener, "*The Extrapolation, Interpolation and Smoothing of Stationary Time Series*," John Wiley and Sons Inc., New York, N. Y. 1949.
- [2] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Trans. ASME, Series D, Journal of Basic Engineering*, Vol. 82, pp. 35-45, 1960.
- [3] R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," *Trans. ASME, Series D, Journal of Basic Engineering*, Vol. 83, pp. 95-108, 1961.
- [4] K. Ito, "On the Existence and Uniqueness of Solutions of Stochastic Integral Equations of the Volterra Type," *Kodai Math. Journal*, No. 2, pp. 158-170, 1979.
- [5] A. V. Balakrishnan, "On stochastic Bang Bang Control," *Lecture Notes in Control and Information Sciences*, vol. 25, New York: Spriger-Verlag, pp. 221-238, 1980.
- [6] M. L. Kleptsina and A. Yu. Veretennikov, "On Filtering Properties of Conditional Laws of Ito-Volterra processes," *Statistics and Control of Stochastics Processes. Steklov Seminar. 1984*, New York: Optimization Software Inc., Publication Division, pp. 179-196, 1985.
- [7] L. E. Shaikhet, "On an Optimal Control Problem of Partly Observable Stochastic Volterra Process," *Problems of Control and Information Theory*, vol. 16, No. 6, pp. 439-448, 1987.
- [8] M. V. Basin, "On Filtering of Ito-Volterra Processes over Discrete-Continuous Observations," *Automation and Remote Control*, vol.53, No. 8, Part 1, pp. 1183-1192, 1992.
- [9] Y. V. Orlov and M. V. Basin, "On Minimax Filtering over Discrete-Continuous Observations," *IEEE Trans. Automat. Control*, vol. AC-40, No. 9, pp. 1623-1626, 1995.
- [10] A. N. Kolmogorov and S. V. Fomin, "*Introductory Real Analysis*," New York: Dover, 1975.
- [11] R. S. Liptser and A. N. Shiriyayev, "*Statistics of Random Processes I: General Theory*," New York, Springer-Verlag, Graduate Texts in Mathematics, 1977.
- [12] A. N. Shiriyayev, "*Probability*," New York, Springer-Verlag, Graduate Texts in Mathematics, 1984.
- [13] T. C. Gard, "*Introduction to Stochastic Differential Equations*," *Monographs and Textbooks in Pure Applied Mathematics*, Marcel Dekker, Inc., 1988.

- [14] J. M. Amillo y F. de Arriaga, "*Análisis Matemático con Aplicaciones a la Computación*," Editorial Mc Graw-Hill, 1987.
- [15] A. S. Cakmak, J. F. Botha and W. G. Gray, "*Computational and Applied Mathematics for Engineering Analysis*," New York, Springer-Verlag, Computational Mechanics Publications Southampton Boston, 1987.
- [16] A. C. Antoulas, "*Mathematical System Theory: the Influence of R. E. Kalman*," New York, Springer-Verlag, 1991.
- [17] T. Kailath, "*Linear Systems*," Prentice-Hall Information and Systems Sciences Series, Thomas Kailath Series Editor, 1980.
- [18] B. O. Anderson and J. B. Moore, "*Optimal Control*," Linear Quadratic Methods, Prentice Hall, Information and Systems Science Series, 1990.
- [19] S. K. Mitter, "Filtering and Stochastic Control: A Historical Perspective," *Control Systems*, 16 (3), pp. 67-76, 1996.
- [20] P. Frost and T. Kailath, "An Innovations Approach to Least Squares Estimation-Part III: Nonlinear Estimation in White Gaussian Noise," *IEEE Trans. on Automatic Control*, AC-16, pp.217-226, June 1971.
- [21] F. Riesz and B. Sz. Nagy, "*Functional Analysis*," (Translated by L. F. Boron), Frederick Ungar Publishing Co., New York 1955.
- [22] H. W. Bode and C. E. Shannon, "A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory," *Proc. IRE*, Vol. 83, pp. 417-425, April 1950.
- [23] L. L. Scharf, "*Statistical Signal Processing*," Detection, Estimation, and Time Series Analysis; Addison Wesley Publishing Company, 1990.
- [24] M. V. Basin and M. A. Villanueva L., "On Filtering Problems over Ito-Volterra Observations," *Proc. American Control Conference 1999* (San Diego, CA, June 1999), pp. id. ACC99-IEEE1166, Track no. FA16-5, 1999.
- [25] S. Bittanti, A. J. Laub, and J. C. Willems, Eds., "*The Riccati Equation*," Springer-Verlag, New York, 1991.

- [26] W. J. Rugh, "*Linear System Theory*," Second Edition, Prentice-Hall Information and Systems Sciences, 1996.
- [27] L. Schwartz, "*Analyse Mathématique*," Hermann, Paris, 1967.
- [28] Y. V. Orlov, "*Theory of Optimal Systems with Generalized Controls*," Nauka, Moscow, 1988.
- [29] A. F. Filippov, "*Differential Equations with Discontinuous Right-Hand Side*," Kluwer Academic Publishers, New York, 1988.
- [30] Y. V. Orlov, "Vibrocorrect Differential Equations With Measure," *Mathematical Notes*, 38, No. 1, 110-119, 1985.
- [31] V. A. Dykhta and G. A. Kolokol'nikova, "Minimum Conditions on a Set of Sequences in a Degenerate Variational Problem," *Mathematical Notes*, 34, No. 5, 735-744, 1983.
- [32] M. V. Basin, M. A. Villanueva L. and Irma R. Valadez G., "On Filtering Problems over Observations with Delays," *Proc. CDC 1999*, pp. id. CDC99-REG0090, 1999.
- [33] R. Alvarado E., "*Procesamiento de Imagenes Digitales*," Universidad Autónoma de Nuevo León, F.I.M.E., Publicación s/n, Diciembre, 1998.
- [34] L. Schwartz, "*Théorie des Distributions*," I, II, Hermann, Paris, 1950, 1951.

