

2021-01

Identificación de Sistemas Lineales con datos cuantizados

ALBORNOZ GONZALEZ, RICARDO MIGUEL

<https://hdl.handle.net/11673/50066>

Repositorio Digital USM, UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE ELECTRÓNICA

Valparaíso, Chile



**UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA**

IDENTIFICACIÓN DE SISTEMAS LINEALES CON DATOS CUANTIZADOS

Tesis de Grado presentada por

Ricardo Miguel Albornoz González

como requisito parcial para optar al título de

Ingeniería Civil Electrónica

y al grado de

Magíster en Ciencias de la Ingeniería Electrónica

Supervisor

Juan Carlos Agüero Vásquez

Co-Supervisor

Rodrigo Carvajal Guerra

Enero, 2021

Título de Tesis:

Identificación de Sistemas Lineales con Datos Cuantizados

Autor:

Ricardo Miguel Albornoz González

Trabajo de Tesis, presentado en cumplimiento parcial de los requisitos para el título de Ingeniero Civil Electrónico y al grado de Magíster en Ciencias de la Ingeniería Electrónica de la Universidad Técnica Federico Santa María.

Supervisor

Dr. Juan C. Agüero

Co-Supervisor

Dr. Rodrigo Carvajal

Revisor Interno y Presidente de la Comisión

Dr. Juan I. Yuz

Revisor Externo

Dr. Boris I. Godoy

*A mis padres Margarita y Ricardo,
mi hermana Catherine y
mi amor Genesis*

Agradecimientos

Sin duda las primeras personas a las que quisiera agradecer es a mis padres, no tan solo por este proceso que finaliza, sino que por hacer de mi la persona que soy. Siempre me motivaron a pensar en grande y luchar por conseguir esos objetivos, no obstante, lo que más agradezco es que me enseñaron a buscar la felicidad. Todos los momentos vividos, sean dulces o agraces contribuyeron a formarme y conseguir cerrar esta importante etapa. Junto a ellos, quiero agradecer a mi hermanita, quien lideró e inauguró el sueño de la educación superior en mi familia. El tenerte a ti (Cati) hizo todo más fácil, me motivó a educarme e hizo crecer mi interés por el conocimiento y las ciencias. Siempre hemos sido los mejores *partners* y confidentes, y aunque la distancia es grande, siempre estamos juntos.

No puedo dejar de agradecer desde lo más profundo de mi corazón a mi Gene, quien ha estado junto a mi desde antes de partir este desafío, me apoyó en cada instante, me acompañó en los momentos satisfactorios y me contuvo en los momentos complejos. Desde luego debo agradecerte la paciencia, pues mi afán por tomar más desafíos y realizar más actividades de las que el tiempo me permite, han reducido nuestros momentos, pero tu comprensión ha sido infinita.

Debo agradecer a mi gran amigo Julio, con quien comenzamos nuestra aventura a los 2 años de edad y seguimos estando tan unidos como durante toda nuestra infancia. A mis amigos del liceo, Poblete, Nea, Mery, Pablo, Belén y Carlita. A quienes se sumaron a ser personas importantes en mi vida durante la universidad, Opitz, Danilo y Panchito.

Debo dar las gracias a mis tíos, Margarita y Rubén, quienes este año me han acogido como un hijo. He sentido ese cariño familiar y su apoyo contribuyó a culminar esta etapa con éxito. Junto a ellos debo agradecer a mis tías, Gloria y Raquel, quienes me han apoyado desde que nací, hasta el día de hoy, pasando por cada una de mis etapas como estudiante.

Asimismo, debo agradecer al profesor Juan Carlos Agüero, quien ha guiado este trabajo de tesis. Su interés en general en las ciencias, su forma de entender nuestra área y hacer todo lo complejo tan simple me impresiona, aunque más de alguna vez me puso en aprietos. Por sobre todo, agradezco encontrar un profesor siempre disponible, con el mejor de los ánimos y preocupado en todo momento (por sobre lo profesional) de las personas que trabajan con él. Adicionalmente, agradecer al profesor Rodrigo, quien nos ha acompañado en este proceso de investigación. A Angel, con quien compartimos los últimos momentos de este proceso, intercambiando conocimiento y discutiendo ideas que sin duda han sido una importante contribución en este trabajo.

Finalmente, agradezco el apoyo del Centro Basal “Advanced Center for Electrical and Electronic Engineering” (AC3E) - FB0008, al proyecto ANID-FONDECYT 1181158 y al Programa de Incentivo a la Iniciación Científica (PIIC) de la USM.

Resumen

El área de identificación de sistemas ha experimentado un crecimiento exponencial en las últimas décadas debido al aumento de novedosas técnicas de Control moderno, las que en su mayoría se basan en modelos.

Por otro lado, el crecimiento tecnológico y la utilización de dispositivos cada vez más diversos conectados a redes, genera un desafío en cuanto a disminuir los costos de comunicación. Para afrontar estas necesidades se realizan operaciones como cuantización de señales (entre otras), las que producen una pérdida de información importante.

La presente Tesis tiene como objetivo principal la identificación de un sistema lineal e invariante en el tiempo, planteado en espacio de estados, donde la señal de salida experimenta un proceso de cuantización. Esta pérdida de información implica desarrollar técnicas distintas a las usuales para obtener estimadores con buenas características.

Acá, se presenta una solución encontrando el estimador de Máxima Verosimilitud para los parámetros del sistema. Dicho estimador, presenta características deseables, tales como, invarianza, eficiencia, consistencia, etc. Para la obtención del estimador de máxima verosimilitud, se estudia e implementa una solución a través del algoritmo iterativo *Expectation-Maximization*.

El rendimiento del algoritmo desarrollado es evaluado a través de simulaciones numéricas, donde es comparado con técnicas clásicas que no consideran cuantización.

Adicionalmente, y debido a las estadísticas necesarias en el algoritmo *Expectation-Maximization*, se desarrollan e implementan técnicas de Filtrado y Suavizado para el sistema en particular, las que también son evaluadas a través de ejemplos numéricos.

Abstract

The area of system identification has experienced exponential growth in recent decades due to the increase of modern control techniques, which are mostly based on models.

On the other hand, technological growth and the use of increasingly diverse devices connected to networks, creates a challenge in terms of reducing communication costs. To address these needs, operations such as signal quantization (among others) are carried out, which produce a significant loss of information.

The main objective of this Thesis is to identify a linear and time-invariant system, written in state space form, where the output signal undergoes a quantization process. This loss of information implies the need for developing techniques, other than the usual ones, to obtain estimators with good characteristics.

Here, a solution is presented by finding the Maximum Likelihood estimator for the system parameters. This estimator has desirable characteristics, such as invariance, efficiency, consistency, etc. To obtain the Maximum Likelihood estimator, a solution is studied and implemented through the iterative Expectation-Maximization algorithm.

The performance of the proposed algorithm is evaluated through numerical simulations, where it is compared with classical techniques that do not consider quantization.

Additionally, and due to the necessary statistics in the Expectation-Maximization algorithm, Filtering and Smoothing techniques are developed and implemented for the system of interest, which are also evaluated through numerical examples.

Índice general

1. Introducción	1
1.1. Definición del Problema	4
1.2. Principales Contribuciones de la Tesis	5
1.3. Publicación asociada	5
1.4. Estructura del Documento	6
2. Problema de Filtraje y Suavizado	7
2.1. Enfoque Bayesiano para el Problema de Filtraje	8
2.2. Enfoque Bayesiano para el Problema de Suavizado	10
2.2.1. Forward-Backward Recursions	11
2.2.2. Generalised Two-Filter Formula	12
3. Filtraje y Suavizado en Sistemas Lineal con Salida Cuantizada	14
3.1. Filtraje	14
3.1.1. Filtro de Partículas	14
3.1.2. Filtro por Suma de Gaussianas	23
3.1.3. Ejemplos Numéricos	31
3.2. Suavizado	41
3.2.1. Suavizador de Partículas	41
3.2.2. Ejemplos Numéricos	44
4. Estimación por Máxima Verosimilitud a través del Algoritmo EM	50
4.1. Estimación por Máxima Verosimilitud	50
4.1.1. Propiedades del estimador ML	51
4.2. Nociones del Algoritmo EM	52
4.2.1. Algoritmo MM	52
4.2.2. Algoritmo EM	53

5. Estimación por Máxima Verosimilitud de Sistemas Dinámicos Lineales con Salida Cuantizada	59
5.1. Resumen del Algoritmo a través de partículas	67
6. Simulaciones	68
7. Conclusiones	81
7.1. Trabajo Futuro	82
A. Propiedades de Sumatorias y Distribuciones Gaussianas	84
B. Transformaciones de Similaridad	88
Referencias	94

Índice de figuras

1.1. Sistema Ciberfísico.	2
1.2. Sistema de Control sobre Redes (NCS).	3
1.3. Sistema en espacio de estados con salida cuantizada.	5
3.1. Distribución Gaussiana bi-dimensional y su representación de Monte Carlo	15
3.2. Esquema de una iteración del Algoritmo 3.2	22
3.3. Cuantizador utilizado en los ejemplos 1 a 6.	31
3.4. Estimación de Estado y Error de Estimación.	32
3.5. Error Cuadrático Medio entre el estado real x_t y la estimación $\hat{x}_{t t}$ de cada uno de los métodos.	33
3.6. Estadísticas con $\Delta = 10$, 10 Gaussianas y 20000 partículas.	34
3.7. Estadísticas con $\Delta = 10$, 6 Gaussianas y 300 partículas.	35
3.8. Función de densidad $p(x_t y_{1:t})$ para distintos instantes de tiempo, usando $\Delta = 10$	36
3.9. Error cuadrático medio en sistema de segundo orden.	37
3.10. Estimación de Estado y Error de Estimación.	38
3.11. Funciones de Densidad de Probabilidad Marginales de cada uno de los estados, para distinto número de Gaussianas y partículas, utilizando $\Delta = 10$ y en el instante $t = 60$	39
3.12. Función de densidad de probabilidad $p(x_t y_{1:t})$ estimada por cada método, en el instante $t = 60$, con $\Delta = 10$, 20 Gaussianas y 10000 partículas.	40
3.13. Estimación y Varianza considerando $\Delta = 10$ y 20000 partículas.	44
3.14. Error Cuadrático Medio incluyendo suavizado.	45
3.15. Función de densidad $p(x_t y_{1:N})$ para distintos tiempos, con $\Delta = 10$	46
3.16. Estimación de Estados, con $\Delta = 10$ y 10000 partículas.	47
3.17. Varianza de la Estimación, con $\Delta = 10$ y 10000 partículas.	48
3.18. Error Cuadrático Medio incluyendo suavizado en sistema de segundo orden.	48

3.19. Función de densidad $p(x_t y_{1:N})$ para cada método, con $\Delta = 10$, 10 Gaussianas y 10000 partículas.	49
6.1. Diagrama de Bode de 200 realizaciones, donde se muestra la respectiva media con línea segmentada.	69
6.2. Boxplot de la estimación de la covarianza del ruido de medición R , obtenida para las 200 realizaciones.	70
6.3. Error de estimación relativo.	70
6.4. Boxplot de los parámetros fijando Q al valor verdadero.	71
6.5. Diagrama de Bode de 200 realizaciones, con salida cuantizada en 2 bits.	72
6.6. Error de estimación relativo usando 2 bits para la salida.	73
6.7. Diagrama de Bode de 200 realizaciones, con salida cuantizada en 1 bits.	73
6.8. Error de estimación relativo usando 1 bits para la salida.	74
6.9. Diagrama de Bode de 200 realizaciones, donde se muestra la respectiva media con línea segmentada, en sistema de segundo orden.	75
6.10. Boxplot de la estimación de la covarianza del ruido de medición R , obtenida para las 200 realizaciones, en sistema de segundo orden.	76
6.11. Estimación de ambos polos para 200 realizaciones, con su respectiva media.	77
6.12. Error de estimación relativo, en sistema de segundo orden.	77
6.13. Cuantizador con pendiente, utilizado en el Ejemplo 7.	78
6.14. Diagrama de Bode de 200 realizaciones, donde se muestra la respectiva media con línea segmentada.	79
6.15. Boxplot de la estimación de la covarianza del ruido de medición R , obtenida para las 200 realizaciones.	80
6.16. Error de estimación relativo.	80

1

Introducción

Los campos de Teoría de Comunicaciones y Teoría de Control fueron tradicionalmente áreas con pocos puntos en común. Por un lado, Teoría de Comunicaciones se ocupa principalmente de la transmisión confiable de la información de un punto a otro, indiferente al propósito específico de la información transmitida y si eventualmente se retroalimenta a la fuente. Por otro lado, la Teoría de Control, se ocupa principalmente del uso de la información en un ciclo de retroalimentación para lograr algún objetivo de rendimiento, y comúnmente asume enlaces de comunicación perfectos entre la planta y los controladores [1].

La Teoría de Control ha sufrido varias fases distintas de desarrollo y ha habido una progresión desde la interconexión mecánica directa, la interconexión neumática e hidráulica, la conexión eléctrica, y finalmente, la conexión en red y/o inalámbrica. La última fase ha abierto desafíos completamente nuevos para la comunidad de diseño de sistemas de control. En ingeniería de sistemas con grandes anchos de banda en los enlaces de comunicación, aún tiene sentido tratar la comunicación y el control como funciones independientes, ya que el análisis y diseño del sistema general es simplificado. Sin embargo, los recientes avances en tecnologías de comunicación y aplicaciones emergentes, tales como redes de sensores, sistemas micro-electromecánicos, telefonía móvil, y redes de control industrial, comenzaron a cambiar la validez del enfoque modular. Este tipo de aplicaciones abrió la puerta a una nueva generación de sistemas de ingeniería, llamados Sistemas Ciberfísicos (CPS por las siglas de la frase en inglés *Cyber-Physical Systems*), los que integran tecnologías de computación, comunicación y control con sistemas físicos [2]. En este tipo de sistemas, el foco está en la interacción entre varios sistemas físicos con control y dispositivos de monitoreo (Estación Base) conectados vía redes de sensores y actuadores. Un sistema ciberfísico típico es ilustrado en la Figura 1.1.

Este tipo de avances tanto tecnológicos como de control han empujado el desarrollo del área llamada Sistemas de Control sobre Redes (NCS por las siglas en inglés de *Networked Control Systems*) [3], el que ha enfrentado el problema de análisis y diseño de control a través de un canal de comunicación en red. En particular, la estación base de un sistema ciberfísico representa el controlador [4]. La Figura 1.2 muestra un diagrama de bloques típico de un NCS, donde los

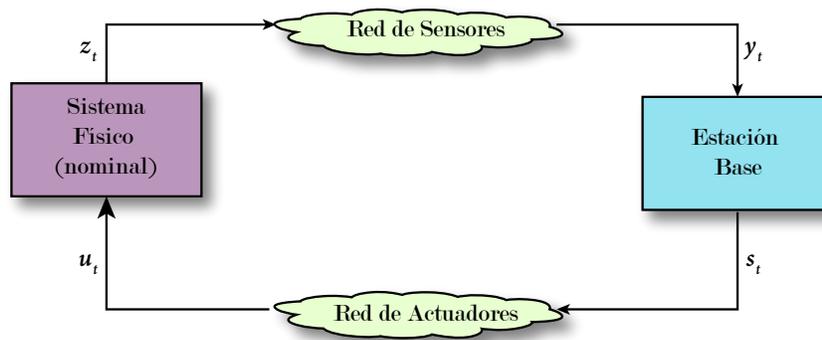


Figura 1.1: Sistema Ciberfísico.

comandos de entrada y las mediciones usadas por el controlador son transmitidos a través de una red. En general, se necesita acondicionar las señales antes de transmitir sobre la red y en algunos casos es necesario reinterpretar las señales al recibirlas (si existió codificación). Esta nueva arquitectura de control incrementa la aplicabilidad y reduce los costos en comparación a soluciones cableadas. Sin embargo, también presenta dificultades, ya que los enlaces de comunicación pueden presentar pérdida de datos, experimentar retardos aleatorios y están sujetos a limitaciones [5], lo que obliga a los diseñadores a cuantizar las señales para reducir los costos de comunicación.

Por otro lado, la mayoría de los métodos modernos de control y procesamiento de datos están basados en modelos [6–8]. Estos modelos que representan sistemas físicos, pueden en principio, ser obtenidos usando principios de la Física y otras ramas de la ciencia. Sin embargo, dado que los sistemas son complejos, es necesario ajustarlos usando datos experimentales, lo que recae en el área de Identificación de Sistemas [9, 10], entre otras cosas. Por consiguiente, es necesario desarrollar nuevas metodologías para obtener modelos de los sistemas físicos que incluyan en alguna medida las dificultades del canal de comunicación. En este caso, un problema de interés y en el cual se enfoca esta tesis, es en la identificación de sistemas en que solo observaciones de salida cuantizada están disponibles [11], debido a las limitaciones descritas anteriormente.

La estimación de este tipo de sistemas es una tarea desafiante, ya que las técnicas de identificación clásicas [9, 10] pueden entregar resultados deficientes, debido a la presencia del cuantizador y la consecuente pérdida de información. Por esto, es necesario modificar la teoría y los algoritmos de identificación clásicos, para incorporar el conocimiento de la cuantización en nuevos algoritmos.

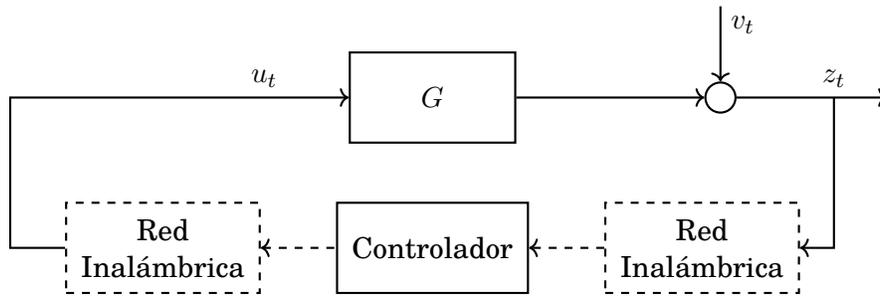


Figura 1.2: Sistema de Control sobre Redes (NCS).

El problema de identificación de sistemas dinámicos usando mediciones cuantizadas ha atraído la atención de varios grupos de investigación [11–29], la razón ya ha quedado clara. Sin embargo, la mayoría de los métodos encontrados en la literatura han sido desarrollados para una clase particular, y muy limitada, de sistemas lineales, planteados en función de transferencia. Por ejemplo, en [30] se desarrolló un estimador de máxima verosimilitud (ML por su nombre en inglés *Maximum Likelihood*) para la estimación de parámetros de un sistema de corriente alterna. En [31] se desarrolló un algoritmo con buenas propiedades de convergencia para identificar un sistema de respuesta a impulso finita (FIR por su nombre en inglés *Finite Impulse Response*). En [32] se encuentra un algoritmo ciego¹ de identificación para sistemas FIR multicanal. En [23] y [28] se desarrolló un algoritmo EM (por su nombre en inglés *Expectation-Maximization*) para identificar sistemas FIR. Por otro lado, en [26] se desarrolló un método de identificación para sistemas OE (*Output Error* en inglés) con salida cuantizada e intermitente. El trabajo en [33] considera ruido coloreado para el diseño conjunto de cuantizador y estimador, basado en el algoritmo recursivo y las ideas desarrolladas en [29], sin embargo, el estimador ML de un sistema general (con ruido coloreado), aún no ha sido desarrollado.

Los trabajos más cercanos a identificar un sistema general con salida cuantizada en un marco de estimación ML son [34] y [35]. En [34] se desarrolla un algoritmo para sistemas autoregresivos con variables exógenas (ARX por su nombre en inglés *AutoRegressive with eXogenous variables*).

Por otro lado, el cambio de enfoque o tendencia que ha experimentado la Teoría de Control hacia técnicas de control moderno basadas en el concepto de estado han comenzado el desarrollo de métodos de identificación directamente en espacio de estados [9, 36], en contraste con la metodología usualmente implementada en sistemas lineales, que consiste en identificar un sistema en función de transferencia y luego obtener su representación en espacio de estados.

En esta línea, en [35] se plantea el problema en espacio de estados, lo que puede abarcar una amplia gama de sistemas lineales. Sin embargo, la solución a través de escenarios produce un crecimiento de dimensión del problema en función del tiempo (algo que también sucede en [34]) y las covarianzas de los ruidos no son estimadas.

El problema a desarrollar en esta tesis se puede considerar una extensión de [35], ya que planteamos el sistema en espacio de estados, sin embargo, reemplazamos el uso de *scenarios* por un algoritmo de suavizado y buscamos estimar tanto las matrices del sistema en espacio de estados como las covarianzas de los ruidos. Este enfoque nos lleva a la necesidad de desarrollar e implementar algoritmos de filtraje y un algoritmo de suavizado. A continuación se define

¹Un método de identificación ciego es identificar los parámetros del sistema solo usando sus salidas bajo algunas suposiciones débiles sobre la entrada y la estructura del sistema.

formalmente el problema de interés.

1.1 Definición del Problema

El problema considerado en esta tesis es obtener el estimador por máxima verosimilitud de los parámetros de un sistema de orden n y salida cuantizada, descrito por las siguientes ecuaciones en espacio de estados

$$x_{t+1} = Ax_t + Bu_t + w_t \quad (1.1a)$$

$$z_t = Cx_t + Du_t + v_t \quad (1.1b)$$

$$y_t = \mathfrak{q}\{z_t\}, \quad (1.1c)$$

donde $u_t \in \mathbb{R}^r$ denota la entrada, $x_t \in \mathbb{R}^n$ el vector de estado, $z_t \in \mathbb{R}^m$ la salida (no cuantizada) e $y_t \in \mathbb{R}^m$ es su versión cuantizada.

El sistema contiene ruido de proceso $w_t \in \mathbb{R}^n$ y ruido de medición $v_t \in \mathbb{R}^m$, siendo ambos secuencias de ruido blanco Gaussianos de media cero y covarianzas Q y R , respectivamente. Además, está sujeto a la distribución del estado inicial x_1 , que también es Gaussiano, pero de media μ_1 y matriz de covarianza P_1 . Los ruidos tanto de medición como de proceso son independientes entre sí, y el último también es independiente del estado inicial.

Por otro lado, la función $\mathfrak{q}\{\cdot\}$ corresponde al cuantizador, definido por

$$\mathfrak{q}\{z_t\} = \begin{cases} \nu_1 & \text{si } z_t \in \Omega_1 \\ \nu_2 & \text{si } z_t \in \Omega_2 \\ \vdots & \vdots \\ \nu_L & \text{si } z_t \in \Omega_L \end{cases}, \quad (1.2)$$

el que cuenta con L niveles de cuantización y donde los intervalos Ω_ℓ son disjuntos y corresponden a un hipercubo m -dimensional con extremos ordenados en los vectores a_t^ℓ y b_t^ℓ , esto es $\Omega_\ell = \{z_t \in \mathbb{R}^m : a_t^\ell \leq z_t < b_t^\ell\}$ con $\ell \in [1, L] \subset \mathbb{N}$. En la Figura 1.3 se puede observar un esquema del sistema en cuestión.

Los parámetros a estimar se coleccionarán en el vector θ de la siguiente forma:

$$\theta \triangleq \begin{bmatrix} \beta \\ \eta \end{bmatrix} \quad (1.3)$$

donde

$$\beta \triangleq \begin{bmatrix} \text{vec}\{\Gamma\} \\ \mu_1 \end{bmatrix}; \quad \eta \triangleq \begin{bmatrix} \text{vec}\{\Pi\} \\ \text{vec}\{P_1\} \end{bmatrix} \quad (1.4)$$

y

$$\Gamma \triangleq \begin{bmatrix} A & B \\ C & D \end{bmatrix}; \quad \Pi \triangleq \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}. \quad (1.5)$$

La estimación se realiza sobre la base de una secuencia de datos entrada-salida observados $\{u_{1:N}, y_{1:N}\}$ donde

$$u_{1:N} \triangleq \{u_1, u_2, \dots, u_N\}; \quad y_{1:N} \triangleq \{y_1, y_2, \dots, y_N\}, \quad (1.6)$$

y de manera análoga podríamos referirnos a otro conjunto de datos.

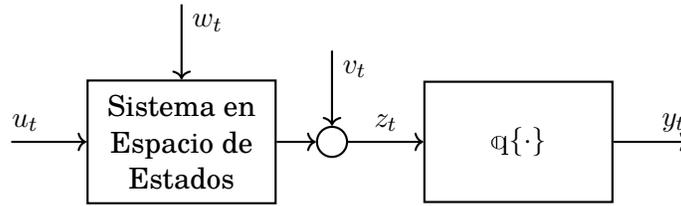


Figura 1.3: Sistema en espacio de estados con salida cuantizada.

1.2 Principales Contribuciones de la Tesis

Las principales contribuciones de la tesis son las siguientes:

1. Se propone un algoritmo de filtraje por suma de Gaussianas para un sistema en lineal e invariante en el tiempo, escrito en tiempo espacio de estados, donde la salida experimenta un proceso de cuantización. El algoritmo propuesto se obtiene de forma natural al considerar una aproximación de la función de masa de probabilidad de salida del sistema, a través de una suma de Gaussianas.
2. Se implementa un algoritmo de filtraje basado en partículas para el sistema de interés, utilizando la función de masa de probabilidad en forma exacta. Ambos algoritmos se evalúan a través de simulaciones numéricas y se comparan con el Filtro de Kalman estándar.
3. Se implementa una solución al problema de suavizado para el sistema de interés basada en partículas. El algoritmo es evaluado a través de simulaciones numéricas y comparado con el respectivo algoritmo de filtraje.
4. Se propone un método de identificación para el sistema de interés. El método propuesto obtiene el estimador de Máxima Verosimilitud a través del algoritmo *Expectation-Maximization*.
5. El algoritmo de identificación propuesto se implementa a través de las soluciones al problema de filtraje y suavizado, basadas en partículas. El desempeño del algoritmo se evalúa a través de simulaciones numéricas.

1.3 Publicación asociada

- R. Albornoz, R. Carvajal, and J. C. Agüero, “A novel Bayesian filtering method for systems with quantized output data”, in *2019 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*. IEEE, 2019, pp. 1–7.

Abstract: In this paper we develop a novel scheme for state estimation of discrete-time linear time-invariant systems with quantized output data. We take a Bayesian approach, therefore, we describe the behavior of the *a posteriori* probability density function of the state. The difficulty of this problem lies in the probability function of the measurable output given the state, which we approach through an approximation by a Gaussian sum, that naturally leads to a Gaussian sum for the *a posteriori* density function.

1.4 Estructura del Documento

La estructura de los siguientes capítulos de esta tesis se describen a continuación.

- **Capítulo 2:** Se establece el problema de Filtrado y Suavizado general a través de un enfoque Bayesiano.
- **Capítulo 3:** Se presenta un estudio al problema de Filtrado y Suavizado Bayesianos para el problema aquí presentado, incluyendo cuantización de salida. Este estudio es motivado por la necesidad encontrada en el algoritmo descrito en el Capítulo 5. Aquí se presenta la primera contribución de este trabajo, la que corresponde a una solución para el problema de filtrado mediante Suma de Gaussianas y una implementación alternativa, mediante Filtro de Partículas (*Particle Filter*). Luego, para el problema de Suavizado, se implementa una solución a través de un Suavizador de Partículas (*Particle Smoother*).
- **Capítulo 4:** Se comienza dando una descripción del estimador de máxima verosimilitud y sus propiedades, para luego pasar al estudio del algoritmo EM y algunas propiedades de convergencia, introducido a través del amplio concepto de algoritmos MM.
- **Capítulo 5:** Aquí se presenta el algoritmo de estimación del sistema con datos cuantizados, el que entrega el estimador de máxima verosimilitud, a través del algoritmo EM, lo que es el objetivo principal de esta tesis.
- **Capítulo 6:** Se muestran una serie de simulaciones, para evaluar el rendimiento del método obtenido, presentando algunas comparaciones con otros métodos de estimación.
- **Capítulo 7:** Finalmente se presentan las conclusiones del estudio, y se discuten alcances de trabajos futuros.

2

Problema de Filtrado y Suavizado

El término *filtrado óptimo* tradicionalmente se refiere a una clase de métodos que pueden ser usados para estimar el estado de un sistema variante en el tiempo que es indirectamente observado a través de mediciones ruidosas. El término *óptimo* en este contexto se refiere a optimalidad estadística.

Existen distintas formas de plantear este problema, sin embargo, acá se utilizará *filtrado Bayesiano* [37], que se refiere a la forma Bayesiana de formular el filtrado óptimo.

Suavizador Bayesiano es a menudo considerado como una clase de métodos dentro del campo de filtrado Bayesiano. Mientras los filtros Bayesianos en su forma básica solo calculan estimaciones del estado actual del sistema dada un historial de mediciones, los suavizadores Bayesianos pueden ser usados para reconstruir estados que sucedieron antes del tiempo actual. Aunque el término *suavizado* a veces se usa en un sentido más general para métodos que generan una representación suave de los datos (a diferencia de brusca), en el contexto de filtrado Bayesiano el término suavizado (Bayesiano) tiene este significado más definido.

El éxito del filtrado lineal óptimo en aplicaciones de ingeniería se debe principalmente al artículo fundacional de Kalman [38, 39], que describe la solución recursiva al problema de filtrado lineal óptimo de sistemas de tiempo discreto. Rápidamente después del aporte de Kalman se descubrió que el Filtro de Kalman pertenecía a la clase de filtros Bayesianos [40–43], teoría que se desarrollaba en paralelo a la de Kalman por Stratonovich [44]. El correspondiente suavizador Bayesiano también fue desarrollado poco tiempo después de la invención del Filtro de Kalman [45, 46].

A pesar de que la derivación original del Filtro de Kalman fue basada en el enfoque de mínimos cuadrados, las mismas ecuaciones pueden ser derivadas desde un análisis de probabilidad Bayesiana puro, y es este el enfoque que se seguirá para el problema de datos cuantizados en este trabajo.

2.1 Enfoque Bayesiano para el Problema de Filtraje

Definición 2.1 (Modelo en espacio de estados probabilístico) Un modelo en espacio de estados probabilístico o modelo de filtraje no lineal, consiste de una secuencia de distribuciones de probabilidad condicional:

$$x_{t+1}|x_t \sim p(x_{t+1}|x_t), \quad (2.1)$$

$$y_t|x_t \sim p(y_t|x_t) \quad (2.2)$$

para $t = 1, 2, \dots$, donde

- $x_t \in \mathbb{R}^n$ es el estado del sistema en el tiempo t .
- $y_t \in \mathbb{R}^m$ es la medición en el tiempo t .
- $p(x_{t+1}|x_t)$ es modelo dinámico que describe la dinámica estocástica del sistema.
- $p(y_t|x_t)$ es el modelo de medición.
- $p(x_1|x_0) = p(x_1)$ es la densidad del estado inicial.

Lema 2.1 El modelo en espacio de estados probabilístico para el sistema en (1.1) es:

$$p(x_{t+1}|x_t) = \mathcal{N}_{x_{t+1}}(Ax_t + Bu_t, Q), \quad (2.3)$$

$$p(y_t|x_t) = \int_{a_t^\ell}^{b_t^\ell} \mathcal{N}_{v_t}(Cx_t + Du_t, R) dv_t. \quad (2.4)$$

donde a_t^ℓ y b_t^ℓ son obtenidos de la medición de y_t , tales que para cualquier valor $z_t \in [a_t^\ell, b_t^\ell]$ se cumple $\mathbb{Q}\{z_t\} = y_t$.

Demostración. La demostración de (2.3) es directa de (1.1a) y que $p(w_t) = \mathcal{N}_{w_t}(0, Q)$, usando transformación de variable aleatoria.

Luego, desde el cuantizador en (1.1c), se observa que la probabilidad de que $y_t = \nu_\ell$ es igual a la probabilidad de la variable aleatoria z_t pertenezca al conjunto $\Omega_\ell = \{z_t : a_t^\ell \leq z_t < b_t^\ell\}$. Esta probabilidad puede ser obtenida desde desde la función de distribución como sigue:

$$\mathbb{P}\{y_t = \nu_\ell|x_t\} = \mathbb{P}\{a_t^\ell \leq z_t < b_t^\ell\}, \quad (2.5)$$

donde $\mathbb{P}\{\cdot\}$ denota probabilidad. Considerando el sistema (1.1), obtenemos que:

$$p(y_t|x_t) = \mathbb{P}\{a_t^\ell \leq v_t - Cx_t - Du_t < b_t^\ell|x_t\}, \quad \text{si } y_t = \nu_\ell. \quad (2.6)$$

Entonces, usando que $p(v_t) = \mathcal{N}_{v_t}(0, R)$ se tiene que

$$p(y_t|x_t) = \mathbb{P}\{a_t^\ell \leq v_t < b_t^\ell\} = \int_{a_t^\ell}^{b_t^\ell} \mathcal{N}_{v_t}(Cx_t + Du_t, R) dv_t, \quad (2.7)$$

donde los límites de la integral corresponden a los extremos de los conjuntos Ω_ℓ . □

Note que (2.3) es una función de densidad de probabilidades (pdf por *Probability Density Function*) pues x_t es una variable aleatoria continua, y (2.4) es comúnmente llamada función de masa de probabilidades (pmf por *Probability Mass Function*), ya que y_t es una variable aleatoria discreta. Más adelante en el documento, se realizan operaciones entre funciones de probabilidades de variable aleatoria continua y de variable aleatoria discreta, lo que no genera ningún inconveniente [47]. En lo que sigue, no será especificado el tipo de variable aleatoria y se podría utilizar simplemente función de probabilidades para referirse a cualquiera de los dos casos. En [47] se utiliza el término *Generalized Probability Density Function* para generalizar estas ideas, el que además incluye funciones de probabilidad que pueden mezclar tanto variables continuas como discretas.

El propósito del filtraje Bayesiano es calcular la distribución *a posteriori* o como es llamada comúnmente *densidad filtrada* del estado x_t en cada instante de tiempo t dado el historial de mediciones hasta el tiempo t :

$$p(x_t|y_{1:t}). \quad (2.8)$$

En el siguiente teorema se enuncian las ecuaciones fundamentales de filtraje Bayesiano.

Teorema 2.1 Las ecuaciones recursivas para calcular la distribución $p(x_t|y_{1:t})$ y la distribución predicha $p(x_{t+1}|y_{1:t})$ en el tiempo t son dadas por el Teorema de Bayes

$$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})}, \quad (2.9)$$

llamado comúnmente paso de *Corrección* o en inglés *Measurement Update* y la ecuación de Chapman - Kolmogorov

$$p(x_{t+1}|y_{1:t}) = \int p(x_{t+1}|x_t)p(x_t|y_{1:t}) dx_t, \quad (2.10)$$

generalmente llamado paso de *Predicción* o *Time Update*. El término $p(y_t|y_{1:t-1})$ es una constante de normalización, que puede ser calculada como

$$p(y_t|y_{1:t-1}) = \int p(y_t|x_t)p(x_t|y_{1:t-1}) dx_t, \quad (2.11)$$

y la recursión entre ambos pasos es inicializada desde la distribución *a priori* $p(x_1|y_0) = p(x_1)$.

Demostración. La distribución de x_t dado y_t e $y_{1:t-1}$, es decir, $y_{1:t}$, puede ser calculada por el Teorema de Bayes

$$\begin{aligned} p(x_t|y_{1:t}) &= \frac{p(y_t|x_t, y_{1:t-1})p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \\ &= \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \end{aligned}$$

lo que corresponde a (2.9). La última igualdad viene de imponer que el sistema posea independencia condicional de las mediciones, esto es, $p(y_t|x_t, y_{1:t-1}) = p(y_t|x_t)$, lo que es evidente tanto en un sistema en espacio de estados clásico como en (1.1).

Luego, la distribución conjunta de x_{t+1} y x_t dado $y_{1:t}$ puede ser calculado como

$$p(x_{t+1}, x_t|y_{1:t}) = p(x_{t+1}|x_t, y_{1:t})p(x_t|y_{1:t}) \quad (2.12)$$

$$= p(x_{t+1}|x_t)p(x_t|y_{1:t}), \quad (2.13)$$

2.2. Enfoque Bayesiano para el Problema de Suavizado

donde la segunda igualdad viene de considerar x_t como un proceso Markoviano [37, 43], que en palabras sencillas quiere decir que la ley de probabilidad del proceso en el futuro, para un estado dado del proceso, no depende de cómo el proceso llegó al estado dado. Ahora, la distribución marginal de x_{t+1} dado $y_{1:t}$ puede ser obtenida integrando la distribución (2.13) sobre x_t ,

$$p(x_{t+1}|y_{1:t}) = \int p(x_{t+1}|x_t)p(x_t|y_{1:t}) dx_t,$$

lo que corresponde a (2.10). □

Definición 2.2 Una vez obtenidas las distribuciones (2.9) y (2.10), estamos interesados en obtener el estimador de estado óptimo, $\hat{x}_{t|t}$, en el sentido que minimice el error cuadrático medio, dado por

$$\hat{x}_{t|t} = \mathbb{E}\{x_t|y_{1:t}\} = \int x_t p(x_t|y_{1:t}) dx_t, \quad (2.14)$$

y, además, su medida de precisión dada por la matriz de covarianza $P_{t|t}$,

$$P_{t|t} = \mathbb{E}\{(x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})^T | y_{1:t}\} = \int (x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})^T p(x_t|y_{1:t}) dx_t. \quad (2.15)$$

Las ecuaciones (2.9) y (2.10) generalmente son imposibles de calcular en forma cerrada, por lo que la implementación de la solución conceptual requiere el almacenamiento de las funciones de densidad de probabilidades completas, lo que equivale a un conjunto de estadísticas de dimensión infinita. En algunos casos, las densidades se pueden caracterizar completamente por un conjunto de estadísticas suficientes. Aún así, la complejidad se puede incrementar en cada iteración y provocar que los elementos que se deban almacenar crezcan exponencialmente. Por esto, es común considerar aproximaciones para obtener soluciones subóptimas. Y en algunas raras excepciones, es posible resolver dichas ecuaciones en forma cerrada, como sucede en el problema resuelto por el Filtro de Kalman, el que en algunos casos llamaremos por su nombre en inglés, *Kalman Filter* (KF).

2.2 Enfoque Bayesiano para el Problema de Suavizado

El propósito del suavizado Bayesiano es calcular la distribución del estado x_t en cada instante de tiempo t dadas las mediciones hasta el tiempo N :

$$p(x_t|y_{1:N}). \quad (2.16)$$

El problema de obtener esta densidad puede ser abordado de dos formas, el primer método es conocido como *Forward-Backward Recursions* y el siguiente como *Generalised Two-Filter Formula*.

2.2.1 Forward-Backward Recursions

Teorema 2.2 Las ecuaciones recursivas hacia atrás que permiten calcular la distribución suavizada $p(x_t|y_{1:N})$ para cualquier $t < N$ son dadas por las siguientes ecuaciones de suavizado:

$$p(x_{t+1}|y_{1:t}) = \int p(x_{t+1}|x_t) p(x_t|y_{1:t}) dx_t, \quad (2.17)$$

$$p(x_t|y_{1:N}) = p(x_t|y_{1:t}) \int \left(\frac{p(x_{t+1}|x_t)p(x_{t+1}|y_{1:N})}{p(x_{t+1}|y_{1:t})} \right) dx_{t+1}, \quad (2.18)$$

donde $p(x_t|y_{1:t})$ es la distribución obtenida en el problema de filtraje en el instante de tiempo t y el término $p(x_{t+1}|y_{1:t})$ es la distribución predicha en el instante de tiempo $t + 1$.

Demostración. La ecuación (2.17) corresponde a la ecuación de Chapman-Kolmogorov y ya fue demostrada en el Teorema 2.1.

Luego, para demostrar (2.18), primero consideramos

$$\begin{aligned} p(x_t|x_{t+1}, y_{1:N}) &= p(x_t|x_{t+1}, y_{1:t}, y_{t+1:N}) \\ &= \frac{p(x_t, x_{t+1}, y_{1:t}, y_{t+1:N})}{p(x_{t+1}, y_{1:t}, y_{t+1:N})} \\ &= \frac{p(y_{t+1:N}|x_t, x_{t+1}, y_{1:t}) p(x_t, x_{t+1}, y_{1:t})}{p(x_{t+1}, y_{1:t}, y_{t+1:N})} \\ &= \frac{p(y_{t+1:N}|x_t, x_{t+1}, y_{1:t}) p(x_t|x_{t+1}, y_{1:t}) p(x_{t+1}, y_{1:t})}{p(x_{t+1}, y_{1:t}, y_{t+1:N})} \\ &= \frac{p(y_{t+1:N}|x_t, x_{t+1}, y_{1:t}) p(x_t|x_{t+1}, y_{1:t})}{p(y_{t+1:N}|x_{t+1}, y_{1:t})}, \end{aligned}$$

y debido a la propiedad de Markov de (1.1), x_t no aporta más información sobre $y_{t+1:N}$, entonces $p(y_{t+1:N}|x_t, x_{t+1}, y_{1:t}) = p(y_{t+1:N}|x_{t+1}, y_{1:t})$, con lo que se cumple

$$p(x_t|x_{t+1}, y_{1:N}) = p(x_t|x_{t+1}, y_{1:t}). \quad (2.19)$$

Usando la regla de Bayes podemos escribir

$$\begin{aligned} p(x_t|x_{t+1}, y_{1:N}) &= p(x_t|x_{t+1}, y_{1:t}) \\ &= \frac{p(x_t, x_{t+1}|y_{1:t})}{p(x_{t+1}|y_{1:t})} \\ &= \frac{p(x_{t+1}|x_t, y_{1:t})p(x_t|y_{1:t})}{p(x_{t+1}|y_{1:t})}, \end{aligned}$$

y como x_{t+1} solo depende de x_t ,

$$= \frac{p(x_{t+1}|x_t)p(x_t|y_{1:t})}{p(x_{t+1}|y_{1:t})}. \quad (2.20)$$

Por otro lado, la densidad conjunta del estado actual y el siguiente, dadas todas las mediciones de la salida $y_{1:N}$ puede ser calculada como,

$$p(x_t, x_{t+1}|y_{1:N}) = p(x_t|x_{t+1}, y_{1:N})p(x_{t+1}|y_{1:N})$$

2.2. Enfoque Bayesiano para el Problema de Suavizado

y usando (2.20) en el primer factor

$$p(x_t, x_{t+1}|y_{1:N}) = \frac{p(x_{t+1}|x_t)p(x_t|y_{1:t})p(x_{t+1}|y_{1:N})}{p(x_{t+1}|y_{1:t})},$$

donde $p(x_{t+1}|y_{1:N})$ es la distribución suavizada en el tiempo $t + 1$. Luego, integrando sobre x_{t+1} se obtiene la distribución marginal de x_t dado $y_{1:N}$, lo que corresponde a (2.18). \square

Con las ecuaciones (2.17) y (2.18) sucede lo mismo que con las ecuaciones de filtraje. En general no son posibles de calcular en forma cerrada, exceptuando el caso muy puntual del suavizador RTS (por sus autores Raunch-Tung-Striebel) [45], donde el sistema es lineal y los ruidos son Gaussianos, es decir, es la solución para el problema de suavización correspondiente al mismo sistema que resuelve el Filtro de Kalman, por esto, a veces es llamado Suavizador de Kalman, nombre que será utilizado en este documento o en su defecto, *Kalman Smoother*.

2.2.2 Generalised Two-Filter Formula

La *two-filter formula* es una alternativa a la técnica de suavizado mediante recursión hacia atrás en (2.17)-(2.18). La principal dificultad que se presenta en (2.18) es el término $p(x_{t+1}|y_{1:t})$ que aparece en el denominador, el que en muchas implementaciones donde se trabajan densidades no Gaussianas, no permite obtener una expresión para $p(x_t|y_{1:N})$. Por lo tanto, es necesaria una fórmula que no contenga una división explícita por una densidad no Gaussiana.

Teorema 2.3 La distribución suavizada $p(x_t|y_{1:N})$ para $t < N$ se obtiene de

$$p(x_t|y_{1:N}) = \frac{p(x_t|y_{1:t-1})p(y_{t:N}|x_t)}{p(y_{t:N}|y_{1:t-1})} \propto p(x_t|y_{1:t-1})p(y_{t:N}|x_t), \quad (2.21)$$

donde $p(y_{t:N}|x_t)$ se puede calcular recursivamente a través de

$$p(y_{t+1:N}|x_t) = \int p(y_{t+1:N}|x_{t+1})p(x_{t+1}|x_t)dx_{t+1} \quad (2.22)$$

$$p(y_{t:N}|x_t) = p(y_t|x_t)p(y_{t+1:N}|x_t), \quad (2.23)$$

que es inicializada en $p(y_N|x_N)$.

Demostración. La identidad en (2.21) viene de forma directa de aplicar el teorema de Bayes como sigue

$$\begin{aligned} p(x_t|y_{1:N}) &= \frac{p(x_t, y_{t:N}|y_{1:t-1})}{p(y_{t:N}|y_{1:t-1})} \\ &= \frac{p(x_t|y_{1:t-1})p(y_{t:N}|x_t, y_{1:t-1})}{p(y_{t:N}|y_{1:t-1})} \\ &= \frac{p(x_t|y_{1:t-1})p(y_{t:N}|x_t)}{p(y_{t:N}|y_{1:t-1})} \end{aligned}$$

donde la última igualdad viene de que $y_{t:N}$ no depende de $y_{1:t-1}$ cuando se dispone del valor del estado x_t . Nótese además que $p(y_{t:N}|y_{1:t-1})$ es una constante de normalización, por lo que se cumple la proporcionalidad en (2.21).

Luego, se tiene

$$\begin{aligned}
 p(y_{t+1:N}|x_t) &= \int p(y_{t+1:N}, x_{t+1}|x_t) dx_{t+1} \\
 &= \int p(y_{t+1:N}|x_{t+1}, x_t)p(x_{t+1}|x_t) dx_{t+1} \\
 &= \int p(y_{t+1:N}|x_{t+1})p(x_{t+1}|x_t) dx_{t+1}
 \end{aligned}$$

donde la última igualdad se debe a que las mediciones $y_{t+1:N}$ no dependen de estados pasados, cuando se dispone del estado x_{t+1} . Por último, (2.23) es directo de separar la medición en el instante t de las mediciones futuras y que éstas últimas, no dependen de la anterior. \square

Es importante destacar que $p(y_{t:N}|x_t)$ obtenida en esta implementación no es una densidad de probabilidad en x_t , por lo que incluso es posible que $\int p(y_{t:N}|x_t)dx_t$ diverja. Sin embargo, esto no es un problema cuando se puede obtener $p(y_{t:N}|x_t)$ en forma exacta, que no es el caso abordado en el siguiente capítulo.

3

Filtraje y Suavizado en Sistemas Lineal con Salida Cuantizada

3.1 Filtraje

3.1.1 Filtro de Partículas

La idea fundamental en este tipo de métodos es evaluar las integrales por un enfoque aleatorizado que emplea la ley fuerte de los grandes números (SLLN, por su nombre inextenso en inglés *Strong Law of Large Numbers*), englobados en los llamados métodos de Monte Carlo, los que son capaces de estimar la densidad involucrada como se explica a continuación.

Si es posible construir un generador de números aleatorios que entregue realizaciones (no correlacionadas) $\{x_t^i, i = 1, \dots, M\}$ con respecto a una densidad de probabilidad objetivo dada $p(x_t)$, una estimación de esta distribución es dada por

$$p_M(x_t) = \frac{1}{M} \sum_{i=1}^M \delta(x_t - x_t^i) \quad (3.1)$$

donde $\delta(x_t - x_t^i)$ denota un delta de Dirac en x_t^i . La Figura 3.1 muestra cómo un conjunto de realizaciones representa una distribución Gaussiana bi-dimensional, idea fundamental de los métodos de Monte Carlo.

En el contexto de filtraje o estimación Bayesiana, la densidad $p(x_t)$ es la densidad *a posteriori* $p(x_{1:t}|y_{1:t})$ (o su densidad marginal $p(x_t|y_{1:t})$), donde el principal objetivo es estimar recursivamente la densidad $p(x_{1:t}|y_{1:t})$ y sus esperanzas

$$I(g) = \mathbb{E} \{g(x_{1:t})\} = \int g(x_{1:t})p(x_{1:t}|y_{1:t}) dx_t \quad (3.2)$$

para alguna función de interés $g : \mathcal{X} \rightarrow \mathbb{R}^{n_g}$ integrable con respecto a $p(x_{1:t}|y_{1:t})$.

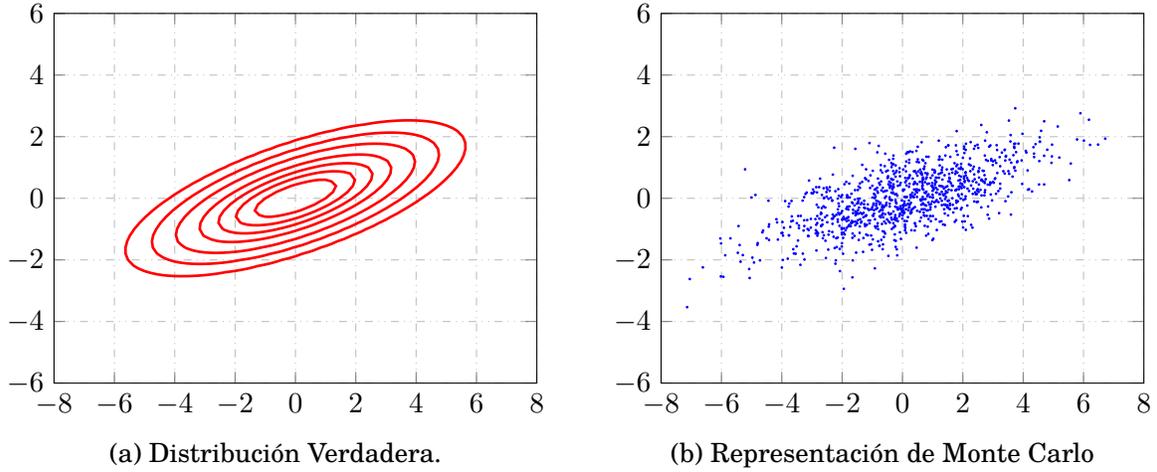


Figura 3.1: Distribución Gaussiana bi-dimensional y su representación de Monte Carlo

Usando (3.1) es sencillo estimar $I(g)$ como

$$I_M(g) = \int g(x_{1:t}) p_M(x_{1:t}|y_{1:t}) dx_t = \frac{1}{M} \sum_{i=1}^M g(x_{1:t}^i). \quad (3.3)$$

Esta estimación es no sesgada y, si la varianza de $g(x_{1:t})$ satisface $\sigma_g^2 \triangleq \mathbb{E}\{g^2(x_{1:t})\} - I^2(g) < +\infty$, entonces la varianza de $I_M(g)$ es igual a $\text{var}(I_M(g)) = \frac{\sigma_g^2}{M}$. Claramente, a partir de la ley de los grandes números,

$$I_M(g) \xrightarrow[N \rightarrow \infty]{a.s.} I(g), \quad (3.4)$$

donde $\xrightarrow{a.s.}$ denota convergencia casi segura. Más aún, si $\sigma_g^2 < +\infty$, el Teorema del Límite Central indica que

$$\sqrt{M} [I_M(g) - I(g)] \xrightarrow[M \rightarrow \infty]{d} \mathcal{N}(0, \sigma_g^2) \quad (3.5)$$

donde \xrightarrow{d} denota convergencia en distribución. La ventaja de este método de Monte Carlo perfecto es clara. Desde un conjunto de muestras aleatorias $\{x_{1:t}^i, i = 1, \dots, M\}$, uno puede estimar fácilmente cualquier cantidad $I(g)$ y la velocidad de convergencia de esta estimación es independiente de la dimensión del integrando. En contraste, cualquier método de integración numérica determinística tiene una velocidad de convergencia que decrece con la dimensión del integrando.

Desafortunadamente, es común que no sea posible obtener muestras desde la distribución *a posteriori* $p(x_{1:t}|y_{1:t})$. En estadística aplicada, son populares los métodos *Markov chain Monte Carlo* (MCMC) para muestrear desde distribuciones de probabilidad complejas [48]. Sin embargo, los métodos MCMC son algoritmos iterativos que no siempre son apropiados para un problema de estimación recursiva. Entonces, una alternativa es aplicar el método conocido como *importance sampling*.

Importance Sampling

Para desarrollar este método introducimos una densidad arbitraria $\pi(x_{1:t}|y_{1:t})$ llamada *importance density* (también conocida como *proposal density*). Asumiendo que queremos evaluar $I(g)$, y siempre que el soporte de $\pi(x_{1:t}|y_{1:t})$ incluya el soporte de $p(x_{1:t}|y_{1:t})$, esto es,

$$p(x_{1:t}|y_{1:t}) > 0 \Rightarrow \pi(x_{1:t}|y_{1:t}) > 0 \quad \forall x_{1:t} \in \mathbb{R}^n, \quad (3.6)$$

obtenemos la identidad,

$$I(g) = \frac{\int g(x_{1:t})\alpha(x_{1:t})\pi(x_{1:t}|y_{1:t}) dx_{1:t}}{\int \alpha(x_{1:t})\pi(x_{1:t}|y_{1:t}) dx_{1:t}} \quad (3.7)$$

donde $\alpha(x_{1:t})$ ¹ son conocidos como *importance weights*, y están dados por

$$\alpha(x_{1:t}) = \frac{p(x_{1:t}|y_{1:t})}{\pi(x_{1:t}|y_{1:t})}. \quad (3.8)$$

Entonces, si podemos generar M muestras i.i.d de $\pi(x_{1:t}|y_{1:t})$ también llamadas partículas $x_{1:t}^i, i = 1, \dots, M$, una estimación de Monte Carlo de $I(g)$ es

$$\hat{I}_M(g) = \frac{\frac{1}{M} \sum_{i=1}^M g(x_{1:t}^i)\alpha(x_{1:t}^i)}{\frac{1}{M} \sum_{j=1}^M \alpha(x_{1:t}^j)} = \sum_{i=1}^M g(x_{1:t}^i)\tilde{\alpha}_t^i, \quad (3.9)$$

donde $\tilde{\alpha}(x_{1:t}^i)$ son los *importance weights* normalizados, dados por

$$\tilde{\alpha}(x_{1:t}^i) = \frac{\alpha(x_{1:t}^i)}{\sum_{j=1}^M \alpha(x_{1:t}^j)}. \quad (3.10)$$

Para M finito, $\hat{I}_M(g)$ es sesgado (razón de dos estimaciones), pero, bajo suposiciones débiles, la ley fuerte de los grandes número aplica, entonces $\hat{I}_M(g) \xrightarrow[N \rightarrow \infty]{a.s.} I(g)$. Considerando suposiciones adicionales, se obtiene un Teorema de Límite Central con velocidad de convergencia independiente de la dimensión del integrando [49].

Es evidente que este método de integración puede ser interpretado como un método donde la distribución $p(x_{1:t}|y_{1:t})$ es aproximada por

$$\hat{p}_M(x_{1:t}|y_{1:t}) = \sum_{i=1}^M \tilde{\alpha}_t^i \delta(x_{1:t} - x_{1:t}^i), \quad (3.11)$$

e $\hat{I}_M(g)$ no es más que la función $g(x_{1:t})$ integrada con respecto a la densidad aproximada $\hat{p}_M(x_{1:t}|y_{1:t})$,

$$\hat{I}_M(g) = \int g(x_{1:t})\hat{p}_M(x_{1:t}|y_{1:t}) dx_{1:t}. \quad (3.12)$$

Importance Sampling es un método de integración de Monte Carlo general, sin embargo, en su forma más simple, no es adecuado para estimación recursiva, ya que se necesita disponer de todas las mediciones $y_{1:t}$ antes de estimar $p(x_{1:t}|y_{1:t})$. Luego, con cada nueva medición y_{t+1} se requiere recalculer los *importance weights* sobre la secuencia de estados completa. Por esto, el costo computacional de esta operación incrementa con el tiempo.

¹En la literatura es común denotar los *importance weights* con la letra w en lugar de α , sin embargo, acá hemos reservado la letra w para el ruido de medición en la representación en espacio de estados.

Sequential Importance Sampling

El método de *importance sampling* puede ser modificado de tal forma que sea posible calcular una estimación $\hat{p}_M(x_{1:t}|y_{1:t})$ de $p(x_{1:t}|y_{1:t})$ sin modificar las trayectorias pasadas $\{x_{1:t-1}^i, i = 1, \dots, M\}$. Esto significa que la *importance density* $\pi(x_{1:t}|y_{1:t})$ en el tiempo t admite como distribución marginal en el tiempo $t - 1$ la *importance function* $\pi(x_{1:t-1}|y_{1:t-1})$, tal que,

$$\pi(x_{1:t}|y_{1:t}) = \pi(x_t|x_{1:t-1}, y_{1:t})\pi(x_{1:t-1}|y_{1:t-1}). \quad (3.13)$$

Iterando, se obtiene

$$\pi(x_{1:t}|y_{1:t}) = \pi(x_1|y_1) \prod_{k=2}^t \pi(x_k|x_{1:k-1}, y_{1:k}). \quad (3.14)$$

Para derivar la ecuación recursiva, primero expresamos la densidad $p(x_{1:t}|y_{1:t})$ en términos de $p(x_{1:t-1}|y_{1:t-1})$, $p(y_t|x_t)$ y $p(x_t|x_{t-1})$ como sigue:

$$p(x_{1:t}|y_{1:t}) = \frac{p(y_t|x_{1:t}, y_{1:t-1})p(x_{1:t}|y_{1:t-1})}{p(y_t|y_{1:t-1})} \quad (3.15a)$$

$$= \frac{p(y_t|x_{1:t}, y_{1:t-1})p(x_t|x_{1:t-1}, y_{1:t-1})p(x_{1:t-1}|y_{1:t-1})}{p(y_t|y_{1:t-1})} \quad (3.15b)$$

$$= \frac{p(y_t|x_t)p(x_t|x_{t-1})}{p(y_t|y_{1:t-1})}p(x_{1:t-1}|y_{1:t-1}) \quad (3.15c)$$

$$\propto p(y_t|x_t)p(x_t|x_{t-1})p(x_{1:t-1}|y_{1:t-1}). \quad (3.15d)$$

Sustituyendo (3.13) y (3.15d) en (3.8), la ecuación de actualización de pesos es,

$$\alpha_t^i \propto \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)p(x_{1:t-1}^i|y_{1:t-1})}{\pi(x_t^i|x_{1:t-1}^i, y_{1:t})\pi(x_{1:t-1}^i|y_{1:t-1})} \quad (3.16a)$$

$$= \alpha_{t-1}^i \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{\pi(x_t^i|x_{1:t-1}^i, y_{1:t})}. \quad (3.16b)$$

Note que de ahora en adelante hemos dejado la dependencia de los pesos sobre $x_{1:t}^i$ utilizada en (3.8) y entonces $\alpha_t^i = \alpha(x_{1:t}^i)$. Del mismo modo $\tilde{\alpha}_t^i = \tilde{\alpha}(x_{1:t}^i)$.

Lema 3.1 La *importance density* que minimiza la varianza de los *importance weights* α_t^i condicionados a $x_{1:t-1}^i$ e $y_{1:t}$ es,

$$\pi^{opt}(x_t|x_{1:t-1}^i, y_{1:t}) = p(x_t|x_{t-1}^i, y_t). \quad (3.17)$$

Demostración. Primero, la esperanza de los *importance weights* condicionada a $x_{1:t-1}^i$ e $y_{1:t}$ es

$$\begin{aligned} \mathbb{E} \{ \alpha_t^i | x_{1:t-1}^i, y_{1:t} \} &= \int \alpha_{t-1}^i \frac{p(y_t|x_t)p(x_t|x_{t-1}^i)}{\pi(x_t|x_{1:t-1}^i, y_{1:t})} \pi(x_t|x_{1:t-1}^i, y_{1:t}) dx_t \\ &= \alpha_{t-1}^i \int p(y_t, x_t|x_{t-1}^i) dx_t \\ &= \alpha_{t-1}^i p(y_t|x_{t-1}^i) \end{aligned}$$

3.1. Filtraje

Luego, de la definición de varianza,

$$\begin{aligned}
& \mathbb{E} \left\{ \left(\alpha_t^i - \mathbb{E} \left\{ \alpha_t^i | x_{1:t-1}^i, y_{1:t} \right\} \right)^2 \middle| x_{1:t-1}^i, y_{1:t} \right\} \\
&= \mathbb{E} \left\{ \left(\alpha_{t-1}^i \frac{p(y_t | x_t) p(x_t | x_{t-1}^i)}{\pi(x_t | x_{1:t-1}^i, y_{1:t})} - \alpha_{t-1}^i p(y_t | x_{t-1}^i) \right)^2 \middle| x_{1:t-1}^i, y_{1:t} \right\} \\
&= (\alpha_{t-1}^i)^2 \left(\int \frac{p^2(y_t, x_t | x_{t-1}^i)}{\pi(x_t | x_{1:t-1}^i, y_{1:t})} dx_t - 2 \int p(y_t | x_t) p(x_t | x_{t-1}^i) p(y_t | x_{t-1}^i) dx_t + p^2(y_t | x_{t-1}^i) \right) \\
&= (\alpha_{t-1}^i)^2 \left(\int \frac{p^2(y_t, x_t | x_{t-1}^i)}{\pi(x_t | x_{1:t-1}^i, y_{1:t})} dx_t - p^2(y_t | x_{t-1}^i) \right)
\end{aligned}$$

lo que claramente se convierte en cero si $\pi(x_t | x_{1:t-1}^i, y_{1:t}) = \frac{p(y_t, x_t | x_{t-1}^i)}{p(y_t | x_{t-1}^i)}$, que es equivalente a $\pi(x_t | x_{1:t-1}^i, y_{1:t}) = p(x_t | x_{t-1}^i, y_t)$. \square

Note que no siempre es posible obtener muestras desde la *importance density* óptima, por lo que en esos casos se emplea una aproximación de $p(x_t | x_{t-1}^i, y_t)$ para $\pi(x_t | x_{1:t-1}^i, y_{1:t})$.

Además, si $\pi(x_t | x_{1:t-1}^i, y_{1:t}) = \pi(x_t | x_{t-1}^i, y_t)$, la *importance density* solo depende de x_{t-1} e y_t . Este caso particular es útil cuando sólo se busca alguna estadística del problema de filtraje, es decir, sobre $p(x_t | y_{1:t})$. En tal caso, sólo se necesita almacenar x_t^i , y se puede descartar tanto la trayectoria $x_{1:t-1}^i$ como el historial de observaciones $y_{1:t-1}$. Los ecuación de actualización de pesos queda dada por

$$\alpha_t^i \propto \alpha_{t-1}^i \frac{p(y_t | x_t^i) p(x_t^i | x_{t-1}^i)}{\pi(x_t^i | x_{t-1}^i, y_t)}, \quad (3.18)$$

y la densidad $p(x_t | y_{1:t})$ puede ser aproximada como

$$p(x_t | y_{1:t}) \approx \hat{p}_M(x_t | y_{1:t}) = \sum_{t=1}^M \tilde{\alpha}_t^i \delta(x_t - x_t^i). \quad (3.19)$$

Algoritmo 3.1 (Sequential Importance Sampling)

- (1) Inicialice las partículas, $\{x_0^i\}_{i=1}^M \sim p(x_0)$ y fije $t = 1$.
- (2) Obtenga muestras $x_t^i \sim \pi(x_t|x_{t-1}^i, y_t)$, $i = 1, \dots, M$.
- (3) Evalúe los *importance weights* de acuerdo a

$$\alpha_t^i \propto \alpha_{t-1}^i \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{\pi(x_t^i|x_{t-1}^i, y_t)}. \quad (3.20)$$

- (4) Normalice los pesos de la forma

$$\tilde{\alpha}_t^i = \frac{\alpha_t^i}{\sum_{j=1}^M \alpha_t^j}. \quad (3.21)$$

- (5) Si $t < N$ incremente $t \rightarrow t + 1$ y retorne al paso 2, de otro modo termine.
-

Un problema que ha sido demostrado [50] en el método SIS es que la varianza de los pesos solo puede incrementar sobre el tiempo, lo que tiene un efecto perjudicial en la precisión y conduce al problema conocido como, *fenómeno de degeneración*. En términos prácticos, significa que después de una cierta cantidad de pasos de la recursión, todos los pesos serán despreciables, excepto uno.

Sequential Importance Resampling

Para evitar el fenómeno de degeneración, se requiere introducir un paso adicional de selección de partículas. La idea clave del método llamado *resampling* es eliminar las partículas que tienen bajos pesos \tilde{w}_t^i y multiplicar las partículas que tienen pesos altos [51]. Más formalmente, el método reemplaza la distribución empírica ponderada $\hat{p}_M(x_t|y_{1:t}) = \sum_{i=1}^M \tilde{\alpha}_t^i \delta(x_t - x_t^i)$ por una medida no ponderada

$$p_M(x_t|y_{1:t}) = \frac{1}{M} \sum_{i=1}^M M_t^i \delta(x_t - x_t^i) \quad (3.22)$$

donde M_t^i es el número de descendientes asociados a la partícula x_t^i y es un número entero tal que $\sum_{i=1}^M M_t^i = M$. Si $M_t^i = 0$, entonces la partícula x_t^i muere. El valor M_t^i es escogido tal que $p_M(x_t|y_{1:t})$ sea cercana a $\hat{p}_M(x_t|y_{1:t})$ en el sentido que, para cualquier función g ,

$$\int g(x_t) p_M(x_t|y_{1:t}) dx_t \approx \int g(x_t) \hat{p}_M(x_t|y_{1:t}) dx_t. \quad (3.23)$$

Después del paso de selección, las partículas sobrevivientes x_t^i , que cumplen $M_t^i > 0$, son aproximadamente distribuidas de acuerdo a $p(x_t|y_{1:t})$.

Existen varias formas diferentes de seleccionar M_t^i , el más popular fue el introducido en [51]. Aquí uno obtiene las partículas sobrevivientes muestreando M veces desde la distribución

3.1. Filtraje

discreta $\hat{p}_M(x_t|y_{1:t})$, esto es equivalente a muestrear el número de descendientes M_t^i de acuerdo a la distribución multinomial de parámetros $\tilde{\alpha}_t^i$.

Al agregar el paso de *resampling* al Algoritmo 3.1 se obtiene el algoritmo genérico llamado *Sequential Importance Resampling* y popularmente conocido como Filtro de Partículas (*Particle Filter*). Hasta aquí, el algoritmo de filtraje depende de una correcta elección de la *importance density* $\pi(x_t|x_{t-1}^i, y_t)$ en (3.18). Según el Lema 3.1 la elección óptima sería $\pi(x_t|x_{t-1}^i, y_t) = p(x_t|x_{t-1}^i, y_t)$, esto es

$$p(x_t|x_{t-1}^i, y_t) = \frac{p(y_t|x_t)p(x_t|x_{t-1}^i)}{p(y_t|x_{t-1}^i)}. \quad (3.24)$$

Sustituyendo (3.24) en (3.18) se tiene

$$\alpha_t^i \propto \alpha_{t-1}^i p(y_t|x_{t-1}^i), \quad (3.25)$$

expresión que no depende de x_t^i , por lo cual, los pesos en t podrían ser calculados antes de propagar las partículas hacia el tiempo t . Para usar la *importance density* óptima, se debería disponer de muestras de $p(x_t|x_{t-1}^i, y_t)$ y evaluar

$$p(y_t|x_{t-1}^i) = \int p(y_t|x_t)p(x_t|x_{t-1}^i) dx_t, \quad (3.26)$$

para luego normalizar los pesos. En algunos casos especiales dichas tareas son posibles, como el caso Gaussiano, esto es $p(y_t|x_{t-1})$ y $p(y_t|x_t)$ densidades Gaussianas. En dicho caso es posible demostrar que $p(x_t|x_{t-1}, y_t)$ y $p(y_t|x_{t-1})$ son también densidades Gaussianas, por cual, obtener una muestra de $p(x_t|x_{t-1}, y_t)$ es sencillo. Sin embargo, en la mayoría de los casos, ambas tareas son complejas.

En particular, en el problema dado por el sistema en (1.1) se generan las densidades (2.3) y (2.4), donde una no es Gaussiana y por ende, no es posible obtener $p(x_t|x_{t-1}^i, y_t)$ en (3.24) en forma cerrada, por lo que no es posible obtener muestras al menos de forma sencilla. De la misma forma, si tuviéramos dichas muestras en el paso $t - 1$, el cálculo de (3.26) tampoco es cerrado (aunque podría ser más sencillo de aproximar). Note que la no Gaussianidad en dicho problema se produce por la cuantización en (1.1c), si ella no estuviera, se cumplirían las condiciones antes expuestas de Gaussianidad, sin embargo, dicho problema es resuelto por el Filtro de Kalman en forma cerrada y no es necesario una implementación a través de métodos de Monte Carlo.

En dichos casos como el aquí expuesto, donde no es muy factible utilizar la *importance density* óptima, utilizamos una densidad subóptima de tal forma que, sea sencillo realizar su muestreo y evaluar los pesos α_t^i . Una elección común es

$$\pi(x_t|x_{t-1}^i, y_t) = p(x_t|x_{t-1}^i), \quad (3.27)$$

que cumple de forma directa con (3.14) de la forma

$$\pi(x_{1:t}|y_{1:t}) = p(x_{1:t}) = p(x_1) \prod_{k=2}^t p(x_k|x_{k-1}). \quad (3.28)$$

Sustituyendo (3.27) en (3.18),

$$\alpha_t^i \propto \alpha_{t-1}^i p(y_t|x_t^i), \quad (3.29)$$

por lo que ya no es posible calcular los pesos antes de propagar las partículas. Note que esta implementación es apropiada para el enfoque clásico de filtraje visto en el Capítulo 2, ya que si x_{t-1} usada en (3.27) es una realización distribuida como $x_{t-1} \sim p(x_{t-1}|y_{1:t-1})$ (obtenida por ejemplo desde una iteración anterior del mismo algoritmo), entonces la densidad no condicionada es dada por

$$p(x_t|y_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1} \quad (3.30)$$

que corresponde al paso de predicción del Teorema 2.1.

Al considerar la elección en (3.27) y realizar el paso de *resampling* en cada iteración, se obtiene el Algoritmo 3.2, que es una de las muchas implementaciones conocidas como filtros de partículas, y a veces es llamado *Bootstrap Filter*.

Algoritmo 3.2 (*Basic Particle Filter or Bootstrap Filter*)

- (1) Inicializar partículas, $\{\tilde{x}_1^i\}_{i=1}^M \sim p(x_1)$ y fije $t = 1$.
- (2) Calcule los pesos $\{\tilde{\alpha}_t^i\}_{i=1}^M$,

$$\tilde{\alpha}_t^i = \frac{p(y_t|\tilde{x}_t^i)}{\sum_{j=1}^M p(y_t|\tilde{x}_t^j)}, \quad i = 1, \dots, M, \quad (3.31)$$

con $p(y_t|\tilde{x}_t^i)$ en (2.4) del Lema 2.1.

- (3) Para cada $j = 1, \dots, M$ genere una nueva partícula x_t^j con reemplazo (*resample*) de acuerdo a,

$$P(x_t^j = \tilde{x}_t^i) = \tilde{\alpha}_t^i, \quad i = 1, \dots, M. \quad (3.32)$$

- (4) Prediga las partículas seleccionando M muestras i.i.d de acuerdo a

$$\tilde{x}_{t+1}^i \sim p(x_{t+1}|x_t^i), \quad i = 1, \dots, M, \quad (3.33)$$

caracterizada por (2.3) en el Lema 2.1.

- (5) Si $t < N$ incremente $t \rightarrow t + 1$ y retorne al paso (2), de otro modo termine.
-

Note que en (3.31) no aparece el peso evaluado en $t-1$, dado que al realizar el paso de *resampling* en cada iteración, todas las partículas tienen el mismo peso. Además, hemos denotado por x_t^i las partículas después del paso de *resampling* y \tilde{x}_t^i a las partículas previas al *resampling*. Se destaca esta diferencia debido a que dicho paso toma una secuencia independiente $\{\tilde{x}_t^i\}$ y entrega una $\{x_t^i\}$ dependiente. Así, la velocidad de convergencia en (3.4) dada por la ley de los grandes números decrece cuando la correlación en $\{x_t^i\}$ crece [52], por lo cual, es preferible realizar aproximaciones del tipo (3.3) desde

$$p(x_t|y_{1:t}) \approx \hat{p}_M(x_t|y_{1:t}) = \sum_{i=1}^M \tilde{\alpha}_t^i \delta(x_t - \tilde{x}_t^i). \quad (3.34)$$

3.1. Filtraje

$i = 1, \dots, M = 10$ partículas

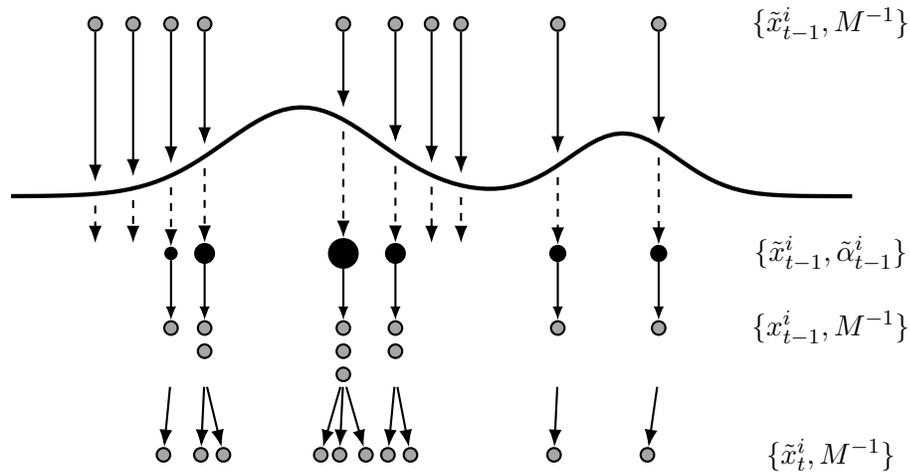


Figura 3.2: Esquema de una iteración del Algoritmo 3.2

En la Figura 3.2, se muestra un diagrama de cómo evolucionan las partículas en el Algoritmo 3.2. En el ejemplo, el algoritmo parte en $t - 1$ con una muestra no ponderada $\{\tilde{x}_{t-1}^i, M^{-1}\}$, que proporciona una aproximación de $p(x_{t-1}|y_{1:t-2})$. Para cada partícula se calcula el *importance weight* usando la información en disponible en $t - 1$. Esto resulta en las partículas ponderadas $\{\tilde{x}_{t-1}^i, \tilde{\alpha}_{t-1}^i\}$, que entregan una aproximación de $p(x_{t-1}|y_{1:t-1})$. Luego, el paso de *resampling* selecciona sólo las partículas más apropiadas para obtener una muestra no ponderada $\{x_{t-1}^i, M^{-1}\}$, que aún es una aproximación de $p(x_{t-1}|y_{1:t-1})$. Finalmente, el paso de predicción introduce una variedad que resulta en el conjunto $\{\tilde{x}_t^i, M^{-1}\}$, que es una aproximación de $p(x_t|y_{1:t-1})$.

3.1.2 Filtro por Suma de Gaussianas

En la literatura se encuentran diversos tipos de filtros por suma de Gaussianas [37, 53–55], donde, en general, la idea también es aproximar la función de densidad $p(x_t|y_{1:t})$ para cada instante t de tal forma de lograr resolver las ecuaciones (2.9) y (2.10).

La gran mayoría de filtros se aplican a sistemas no lineales donde los términos $Ax_t + Bu_t$ y $Cx_t + Du_t$ de un sistema lineal en espacio de estados estándar, se reemplazan por funciones no lineales $f(x_t, u_t)$ y $h(x_t, u_t)$, respectivamente, y tanto el ruido de proceso w_t , como el ruido de medición v_t son aditivos. Para dicho sistema, la idea es escoger una función $p(x_t|y_{1:t}) = \sum_{i=1}^m \alpha_{t|t}^i \mathcal{N}(\hat{x}_{t|t}^i, P_{t|t}^i)$ (idealmente en $p(x_1)$) y para cada Gaussiana ejecutar un Filtro de Kalman Extendido, esto es, aplicar una linealización en las funciones $f(\cdot)$ y $h(\cdot)$ para propagar las covarianzas y utilizar directamente la función no lineal para propagar las medias. Así entonces, se obtiene un filtro subóptimo con ecuaciones iterativas explícitas [55]. Las ventajas en ese tipo de filtro por suma de Gaussianas son, en primer lugar, que solo se debe escoger una aproximación por suma de Gaussianas en el instante de tiempo inicial, para luego propagar a través del tiempo los pesos, medias y covarianza. En segundo lugar, la cantidad de términos en la suma es fija. Las desventajas y principales limitaciones que se tienen para aplicar dicho filtro al sistema en (1.1) es que, en la ecuación de medición $y_t = \mathfrak{Q}\{Cx_t + Du_t + v_t\}$, el ruido no es aditivo y, más grave aún, dada la características del cuantizador, la función no es diferenciable y por ende no es posible linealizar la ecuación de salida. En este trabajo de tesis, se aborda el problema desde el punto de vista Bayesiano, por lo que intentamos resolver las ecuaciones del Teorema 2.1. Las ideas básicas presentadas en esta sección fueron publicadas en [56].

Al igual que con *Particle Filtering*, la idea básica es intentar aproximar la densidad *a posteriori*, la que normalmente no puede ser caracterizada por una cantidad finita de estadísticas. En este caso, se quiere realizar la aproximación mediante una suma de Gaussianas, basados en el siguiente Lema.

Teorema 3.1 (Wiener-Tauberian) Cualquier densidad de probabilidad de una variable aleatoria $x \in \mathbb{R}^n$, $p(x)$, con soporte compacto, puede ser aproximada tanto como se desee en el espacio $L_1(\mathbb{R}^n)$ por una distribución de la forma

$$p_A(x) = \sum_{i=1}^{K_v} \alpha_i \mathcal{N}(\hat{x}_i, P_i) \quad (3.35)$$

para algún entero K_v , escalares positivos α_i con $\sum_{i=1}^{K_v} \alpha_i = 1$, vectores n -dimensional \hat{x}_i y matrices semidefinidas positivas P_i .

Demostración. La demostración de este Teorema se encuentra en [57]. □

El Teorema anterior nos dice que podemos aproximar cualquier función de densidad mediante una suma de Gaussianas con la precisión que deseemos. Sin embargo, no nos indica como escoger los parámetros α_i , \hat{x}_i y P_i . El problema de escoger los parámetros α_i , \hat{x}_i y P_i para obtener la “mejor” aproximación p_A de alguna función de densidad p puede ser estudiado considerando la norma L^k . Entonces, la distancia entre p y p_A puede ser definida como

$$\|p - p_A\| = \int \left| p(x) - \sum_{i=1}^{K_v} \alpha_i \mathcal{N}(\hat{x}_i, P_i) \right|^k dx. \quad (3.36)$$

3.1. Filtraje

Así, uno puede intentar escoger α_i , \hat{x}_i y P_i ($i = 1, 2, \dots, K_v$) tal que la distancia $\|p - p_A\|$ sea minimizada.

Existen otras normas y formulaciones del problema que pueden ser considerados. En muchos casos, puede ser deseable que la aproximación coincida con algunos momentos, por ejemplo, la media y la varianza de la densidad verdadera. Si esto fuera un requerimiento, entonces se podrían considerar como restricciones sobre el problema de minimización.

Es claro entonces que es posible aproximar una función de densidad con una suma de Gaussianas con un número finito de términos, lo que se convierte en un problema de optimización. Sin embargo, un problema de optimización de las características de (3.36), generalmente es costoso por si mismo, por lo cual, si se quisiera utilizar en un algoritmo de filtraje, la solución no es muy práctica. Más aún, si se tuviera que resolver dicho problema en cada instante de tiempo, el problema se volvería prácticamente intratable desde el punto de vista de recursos y de ejecución en tiempo real. Por esto, sería conveniente poder realizar alguna aproximación mediante suma de Gaussianas, sin tener que resolver dicho problema de optimización.

En este punto, es útil recordar cómo es posible resolver (2.9) y (2.10) en forma cerrada en el problema que resuelve el Filtro de Kalman. Esto es posible dado que tanto $p(x_{t+1}|x_t)$, como $p(y_t|x_t)$, son densidades Gaussianas, además de la suposición en el estado inicial $p(x_1)$. En el caso de cuantización en los datos, la densidad $p(y_t|x_t)$ en (2.4) claramente es no Gaussiana producto de la presencia del cuantizador, pero afortunadamente, es la integral de una densidad Gaussiana. Así entonces, desde las ideas básicas del cálculo integral, es posible realizar una aproximación de la integral mediante una sumatoria y obtener la estructura de una suma de Gaussianas apropiadas que permita resolver (2.9). Esto se establece en el siguiente Lema.

Lema 3.2 La función de masa de probabilidad

$$p(y_t|x_t) = \int_{a_t^\ell}^{b_t^\ell} \mathcal{N}_{v_t}(Cx_t + Du_t, R) dv_t$$

puede ser aproximada usando una Suma de Riemann de K_v términos considerando la regla del punto medio como

$$p(y_t|x_t) \approx \sum_{i=1}^{K_v} \frac{b_t^\ell - a_t^\ell}{K_v} \mathcal{N}_{a_t^\ell} (Cx_t + Du_t + \mu_t^i, R) \quad (3.37)$$

donde

$$\mu_t^i = - \left(\frac{2i-1}{2} \right) \left(\frac{b_t^\ell - a_t^\ell}{K_v} \right), \quad i = 1, \dots, K_v, \quad \ell = 1, \dots, L. \quad (3.38)$$

Demostración. Al usar una Suma de Riemann de K_v se debe dividir el intervalo $[a_t^\ell, b_t^\ell]$ en K_v subintervalos de ancho $\Delta v = \frac{b_t^\ell - a_t^\ell}{K_v}$, y entonces

$$p(y_t|x_t) \approx \sum_{i=1}^{K_v} \mathcal{N}_{v_i}(Cx_t + Du_t, R) \Delta v. \quad (3.39)$$

Luego, la función es evaluada en el punto medio de cada subintervalo

$$v_i = a_t^\ell - Cx_t - Du_t + \left(\frac{2i-1}{2} \right) \left(\frac{b_t^\ell - a_t^\ell}{K_v} \right), \quad (3.40)$$

con lo que es posible escribir la aproximación de la forma en (3.37) □

Observación Las ecuaciones del Teorema que se presenta a continuación se basan en la aproximación del Lema anterior. Sin embargo, dicha aproximación se puede realizar a través de la mayoría de métodos numéricos de cálculo de integrales, como Cuadratura de Gauss-Legendre, Regla de Simpson, etc., y solo se deberían ajustar las ecuaciones. Se deja como trabajo futuro encontrar el método que obtenga los mejores resultados.

Teorema 3.2 (Filtro por Suma de Gaussianas): Sea $p(x_t|y_{1:t-1})$ una función de densidad de probabilidades *a priori*, compuesta de una mezcla normalizada de M_t Gaussianas, tal que,

$$p(x_t|y_{1:t-1}) = \sum_{j=1}^{M_t} \tilde{\alpha}_{t-1}^j \mathcal{N}_{x_t}(\hat{x}_{t|t-1}^j, P_{t|t-1}^j). \text{ Las funciones } p(y_t|x_t) \text{ y } p(x_{t+1}|x_t) \text{ están dadas en (2.4)}$$

y (2.3), respectivamente, y $p(y_t|x_t)$ es aproximada como en (3.37). Entonces, las recursiones de filtraje para el sistema en (1.1) con $\ell = 1, \dots, L$ niveles de cuantización, están dadas a continuación.

■ Corrección:

$$p(x_t|y_{1:t}) = \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \mathcal{N}_{x_t}(\hat{x}_{t|t}^{ij}, P_{t|t}^j) \quad (3.41)$$

donde

$$\hat{x}_{t|t}^{ij} = \hat{x}_{t|t-1}^j + K_t^j (a_t^\ell - C \hat{x}_{t|t-1}^j - \mu_t^i) \quad (3.42)$$

$$P_{t|t}^j = (I - K_t^j C) P_{t|t-1}^j \quad (3.43)$$

$$K_t^j = P_{t|t-1}^j C^T (R + C P_{t|t-1}^j C^T)^{-1} \quad (3.44)$$

$$\alpha_t^{ij} = \alpha_{t-1}^j \mathcal{N}_{a_t^\ell} (C \hat{x}_{t|t-1}^j + D u_t + \mu_t^i, R + C P_{t|t-1}^j C^T) \quad (3.45)$$

$$\tilde{\alpha}_t^{ij} = \frac{\alpha_t^{ij}}{\sum_{\ell=1}^m \sum_{s=1}^{M_t} \alpha_t^{\ell s}} \quad (3.46)$$

$$\mu_t^i = - \left(\frac{2i-1}{2} \right) \left(\frac{b_t^\ell - a_t^\ell}{K_v} \right) \quad (3.47)$$

■ Predicción:

$$p(x_{t+1}|y_{1:t}) = \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \mathcal{N}_{x_{t+1}}(\hat{x}_{t+1|t}^{ij}, P_{t+1|t}^j) \quad (3.48)$$

$$= \sum_{r=1}^{K_v \cdot M_t} \tilde{\alpha}_{t|t}^r \mathcal{N}_{x_{t+1}}(\hat{x}_{t+1|t}^r, P_{t+1|t}^r) \quad (3.49)$$

con $r = M_t * (i - 1) + j$ para $i = 1, \dots, K_v$ y $j = 1, \dots, M_t$, donde

$$\hat{x}_{t+1|t}^r = A \hat{x}_{t|t}^{ij} + B u_t \quad (3.50)$$

$$P_{t+1|t}^r = Q + A P_{t|t}^j A^T. \quad (3.51)$$

y fijamos $j = r$ y $M_{t+1} = K_v \cdot M_t$ para volver al paso anterior.

3.1. Filtraje

La recursión se inicializa ajustando los parámetros de la distribución *a priori*

$$p(x_1|y_0) = \sum_{j=1}^{M_0} \tilde{\alpha}_0^j \mathcal{N}_{x_t}(\hat{x}_{1|0}^j, P_{1|0}^j). \quad (3.52)$$

Demostración. Para el paso de corrección de utilizamos (2.9) sin considerar la constante de normalización, entonces,

$$\begin{aligned} p(x_t|y_{1:t}) &\propto p(y_t|x_t)p(x_t|y_{1:t-1}) \\ &= \sum_{i=1}^{K_v} \mathcal{N}_{a_t^i}(Cx_t + Du_t + \mu_t^i, R) \sum_{j=1}^{M_t} \tilde{\alpha}_{t-1}^j \mathcal{N}_{x_t}(\hat{x}_{t|t-1}^j, P_{t|t-1}^j) \\ &= \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_{t-1}^j \mathcal{N}_{a_t^i}(Cx_t + Du_t + \mu_t^i, R) \mathcal{N}_{x_t}(\hat{x}_{t|t-1}^j, P_{t|t-1}^j) \\ &= \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_{t-1}^j \mathcal{N}_{a_t^i} \left(C\hat{x}_{t|t-1}^j + Du_t + \mu_t^i, R + CP_{t|t-1}^j C^T \right) \mathcal{N}_{x_t} \left(\hat{x}_{t|t}^{ij}, P_{t|t}^j \right) \end{aligned} \quad (3.53)$$

El último paso es directo de aplicar la propiedad presentada en el Lema A.2 del Apéndice A, lo que además define (3.42), (3.43) y (3.44). La demostración de (3.41) se completa al definir los nuevos coeficientes y realizar la respectiva normalización. Para el paso de predicción empleamos (2.10), esto es

$$\begin{aligned} p(x_{t+1}|y_{1:t}) &= \int p(x_{t+1}|x_t)p(x_t|y_{1:t}) dx_t \\ &= \int \mathcal{N}_{x_{t+1}}(Cx_t + Bu_t, Q) \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \mathcal{N}_{x_t}(\hat{x}_{t|t}^{ij}, P_{t|t}^j) dx_t \\ &= \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \int \mathcal{N}_{x_{t+1}}(Cx_t + Bu_t, Q) \mathcal{N}_{x_t}(\hat{x}_{t|t}^{ij}, P_{t|t}^j) dx_t \end{aligned}$$

Luego, utilizando nuevamente la propiedad en el Lema A.2 se obtiene (3.50) y (3.51), que definen (3.48). Por último, para llevar la sumatoria doble en i y j a una sumatoria simple en r usamos el Lema A.1 presentado en el Apéndice A. \square

Teorema 3.3 (Estimador subóptimo) Con la distribución *a posteriori* en (3.41) el estimador en (2.14) y la medida de desempeño subóptimos en (2.15) están dados por

$$\hat{x}_{t|t} = \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_{t|t}^{ij} \hat{x}_{t|t}^{ij} \quad (3.54)$$

$$P_{t|t} = \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_{t|t}^{ij} \left(P_{t|t}^j + \left(\hat{x}_{t|t}^{ij} - \hat{x}_{t|t} \right) \left(\hat{x}_{t|t}^{ij} - \hat{x}_{t|t} \right)^T \right). \quad (3.55)$$

Demostración. La demostración del estimador se obtiene directamente al utilizar (2.14) junto

a (3.41)

$$\begin{aligned}
 \hat{x}_{t|t} &= \mathbb{E} \{x_t | y_{1:t}\} = \int x_t p(x_t | y_{1:t}) dx_t \\
 &= \int x_t \left(\sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \mathcal{N}_{x_t} \left(\hat{x}_{t|t}^{ij}, P_{t|t}^j \right) \right) dx_t \\
 &= \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \left(\int x_t \mathcal{N}_{x_t} \left(\hat{x}_{t|t}^{ij}, P_{t|t}^j \right) dx_t \right) \\
 &= \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \hat{x}_{t|t}^{ij}.
 \end{aligned}$$

Luego, para la medida de desempeño subóptima utilizamos (2.15),

$$\begin{aligned}
 P_{t|t} &= \mathbb{E} \{ (x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})^T | y_{1:t} \} = \int (x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})^T p(x_t | y_{1:t}) dx_t \\
 &= \int (x_t - \hat{x}_{t|t})(x_t - \hat{x}_{t|t})^T \left(\sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \mathcal{N}_{x_t} \left(\hat{x}_{t|t}^{ij}, P_{t|t}^j \right) \right) dx_t \\
 &= \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \int \left(x_t x_t^T - x_t \hat{x}_{t|t}^T - \hat{x}_{t|t} x_t^T + \hat{x}_{t|t} \hat{x}_{t|t}^T \right) \mathcal{N}_{x_t} \left(\hat{x}_{t|t}^{ij}, P_{t|t}^j \right) dx_t \\
 &= \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \left(P_{t|t}^j + \hat{x}_{t|t}^{ij} (\hat{x}_{t|t}^{ij})^T - \hat{x}_{t|t}^{ij} \hat{x}_{t|t}^T - \hat{x}_{t|t} (\hat{x}_{t|t}^{ij})^T + \hat{x}_{t|t} \hat{x}_{t|t}^T \right) \\
 &= \sum_{i=1}^{K_v} \sum_{j=1}^{M_t} \tilde{\alpha}_t^{ij} \left(P_{t|t}^j + \left(\hat{x}_{t|t}^{ij} - \hat{x}_{t|t} \right) \left(\hat{x}_{t|t}^{ij} - \hat{x}_{t|t} \right)^T \right)
 \end{aligned}$$

□

Si analizamos el Teorema 3.2, es posible percatarse que en el instante de tiempo t comenzamos con una densidad $p(x_t | y_{1:t-1})$ compuesta por una suma de M_t Gaussianas y de inmediato en el paso de corrección se obtiene una densidad compuesta por $K_v \cdot M_t$ Gaussianas, lo que produce que en el paso de predicción se obtenga de igual forma una suma de $K_v \cdot M_t$ términos, la que debería ser utilizada como densidad *a priori* en el siguiente instante de tiempo. Así entonces, es claro que el número de términos crecerá exponencialmente en función del tiempo t , lo que produce un aumento de complejidad inabordable. Para resolver este problema, estudiamos técnicas de reducción de suma de Gaussianas.

Reducción de Sumas de Gaussianas

Para afrontar el problema de aumento de complejidad, existen variados enfoques y numerosos métodos de reducción de Suma de Gaussianas.

El enfoque más simple es conocido como *Poda (Pruning)*, el que consiste en descartar las componentes de la suma de Gaussianas con peso probabilístico despreciable y mantener las restantes en una densidad de suma de Gaussianas. Esto puede ser realizado adoptando cualquiera de los siguientes métodos:

3.1. Filtrado

- Retener las componentes de la suma con los pesos mayores y descartar las restantes.
- Descartar el conjunto de componentes con los pesos más pequeños tal que el total de pesos descartado no exceda un umbral ϵ .
- Retener todas las componentes con pesos mayor o igual a un umbral ϵ y descartar las componentes con pesos menores que ϵ , donde ϵ es escogido dependiendo del problema.

El enfoque anterior, aunque es simple, no es óptimo y generalmente no consigue un buen desempeño en la reducción, perdiendo características importantes de la suma de Gaussianas original. De esta forma es mejor optar por un enfoque de *unión* (*Joining*) de componentes, el que consiste en medir la similitud de dos componentes de la suma de Gaussianas y unir las en una sola, conservando en algunos casos, los dos primeros momentos.

En particular, nos enfocamos en el algoritmo de Kullback-Leibler para reducción de Gaussianas, que fue propuesto en [58], donde la idea es reducir la siguiente suma de Gaussianas

$$p(x) = \sum_{i=1}^I \alpha_i \mathcal{N}_x(\mu_i, P_i), \quad \sum_{i=1}^I \alpha_i = 1 \quad (3.56)$$

en otra suma de Gaussianas

$$g(x) = \sum_{\ell=1}^L \gamma_\ell \mathcal{N}_x(\nu_\ell, \Omega_\ell), \quad \sum_{\ell=1}^L \gamma_\ell = 1 \quad (3.57)$$

donde $1 \leq L \leq I$. El resultado de realizar la unión de dos componentes de $p(x)$ es dado por

$$\alpha_{ij} \mathcal{N}_x(\mu_{ij}, P_{ij}) = f[\alpha_i \mathcal{N}_x(\mu_i, P_i), \alpha_j \mathcal{N}_x(\mu_j, P_j)] \quad (3.58)$$

donde

$$\begin{aligned} \alpha_{ij} &= \alpha_i + \alpha_j \\ \mu_{ij} &= \frac{1}{\alpha_i + \alpha_j} \{\alpha_i \mu_i + \alpha_j \mu_j\} \\ P_{ij} &= \frac{1}{\alpha_i + \alpha_j} \left\{ \alpha_i P_i + \alpha_j P_j + \frac{\alpha_i \alpha_j}{\alpha_i + \alpha_j} (\mu_i - \mu_j) (\mu_i - \mu_j)^T \right\}. \end{aligned} \quad (3.59)$$

Unir las componentes de la forma anterior es un factor común en la mayoría de métodos de reducción de Gaussianas basados en unión de componentes. Por el contrario, la forma de escoger qué componentes unir es lo que diferencia la mayoría de los métodos. En [58] se propone unir las componentes i y j con $i \neq j$ tal que minimicen una medida de similaridad $B(i, j)$ dada por la siguiente expresión:

$$B(i, j) = \frac{1}{2} [\alpha_{ij} \log \det P_{ij} - \alpha_i \log \det P_i - \alpha_j \log \det P_j]. \quad (3.60)$$

La función $B(i, j)$ tiene las propiedades, (i) $B(i, j) = B(j, i)$, (ii) $B(i, i) = 0$, con lo que las combinaciones posibles se reducen a $0.5I(I - 1)$.

En [59] se propone un algoritmo que sigue el método de reducción de Kullback-Leibler, pero modificado de tal forma que el usuario pueda escoger la cantidad mínima y máxima de componentes que se deseen después de la reducción y un umbral λ que no debe ser excedido por la cota $B(i, j)$. Este algoritmo se muestra a continuación.

Algoritmo 3.3 Reducción de componentes de una suma de Gaussianas

- (1) Fije la cantidad mínima L_b y máxima U_b de componentes que tendrá la suma reducida, el umbral λ y los datos de la suma que será reducida $\{\alpha_i, \mu_i, P_i\}_{i=1}^I$.
 - (2) Fije $k = I$ y defina un conjunto de enteros $S \triangleq \{1, \dots, I\}$.
 - (3) Calcule las entradas de la matriz $B(i, j)$ usando (3.60) para $i \in S$ y $j \in S$ y notando que $B(i, i) = 0$ y $B(i, j) = B(j, i)$.
 - (4) Encuentre $(i^*, j^*) = \arg \min_{i, j} B(i, j)$.
 - (5) Una los componentes i^* y j^* de la suma usando (3.59) tal que la i^* -ésima componente $\{\alpha_{i^*}, \mu_{i^*}, P_{i^*}\}$ de la data es remplazada con $\{\alpha_{i^*j^*}, \mu_{i^*j^*}, P_{i^*j^*}\}$.
 - (6) Elimine j^* desde el conjunto S tal que S es reemplazado por $S \setminus j^*$.
 - (7) Ajuste $k = k - 1$.
 - (8) Re-calcule las entradas de la matriz $B(i, j)$ usando (3.60) para $i \in S$ y $j \in S$, notando que $B(i, i) = 0$ y $B(i, j) = B(j, i)$ y que sólo la i^* -ésima fila y columna han cambiado.
 - (9) Si $k > U_b$, o si $k > L_b$ y $\min_{i, j} B(i, j) < \lambda$ retorne al paso (4), de lo contrario, termine.
-

Resumen del Filtro por Suma de Gaussianas

En el algoritmo que sigue a continuación se resumen los pasos a seguir para obtener las estadísticas de filtraje a través del Filtro por Suma de Gaussianas para un sistema con datos cuantizados como el de (1.1).

Algoritmo 3.4 (Filtro por Suma de Gaussianas)

- (1) Fije el número de Gaussianas K_v para aproximar $p(y_t|x_t)$, el número máximo L_b y mínimo U_b de Gaussianas a mantener, y el umbral λ para la reducción de Gaussianas.
- (2) Fije la densidad $p(x_1|y_0)$ como una suma de Gaussianas o como una única Gaussiana, y entonces M_1 .
- (3) Fije $t = 1$.
- (4) Con la medición y_t calcule $\hat{x}_{ij,t|t}$, $P_{j,t|t}$ y $\alpha_{j,t|t}$ para $i = 1, \dots, K_v$ y $j = 1, \dots, M_t$, a través de (3.42) - (3.47) en el Teorema 3.2.
- (5) Calcule las estadísticas de interés $\hat{x}_{t|t}$ y $P_{t|t}$ a través de (3.54) y (3.55) en el Teorema 3.3.
- (6) Calcule las estadísticas predichas $\hat{x}_{r,t+1|t}$ y $P_{r,t+1|t}$ con $r = M_t(i - 1) + j$, a través de (3.50) y (3.51) en el Teorema 3.2.

3.1. Filtraje

- (7) Si el número de Gaussianas $K_v \cdot M_t$ es mayor a U_b , realice una reducción de Gaussianas a través del Algoritmo 3.3.
 - (8) Si $t < N$, incremente $t \rightarrow t + 1$ y retorne al paso (4), de otro modo termine.
-

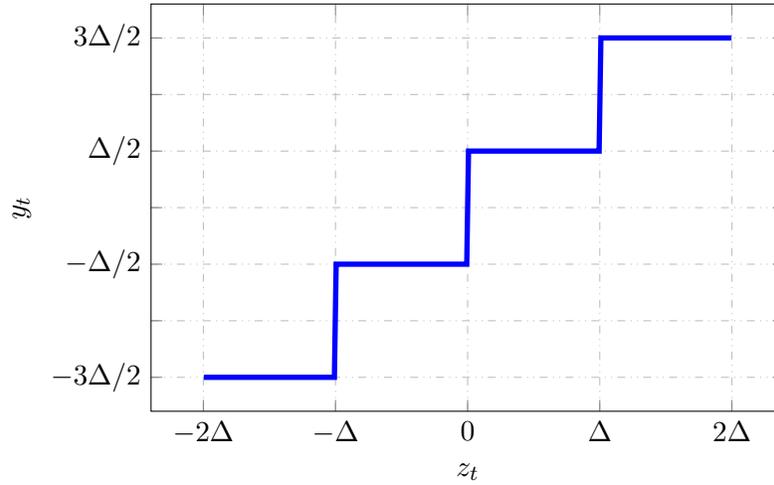


Figura 3.3: Cuantizador utilizado en los ejemplos 1 a 6.

3.1.3 Ejemplos Numéricos

Ejemplo 1

Considere el sistema

$$\begin{aligned} x_{t+1} &= 0.9x_t + 0.5u_t + w_t \\ z_t &= 2x_t + v_t \\ y_t &= \mathfrak{q}\{z_t\} \end{aligned} \quad (3.61)$$

donde $w_t \sim \mathcal{N}(0, 1)$ y $v_t \sim \mathcal{N}(0, 1)$.

Para este ejemplo numérico se simuló el sistema (3.61) usando $N = 100$, donde $t = 1, \dots, N$, $u_{1:N}$ es una secuencia de ruido Gaussiano de media 0 y varianza 0.5, y un cuantizador de la forma que se muestra en la Figura 3.3. En todos los casos descritos a continuación se ejecutó el Filtro de Kalman (KF), el Filtro por Suma de Gaussianas (GSF) del Algoritmo 3.4 y el Filtro de Partículas (PF) descrito en el Algoritmo 3.2.

En la Figura 3.4 se muestra, a mano izquierda, la estimación $\hat{x}_{t|t}$ obtenida mediante cada uno de los tres métodos descritos, junto con el estado real x_t , y a mano derecha el error de estimación. De arriba hacia abajo se repiten los resultados usando distintos valores de Δ para el cuantizador utilizado.

De los resultados se observa que para Δ pequeño, los resultados a simple vista son similares para todos los métodos. Sin embargo, las diferencias con el resultado obtenido a través del Filtro de Kalman se incrementan a medida que crece el nivel de cuantización. Lo anterior se debe a que el Filtro de Kalman no incorpora información de la cuantización, a diferencia de los dos métodos desarrollados, los que resultan en estimaciones muy similares en todas las simulaciones.

Aunque al inspeccionar el error en la Figura 3.4 resulta un poco más fácil ver las diferencias en los métodos que al ver directamente las estimaciones del estado, se prefirió buscar otra métrica que permita la comparación del desempeño de los métodos. Además, se quiere a su vez,

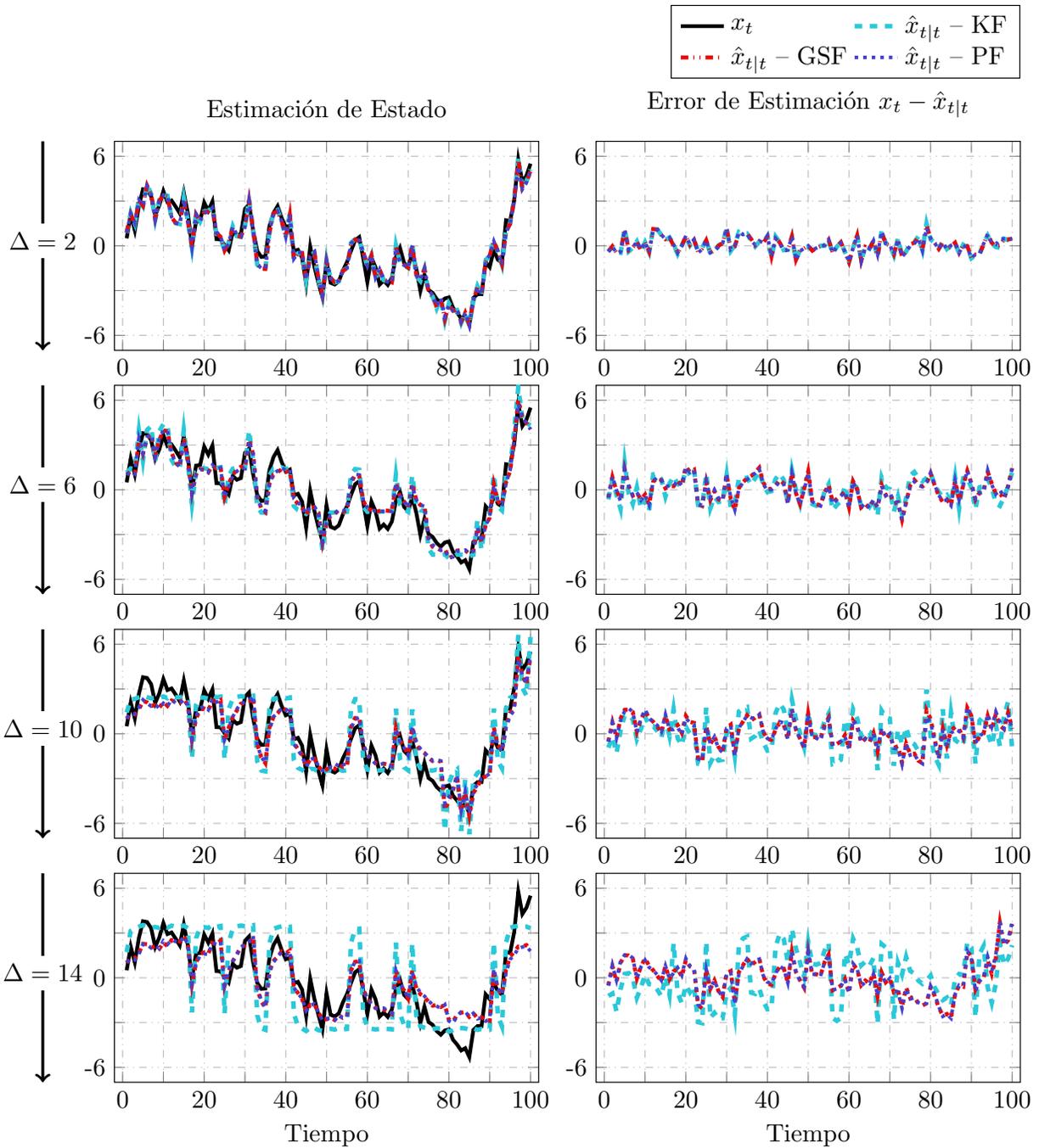


Figura 3.4: Estimación de Estado y Error de Estimación.

comparar el desempeño del Filtro por Suma de Gaussianas para distinto número de Gaussianas y el del Filtro de Partículas para distinto número de partículas.

Por esto, en la Figura 3.5 se muestra el error cuadrático medio (MSE, por la siglas de *Mean Squared Error*), calculado de la forma

$$MSE = \frac{1}{N} \sum_{t=1}^N (x_t - \hat{x}_{t|t})^2. \quad (3.62)$$

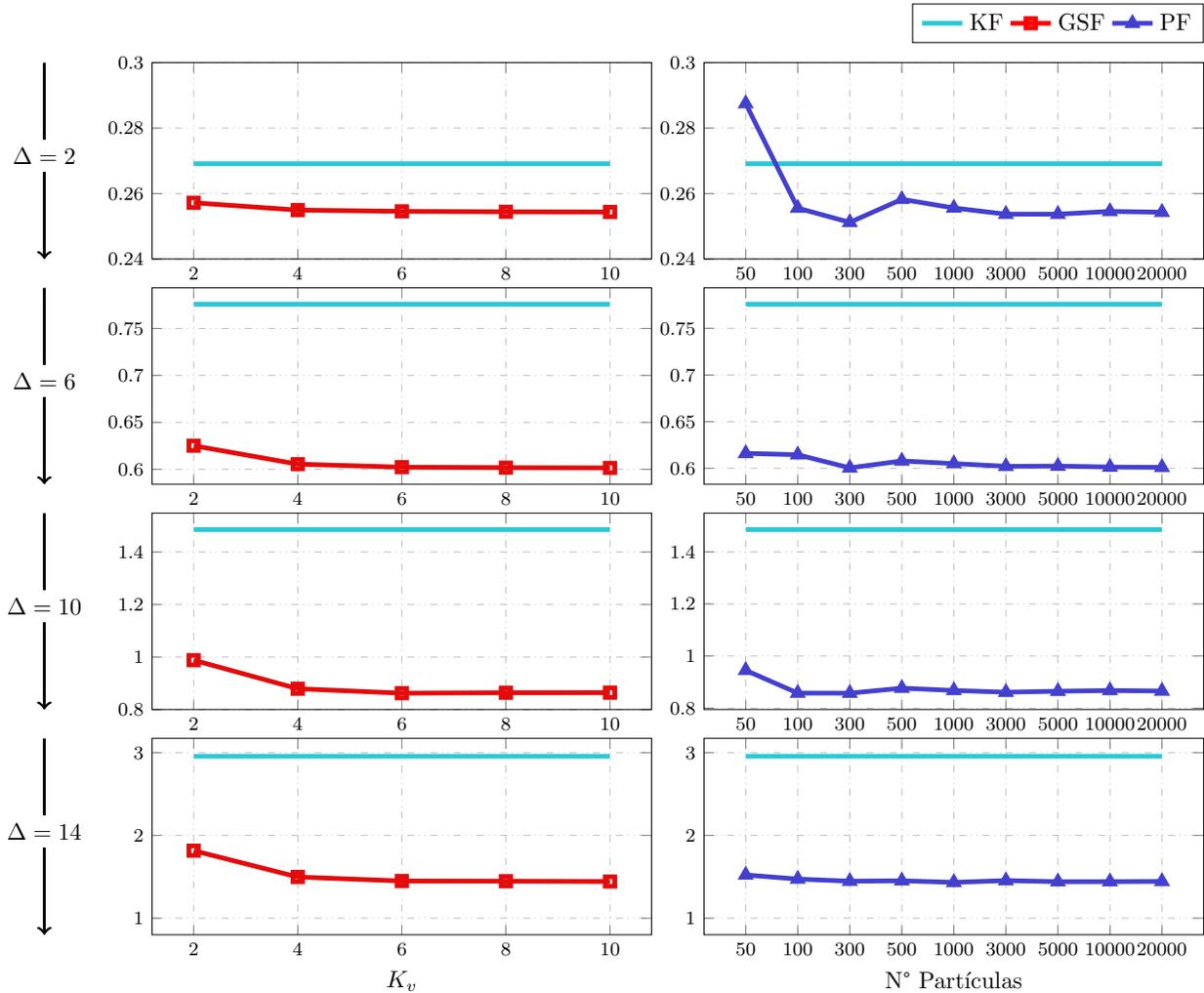
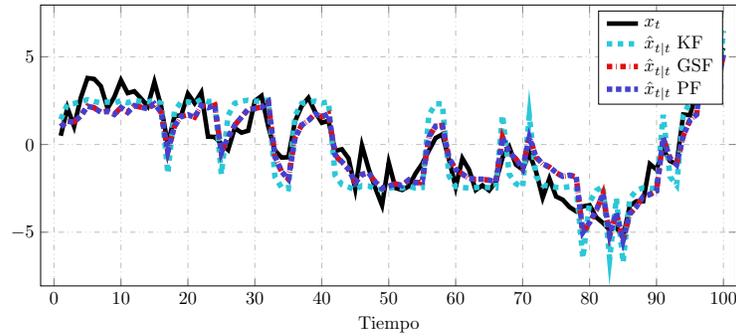
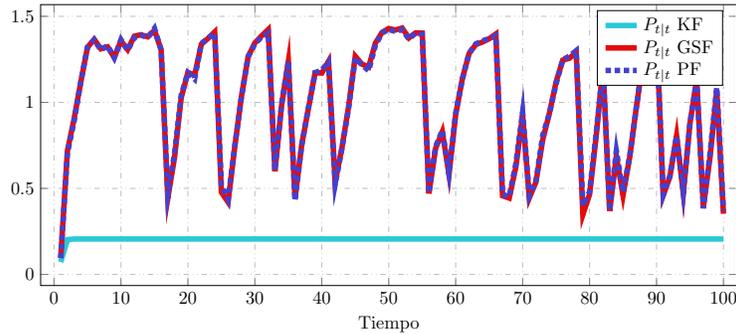


Figura 3.5: Error Cuadrático Medio entre el estado real x_t y la estimación $\hat{x}_{t|t}$ de cada uno de los métodos.

Para el caso del Filtro por Suma de Gaussianas se ha considerado $U_b = K_v$ y $L_b = 0.9U_b$. Esta elección se realizó ya que durante las simulaciones se evidenció que mantener un número de Gaussianas U_b mucho mayor que el número de Gaussianas K_v utilizadas para aproximar $p(y_t|x_t)$ no mostró beneficios.

A la izquierda de la Figura 3.5 se muestra el error cuadrático medio en función del número de Gaussianas, para el Filtro por Suma de Gaussianas y a la derecha se evidencia dicha métrica

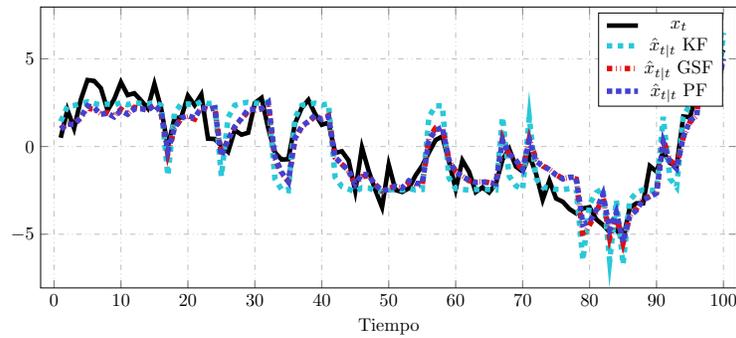
(a) Estimación de Estado $\hat{x}_{t|t}$ para cada método.(b) Covarianza $P_{t|t}$ para cada método..Figura 3.6: Estadísticas con $\Delta = 10$, 10 Gaussianas y 20000 partículas.

para el Filtro de Partículas, en función del número de partículas. Además, en todos los gráficos se incluyó el MSE obtenido por el Filtro de Kalman para efectos de comparación. Como se comentó anteriormente, dado que el Filtro de Kalman no incorpora información acerca de la cuantización, el MSE es peor que el que se obtiene con los métodos desarrollados en esta tesis y empeora a medida que la cuantización se hace más gruesa.

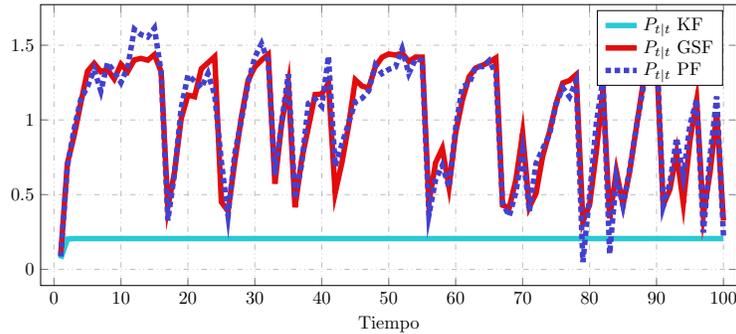
En prácticamente todas las gráficas de la Figura 3.5, se aprecia como se obtiene un mejor resultado incluyendo la información del cuantizador, es decir, usando el Filtro por Suma de Gaussianas o el Filtro de Partículas.

Luego, al comparar ambos métodos que incluyen cuantización, se puede observar que convergen al mismo resultado, estacionándose en los mismos valores de error cuadrático medio. Sin embargo, es claro que el número de Gaussianas necesarias para lograr un buen resultado, es significativamente menor que el número necesario de partículas. Sin embargo, el número de Gaussianas o partículas no es el único elemento a considerar al momento de seleccionar una técnica de filtraje, ya que el algoritmo de reducción de Gaussianas produce un gran aumento en el costo computacional.

Por último, se destaca que cuando el paso de cuantización es pequeño, prácticamente con el mínimo número de Gaussianas se logra conseguir un buen resultado. Sin embargo, cuando el nivel de cuantización es grande, se evidencia la necesidad de más Gaussianas. Por el contrario, para el Filtro de Partículas es más complejo conseguir buenos resultados cuando el paso de cuantización es pequeño, necesitando más partículas para converger que cuando el nivel de cuantización crece.



(a) Estimación de Estado $\hat{x}_{t|t}$ para cada método.



(b) Covarianza $P_{t|t}$ para cada método..

Figura 3.7: Estadísticas con $\Delta = 10$, 6 Gaussianas y 300 partículas.

Con el fin de observar las diferencias entre las estimaciones y el estado real de mejor forma, se presenta en la Figura 3.6a una de las representaciones de la Figura 3.4 en mayor tamaño. En este caso, se ha escogido un valor de $\Delta = 10$ y para asegurar la convergencia de los métodos se escogieron 10 Gaussianas y 20000 partículas. En la gráfica en cuestión se observa claramente cómo los resultados del Filtro por Suma de Gaussianas y el Filtro de Partículas son prácticamente idénticos y bastante distante de los resultados obtenidos por el Filtro de Kalman.

En la Figura 3.6b se muestra la covarianza o varianza (caso escalar) de la estimación $P_{t|t}$, donde se aprecia una notable diferencia entre el Filtro de Kalman y los métodos aquí presentados, los que además se muestran evidentemente iguales. Es importante destacar en este caso que el Filtro de Kalman estima una densidad Gaussiana, por lo que la varianza que se muestra no indica una estimación mejor, sino que una varianza errónea, al estimar una función de densidad de probabilidad errónea.

En la Figura 3.7 se muestran las mismas estadísticas que en la Figura 3.6, pero ahora se ha reducido el número de Gaussianas y de partículas. El número se ha escogido considerando donde las tendencias de la Figura 3.5 muestran el comienzo de la convergencia. Por esto, se ha escogido utilizar 6 Gaussianas por un lado, y 300 partículas por otro lado. Estas gráficas muestran cómo el resultado en la estimación de estado se mantiene prácticamente idéntico y no se observan diferencia entre los métodos presentados, lo cual respalda la convergencia mostrada en la Figura 3.5, para los parámetros escogidos. Sin embargo, la Figura 3.7b muestra que los resultados para el segundo momento presentan algunas diferencias en ciertos instantes, por lo cual, la simplificación en las aproximaciones repercuten en mayor medida en los momen-

3.1. Filtraje

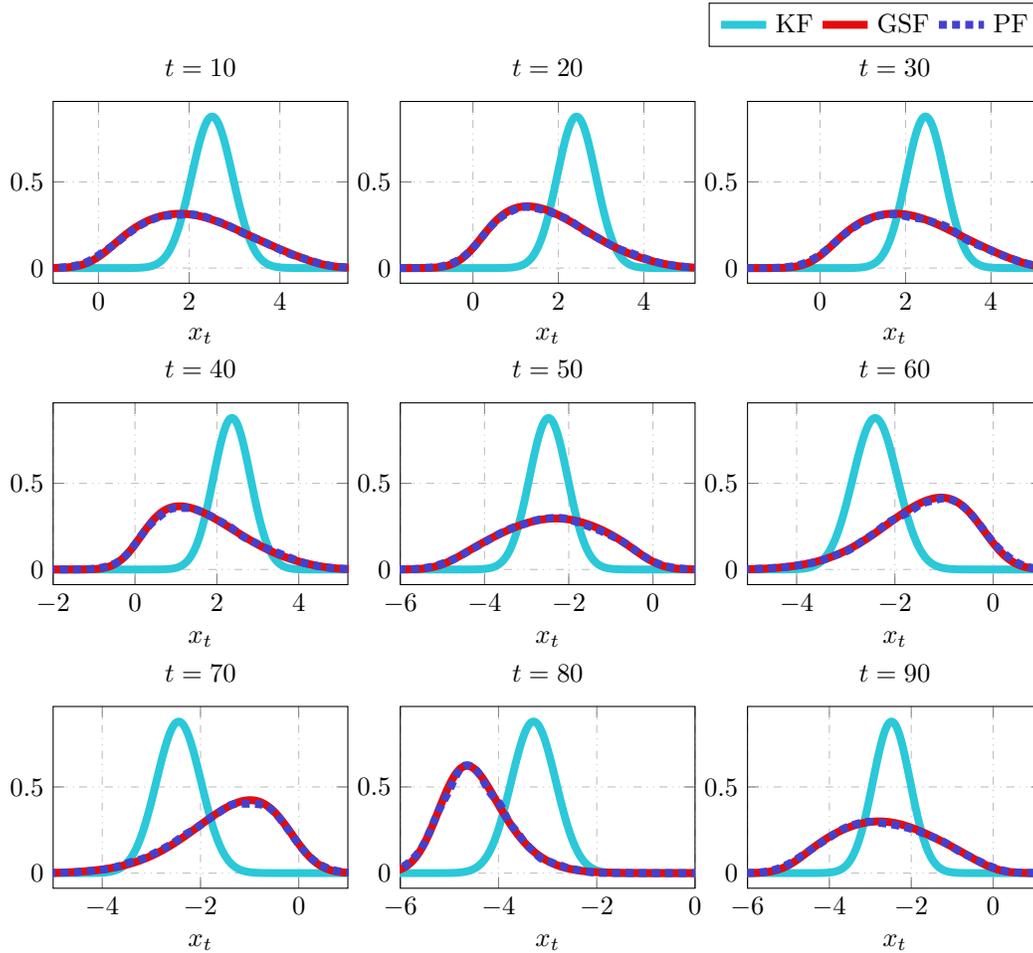


Figura 3.8: Función de densidad $p(x_t | y_{1:t})$ para distintos instantes de tiempo, usando $\Delta = 10$.

tos de orden mayor. A pesar de esto, dependiendo el objetivo buscado en el cálculo de dichos momentos, se considera un resultado bastante aceptable.

Por último, y para reforzar la idea expresada en la Figura 3.6b, donde se indicó un error en la varianza de la estimación dada por el Filtro de Kalman, en la Figura 3.8 se muestran las funciones de densidad de probabilidad obtenidas por cada uno de los métodos, para distintos instantes de tiempo. Es claro entonces que, a pesar de que el Filtro de Kalman no muestra un error exagerado en la media, la densidad estimada es completamente errónea, lo que claramente repercute en la varianza, como se observó anteriormente. Por otro lado, los métodos acá presentados evidentemente obtienen el mismo resultado.

Ejemplo 2

Considere el sistema

$$\begin{aligned} x_{t+1} &= \begin{bmatrix} 0.9 & 0 \\ 0.3 & 0.5 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u_t + w_t \\ z_t &= [1 \quad 0.1] x_t + 0.8u_t + v_t \\ y_t &= \mathfrak{Q}\{z_t\} \end{aligned} \tag{3.63}$$

donde $w_t \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}\right)$ y $v_t \sim \mathcal{N}(0, 0.1)$. Al igual que en el ejemplo anterior se considera $N = 100$, un secuencia de entrada de ruido blanco Gaussiano de media 0 y varianza 2, y el cuantizador de la Figura 3.3.

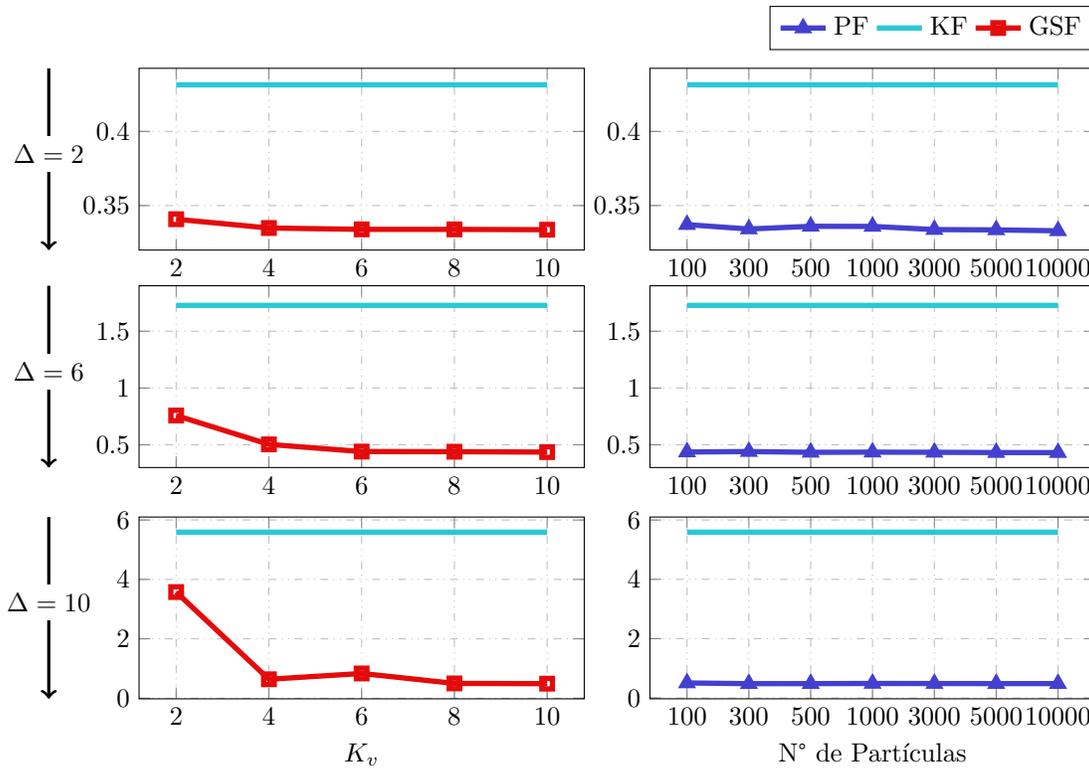


Figura 3.9: Error cuadrático medio en sistema de segundo orden.

En la Figura 3.9 se muestra la suma del error cuadrático medio de cada estados, es decir,

$$MSE = \frac{1}{N} \sum_{t=1}^N (x_t - \hat{x}_{t|t})^T (x_t - \hat{x}_{t|t}) . \tag{3.64}$$

En la Figura 3.9 se aprecia claramente cómo el error cuadrático medio de los métodos desarrollados converge y presenta una mejora sustancial en cuanto al Filtro de Kalman estándar, el cual no tiene conocimiento del nivel de cuantización. De la misma forma, la mejoría con respecto al Filtro de Kalman crece a medida que aumenta el nivel de cuantización Δ .

3.1. Filtraje

A continuación, en la Figura 3.10 se muestra la estimación de cada uno de los estados junto con el error de estimación bajo el gráfico respectivo, para tres niveles distintos de cuantización. Al observar las gráficas es evidente como los métodos que si consideran el nivel de cuantización muestran un rendimiento mejor que el Filtro de Kalman tradicional.

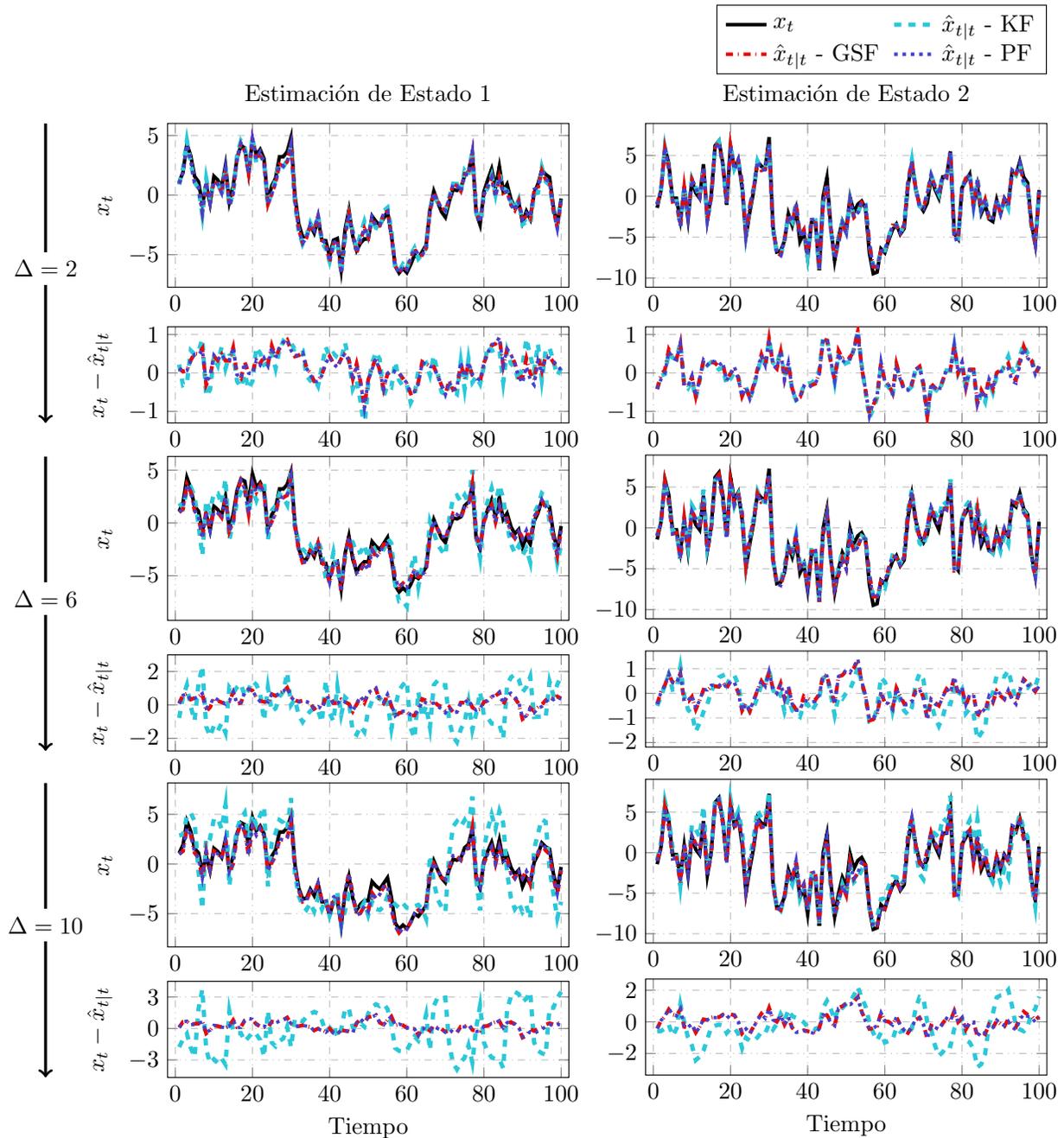


Figura 3.10: Estimación de Estado y Error de Estimación.

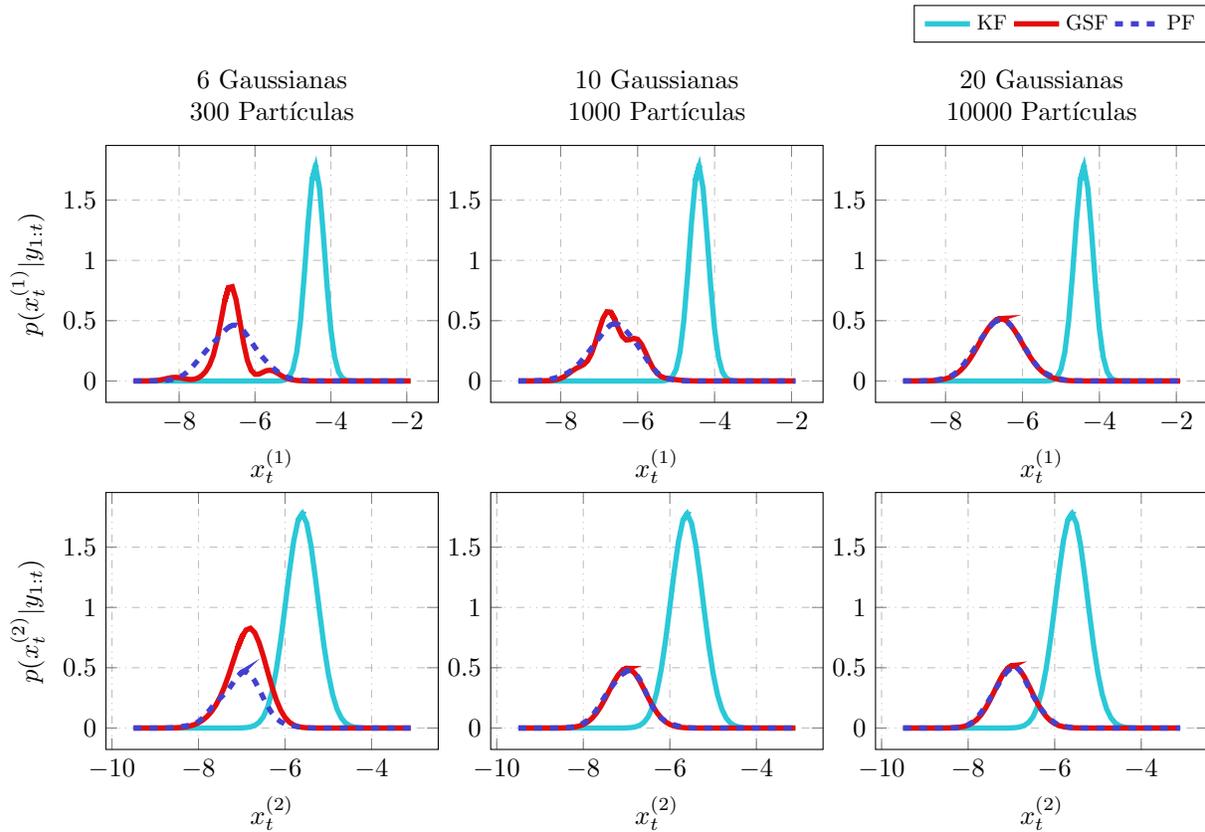


Figura 3.11: Funciones de Densidad de Probabilidad Marginales de cada uno de los estados, para distinto número de Gaussianas y partículas, utilizando $\Delta = 10$ y en el instante $t = 60$.

En la Figura 3.11 se pueden ver las funciones de densidad de probabilidades marginales en el instante de tiempo $t = 60$ para cada uno de los estados, utilizando $\Delta = 10$. En la parte superior de la figura se muestra el primer estado, mientras las figuras inferiores, se presenta el segundo estado. De izquierda a derecha, aumenta el número de Gaussianas y partículas como se indica la misma figura.

La Figura 3.11, muestra que tanto la media como la varianza de las funciones de densidad de probabilidad marginales obtenidas por el filtro de Kalman son totalmente distintas a la de los métodos desarrollados. Por otra parte, el Filtro por Suma de Gaussianas y el Filtro de Partículas, en todos los casos coinciden en media. Sin embargo, la forma de las funciones de densidad de probabilidad estimadas se diferencia bastante al considerar un número reducido de Gaussianas o de Partículas. A diferencia de lo que ocurrió en la figura 3.9, se debió llegar a 20 Gaussianas para obtener una similitud en forma con lo obtenido con el Filtro de Partículas usando 10000 elementos. Esto se debe a que el MSE sólo considera el momento de primer orden, el que no presenta mayores dificultades al ser estimado.

Para finalizar este ejemplo, en la Figura 3.12 se muestra la función de densidad de probabilidad $p(x_t|y_{1:t})$ en el instante de tiempo $t = 60$, estimada por cada uno de los métodos utilizados. Las observaciones son similares a las de la figura 3.11, aunque ahora se puede apreciar las funciones de densidad completas, incluyendo la correlación entre estados. De esta forma, es

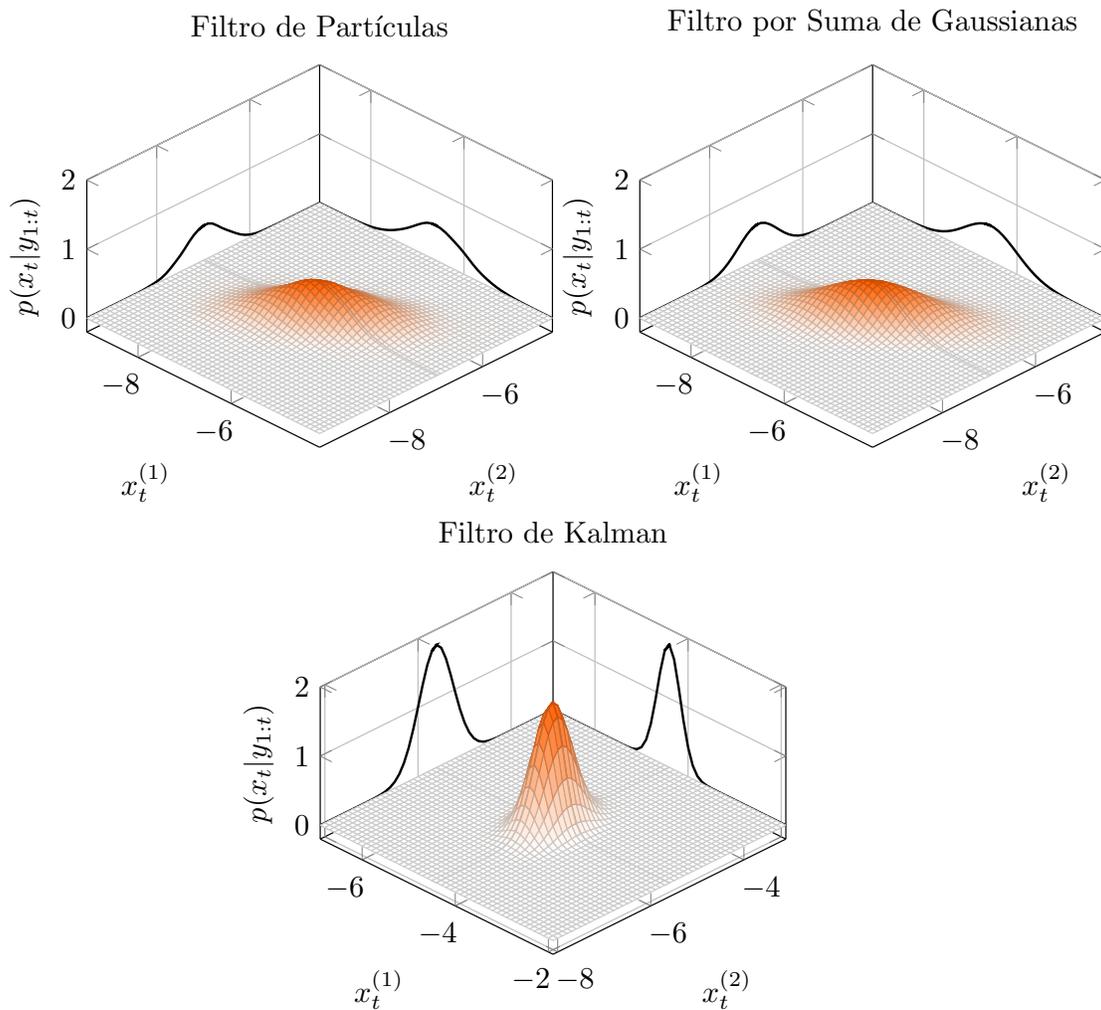


Figura 3.12: Función de densidad de probabilidad $p(x_t|y_{1:t})$ estimada por cada método, en el instante $t = 60$, con $\Delta = 10$, 20 Gaussianas y 10000 partículas.

evidente la convergencia al mismo resultado cuando el número de Gaussianas y el número de partículas crece.

3.2 Suavizado

3.2.1 Suavizador de Partículas

El mismo enfoque que se utilizó para obtener el Algoritmo 3.2 puede ser extendido para aproximar la densidad en (2.16), donde el objetivo es buscar una aproximación de la forma

$$p(x_t|y_{1:N}) \approx p_M(x_t|y_{1:N}) = \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \delta(x_t - \tilde{x}_t^i). \quad (3.65)$$

Como se explicó en la Sección 3.1.1, las partículas que aproximan (3.34) se pueden usar vía *Importance Sampling* para aproximar (2.18) de acuerdo a

$$p(x_{t+1}|y_{1:t}) \approx \sum_{i=1}^M \tilde{\alpha}_t^i p(x_{t+1}|\tilde{x}_t^i). \quad (3.66)$$

Inicialmente se destaca que para el caso particular $t = N$, la función de densidad de probabilidad suavizada $p(x_t|y_{1:N})$, y la función de densidad de probabilidad filtrada $p(x_t|y_{1:t})$ son iguales. Por lo tanto, los pesos en (3.65) se pueden inicializar con $\tilde{\alpha}_{N|N}^i = \tilde{\alpha}_N^i$. Además, como las partículas \tilde{x}_N^i son idénticas para filtrado y suavizado, inmediatamente se tiene una aproximación de $p(x_N|y_{1:N})$.

Para obtener una aproximación en todo instante de tiempo, se debe resolver recursión dada en el Teorema 2.2. Luego, como se tiene una aproximación inicial en $t = N$, se puede asumir que se tiene disponible una aproximación como en (3.65) para el tiempo $t + 1$, y junto a (3.66) aproximar la integral en (2.18) de la forma

$$\int \frac{p(x_{t+1}|x_t) p(x_{t+1}|y_{1:N})}{p(x_{t+1}|y_{1:t})} dx_{t+1} \approx \sum_{j=1}^M \frac{\tilde{\alpha}_{t+1|N}^j p(\tilde{x}_{t+1}^j|x_t)}{\sum_{\ell=1}^M \tilde{\alpha}_t^\ell p(\tilde{x}_{t+1}^j|\tilde{x}_t^\ell)}, \quad (3.67)$$

el que es un término dependiente de x_t . Para completar la aproximación de la expresión en (2.18), solo falta agregar el término $p(x_t|y_{1:t})$, que puede ser aproximado mediante (3.34). La expresión en (3.34) tiene un valor $\tilde{\alpha}_t^i$ sólo en \tilde{x}_t^i , por lo que cada partícula en \tilde{x}_t^i de la aproximación (3.65) tendrá un peso dado por

$$\tilde{\alpha}_{t|N}^i = \tilde{\alpha}_t^i \sum_{j=1}^M \frac{\tilde{\alpha}_{t+1|N}^j p(\tilde{x}_{t+1}^j|\tilde{x}_t^i)}{\sum_{\ell=1}^M \tilde{\alpha}_t^\ell p(\tilde{x}_{t+1}^j|\tilde{x}_t^\ell)}. \quad (3.68)$$

Siguiendo dicha estrategia para cada \tilde{x}_t^i , la densidad suavizada $p(x_t|y_{1:N})$ queda representada por

$$p(x_t|y_{1:N}) \approx \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \delta(x_t - \tilde{x}_t^i), \quad (3.69a)$$

$$\tilde{\alpha}_{t|N}^i = \tilde{\alpha}_t^i \sum_{j=1}^M \tilde{\alpha}_{t+1|N}^j \frac{p(\tilde{x}_{t+1}^j|\tilde{x}_t^i)}{v_t^j} \quad (3.69b)$$

$$v_t^j \triangleq \sum_{\ell=1}^M \tilde{\alpha}_t^\ell p(\tilde{x}_{t+1}^j|\tilde{x}_t^\ell). \quad (3.69c)$$

3.2. Suavizado

Las nociones anteriores se pueden resumir en el siguiente algoritmo de suavizado.

Algoritmo 3.5 (*Basic Particle Smoother*)[60]

- (1) Ejecute el filtro de partículas (Algoritmo 3.2) y almacene las partículas $\{\tilde{x}_t^i\}_{i=1}^M$ y sus pesos $\{\tilde{\alpha}_t^i\}_{i=1}^M$, para $t = 1, \dots, N$.
- (2) Inicialice los pesos suavizados comenzando por el peso filtrado final $\{\tilde{\alpha}_t^i\}$ en el tiempo $t = N$,

$$\tilde{\alpha}_{N|N}^i = \tilde{\alpha}_N^i, \quad i = 1, \dots, M \quad (3.70)$$

y fije $t = N - 1$.

- (3) Calcule los pesos suavizados $\{\tilde{\alpha}_{t|N}^i\}_{i=1}^M$ usando los pesos filtrados $\{\tilde{\alpha}_t^i\}_{i=1}^M$ y las partículas $\{\tilde{x}_t^i, \tilde{x}_{t+1}^i\}_{i=1}^M$, a través de las ecuaciones en (3.69b)-(3.69c), con la función de densidad de probabilidad (2.1) en el Lema 2.1.
- (4) Si lo requiere, obtenga las estadísticas necesarias de acuerdo a

$$\mathbb{E} \{g(x_t) | y_{1:N}\} \approx \sum_{i=1}^M \tilde{\alpha}_{t|N}^i g(\tilde{x}_t^i). \quad (3.71)$$

- (5) Si $t > 1$, decremente $t \rightarrow t - 1$ y retorne al paso (3). De otro modo termine.

En muchos casos se requiere calcular estadísticas de $p(x_{t+1}, x_t | y_{1:N})$, por lo cual, enunciamos el siguiente Lema.

Lema 3.3 Es posible aproximar $\mathbb{E} \{g(x_{t+1}, x_t) | y_{1:N}\}$ usando los resultados de los Algoritmos 3.2 y 3.5, como

$$\mathbb{E} \{g(x_{t+1}, x_t) | y_{1:N}\} \approx \sum_{i=1}^M \sum_{j=1}^M \tilde{\alpha}_{t|N}^{ij} g(\tilde{x}_{t+1}^j, \tilde{x}_t^i) \quad (3.72)$$

con

$$\tilde{\alpha}_{t|N}^{ij} \triangleq \tilde{\alpha}_t^i \tilde{\alpha}_{t+1|N}^j \frac{p(\tilde{x}_{t+1}^j | \tilde{x}_t^i)}{\sum_{\ell=1}^M w_t^\ell p(\tilde{x}_{t+1}^j | \tilde{x}_t^\ell)}. \quad (3.73)$$

Demostración. Primero, vemos que

$$p(x_{t+1}, x_t | y_{1:N}) = p(x_t | x_{t+1}, y_{1:t}) p(x_{t+1} | y_{1:N}) \quad (3.74)$$

$$= \frac{p(x_{t+1} | x_t) p(x_t | y_{1:t})}{p(x_{t+1} | y_{1:t})} p(x_{t+1} | y_{1:N}) \quad (3.75)$$

donde hemos utilizado $p(x_t|x_{t+1}, y_{1:N}) = p(x_t|x_{t+1}, y_{1:t})$ en (2.19). De esta forma, la esperanza requerida cumple

$$\mathbb{E} \{g(x_{t+1}, x_t)|y_{1:N}\} = \int \int g(x_{t+1}, x_t) \frac{p(x_{t+1}|x_t) p(x_t|y_{1:t})}{p(x_{t+1}|y_{1:t})} p(x_{t+1}|y_{1:N}) dx_t dx_{t+1} \quad (3.76a)$$

$$= \int \left(\int g(x_{t+1}, x_t) p(x_{t+1}|x_t) p(x_t|y_{1:t}) dx_t \right) \frac{p(x_{t+1}|y_{1:N})}{p(x_{t+1}|y_{1:t})} dx_{t+1}, \quad (3.76b)$$

por lo que usando (3.34) se obtiene

$$\mathbb{E} \{g(x_{t+1}, x_t)|y_{1:N}\} \approx \sum_{i=1}^M \tilde{\alpha}_t^i \int g(x_{t+1}, \tilde{x}_t^i) p(x_{t+1}|\tilde{x}_t^i) \frac{p(x_{t+1}|y_{1:N})}{p(x_{t+1}|y_{1:t})} dx_{t+1}. \quad (3.76c)$$

Introduciendo (3.65) se consigue

$$\mathbb{E} \{g(x_{t+1}, x_t)|y_{1:N}\} = \sum_{i=1}^M \sum_{j=1}^M \tilde{\alpha}_t^i \tilde{\alpha}_{t+1|N}^j g(\tilde{x}_{t+1}^j, \tilde{x}_t^i) \frac{p(\tilde{x}_{t+1}^j|\tilde{x}_t^i)}{p(\tilde{x}_{t+1}^j|y_{1:t})}. \quad (3.76d)$$

Por último, podemos aproximar el denominador en la expresión anterior de acuerdo a

$$p(\tilde{x}_{t+1}^j|y_{1:t}) = \int p(\tilde{x}_{t+1}^j|x_t) p(x_t|y_{1:t}) dx_t \quad (3.77a)$$

$$\approx \sum_{\ell=1}^M \tilde{\alpha}_t^\ell p(\tilde{x}_{t+1}^j|\tilde{x}_t^\ell). \quad (3.77b)$$

Finalmente, definiendo (3.73) se completa la prueba. \square

Observando (3.72), podemos considerar la aproximación

$$p(x_t, x_{t+1}|y_{1:N}) \approx \sum_{i=1}^M \sum_{j=1}^M \tilde{\alpha}_t^{ij} \delta \left((x_t, x_{t+1}) - (\tilde{x}_t^i, \tilde{x}_{t+1}^j) \right). \quad (3.78)$$

Esta implementación fue presentada en [50] y no es la única implementación que existe de *Particle Smoothing*. Fue escogida debido a la claridad con que se puede expresar su derivación y las aproximaciones utilizadas. Otras alternativas a este algoritmo que proporcionan muestras estocásticas que aproximan la densidad suavizada se presentan en [61–66]. En particular, [61, 62] son implementaciones a través de la *Generalized Two-Filter Formula*, presentada en el Lema 2.3, donde se define un conjunto de densidades artificiales para superar el problema de no integrabilidad de $p(y_{t:N}|x_t)$ mencionado en la Sección 2.2.2.

3.2.2 Ejemplos Numéricos

Ejemplo 3

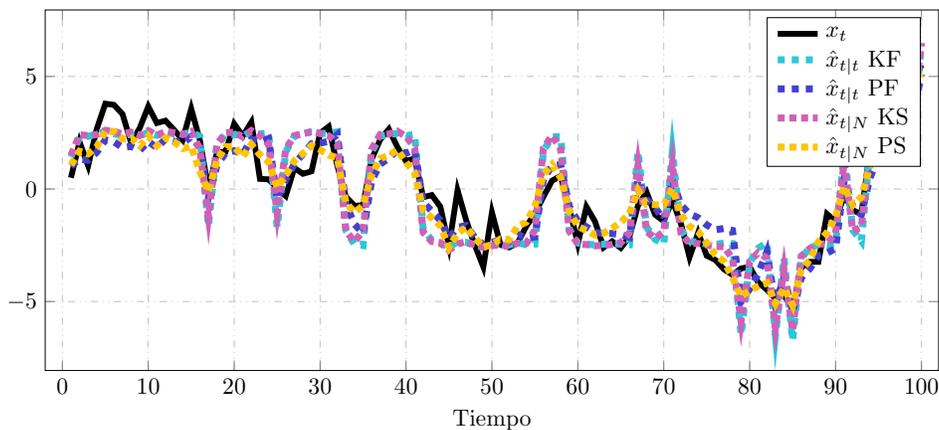
Considere el mismo sistema utilizado en el Ejemplo 1, dado a continuación:

$$\begin{aligned}x_{t+1} &= 0.9x_t + 0.5u_t + w_t \\z_t &= 2x_t + v_t \\y_t &= \mathbb{Q}\{z_t\}\end{aligned}\tag{3.79}$$

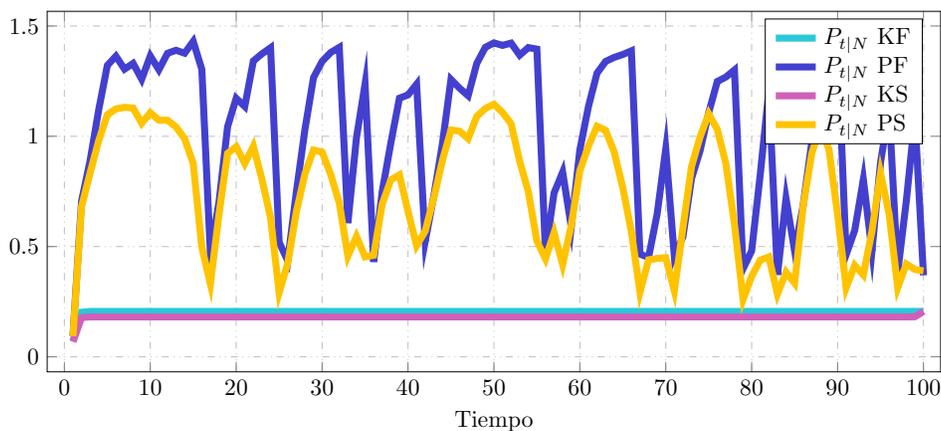
donde $w_t \sim \mathcal{N}(0, 1)$ y $v_t \sim \mathcal{N}(0, 1)$. Además, considere las mismas condiciones de simulación explicitadas en el Ejemplo 1.

A diferencia de los ejemplos de Filtraje, se considera sólo una implementación que incorpora información del proceso de cuantización, la que fue presentada en la Sección 3.2.1, y es basada en Partículas.

El actual ejemplo y el siguiente, se enfocan en mostrar las diferencias entre los resultados obtenidos a través de Filtraje y Suavizado, por lo que se elaboran gráficas incluyendo Filtro de Kalman, Filtro de Partículas, Suavizador de Kalman y Suavizador de Partículas.



(a) Estimación de Estado



(b) Varianza de la estimación

Figura 3.13: Estimación y Varianza considerando $\Delta = 10$ y 20000 partículas.

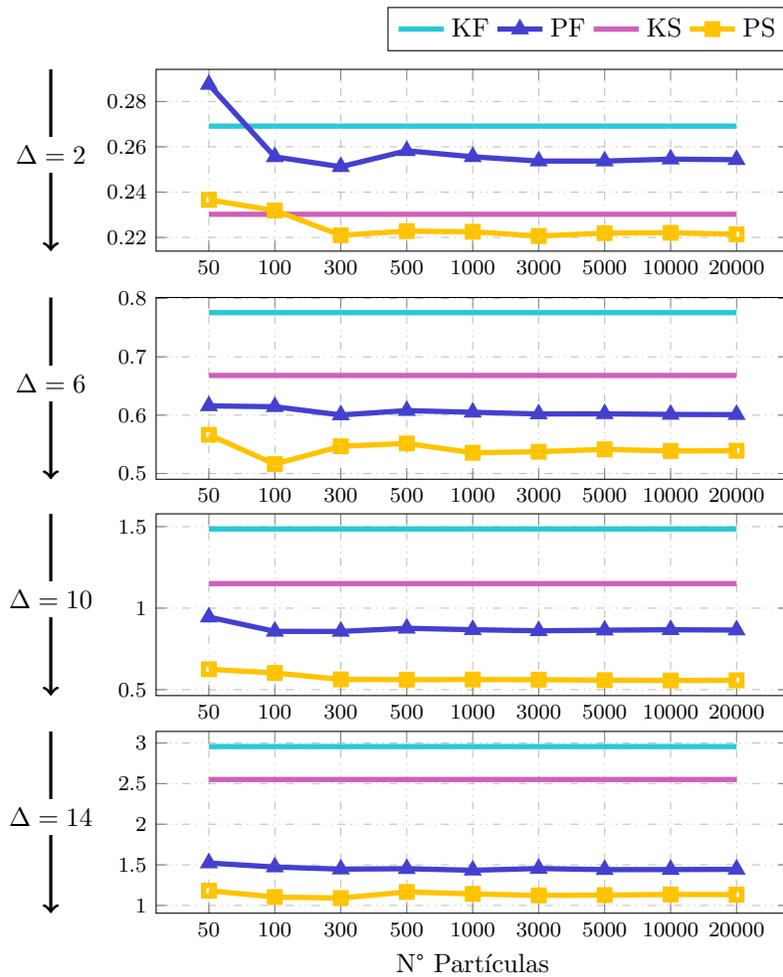


Figura 3.14: Error Cuadrático Medio incluyendo suavizado.

La Figura 3.13 muestra los resultados obtenidos usando $\Delta = 10$ y considerando un número exagerado de 20000 partículas, para asegurar convergencia.

Se aprecia en la Figura 3.13, que los resultados en cuanto a la estimación de estado es mejorada a través de los suavizadores, para cada uno de los casos. Sin embargo, la diferencia a simple vista es difícil de observar. Por el contrario, para el caso de la varianza de la estimación, la mejoría introducida a través del Suavizador de Partículas, con respecto al Filtro de Partículas, es evidente. Adicionalmente, se aprecia que el Suavizador de Kalman obtiene una varianza menor a la del respectivo filtro, no obstante, como se mencionó anteriormente, es una varianza errónea, pues estima una densidad Gaussiana que no es correcta. Se recuerda además que el Filtro y Suavizador de Kalman calculan las Covarianzas *off line*, es decir, no utilizan las mediciones y_t .

En la Figura 3.14, se muestra el error cuadrático medio en función del número de partículas y del nivel de cuantización, tal como se hizo en el Ejemplo 1. Se aprecia que el MSE en función del número de partículas tiene un comportamiento similar al del Filtro de Partículas. Sin embargo, en este caso se destaca cómo el error cuadrático medio obtenido, siempre es me-

3.2. Suavizado

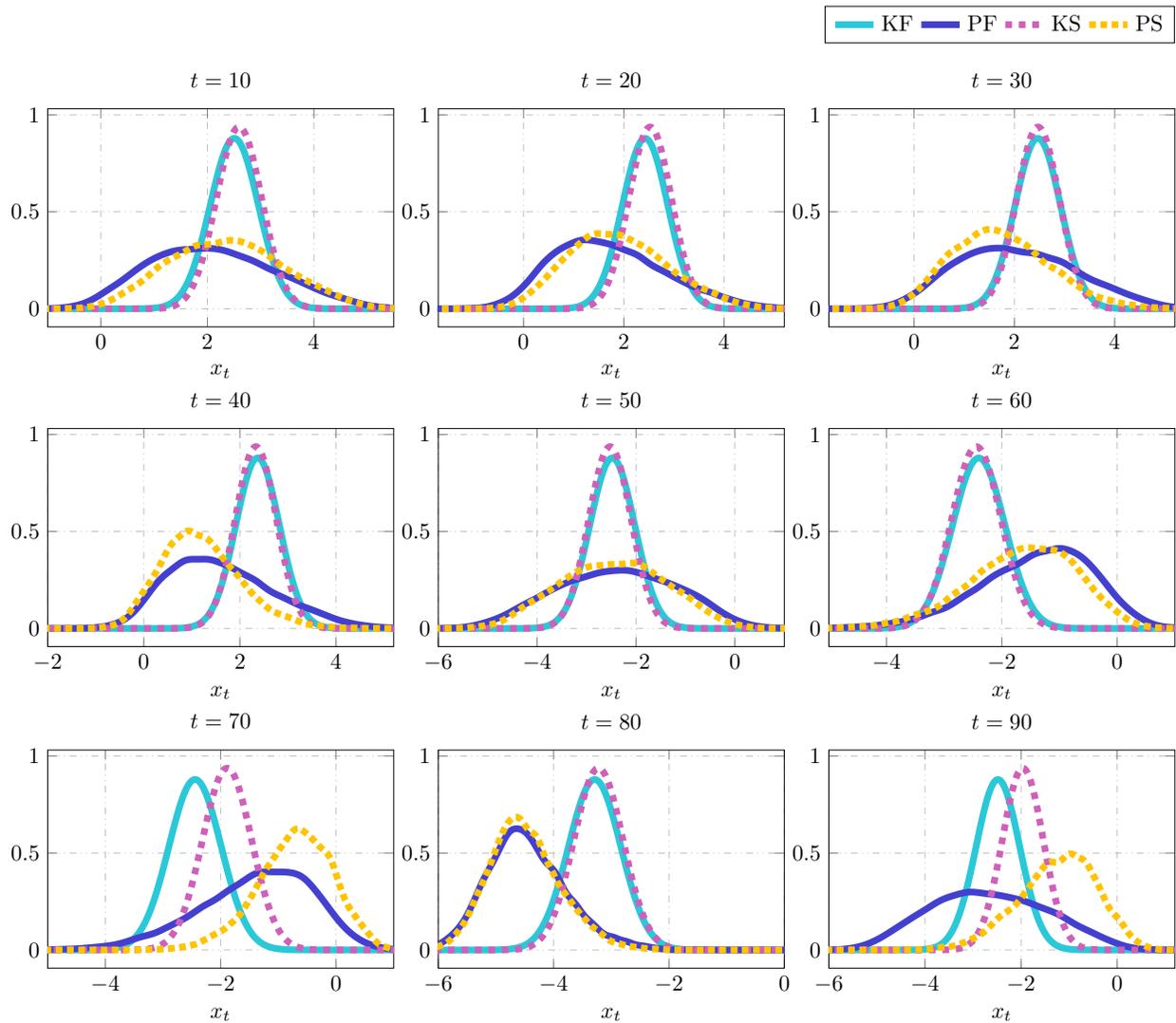


Figura 3.15: Función de densidad $p(x_t | y_{1:N})$ para distintos tiempos, con $\Delta = 10$.

nor al de su respectivo filtro, siendo el del Suavizador de Partículas el menor de todos. Esto evidencia la mejoría de considerar la cuantización en la estimación de estado, y además, como un suavizador, que dispone de mediciones futuras, mejora los resultados de un filtro.

Por último, en la Figura 3.15 se pueden observar las estimaciones de la función de densidad de probabilidad $p(x_t | y_{1:N})$ a través de cada método.

La Figura 3.15 destaca cómo la función de densidad de probabilidad cambia al considerar la etapa de cuantización, por lo que los métodos basados en partículas presentan funciones densidad evidentemente distintas a las de los métodos llamados de Kalman. Lo anterior, nuevamente justifica los resultados presentados en la Figura 3.13b. Por último, se puede ver cómo el suavizado produce correcciones importantes, con respecto al filtraje, en algunos instantes de tiempo, y en otros casos, las correcciones son despreciable.

Ejemplo 4

Nuevamente considere el sistema dado en el Ejemplo 2, expresado por:

$$\begin{aligned} x_{t+1} &= \begin{bmatrix} 0.9 & 0 \\ 0.3 & 0.5 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u_t + w_t \\ z_t &= [1 \quad 0.1] x_t + 0.8u_t + v_t \\ y_t &= \mathbb{Q}\{z_t\} \end{aligned} \tag{3.80}$$

donde $w_t \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}\right)$ y $v_t \sim \mathcal{N}(0, 0.1)$. Y además, considere las mismas condiciones de simulación que en el Ejemplo 2.

En la Figura 3.16 se muestra la estimación de ambos estados para el caso en que $\Delta = 10$, usando 10000 partículas. Es posible apreciar la ventaja de los métodos de partículas, aunque, al igual que antes, la ventaja del Suavizador no es fácil de apreciar a simple vista.

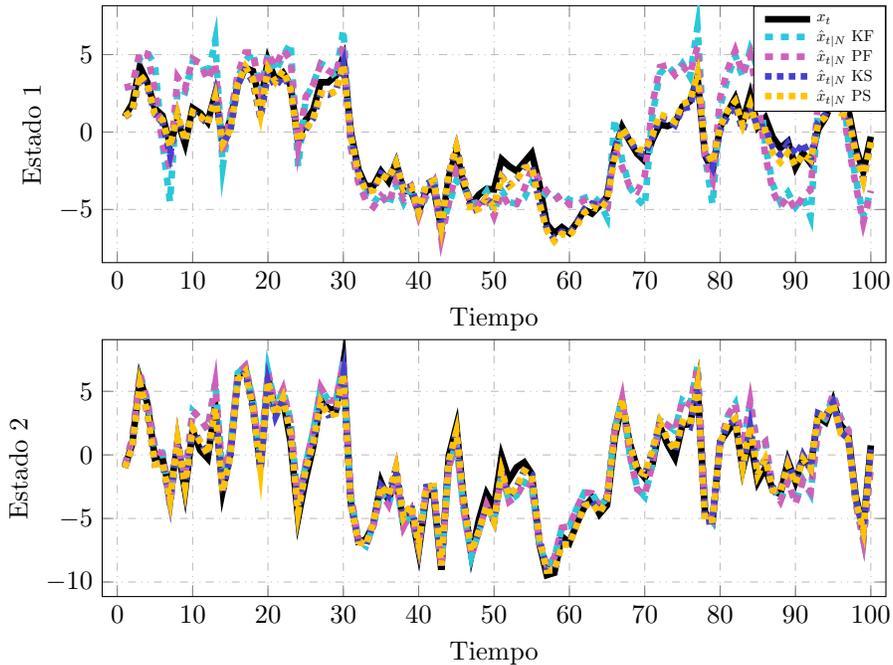


Figura 3.16: Estimación de Estados, con $\Delta = 10$ y 10000 partículas.

La Figura 3.17 presenta la varianza de la estimación $\text{trace}(P_{t|N})$, bajo las mismas condiciones recién expuestas. Nuevamente se puede ver que la mayor ventaja del suavizador corresponde a la disminución clara en la varianza de la estimación.

3.2. Suavizado

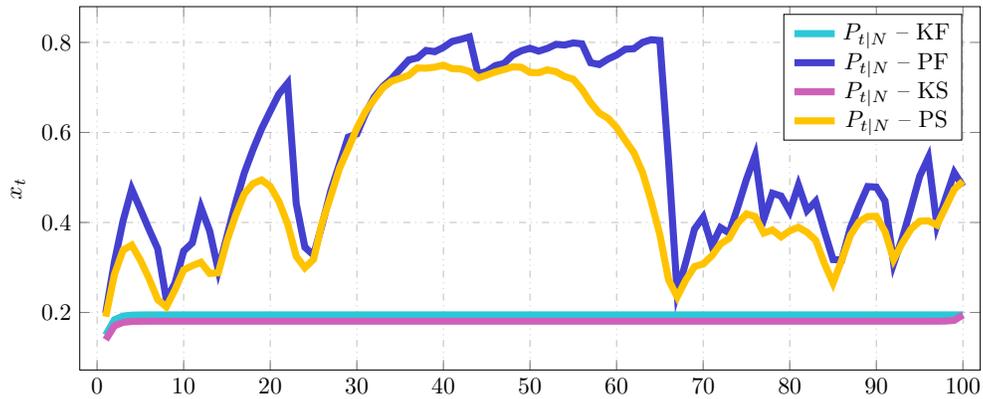


Figura 3.17: Varianza de la Estimación, con $\Delta = 10$ y 10000 partículas.

La Figura 3.18 muestra el error cuadrático medio en función del número de partículas para los métodos ya mencionados. Se aprecia cómo los métodos basados en partículas son evidentemente mejor que los que no consideran la cuantización, aunque a medida de que el nivel de cuantización crece, la mejoría del Suavizador de Partículas con respecto al Filtro de Partículas se reduce. Sin embargo, la Figura 3.18, que correspondía a la varianza para el caso $\Delta = 10$ mostró que de igual forma la varianza de la estimación era reducida por el suavizador.

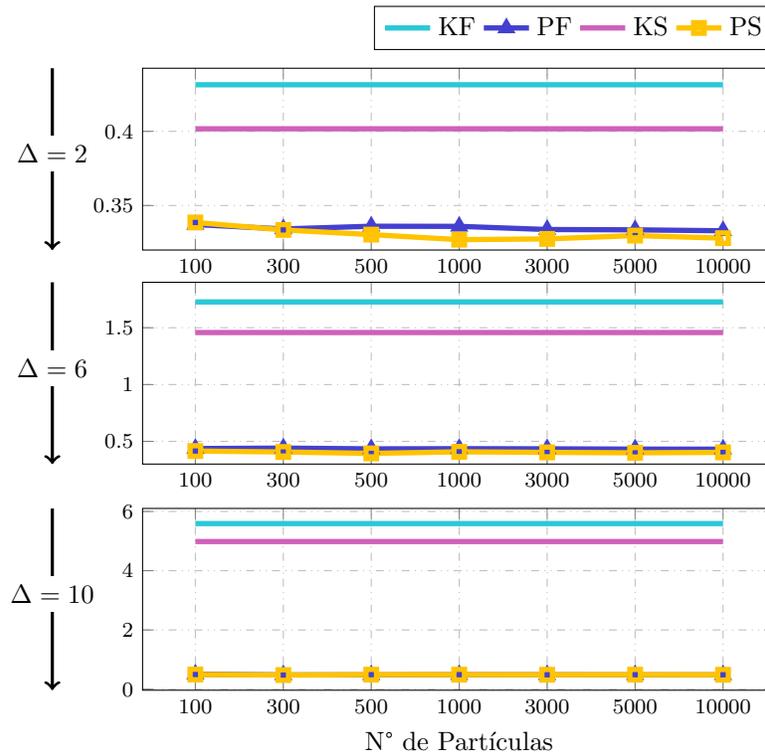


Figura 3.18: Error Cuadrático Medio incluyendo suavizado en sistema de segundo orden.

Por último, en la Figura 3.19 se puede observar nuevamente que el no considerar el proceso de cuantización produce una función de densidad de probabilidad $p(x_t|y_{1:N})$ errónea, con una

covarianza evidentemente menor que la real.

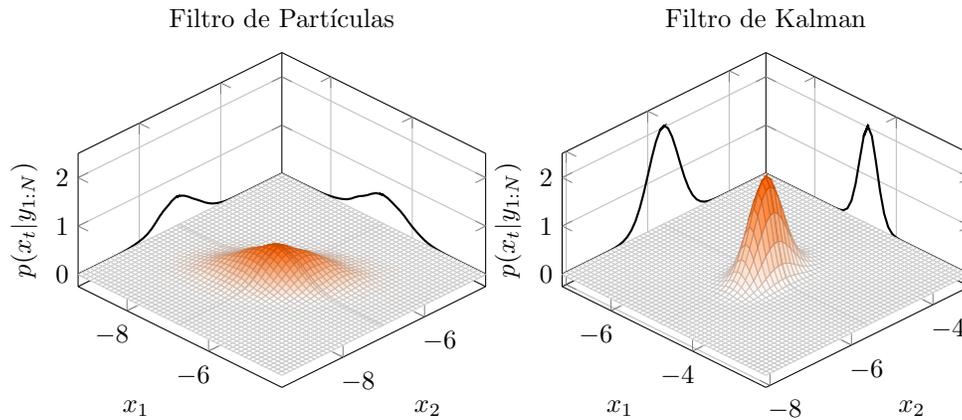


Figura 3.19: Función de densidad $p(x_t|y_{1:N})$ para cada método, con $\Delta = 10$, 10 Gaussianas y 10000 partículas.

Es importante destacar que en algunas aplicaciones no es la estimación del estado la única estimación de interés. Los resultados de este capítulo demuestran la ventaja de usar los métodos que consideren el proceso de cuantización, sobre todo al momento de estimar momentos de orden superior al primer momento. Como se verá más adelante, en la aplicación de interés de esta tesis no se requiere tan solo el momento de primer orden, por lo que se desea disponer de la función de densidad suavizada $p(x_t|y_{1:N})$, bien caracterizada.

4

Estimación por Máxima Verosimilitud a través del Algoritmo EM

4.1 Estimación por Máxima Verosimilitud

El método de estimación por máxima verosimilitud (ML por las siglas de su nombre en inglés *Maximum Likelihood*) fue desarrollado en el contexto de teoría estadística por Fisher en 1921 [67].

Asumimos que el mecanismo de generación de datos puede ser representado por una familia paramétrica de distribuciones de probabilidad $\mathcal{P} = \{p(\cdot|\theta) : \theta \in \Theta\}$ donde θ es un parámetros y $\Theta \subset \mathbb{R}^d$.

Para un conjunto de datos dados $y_{1:N}$ podemos considerar $p(y_{1:N}|\theta)$ como una función de θ en Θ , y referirnos a $p(y_{1:N}|\cdot)$ como la Función de Verosimilitud en Θ .

El principio de estimación ML nos conduce a tomar nuestro estimador de θ tal que, dentro de un espacio de parámetros, maximice la función de verosimilitud, lo que se formaliza en la siguiente definición.

Definición 4.1 El estimador ML del vector de parámetros θ es tal que,

$$\hat{\theta}_{ML} = \arg \max_{\theta \in \Theta} p(y_{1:N}|\theta), \quad (4.1)$$

donde $L(\theta) = p(y_{1:N}|\theta)$ es la función de verosimilitud.

Observación Es posible demostrar que para modelos lineales con perturbaciones Gaussianas, el estimador ML se reduce al estimador de mínimos cuadrados [68].

4.1.1 Propiedades del estimador ML

En los teoremas que se presentan a continuación se destacan algunas de las propiedades más importantes del estimador ML.

Teorema 4.1 (Principio de invarianza) Si $\hat{\theta}$ es un estimador ML de $\theta \in \Theta \subset \mathbb{R}^d$, entonces $g(\hat{\theta})$ es un estimador ML de $g(\theta)$, donde g es una función que cumple con $g : \Theta \rightarrow Z \subset \mathbb{R}^n$ y $n \leq d$.

Teorema 4.2 Cuando existe un estimador no sesgado que logra la cota inferior de Cramér-Rao (es decir, un estimador eficiente y no sesgado), entonces es también el estimador ML.

Teorema 4.3 (Consistencia) Sea $\hat{\theta}_N$ un estimador ML de θ basado en N variables aleatorias independientes e idénticamente distribuidas (i.i.d) $y = (x_1, \dots, x_N)^T$, entonces $\hat{\theta}_N$ converge a θ casi seguramente (a.s por las siglas de *almost surely* en inglés), es decir,

$$\hat{\theta}_N \xrightarrow{a.s} \theta. \quad (4.2)$$

Teorema 4.4 (Normalidad Asintótica) Sea $\hat{\theta}_N$ un estimador ML de θ basado en N variables aleatorias i.i.d $y = (x_1, \dots, x_N)$, entonces $\hat{\theta}_N$ converge en distribución a una variable aleatoria normal, es decir,

$$\sqrt{N} (\hat{\theta}_N - \theta) \xrightarrow{d} \beta, \quad (4.3)$$

donde

$$\beta \sim \mathcal{N}(0, \bar{M}_\theta^{-1}), \quad (4.4)$$

y \bar{M}_θ es el promedio de la matriz de información de Fisher por muestra, esto es,

$$\bar{M}_\theta^{-1} = \frac{1}{N} M_\theta \quad (4.5)$$

donde M_θ es la matriz de información de Fisher, la que corresponde a

$$M_\theta = \mathbb{E} \left\{ \left[\frac{\partial \log p(y|\theta)}{\partial \theta} \right]^T \left[\frac{\partial \log p(y|\theta)}{\partial \theta} \right] \right\}. \quad (4.6)$$

Teorema 4.5 (Eficiencia) Dentro de la clase de estimadores asintóticamente normales y uniformemente consistentes, $\hat{\theta}$ es eficiente en el sentido que alcanza asintóticamente la cota inferior de Cramér - Rao.

A menudo, es útil definir la función de log-verosimilitud $l(\theta) = \log p(y_{1:N}|\theta)$, la que tiene su máximo en el mismo θ y entonces se busca optimizar $l(\theta)$. Sin embargo, es común en la práctica que la función de verosimilitud o log-verosimilitud no pueda ser maximizada analíticamente. En tales casos puede ser posible calcular el estimador ML iterativamente mediante el uso de un procedimiento de maximización de Newton-Raphson o alguna variante, siempre que el número total d de parámetros en el modelo no sea demasiado grande. En algunos casos, obtener el Jacobiano necesario en el método de Newton-Raphson es demasiado costoso o incluso imposible. Por esto, uno de los métodos más utilizados por la simplicidad en sus cálculos es el Algoritmo EM, el que se utiliza en este trabajo.

4.2 Nociones del Algoritmo EM

El algoritmo popularmente conocido como EM por su nombre en inglés *Expectation - Maximization algorithm*, es ampliamente utilizado en situaciones donde, hay variables estocásticas que no son medidas directamente, tales como, mediciones con pérdida de datos, estados en un sistema en espacio de estados, etc.

La literatura estadística desde principios del siglo XX, está repleta de métodos en el espíritu del algoritmo EM o que en realidad son algoritmos EM en contextos especiales. En [69] se captura la esencia y espíritu de varios artículos históricos relacionados a EM. Sin embargo, la formulación del algoritmo EM en su generalidad actual y su respectivo nombre se encuentran en [70], artículo que también da una variedad de ejemplos de su aplicabilidad y algunas propiedades básicas bastante generales. En [71], se encuentra un análisis de convergencia y en [72], es posible encontrar prácticamente todo sobre el algoritmo hasta esa época.

Posterior a los trabajos ya mencionados, en [73–75] se estableció el llamado algoritmo MM, el que más bien es un principio sobre como construir algoritmos de optimización en el que se explotan los argumentos de convexidad. Gracias a estos trabajos el algoritmo EM se puede presentar como un caso particular del algoritmo MM, que es el camino que aquí se ha escogido seguir, por lo cual, primero se presentarán los principios del algoritmo MM como en [75], para posteriormente presentar el algoritmo EM y finalmente estudiar sus propiedades como en [71, 72].

4.2.1 Algoritmo MM

El algoritmo MM debe su nombre al juego de palabras en inglés *majorize/minimize* y *minorize/maximize*, si se busca resolver un problema de minimización o maximización respectivamente. En lo que sigue de esta sección, nos referiremos a estos conceptos en español como *mayorizar* en un problema de minimización o *minorizar* en un problema de maximización. Esta clase de algoritmos opera creando una función sustituta o auxiliar que mayoriza o minoriza la función objetivo. Cuando la función sustituta es optimizada, la función objetivo crece o disminuye según sea necesario. En el contexto de buscar un estimador ML, queremos maximizar la función de verosimilitud, por lo que describimos el algoritmo en este caso.

Definición 4.2 Una función $s(\theta, \hat{\theta}^k)$ se dice minorizar la función $l(\theta)$ en $\hat{\theta}^k$ si se cumplen las siguientes condiciones:

1. $s(\hat{\theta}^k, \hat{\theta}^k) = l(\hat{\theta}^k)$
2. $s(\theta, \hat{\theta}^k) \leq l(\theta), \forall \theta$.

En otras palabras, la superficie $\theta \mapsto s(\theta, \hat{\theta}^k)$ se encuentra siempre por debajo de $l(\theta)$ y es tangente¹ a ella en el punto $\theta = \hat{\theta}^k$.

Observación Note que una función $s(\theta, \hat{\theta}^k)$ se dice mayorizar la función $l(\theta)$ en $\hat{\theta}^k$ siempre que $-s(\theta, \hat{\theta}^k)$ minorice $-l(\theta)$.

Normalmente $\hat{\theta}^k$ representa la iteración actual en una búsqueda de optimizar $l(\theta)$. En un problema de maximización, nosotros maximizamos la función sustituta $s(\theta, \hat{\theta}^k)$ en lugar de $l(\theta)$. Así, definimos la siguiente iteración del algoritmo MM como sigue.

¹Para hablar de tangencia es necesario $\left. \frac{\partial s(\theta, \hat{\theta}^k)}{\partial \theta} \right|_{\hat{\theta}^k} = \left. \frac{\partial l(\theta)}{\partial \theta} \right|_{\hat{\theta}^k}$, condición que a veces es impuesta.

Definición 4.3 Si $\hat{\theta}^k$ es la iteración actual del algoritmo MM, la siguiente iteración está dada por:

$$\hat{\theta}^{k+1} = \arg \max_{\theta} s(\theta, \hat{\theta}^k). \quad (4.7)$$

Lema 4.1 Suponga que se itera en el algoritmo MM, según la Definición 4.3, de tal forma que se escoge una secuencia de parámetros $\{\hat{\theta}^k\}$, tal que

$$s(\hat{\theta}^{k+1}, \hat{\theta}^k) \geq s(\theta, \hat{\theta}^k), \quad \forall \theta \in \Theta \quad (4.8)$$

entonces se cumple

$$l(\hat{\theta}^{k+1}) \geq l(\hat{\theta}^k). \quad (4.9)$$

Demostración. La demostración es sencilla, basta con utilizar la segunda propiedad de la Definición 4.2, luego la desigualdad en (4.8) y finalmente la primera propiedad en la Definición 4.2, es decir,

$$\begin{aligned} l(\hat{\theta}^{k+1}) &\geq s(\hat{\theta}^{k+1}, \hat{\theta}^k) \\ &\geq s(\hat{\theta}^k, \hat{\theta}^k) \\ &= l(\hat{\theta}^k). \end{aligned}$$

□

El lema anterior demuestra que para encontrar un nuevo parámetro $\hat{\theta}^{k+1}$ que aumente el valor de la función $l(\theta)$ en cada iteración, basta con maximizar la función $s(\theta, \hat{\theta}^k)$.

4.2.2 Algoritmo EM

El algoritmo EM es un algoritmo que proporciona un procedimiento sistemático para encontrar la función auxiliar o sustituta del algoritmo MM y que permite calcular el estimador ML de forma iterativa.

Este algoritmo introduce la idea de tres conjuntos de datos, en primer lugar, un *conjunto de datos incompletos* $y_{1:N}$ que es medible y corresponde al conjunto de datos que define el problema de estimación por máxima verosimilitud. Luego, se define un conjunto de variables o *datos ocultos* $x_{1:N}$ que aportan información pero no están disponibles, lo que finalmente da origen a un conjunto de *datos completos* $\{x_{1:N}, y_{1:N}\}$, compuesto por los dos conjuntos anteriores. Estos conjuntos son escogidos de tal forma que la función de verosimilitud de los datos completos $p(x_{1:N}, y_{1:N}|\theta)$ sea más fácil de optimizar.

Lema 4.2 La función log-verosimilitud puede ser reescrita como

$$l_N(\theta) = \mathcal{Q}(\theta, \hat{\theta}^k) - \mathcal{H}(\theta, \hat{\theta}^k), \quad (4.10)$$

donde

$$\mathcal{Q}(\theta, \hat{\theta}^k) = \mathbb{E} \left\{ \log p(x_{1:N}, y_{1:N}|\theta) | y_{1:N}, \hat{\theta}^k \right\} \quad (4.11)$$

$$\mathcal{H}(\theta, \hat{\theta}^k) = \mathbb{E} \left\{ \log p(x_{1:N} | y_{1:N}, \theta) | y_{1:N}, \hat{\theta}^k \right\}. \quad (4.12)$$

4.2. Nociones del Algoritmo EM

Demostración. Usando el Teorema de Bayes es posible escribir

$$p(x_{1:N}, y_{1:N}|\theta) = p(x_{1:N}|y_{1:N}, \theta)p(y_{1:N}|\theta) \quad (4.13)$$

donde se puede aplicar la función logaritmo, multiplicar la ecuación por $p(x_{1:N}|y_{1:N}, \hat{\theta}^k)$ e integrar sobre $x_{1:N}$, entonces

$$\begin{aligned} \log p(x_{1:N}, y_{1:N}|\theta) &= \log p(x_{1:N}|y_{1:N}, \theta) + \log p(y_{1:N}|\theta) \\ \log [p(x_{1:N}, y_{1:N}|\theta)] p(x_{1:N}|y_{1:N}, \hat{\theta}^k) &= \log [p(x_{1:N}|y_{1:N}, \theta)] p(x_{1:N}|y_{1:N}, \hat{\theta}^k) \\ &\quad + \log [p(y_{1:N}|\theta)] p(x_{1:N}|y_{1:N}, \hat{\theta}^k) \\ \int \log [p(x_{1:N}, y_{1:N}|\theta)] p(x_{1:N}|y_{1:N}, \hat{\theta}^k) dx_{1:N} &= \int \log [p(x_{1:N}|y_{1:N}, \theta)] p(x_{1:N}|y_{1:N}, \hat{\theta}^k) dx_{1:N} \\ &\quad + \int \log [p(y_{1:N}|\theta)] p(x_{1:N}|y_{1:N}, \hat{\theta}^k) dx_{1:N}, \end{aligned} \quad (4.14)$$

y notando que

$$\int \log [p(y_{1:N}|\theta)] p(x_{1:N}|y_{1:N}, \hat{\theta}^k) dx_{1:N} = \log [p(y_{1:N}|\theta)] \int p(x_{1:N}|y_{1:N}, \hat{\theta}^k) dx_{1:N} = l(\theta), \quad (4.15)$$

la ecuación en (4.14) se convierte en

$$\mathcal{Q}(\theta, \hat{\theta}^k) = \mathcal{H}(\theta, \hat{\theta}^k) + l(\theta), \quad (4.16)$$

lo que completa la demostración. \square

Considerando el Lema 4.2 recién expuesto, es posible definir la función sustituta o auxiliar del algoritmo MM como

$$s(\theta, \hat{\theta}^k) = \mathcal{Q}(\theta, \hat{\theta}^k) + l_N(\hat{\theta}^k) - \mathcal{Q}(\hat{\theta}^k, \hat{\theta}^k). \quad (4.17)$$

Lema 4.3 La función $s(\theta, \hat{\theta}^k)$ en (4.17) satisface las siguientes propiedades:

1. $s(\hat{\theta}^k, \hat{\theta}^k) = l_N(\hat{\theta}^k)$
2. $s(\theta, \hat{\theta}^k) \leq l_N(\theta)$
3. $\left. \frac{\partial s(\theta, \hat{\theta}^k)}{\partial \theta} \right|_{\hat{\theta}^k} = \left. \frac{\partial l_N(\theta)}{\partial \theta} \right|_{\hat{\theta}^k}$

donde las primeras dos condiciones son las requeridas por el algoritmo MM en la Definición 4.2, entonces $s(\theta, \hat{\theta}^k)$ en (4.17) minoriza a $l_N(\theta)$.

Demostración. La primera propiedad se obtiene directamente al evaluar (4.17) en $\hat{\theta}^k$.

Para la segunda propiedad es posible escribir la función de verosimilitud como

$$\begin{aligned} l_N(\theta) &= \mathcal{Q}(\theta, \hat{\theta}^k) - \mathcal{H}(\theta, \hat{\theta}^k) \\ &= \mathcal{Q}(\theta, \hat{\theta}^k) - \mathcal{Q}(\hat{\theta}^k, \hat{\theta}^k) + l_N(\hat{\theta}^k) + \mathcal{Q}(\hat{\theta}^k, \hat{\theta}^k) - l_N(\hat{\theta}^k) - \mathcal{H}(\theta, \hat{\theta}^k), \end{aligned}$$

donde es posible utilizar (4.17) y (4.10) del Lema 4.2,

$$l_N(\theta) = s(\theta, \hat{\theta}^k) + \mathcal{H}(\hat{\theta}^k, \hat{\theta}^k) - \mathcal{H}(\theta, \hat{\theta}^k), \quad (4.18)$$

luego, la desigualdad $l_N(\theta) \geq s(\theta, \hat{\theta}^k)$ se mantiene si

$$\mathcal{H}(\theta, \hat{\theta}^k) - \mathcal{H}(\hat{\theta}^k, \hat{\theta}^k) \leq 0 \quad (4.19)$$

en (4.18), lo se demuestra a continuación. Primero, según la definición de $\mathcal{H}(\theta, \hat{\theta}^k)$,

$$\begin{aligned} \mathcal{H}(\theta, \hat{\theta}^k) - \mathcal{H}(\hat{\theta}^k, \hat{\theta}^k) &= \mathbb{E} \left\{ \log p(x_{1:N}|y_{1:N}, \theta) \middle| y_{1:N}, \hat{\theta}^k \right\} - \mathbb{E} \left\{ \log p(x_{1:N}|y_{1:N}, \hat{\theta}^k) \middle| y_{1:N}, \hat{\theta}^k \right\} \\ &= \mathbb{E} \left\{ \log \left[\frac{p(x_{1:N}|y_{1:N}, \theta)}{p(x_{1:N}|y_{1:N}, \hat{\theta}^k)} \right] \middle| y_{1:N}, \hat{\theta}^k \right\} \end{aligned}$$

ahora, utilizando la desigualdad de Jensen para una función cóncava como la función logaritmo, se tiene

$$\begin{aligned} &\leq \log \mathbb{E} \left\{ \frac{p(x_{1:N}|y_{1:N}, \theta)}{p(x_{1:N}|y_{1:N}, \hat{\theta}^k)} \middle| y_{1:N}, \hat{\theta}^k \right\} \\ &= \log \int \frac{p(x_{1:N}|y_{1:N}, \theta)}{p(x_{1:N}|y_{1:N}, \hat{\theta}^k)} p(x_{1:N}|y_{1:N}, \hat{\theta}^k) dx_{1:N} \\ &= \log \int p(x_{1:N}|y_{1:N}, \theta) dx_{1:N} \\ &= 0, \end{aligned}$$

lo que concluye la prueba de la segunda propiedad.

Luego, para la tercera propiedad notamos que

$$\begin{aligned} \frac{\partial s(\theta, \hat{\theta}^k)}{\partial \theta} &= \frac{\partial \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \theta} \\ \frac{\partial l_N(\theta)}{\partial \theta} &= \frac{\partial \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \theta} - \frac{\partial \mathcal{H}(\theta, \hat{\theta}^k)}{\partial \theta}. \end{aligned}$$

Por ende, la demostración de la tercera propiedad es equivalente a demostrar

$$\frac{\partial \mathcal{H}(\theta, \hat{\theta}^k)}{\partial \theta} \bigg|_{\hat{\theta}^k} = 0, \quad (4.20)$$

lo que sigue desde

$$\begin{aligned} \frac{\partial \mathcal{H}(\theta, \hat{\theta}^k)}{\partial \theta} &= \frac{\partial}{\partial \theta} \mathbb{E} \left\{ \log p(x_{1:N}|y_{1:N}, \theta) \middle| y_{1:N}, \hat{\theta}^k \right\} \\ &= \mathbb{E} \left\{ \frac{\partial \log p(x_{1:N}|y_{1:N}, \theta)}{\partial \theta} \middle| y_{1:N}, \hat{\theta}^k \right\} \\ &= \mathbb{E} \left\{ \frac{1}{p(x_{1:N}|y_{1:N}, \theta)} \frac{\partial p(x_{1:N}|y_{1:N}, \theta)}{\partial \theta} \middle| y_{1:N}, \hat{\theta}^k \right\} \\ &= \int \frac{1}{p(x_{1:N}|y_{1:N}, \theta)} \frac{\partial p(x_{1:N}|y_{1:N}, \theta)}{\partial \theta} p(x_{1:N}|y_{1:N}, \hat{\theta}^k) dx_{1:N}, \end{aligned}$$

4.2. Nociones del Algoritmo EM

y evaluando en $\hat{\theta}^k$,

$$\begin{aligned}
 \left. \frac{\partial \mathcal{H}(\theta, \hat{\theta}^k)}{\partial \theta} \right|_{\hat{\theta}^k} &= \int \frac{1}{p(x_{1:N}|y_{1:N}, \hat{\theta}^k)} \left. \frac{\partial p(x_{1:N}|y_{1:N}, \theta)}{\partial \theta} \right|_{\hat{\theta}^k} p(x_{1:N}|y_{1:N}, \hat{\theta}^k) dx_{1:N} \\
 &= \int \left. \frac{\partial p(x_{1:N}|y_{1:N}, \theta)}{\partial \theta} \right|_{\hat{\theta}^k} dx_{1:N} \\
 &= \left(\frac{\partial}{\partial \theta} \int p(x_{1:N}|y_{1:N}, \theta) dx_{1:N} \right) \Big|_{\hat{\theta}^k} \\
 &= 0,
 \end{aligned}$$

lo que completa la demostración. \square

Ahora, considerando el Lema 4.3 tenemos que $s(\theta, \hat{\theta}^k)$ en (4.17) es una función sustituta de algún algoritmo MM y por ende, cumple el Lema 4.1, donde la función $l(\theta)$ correspondería a la función log-verosimilitud $l_N(\theta)$. Por otro lado, es fácil ver en (4.17) que maximizar $s(\theta, \hat{\theta}^k)$ es equivalente a maximizar $\mathcal{Q}(\theta, \hat{\theta}^k)$. Así, el algoritmo EM para obtener el estimador ML se resume en el siguiente algoritmo.

Algoritmo 4.1 El algoritmo iterativo EM para obtener el estimador ML se puede llevar a cabo a través de los siguientes pasos:

1. Obtener una estimación inicial $\hat{\theta}_0$.
2. **Paso E:** Obtener la función $\mathcal{Q}(\theta, \hat{\theta}^k)$ a través de (4.11).
3. **Paso M:** Calcular la estimación $\hat{\theta}^{k+1} = \arg \max_{\theta \in \Theta} \mathcal{Q}(\theta, \hat{\theta}^k)$.
4. Volver al paso 2, incrementando el índice $k \rightarrow k + 1$ hasta la convergencia, en tal caso detener.

En la práctica es común que la solución del paso M exista en forma cerrada, sin embargo, para los casos en que no sea factible encontrar el θ que maximice la función $\mathcal{Q}(\theta, \hat{\theta}^k)$, [70] define un algoritmo EM generalizado (algoritmo GEM), el que se enuncia a continuación.

Definición 4.4 (GEM) Se define el algoritmo GEM generalizado como el algoritmo EM para el cual el paso M requiere que $\hat{\theta}^{k+1}$ se escogido tal que

$$\mathcal{Q}(\hat{\theta}^{k+1}, \hat{\theta}^k) \geq \mathcal{Q}(\hat{\theta}^k, \hat{\theta}^k). \quad (4.21)$$

Definición 4.5 (Mapeo EM) Cualquier secuencia del algoritmo EM (GEM) implícitamente define un mapeo $\theta \rightarrow \mathcal{M}(\theta)$, desde el espacio de parámetro Θ al mismo espacio, tal que

$$\hat{\theta}^{k+1} = \mathcal{M}(\hat{\theta}^k). \quad (4.22)$$

La definición anterior indica que si $\hat{\theta}^k$ converge a algún punto θ^* y $\mathcal{M}(\theta)$ es continua, entonces θ^* satisface

$$\theta^* = \mathcal{M}(\theta^*). \quad (4.23)$$

Así θ^* es un punto fijo del mapa \mathcal{M} .

Es importante notar que la condición para la iteración del algoritmo EM normalmente es que $\hat{\theta}^{k+1} = \mathcal{M}_{EM}(\hat{\theta}^k)$ maximice $\mathcal{Q}(\theta, \hat{\theta}^k)$ en cada iteración, que es una condición más estricta, y por ende, un caso particular del algoritmo GEM. Además, es importante destacar que el algoritmo GEM también cumple con el Lema 4.1, pues éste se basó en la condición (4.8), que es equivalente a (4.21).

Convergencia del algoritmo EM

A través de los resultados expuestos, vimos que tanto el algoritmo EM como GEM cumplen con el Lema 4.1 para la función log-verosimilitud $l_N(\theta)$, por ende, la función de verosimilitud es no decreciente después de una iteración del algoritmo GEM, más aún, la función de verosimilitud crecerá si la desigualdad (4.21) es estricta. Así, para una secuencia acotada de valores de la función verosimilitud $\{L(\hat{\theta}^k)\}$, $L(\hat{\theta}^k)$ converge monótonamente a algún L^* . Sin embargo, no hay garantía que L^* sea el máximo global de $L(\theta)$ sobre Θ para el algoritmo EM, incluso si una maximización global de $\mathcal{Q}(\theta, \hat{\theta}^k)$ se llevó a cabo en el paso M. Inclusive, el problema de convergencia podría ser menos satisfactorio y θ^* podría ser si quiera un máximo local, entonces, el algoritmo EM puede converger a un punto silla. A continuación, se presentan algunos teoremas de convergencia que se pueden encontrar en [71, 72]. Comenzamos enunciando el Teorema de convergencia principal dado por [71] para el algoritmo GEM.

Teorema 4.6 Sea $\{\hat{\theta}^k\}$ una secuencia del algoritmo GEM generada por $\hat{\theta}^{k+1} \in \mathcal{M}(\hat{\theta}^k)$. Suponga que

- (i) $\mathcal{M}(\hat{\theta}^k)$ es cerrado sobre el complemento de \mathcal{S} , el conjunto de puntos estacionarios en el interior de Θ .
- (ii) $L(\hat{\theta}^{k+1}) > L(\hat{\theta}^k)$ para todo $\hat{\theta}^k \notin \mathcal{S}$.

Entonces, todos los puntos límites de $\{\hat{\theta}^k\}$ son puntos estacionarios de $L(\theta)$ y $L(\hat{\theta}^k)$ converge monótonamente a $L^* = L(\theta^*)$ para algún punto estacionario $\theta^* \in \mathcal{S}$.

Demostración. La demostración se puede encontrar en [71]. □

Observación El Teorema 4.6 se mantiene si \mathcal{S} es reemplazado por \mathcal{J} , donde \mathcal{J} es el conjunto de máximos locales en el interior de Θ . En tal caso, todos los puntos límites de $\{\hat{\theta}^k\}$ son máximos locales de $L(\theta)$ y $L(\hat{\theta}^k)$ converge monótonamente a $L^* = L(\theta^*)$ para algún máximo local $\theta^* \in \mathcal{J}$ [71].

Para el algoritmo EM, [71] señala que una condición suficiente para la clausura del mapeo EM, \mathcal{M} , es que

$$\mathcal{Q}(\theta, \phi) \text{ sea continua en } \theta \text{ y } \phi. \quad (4.24)$$

Esta condición es muy débil y debería mantenerse en la mayoría de las situaciones prácticas. Para convergencia a valores estacionarios resulta ser la única condición de regularidad requerida. El siguiente Teorema es más útil en el sentido que cumple un amplio rango de aplicaciones estadísticas.

Teorema 4.7 Suponga que $\mathcal{Q}(\theta, \phi)$ satisface la condición de continuidad (4.24). Entonces, todos los puntos límites de cualquier secuencia $\{\hat{\theta}^k\}$ de un algoritmo EM son puntos estacionarios de $L(\theta)$, y $L(\hat{\theta}^k)$ converge monótonamente a algún valor $L^* = L(\theta^*)$ para algún punto estacionario θ^* .

4.2. Nociones del Algoritmo EM

Demostración. El Teorema sigue inmediatamente desde el Teorema 4.6, ya que la condición (i) viene de la suposición de continuidad en (4.24) y la condición (ii) se cumple automáticamente por una secuencia EM. Para más detalles consultar [71]. \square

Observación En este caso, el mismo argumento no es aplicable si \mathcal{S} es reemplazado por \mathcal{J} . Así, para garantizar convergencia a un máximo local, necesitamos condiciones adicionales como (4.25) en el teorema que sigue. Debido a que (4.25) se mantiene para cualquier $\theta \notin \mathcal{S}$, (4.19) y (4.25) implican (ii) en el Teorema 4.6 para todo $\hat{\theta}^k \notin \mathcal{J}$. Por lo tanto, el teorema que sigue es un caso especial del Teorema 4.6.

Teorema 4.8 Suponga que $\mathcal{Q}(\theta, \hat{\theta}^k)$ satisface la condición de continuidad (4.24) y

$$\sup_{\theta \in \Theta} \mathcal{Q}(\theta, \hat{\theta}^k) > \mathcal{Q}(\hat{\theta}^k, \hat{\theta}^k) \text{ para cualquier } \theta \in \mathcal{S} \setminus \mathcal{J} \quad (4.25)$$

Entonces, todos los puntos límites de cualquier secuencia $\{\hat{\theta}^k\}$ de un algoritmo EM son máximos locales de $L(\theta)$ y $L(\hat{\theta}^k)$ converge monótonamente a $L^* = L(\theta^*)$ para algún máximo local θ^* .

Como (4.25) es típicamente difícil de verificar, la utilidad del Teorema 4.8 es algo limitada. A pesar de esto, en [71] se presenta una clase de densidades de probabilidad exponenciales que satisfacen la condición propuesta.

5

Estimación por Máxima Verosimilitud de Sistemas Dinámicos Lineales con Salida Cuantizada

En este capítulo, es donde, provistos de las herramientas y aportes descritos en los capítulos 3 y 4 se resuelve el problema planteado en la sección 1.1.

Por conveniencia reescribimos el sistema (1.1) de la siguiente forma

$$\begin{aligned}\xi_t &= \Gamma\zeta_t + \eta_t \\ y_t &= \mathbb{Q}\{z_t\},\end{aligned}\tag{5.1}$$

donde

$$\xi_t = \begin{bmatrix} x_{t+1} \\ z_t \end{bmatrix}, \quad \zeta_t = \begin{bmatrix} x_t \\ u_t \end{bmatrix}, \quad \eta_t = \begin{bmatrix} w_t \\ v_t \end{bmatrix},\tag{5.2}$$

y Γ fue definido en (1.5). Así, utilizando las nuevas variables es claro que,

$$p(\eta_t) = \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Pi\right),\tag{5.3}$$

donde Π fue definido en (1.5). Luego, considerando (5.1), es directo obtener que

$$p(\xi_t|\zeta_t) = \mathcal{N}(\Gamma\zeta_t, \Pi),\tag{5.4}$$

donde hemos omitido la dependencia de los parámetros, omisión que se mantendrá en lo que resta del capítulo.

Por otro lado, queremos obtener el estimador ML usando el algoritmo EM descrito en el capítulo anterior. Para esto, uno de los primeros pasos es definir los conjuntos de *datos completos*, *datos incompletos* y *datos ocultos*, paso clave en la potencialidad de simplificar el problema de obtener el estimador ML. En este caso, hemos decidido tomar un camino distinto al de [35] y quitamos el conjunto $y_{1:N}$ de los datos completos. Entonces, los conjuntos son:

- **Datos Completos:** $\{x_{1:N+1}, z_{1:N}\}$.
- **Datos Incompletos:** $\{y_{1:N}\}$.

Luego, se quiere obtener la función $\mathcal{Q}(\theta, \hat{\theta}^k)$ del algoritmo, para lo cual, encontramos primero la función de log-verosimilitud de los datos completos a través del siguiente Lema.

Lema 5.1 La función de log-verosimilitud de los datos completos $\log [p(x_{1:N+1}, z_{1:N}|\theta)]$ es dada por:

$$\begin{aligned} \log [p(x_{1:N+1}, z_{1:N}|\theta)] &= -\frac{1}{2} \log \det\{2\pi P_1\} - \frac{1}{2}(x_1 - \mu_1)^T P_1^{-1}(x_1 - \mu_1) \\ &\quad - \frac{N}{2} \log \det\{2\pi \Pi\} - \frac{1}{2} \sum_{t=1}^N (\xi_t - \Gamma \zeta_t)^T \Pi^{-1} (\xi_t - \Gamma \zeta_t). \end{aligned} \quad (5.5)$$

Demostración. Usando la regla de Bayes en la función de verosimilitud de los datos completos se tiene,

$$\begin{aligned} p(x_{1:N+1}, z_{1:N}|\theta) &= p(x_{N+1}, z_N | x_{1:N}, z_{1:N-1}, \theta) p(x_{1:N}, z_{1:N-1} | \theta) \\ &= p(x_{N+1}, z_N | x_N, \theta) p(x_N, z_{N-1} | x_{N-1}, \theta) p(x_{1:N-1}, z_{1:N-2}, \theta) \\ &= p(x_1 | \theta) \prod_{t=1}^N p(x_{t+1}, z_t | x_t, \theta). \end{aligned}$$

Considerando que $p(x_{t+1}, z_t | x_t, \theta) = p(\xi_t | \zeta_t, \theta)$ es Gaussiana, al igual que el estado inicial x_1 y aplicando la función logaritmo

$$\begin{aligned} \log [p(x_{1:N+1}, z_{1:N}|\theta)] &= -\frac{1}{2} \log \det\{2\pi P_1\} - \frac{1}{2}(x_1 - \mu_1)^T P_1^{-1}(x_1 - \mu_1) \\ &\quad - \frac{N}{2} \log \det\{2\pi \Pi\} - \frac{1}{2} \sum_{t=1}^N (\xi_t - \Gamma \zeta_t)^T \Pi^{-1} (\xi_t - \Gamma \zeta_t). \end{aligned}$$

□

Lema 5.2 La función auxiliar $\mathcal{Q}(\theta, \hat{\theta}^k)$ del algoritmo EM en (4.11) para el sistema (5.1) puede ser escrita como:

$$\begin{aligned} \mathcal{Q}(\theta, \hat{\theta}^k) &= -\frac{1}{2} \log \det\{2\pi P_1\} - \frac{N}{2} \log \det\{2\pi \Pi\} \\ &\quad - \frac{1}{2} \text{trace} \{P_1^{-1} [P_{1|N} + (\hat{x}_{1|N} - \mu_1)(\hat{x}_{1|N} - \mu_1)^T]\} \\ &\quad - \frac{N}{2} \text{trace} \{\Pi^{-1} [\Phi - \Psi \Gamma^T - \Gamma \Psi^T + \Gamma \Sigma \Gamma^T]\}. \end{aligned} \quad (5.6)$$

donde

$$\hat{x}_{1|N} = \mathbb{E} \left\{ x_1 | y_{1:N}, \hat{\theta}^k \right\} \quad (5.7)$$

$$P_{1|N} = \mathbb{E} \left\{ (x_1 - \hat{x}_{1|N})(x_1 - \hat{x}_{1|N})^T | y_{1:N}, \hat{\theta}^k \right\} \quad (5.8)$$

$$\Phi = \frac{1}{N} \sum_{t=1}^N \mathbb{E} \left\{ \xi_t \xi_t^T | y_{1:N}, \hat{\theta}^k \right\} \quad (5.9)$$

$$\Psi = \frac{1}{N} \sum_{t=1}^N \mathbb{E} \left\{ \xi_t \zeta_t^T | y_{1:N}, \hat{\theta}^k \right\} \quad (5.10)$$

$$\Sigma = \frac{1}{N} \sum_{t=1}^N \mathbb{E} \left\{ \zeta_t \zeta_t^T | y_{1:N}, \hat{\theta}^k \right\}. \quad (5.11)$$

Demostración. La función auxiliar $\mathcal{Q}(\theta, \hat{\theta}^k)$ está dada por

$$\mathcal{Q}(\theta, \hat{\theta}^k) = \mathbb{E} \left\{ \log p(x_{1:N+1}, z_{1:N} | \theta) | y_{1:N}, \hat{\theta}^k \right\}$$

y la función de log-verosimilitud del conjunto de datos completos fue dada en el Lema 5.1,

$$\begin{aligned} \mathcal{Q}(\theta, \hat{\theta}^k) = & \mathbb{E} \left\{ -\frac{1}{2} \log \det \{2\pi P_1\} - \frac{1}{2} (x_1 - \mu_1)^T P_1^{-1} (x_1 - \mu_1) \right. \\ & \left. - \frac{N}{2} \log \det \{2\pi \Pi\} - \frac{1}{2} \sum_{t=1}^N (\xi_t - \Gamma \zeta_t)^T \Pi^{-1} (\xi_t - \Gamma \zeta_t) \middle| y_{1:N}, \hat{\theta}^k \right\}, \end{aligned}$$

usando la función traza,

$$\begin{aligned} = & -\frac{1}{2} \log \det \{2\pi P_1\} - \frac{N}{2} \log \det \{2\pi \Pi\} \\ & + \text{trace} \left\{ \mathbb{E} \left\{ -\frac{1}{2} (x_1 - \mu_1)^T P_1^{-1} (x_1 - \mu_1) \right. \right. \\ & \left. \left. - \frac{1}{2} \sum_{t=1}^N (\xi_t - \Gamma \zeta_t)^T \Pi^{-1} (\xi_t - \Gamma \zeta_t) \middle| y_{1:N}, \hat{\theta}^k \right\} \right\}, \end{aligned}$$

intercambiando el operador traza con el operador esperanza,

$$\begin{aligned} = & -\frac{1}{2} \log \det \{2\pi P_1\} - \frac{N}{2} \log \det \{2\pi \Pi\} \\ & - \frac{1}{2} \mathbb{E} \left\{ \text{trace} \left\{ (x_1 - \mu_1)^T P_1^{-1} (x_1 - \mu_1) \right\} \middle| y_{1:N}, \hat{\theta}^k \right\} \\ & - \frac{1}{2} \sum_{t=1}^N \mathbb{E} \left\{ \text{trace} \left\{ (\xi_t - \Gamma \zeta_t)^T \Pi^{-1} (\xi_t - \Gamma \zeta_t) \right\} \middle| y_{1:N}, \hat{\theta}^k \right\}, \end{aligned}$$

y a través de la propiedad cíclica de la traza,

$$\begin{aligned}
&= -\frac{1}{2} \log \det\{2\pi P_1\} - \frac{N}{2} \log \det\{2\pi \Pi\} \\
&\quad - \frac{1}{2} \mathbb{E} \left\{ \text{trace} \left\{ P_1^{-1} (x_1 - \mu_1)(x_1 - \mu_1)^T \right\} \middle| y_{1:N}, \hat{\theta}^k \right\} \\
&\quad - \frac{1}{2} \sum_{t=1}^N \mathbb{E} \left\{ \text{trace} \left\{ \Pi^{-1} (\xi_t - \Gamma \zeta_t)(\xi_t - \Gamma \zeta_t)^T \right\} \middle| y_{1:N}, \hat{\theta}^k \right\} \\
&= -\frac{1}{2} \log \det\{2\pi P_1\} - \frac{N}{2} \log \det\{2\pi \Pi\} \\
&\quad - \frac{1}{2} \text{trace} \left\{ P_1^{-1} \mathbb{E} \left\{ (x_1 - \mu_1)(x_1 - \mu_1)^T \middle| y_{1:N}, \hat{\theta}^k \right\} \right\} \\
&\quad - \frac{N}{2} \text{trace} \left\{ \Pi^{-1} \frac{1}{N} \sum_{t=1}^N \mathbb{E} \left\{ (\xi_t - \Gamma \zeta_t)(\xi_t - \Gamma \zeta_t)^T \middle| y_{1:N}, \hat{\theta}^k \right\} \right\},
\end{aligned}$$

donde basta manipulación algebraica y definir (5.7) a (5.11) para obtener (5.6), lo que completa la demostración. \square

Note que de acuerdo a las definiciones en (5.2) y (5.7) a (5.11), varias de las cantidades que componen $\mathcal{Q}(\theta, \hat{\theta}^k)$ pueden ser evaluadas desde los elementos $\hat{x}_{t|N} = \mathbb{E} \{x_t | y_{1:N}\}$, $\mathbb{E} \{x_t x_t^T | y_{1:N}\}$, y $\mathbb{E} \{x_t x_{t-1}^T | y_{1:N}\}$, que de acuerdo a la estructura del modelo de interés en (1.1) y las suposiciones sobre los ruidos de la Sección 1.1, pueden ser calculados usando las ideas vistas en el Capítulo 3, en particular, el Algoritmo 3.5. Sin embargo, aún resta obtener las estadísticas asociadas a z_t , para lo que enunciamos los lemas que siguen.

Lema 5.3 La densidad de la salida no cuantizada z_t dada las mediciones $y_{1:N}$ y el estado actual x_t está dada por

$$p(z_t | x_t, y_{1:N}) = \frac{\mathcal{N}(Cx_t + Du_t, R) \mathbb{I}\{y_t = \mathfrak{q}_1\{z_t\}\}}{\int_{\mathfrak{q}_1^{-1}\{y_t\}} \mathcal{N}(Cx_t + Du_t, R) dz_t} \quad (5.12)$$

donde $\mathbb{I}\{\cdot\}$ corresponde a la función indicatriz que es igual a 1 cuando la variable en su dominio pertenece a un conjunto o 0 si no pertenece. En este caso, $y_t = \mathfrak{q}_1\{z_t\}$ es equivalente a $z_t \in [a_t^\ell, b_t^\ell]$, donde los extremos del hipercubo son entregados a través de la medición y_t .

Demostración. Primero que todo notamos que la variable z_t depende sólo del estado x_t y la medición actual y_t , por ende, $p(z_t | x_t, y_{1:N}) = p(z_t | x_t, y_t)$. Luego,

$$p(z_t | x_t, y_{1:N}) = p(z_t | x_t, y_t) \quad (5.13)$$

$$= \frac{p(z_t, y_t | x_t)}{p(y_t | x_t)} \quad (5.14)$$

$$= \frac{p(z_t | x_t) p(y_t | x_t, z_t)}{p(y_t | x_t)} \quad (5.15)$$

donde $p(z_t | x_t) = \mathcal{N}(Cx_t + Du_t, R)$, $p(y_t | x_t, z_t) = p(y_t | z_t)$ es una indicatriz y el denominador no depende de z_t por lo que es una constante de normalización, dada por ejemplo, por la integral del numerador, lo que completa la prueba. \square

Lema 5.4 Utilizando (3.65) es posible aproximar los momentos en (5.6) asociados a z_t como

$$\mathbb{E} \{z_t | y_{1:N}\} \approx \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \tilde{z}_t^i \quad (5.16)$$

$$\mathbb{E} \{z_t x_t^T | y_{1:N}\} \approx \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \tilde{z}_t^i (\tilde{x}_t^i)^T \quad (5.17)$$

$$\mathbb{E} \{x_{t+1} z_t^T | y_{1:N}\} \approx \sum_{i=1}^M \sum_{j=1}^M \tilde{\alpha}_{t|N}^{ij} \tilde{x}_{t+1}^j (\tilde{z}_t^i)^T \quad (5.18)$$

$$\mathbb{E} \{z_t z_t^T | y_{1:N}\} \approx \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \mathcal{M}_2^i(z_t) \quad (5.19)$$

donde $\tilde{z}_t^i = \mathbb{E} \{z_t | \tilde{x}_t^i, y_t\}$ y $\mathcal{M}_2^i(z_t) = \mathbb{E} \{z_t z_t^T | \tilde{x}_t^i, y_t\}$, con z_t distribuido de acuerdo a (5.12).

Demostración. Primero consideramos la aproximación en (3.65) dada por

$$p(x_t | y_{1:N}) \approx \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \delta(x_t - \tilde{x}_t^i).$$

Así, para (5.16)

$$\begin{aligned} \mathbb{E} \{z_t | y_{1:N}\} &= \int z_t p(z_t | y_{1:N}) dz_t \\ &= \int \int z_t p(z_t | x_t, y_t) p(x_t | y_{1:N}) dz_t dx_t, \\ &\approx \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \int z_t p(z_t | \tilde{x}_t^i, y_t) dz_t \\ &= \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \tilde{z}_t^i. \end{aligned}$$

Luego, para (5.19) basta cambiar z_t por $z_t z_t^T$ y tomar el mismo procedimiento. De manera similar, para (5.17)

$$\begin{aligned} \mathbb{E} \{z_t x_t^T | y_{1:N}\} &= \int \int z_t x_t^T p(z_t, x_t | y_{1:N}) dz_t dx_t \\ &= \int \int z_t x_t^T p(z_t | x_t, y_t) p(x_t | y_{1:N}) dz_t dx_t, \\ &\approx \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \left(\int z_t p(z_t | \tilde{x}_t^i, y_t) dz_t \right) (\tilde{x}_t^i)^T \\ &= \sum_{i=1}^M \tilde{\alpha}_{t|N}^i \tilde{z}_t^i (\tilde{x}_t^i)^T. \end{aligned}$$

Finalmente, para (5.18)

$$\begin{aligned}
\mathbb{E} \{x_{t+1} z_t^T | y_{1:N}\} &= \int \int x_{t+1} z_t^T p(x_{t+1}, z_t | y_{1:N}) dx_{t+1} dz_t \\
&= \int \int \int x_{t+1} z_t^T p(x_{t+1}, x_t, z_t | y_{1:N}) dx_{t+1} dz_t dx_t \\
&= \int \int \int x_{t+1} z_t^T p(z_t | x_{t+1}, x_t, y_{1:N}) p(x_{t+1}, x_t | y_{1:N}) dx_{t+1} dz_t dx_t
\end{aligned}$$

donde podemos usar (3.78) para aproximar la doble integral en x_{t+1} y x_t . Además, destacamos que $p(z_t | x_{t+1}, x_t, y_{1:N}) = p(z_t | x_t, y_t)$, con lo que sigue

$$\begin{aligned}
\mathbb{E} \{x_{t+1} z_t^T | y_{1:N}\} &\approx \sum_{i=1}^M \sum_{j=1}^M \tilde{\alpha}_{t|N}^{ij} \tilde{x}_{t+1}^j \left(\int z_t^T p(z_t | \tilde{x}_t^i, y_t) dz_t \right) \\
&= \sum_{i=1}^M \sum_{j=1}^M \tilde{\alpha}_{t|N}^{ij} \tilde{x}_{t+1}^j (\tilde{z}_t^i)^T
\end{aligned}$$

lo que completa la prueba. \square

Con el Lema 5.4 es posible caracterizar de forma completa la función $\mathcal{Q}(\theta, \hat{\theta}^k)$ en (5.6). Por consiguiente, en el Lema a continuación, nos enfocamos en el paso de maximización.

Lema 5.5 Si Σ satisface $\Sigma > 0$ y θ es particionado como $\theta = \begin{bmatrix} \beta \\ \eta \end{bmatrix}$, donde β parametriza Γ y μ_1 , y η parametriza Π y P_1 , como en (1.4). Entonces,

$$\hat{\beta} = \arg \max_{\theta \in \Theta} \mathcal{Q}(\theta, \hat{\theta}^k) \tag{5.20}$$

es dado por

$$\Gamma = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \Psi \Sigma^{-1}, \quad \mu_1 = \hat{x}_{1|N}. \tag{5.21}$$

Además, Π y P_1 dado por

$$\Pi = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} = \Phi - \Psi \Sigma^{-1} \Psi^T, \quad P_1 = P_{1|N}, \tag{5.22}$$

forman un punto estacionario de $\mathcal{Q}(\theta, \hat{\theta}^k)$ con respecto a η , y ambos Π y P_1 definidos en (5.22) son semidefinidos positivos garantizados.

Finalmente, si Π y P_1 dados por (5.22) son positivos definidos, la secuencia de entrada $\{u_t\}$ es persistentemente excitante (en el sentido de (5.32)) y $\hat{\theta}^k$ implica un sistema controlable y observable, entonces el punto Π, P_1 es más que un punto estacionario, es un máximo global de $\mathcal{Q}(\theta, \hat{\theta}^k)$ con respecto a η .

Demostración. Reescribiendo $\mathcal{Q}(\theta, \hat{\theta}^k)$ como

$$\begin{aligned}
\mathcal{Q}(\theta, \hat{\theta}^k) &= -\frac{1}{2} \log \det\{2\pi P_1\} - \frac{N}{2} \log \det\{2\pi \Pi\} \\
&\quad - \frac{1}{2} \text{trace} \{P_1^{-1} [P_{1|N} + (\hat{x}_{1|N} - \mu_1)(\hat{x}_{1|N} - \mu_1)^T]\} \\
&\quad - \frac{N}{2} \text{trace} \{\Pi^{-1} [(\Gamma - \Psi \Sigma^{-1}) \Sigma (\Gamma - \Psi \Sigma^{-1})^T + \Phi - \Psi \Sigma^{-1} \Psi^T]\},
\end{aligned} \tag{5.23}$$

es evidente que el máximo global con respecto a β es dado por (5.21).

Luego, obteniendo la derivada de $\mathcal{Q}(\theta, \hat{\theta}^k)$ con respecto a Π y P_1 , evaluada en (5.21),

$$\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \Pi^{-1}} = \frac{N}{2} \Pi - \frac{N}{2} [\Phi - \Psi \Sigma^{-1} \Psi^T] \quad (5.24)$$

$$\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial P_1^{-1}} = \frac{1}{2} P_1 - \frac{1}{2} P_{1|N} \quad (5.25)$$

Así, es claro que $\Pi = \Phi - \Psi \Sigma^{-1} \Psi^T$ y $P_1 = P_{1|N}$ es un punto estacionario. Además,

$$\begin{bmatrix} \Phi & \Psi^T \\ \Psi & \Sigma \end{bmatrix} = \mathbb{E} \left\{ \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix} \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix}^T \right\} \geq 0, \quad (5.26)$$

Por lo tanto, el punto estacionario Π dado por (5.22) es el complemento de Schur de Σ en $\mathbb{E} \left\{ \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix} \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix}^T \right\}$ y por ende, por construcción es no negativo.

Por otro lado, para demostrar que el punto estacionario es un máximo global, buscamos el Hessiano de $\mathcal{Q}(\theta, \hat{\theta}^k)$, así comenzamos obteniendo

$$\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \Gamma} = N \Pi^{-1} (\Psi - \Gamma \Sigma). \quad (5.27)$$

Usando las propiedades del producto de Kronecker, se obtiene

$$\frac{\partial \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \text{vec} \{\Gamma\}} = N \text{vec} \{ \Pi^{-1} \Psi \} - N (\Sigma \otimes \Pi^{-1}) \text{vec} \{ \Gamma \} \quad (5.28)$$

$$= N [(\Psi - \Gamma \Sigma)^T \otimes I] \text{vec} \{ \Pi^{-1} \}, \quad (5.29)$$

el que es independiente de μ y P_1 , por lo cual la derivada mixta con respecto a esos parámetros es cero, en cambio,

$$\left. \frac{\partial^2 \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \text{vec} \{\Gamma\} \partial \text{vec} \{\Gamma\}^T} \right|_{\theta = \hat{\theta}^{k+1}} = -N \Sigma \otimes \Pi^{-1} \quad (5.30)$$

y utilizando (5.21),

$$\left. \frac{\partial^2 \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \text{vec} \{\Gamma\} \partial \text{vec} \{\Pi\}^T} \right|_{\theta = \hat{\theta}^{k+1}} = (\Psi - \Psi \Sigma^{-1} \Sigma) \left. \frac{\partial \text{vec} \{ \Pi^{-1} \}}{\partial \text{vec} \{ \Pi \}^T} \right|_{\theta = \hat{\theta}^{k+1}} = 0.$$

Ahora calculamos las derivadas parciales de $\mathcal{Q}(\theta, \hat{\theta}^k)$ con respecto a los parámetros de Π

$$\begin{aligned} \frac{\partial^2 \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \text{vec} \{ \Pi \} \partial \text{vec} \{ \Pi \}^T} &= \frac{N}{2} \Pi^{-1} \otimes \Pi^{-1} \\ &\quad - \frac{N}{2} \left(([\Phi - \Psi \Gamma^T - \Gamma \Psi^T + \Gamma \Sigma \Gamma^T] \Pi^{-1})^T \otimes I \right) (\Pi^{-1} \otimes \Pi^{-1}) \\ &\quad - \frac{N}{2} (I \otimes (\Pi^{-1} [\Phi - \Psi \Gamma^T - \Gamma \Psi^T + \Gamma \Sigma \Gamma^T])) (\Pi^{-1} \otimes \Pi^{-1}). \end{aligned}$$

Sin embargo, de acuerdo a (5.21)

$$\Phi - \Psi\Gamma^T - \Gamma\Psi^T + \Gamma\Sigma\Gamma^T = \Phi - \Psi\Sigma^{-1}\Psi^T = \Pi,$$

por lo que,

$$\left. \frac{\partial^2 \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \text{vec}\{\Pi\} \partial \text{vec}\{\Pi\}^T} \right|_{\theta=\hat{\theta}^{k+1}} = -\frac{N}{2}\Pi^{-1} \otimes \Pi^{-1}.$$

Los elementos del Hessiano de $\mathcal{Q}(\theta, \hat{\theta}^k)|_{\theta=\hat{\theta}^{k+1}}$ asociados con los parámetros P_1 y μ son calculados de manera similar, obteniendo

$$\left. \frac{\partial^2 \mathcal{Q}(\theta, \hat{\theta}^k)}{\partial \theta \partial \theta^T} \right|_{\theta=\hat{\theta}^{k+1}} = \begin{bmatrix} -N\Sigma \otimes \Pi^{-1} & 0 & 0 & 0 \\ 0 & -P_{1|N} & 0 & 0 \\ 0 & 0 & -\frac{N}{2}\Pi^{-1} \otimes \Pi^{-1} & 0 \\ 0 & 0 & 0 & -\frac{1}{2}P_{1|N}^{-1} \otimes P_{1|N}^{-1} \end{bmatrix} < 0, \quad (5.31)$$

donde las condiciones P_1 positivo definido y que $\hat{\theta}^k$ implique un sistema controlable y observable son requeridas para demostrar que $P_{1|N}$ es positivo definido [76]. La relación (5.31) demuestra que Π y P_1 en (5.22) corresponden a un punto que maximiza $\mathcal{Q}(\theta, \hat{\theta}^k)$ globalmente, por ende $\hat{\theta}^{k+1}$ escogido por (5.21) y (5.22) es un máximo global de $\mathcal{Q}(\theta, \hat{\theta}^k)$. \square

La condición de $\{u_t\}$ persistentemente excitante en Lema 5.5 es requerida realmente para demostrar que la elección de θ^{k+1} está bien definida y es única, debido a la positividad de Σ y unicidad de su inversa. Esta relación se resume en el siguiente Teorema.

Teorema 5.1 Suponga que $\hat{\theta}^k$ parametriza un sistema controlable y observable con $\Pi, P_1 > 0$, y que para el largo de datos dados N , la secuencia de entrada $\{u_t\}$ satisface

$$\sum_{t=1}^N u_t u_t^T > 0. \quad (5.32)$$

Entonces, Σ en (5.11) es positivo definido y $\hat{\theta}^{k+1}$ es únicamente definido.

Demostración. Las siguientes definiciones

$$\begin{aligned} \mathcal{A} &\triangleq \frac{1}{N} \sum_{t=1}^N \hat{x}_{t|N} \hat{x}_{t|N}^T, & \mathcal{B} &\triangleq \frac{1}{N} \sum_{t=1}^N \hat{x}_{t|N} u_t^T, \\ \mathcal{C} &\triangleq \frac{1}{N} \sum_{t=1}^N u_t u_t^T, & \mathcal{D} &\triangleq \frac{1}{N} \sum_{t=1}^N P_{t|N}, \end{aligned}$$

permiten la expresión

$$\Sigma = \frac{1}{N} \sum_{t=1}^N \mathbb{E} \left\{ \begin{bmatrix} x_t \\ u_t \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix}^T \middle| y_{1:N} \right\} = \begin{bmatrix} \mathcal{A} + \mathcal{D} & \mathcal{B} \\ \mathcal{B}^T & \mathcal{C} \end{bmatrix}.$$

Evidentemente

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^T & \mathcal{C} \end{bmatrix} \geq 0,$$

ya que es el producto de $\begin{bmatrix} \hat{x}_{t|N} \\ u_t \end{bmatrix}$, por su traspuesto. Si además exigimos que $\mathcal{C} > 0$, entonces se cumple

$$\begin{bmatrix} \mathcal{A} - \mathcal{B}\mathcal{C}^{-1}\mathcal{B}^T & 0 \\ 0 & \mathcal{C} \end{bmatrix} = \begin{bmatrix} I & -\mathcal{B}\mathcal{C}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^T & \mathcal{C} \end{bmatrix} \begin{bmatrix} I & 0 \\ -\mathcal{C}^{-1}\mathcal{B}^T & I \end{bmatrix} \geq 0.$$

Luego, como argumentamos en la demostración del Lema 5.5, $\mathcal{D} > 0$. Entonces,

$$\begin{bmatrix} \mathcal{A} + \mathcal{D} - \mathcal{B}\mathcal{C}^{-1}\mathcal{B}^T & 0 \\ 0 & \mathcal{C} \end{bmatrix} > 0,$$

y así

$$\begin{bmatrix} \mathcal{A} + \mathcal{D} & \mathcal{B} \\ \mathcal{B}^T & \mathcal{C} \end{bmatrix} = \begin{bmatrix} I & \mathcal{B}\mathcal{C}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathcal{A} + \mathcal{D} - \mathcal{B}\mathcal{C}^{-1}\mathcal{B}^T & 0 \\ 0 & \mathcal{C} \end{bmatrix} \begin{bmatrix} I & 0 \\ \mathcal{C}^{-1}\mathcal{B}^T & I \end{bmatrix} > 0,$$

lo que completa la demostración. □

5.1 Resumen del Algoritmo a través de partículas

A continuación se resumen las ideas de este capítulo y se enlazan los cálculos necesarios desde los capítulos anteriores, a través de un algoritmo EM para el sistema (1.1), el que presenta su salida cuantizada.

Algoritmo 5.1 (EM para Sistemas con Salida Cuantizada)

- (1) Inicialice las estimaciones en $\hat{\theta}^0 = \left[\text{vec}\{\Gamma\}^T \quad \mu_1^T \quad \text{vec}\{\Pi\}^T \quad \text{vec}\{P_1\}^T \right]^T$ y fije $k = 0$.
 - (2) Usando la estimación $\hat{\theta}^k$, ejecute el Algoritmo 3.2 para obtener las secuencias $\{\tilde{\alpha}_t^i\}_{i=1}^M$ y $\{\tilde{x}_t^i\}_{i=1}^M$ que aproximan $p(x_t|y_{1:t})$, para $t = 1, \dots, N$.
 - (3) A través de $\hat{\theta}^k$, $\{\tilde{\alpha}_t^i\}_{i=1}^M$ y $\{\tilde{x}_t^i\}_{i=1}^M$ ejecute el Algoritmo 3.5 para obtener los pesos $\{\tilde{\alpha}_{t|N}^i\}_{i=1}^M$, que junto con las partículas $\{\tilde{x}_t^i\}_{i=1}^M$ aproximan $p(x_t|y_{1:N})$.
 - (4) Obtenga Φ , Ψ y Σ definidos en las ecuaciones (5.9) - (5.11), usando $\{\tilde{\alpha}_{t|N}^i\}_{i=1}^M$ y $\{\tilde{x}_t^i\}_{i=1}^M$ a través de métodos de Monte Carlo y el Lema 5.4.
 - (5) Maximice $\mathcal{Q}(\theta, \hat{\theta}^k)$ con respecto a θ escogiendo A , B , C , D , Q , R , P_1 y μ_1 de acuerdo a las ecuaciones (5.21) y (5.22) para obtener una nueva estimación $\hat{\theta}^{k+1}$.
 - (6) Si la secuencia θ^k convergió finalice, de otro modo, incremente $k \rightarrow k + 1$ y retorne al paso (2).
-

6

Simulaciones

Ejemplo 5: Sistema SISO de primer orden

Considere el sistema dado por

$$\begin{aligned}x_{t+1} &= 0.7x_t + 2u_t + w_t \\z_t &= 0.6x_t + u_t + v_t \\y_t &= \mathbb{q}\{z_t\},\end{aligned}\tag{6.1}$$

donde $\mathbb{q}\{\cdot\}$ corresponde al cuantizador representado en la Figura 3.3, $w_t \sim \mathcal{N}(0, 0.8)$ y $v_t \sim \mathcal{N}(0, 1)$.

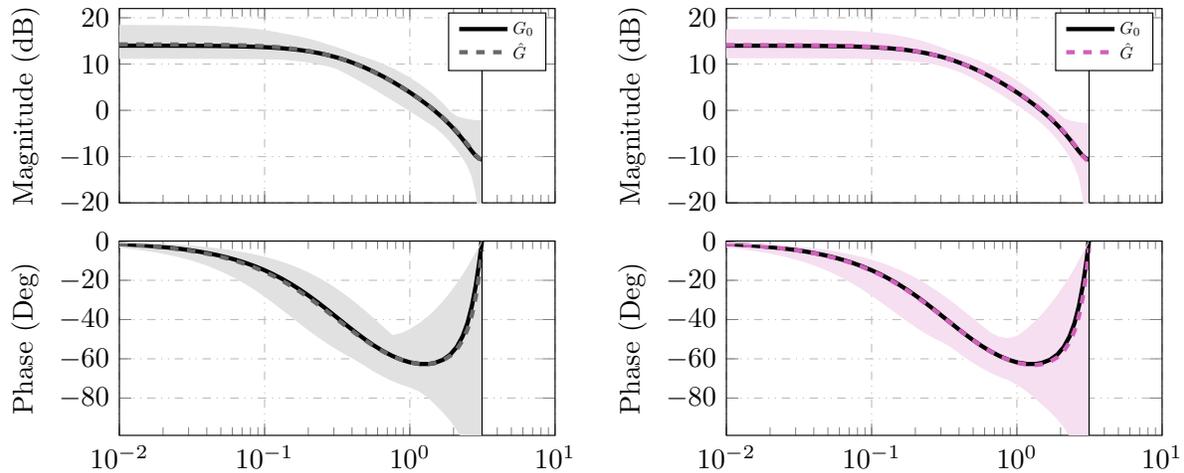
La simulación se ha ejecutado considerando $N = 1000$ y $u_{1:N}$ es una secuencia de ruido Gaussiano de media 0 y covarianza 0.1. Con los conjuntos de datos $u_{1:N}$ e $y_{1:N}$ se ejecutó el Algoritmo 5.1, el algoritmo en [76] que corresponde a un algoritmo EM usando Filtro y Suavizador de Kalman y el método de Subespacios [9]. Los últimos dos métodos no incorporan información de cuantización, y por lo tanto, suponen un sistema lineal e invariante en el tiempo escrito en espacio de estados de la forma clásica. En algunos casos se etiquetará el Algoritmo 5.1 como EM-PS, y el de [76] como EM-KS.

Tanto para el Algoritmo 5.1 como para el de [76] se consideró una estimación inicial $\hat{\theta}_0$ aleatoria escogida de tal forma que pertenezca a un intervalo igual al 50% de la correspondiente entrada en el vector de parámetros verdadero θ^* y se ejecutaron 300 iteraciones del algoritmo EM.

Para el estudio de Monte Carlo se realizaron 200 realizaciones, tanto de ruidos como de la estimación inicial θ_0 .

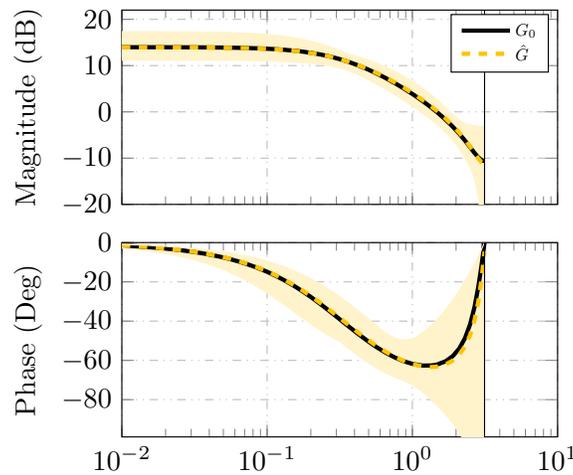
La Figura 6.1 muestra el diagrama de Bode para los tres algoritmos ejecutados. En línea continua negra se muestra el diagrama de Bode del sistema real, el área sombreada representa el espacio que abarcaron las 200 realizaciones, y por ende se puede interpretar como una medida de varianza del estimador, y finalmente, en línea segmentada se muestra la media de las 200 realizaciones. Las gráficas muestran claramente como todos los métodos logran estimar de

buena forma el sistema, obteniendo incluso área sombreada muy similares para las distintas realizaciones.



(a) Método de Subespacios.

(b) Algoritmo EM usando Kalman Smoother.



(c) Algoritmo EM usando Particle Smoother.

Figura 6.1: Diagrama de Bode de 200 realizaciones, donde se muestra la respectiva media con línea segmentada.

En el Lema B.1 del Apéndice B se muestra que la representación en espacio de estados no es única. Debido a esta característica, se graficó el diagrama de Bode en lugar de analizar cada parámetro por separado. Sin embargo, las características de los ruidos no se abordan aquellos diagramas, por esto, la Figura 6.2 muestra un *Box Plot* con los resultados obtenidos por las 200 estimaciones de la covarianza del ruido de medición R , la cual es invariante a transformaciones de similitud (no así la covarianza del ruido de proceso Q), como se muestra en el Lema B.1.

En la Figura 6.2, se muestra claramente como el método desarrollado presenta una gran ventaja en cuanto a la estimación de la covarianza del ruido de medición.

Adicionalmente, se ha obtenido el error de estimación relativo e medido por la norma dos

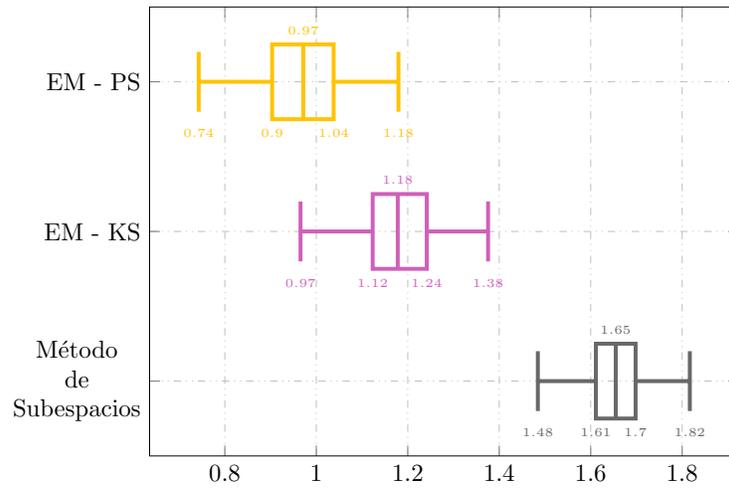


Figura 6.2: Boxplot de la estimación de la covarianza del ruido de medición R , obtenida para las 200 realizaciones.

de la diferencia entre el sistema verdadero $G(\omega)$ y la estimación $\hat{G}(\omega)$, es decir,

$$e = \frac{\int_{-\pi}^{\pi} \text{trace} \left\{ \left\| G(\omega) - \hat{G}(\omega) \right\|^2 \right\} d\omega}{\int_{-\pi}^{\pi} \text{trace} \left\{ \left\| G(\omega) \right\|^2 \right\} d\omega}. \quad (6.2)$$

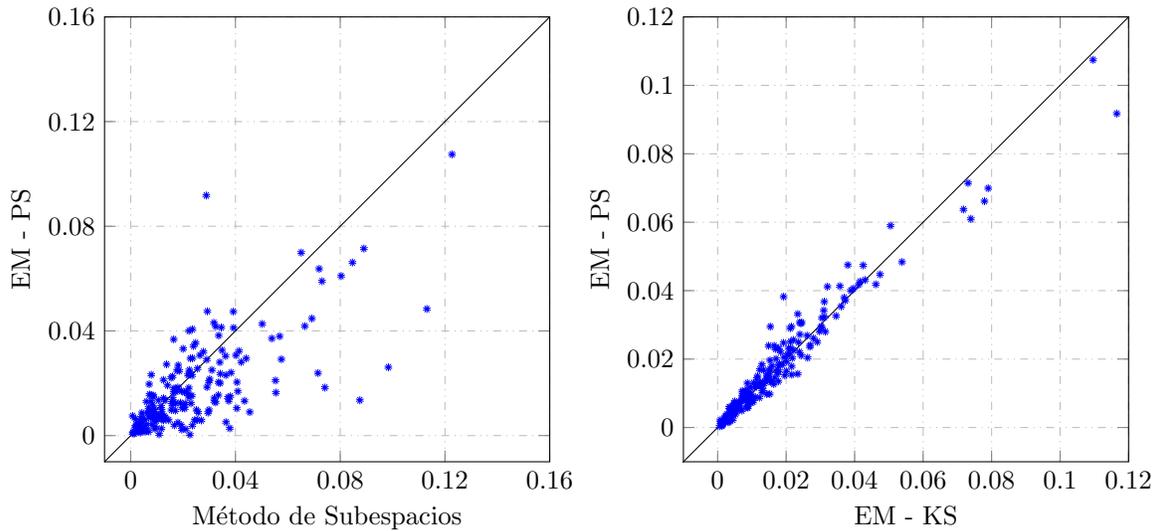
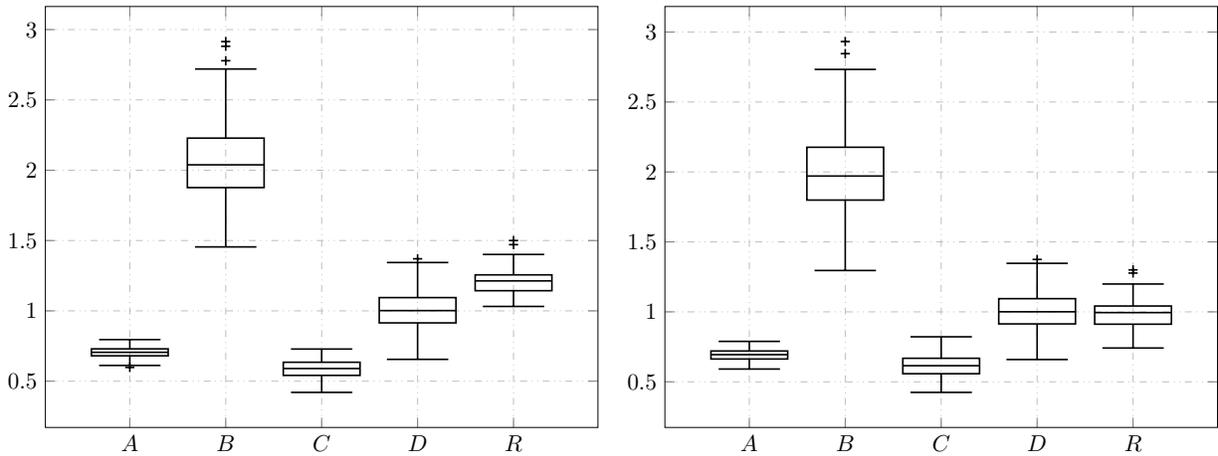


Figura 6.3: Error de estimación relativo.

A la izquierda de la Figura 6.3 se muestra el error e en (6.2), generado por la estimación del Algoritmo 5.1 versus el error obtenido por la estimación a través del método de Subespacios. A la derecha de la Figura 6.3 se presenta el error generado por la estimación del Algoritmo



(a) Algoritmo EM usando Kalman Smoother. (b) Algoritmo EM usando Particle Smoother.

Figura 6.4: Boxplot de los parámetros fijando Q al valor verdadero.

5.1 versus el que entrega la estimación obtenida por el algoritmo EM usando Suavizador de Kalman. A pesar de que no se observan grandes diferencias entre los tres algoritmos, se aprecia que el método de Subespacios posee una leve tendencia a mayores errores.

Debido a la condición ya mencionada de que la representación en espacio de estados de un sistema no es única y con el fin de evaluar la estimación de parámetros en el presente caso de un sistema de primer orden, se han ejecutado ambos algoritmos EM fijando la covarianza del ruido de proceso Q al valor verdadero. De esta forma se debe fijar T en (B.1), y así, el algoritmo debe converger a la representación utilizada en la simulación. En este caso, se han realizado 500 iteraciones del algoritmo EM, debido a que el algoritmo debe converger a una representación en particular. Además, se han mantenido las 200 realizaciones.

En la Figura 6.4 se muestran los *Box Plot* de todos los parámetros para ambos métodos. En 6.4a el resultado usando EM con Suavizador de Kalman y en 6.4b usando EM con Suavizador

	Parámetro	<i>Lower Whisker</i>	Q_1	Mediana (Q_2)	Q_3	<i>Upper Whisker</i>	Media
EM-KS	A	0.6120	0.6805	0.7052	0.7298	0.7958	0.7039
	B	1.4547	1.8764	2.0383	2.2282	2.7196	2.0633
	C	0.4205	0.5412	0.5900	0.6347	0.7285	0.5876
	D	0.6552	0.9145	1.0019	1.0943	1.3437	1.0047
	R	1.0315	1.1433	1.2131	1.2560	1.4012	1.2057
EM-PS	A	0.5913	0.6629	0.6942	0.7202	0.7889	0.6913
	B	1.2956	1.7990	1.9707	2.1765	2.7329	2.0007
	C	0.4245	0.5578	0.6151	0.6683	0.8217	0.6134
	D	0.6592	0.9138	1.0005	1.0944	1.3470	1.0047
	R	0.7415	0.9119	0.9947	1.0418	1.1991	0.9823

Tabla 6.1: Estadísticas asociadas a los Boxplot de la Figura 6.4.

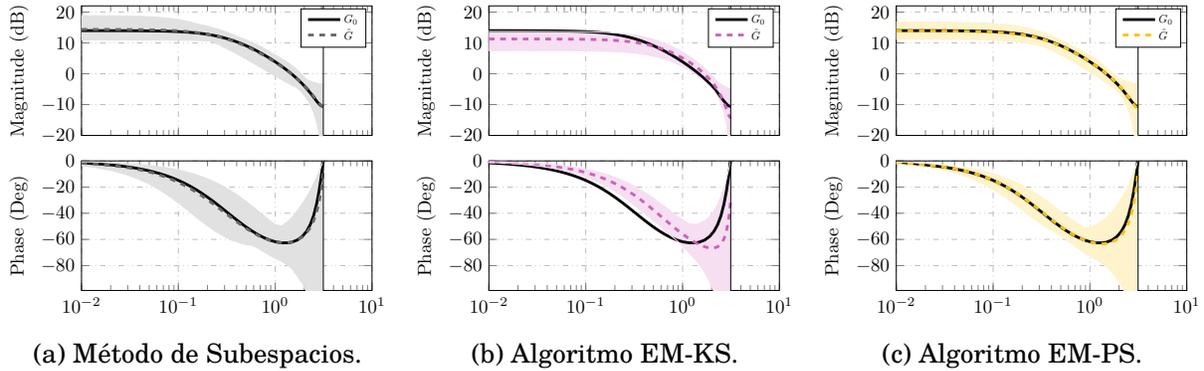


Figura 6.5: Diagrama de Bode de 200 realizaciones, con salida cuantizada en 2 bits.

de Partículas. Además, en la Tabla 6.1 se encuentran las estadísticas asociadas a los *Box Plot* de la Figura 6.4.

Se destaca que efectivamente al fijar uno de los parámetros sujetos a una transformación de similitud se puede obtener la representación en espacio de estados buscada. De esta forma, es posible evaluar cada parámetro (para el caso particular de sistemas de primer orden). Tanto la Figura 6.4, como la Tabla 6.1, muestran que ambos métodos se comportan de buena manera estimando los parámetros del sistema A , B , C y D , notando leves diferencias en B y C , que son justamente los otros parámetros que están sujetos transformación.

Por otro lado, se observa nuevamente una gran diferencia en la estimación de la covarianza del ruido de medición R . El Algoritmo 5.1 obtiene sin dudas mejores resultados que el método tradicional. Además, se aprecia en este caso una mejor estimación aún, que cuando se estimaron todos los parámetros, pasando de una mediana de 0.9716 a una de 0.9947, bastante cercano al valor verdadero de 1.

Es necesario comentar que en las simulaciones se han presentado dificultades en la estimación de la covarianza del ruido de medición R cuando el nivel de cuantización crece. En la práctica, dicha covarianza se acerca a cero a través de las iteraciones del algoritmo EM, lo que produce que los demás parámetros también se alejen del valor verdadero. A pesar de intentar establecer un mínimo a la covarianza R para que no llegue a cero, los demás parámetros se alejan del valor real cuando R a disminuido demasiado.

Debido a esta dificultad y con la intención de ver la diferencia de los métodos para cuantizaciones más gruesas, se han realizado dos simulaciones con niveles de cuantización mayores considerando las covarianzas de ambos ruidos como conocidas. En el primer caso se considera Δ de tal forma de utilizar 2 bits para representar la salida, y en el segundo caso 1 bits.

En la Figura 6.5, se muestra el diagrama de Bode obtenido por los 3 métodos en las 200 realizaciones ejecutadas. Es posible observar que el Algoritmo EM usando Suavizador de Kalman es el que más dista del sistema real, teniendo un sesgo en magnitud y en fase para la mayoría de las frecuencias. Luego, el método de Subespacios presenta un buen desempeño al observar la media de las realizaciones, sin embargo, se aprecia un crecimiento considerable en la varianza del estimador. Por el contrario, se aprecia que el algoritmo desarrollado en este trabajo no presenta mayor variación, ya sea en cuanto a media, como en varianza.

En la Figura 6.6 se muestra nuevamente el error definido en (6.2), donde ahora se evidencia una diferencia clara entre el método propuesto y los otros dos métodos utilizados. Los puntos

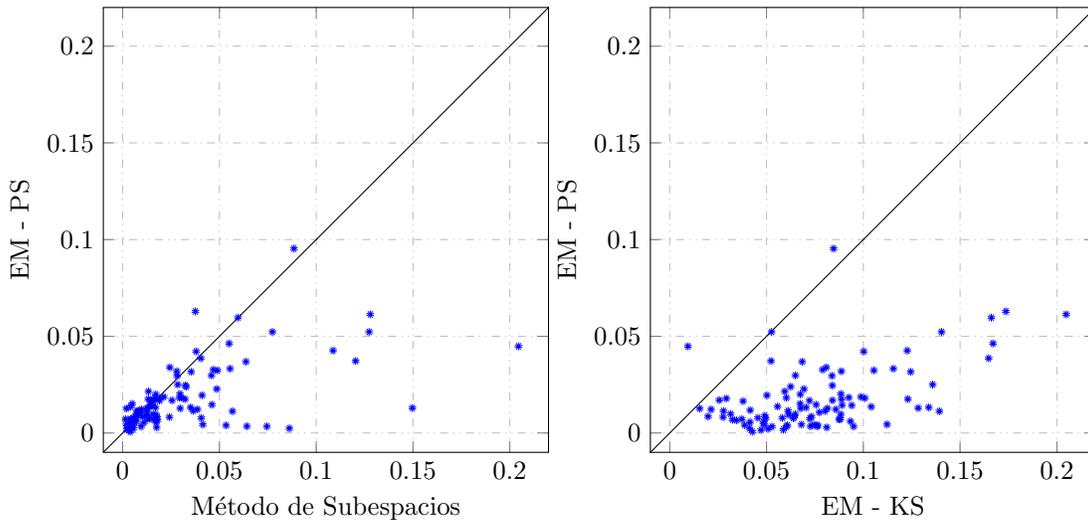


Figura 6.6: Error de estimación relativo usando 2 bits para la salida.

en los gráficos presentan una clara tendencia hacia abajo y la derecha, lo que implica un menor error para el método propuesto. Además, se aprecia que el algoritmo EM usando Suavizador de Kalman presenta errores mayores incluso que el método de Subespacios.

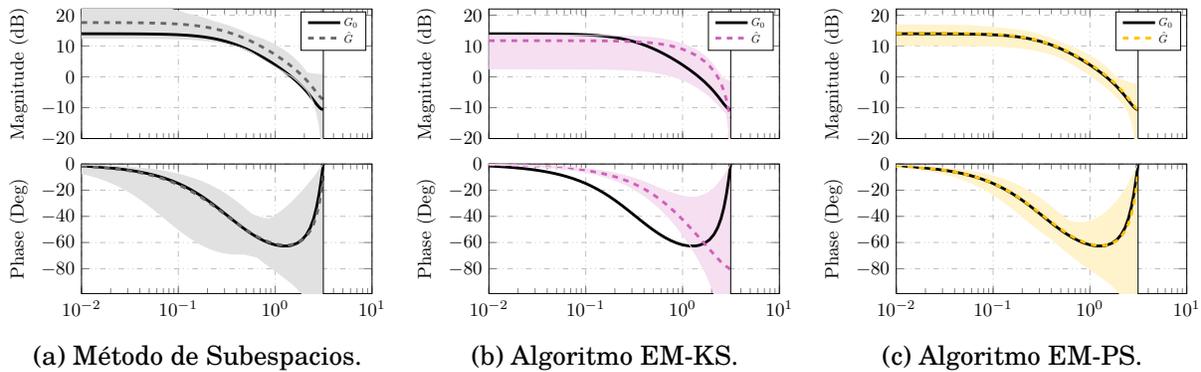


Figura 6.7: Diagrama de Bode de 200 realizaciones, con salida cuantizada en 1 bits.

Como último caso de estudio del sistema en (6.1), la Figura 6.7 se muestra el diagrama de Bode obtenido con los tres métodos utilizados cuando la salida ha sido cuantizada usando 1 bit. En este caso se aprecia un sesgo evidente en la fase obtenida por el algoritmo EM usando Suavizador de Kalman y un desmedro evidente en la magnitud, donde la magnitud obtenida es menor para frecuencias bajas y mayor para frecuencias altas. Por el contrario, los resultados obtenidos a través del método de Subespacio muestran que la media de la fase representa de gran manera el sistema real, aunque con un gran aumento en la varianza. Sin embargo, el resultado en magnitud es evidentemente peor, ya que, presenta un sesgo constante para todas las frecuencias. Por otro lado, el Algoritmo 5.1 no presenta complicaciones, representando tanto la magnitud como la fase del sistema real de gran manera.

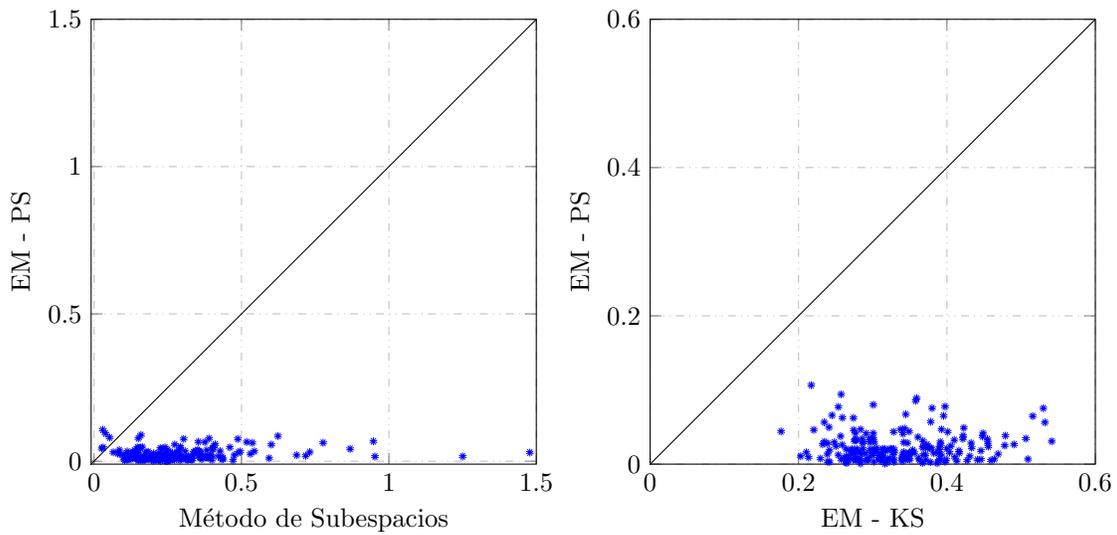


Figura 6.8: Error de estimación relativo usando 1 bits para la salida.

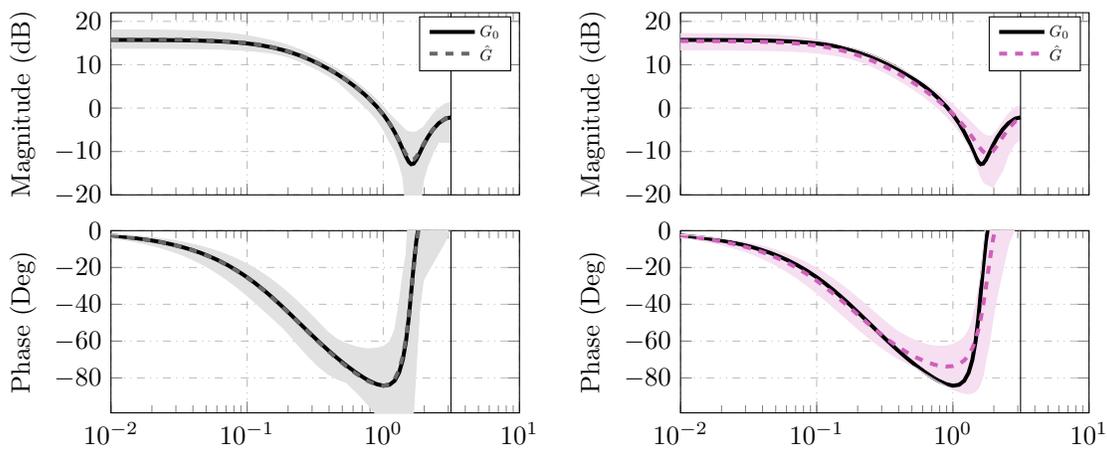
Por último, la Figura 6.8 que muestra el error en (6.2) para la salida usando 1 bits corrobora las apreciaciones obtenidas a través de los diagramas de Bode y muestra que la ventaja del método propuesto es evidente. Además, también muestra que el método de Subespacios tiene un peor desempeño que el conseguido con algoritmo EM usando Suavizador de Kalman, esto debido a que el error definido considera solo la magnitud de los sistemas.

Ejemplo 6: Sistema SISO de segundo orden

Considere el sistema dado por

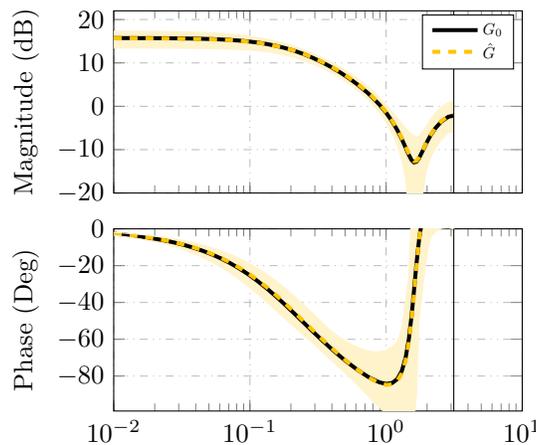
$$\begin{aligned} x_{t+1} &= \begin{bmatrix} 0.8 & 0 \\ 0.3 & -0.1 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ -0.8 \end{bmatrix} u_t + w_t \\ z_t &= [1 \quad 0.5] x_t + 0.8u_t + v_t \\ y_t &= \mathfrak{q}\{z_t\} \end{aligned} \tag{6.3}$$

donde $\mathfrak{q}\{\cdot\}$ corresponde al cuantizador representado en la Figura 3.3, $w_t \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$ y $v_t \sim \mathcal{N}(0, 2)$. La simulación consideró $N = 1000$, $u_{1:N}$ una secuencia de ruido blanco Gaussiano de media 0 y covarianza 1, y una estimación inicial escogida de la misma forma que en el ejemplo anterior.



(a) Método de Subespacios.

(b) Algoritmo EM usando Suavizador de Kalman.



(c) Algoritmo EM usando Suavizador de Partículas.

Figura 6.9: Diagrama de Bode de 200 realizaciones, donde se muestra la respectiva media con línea segmentada, en sistema de segundo orden.

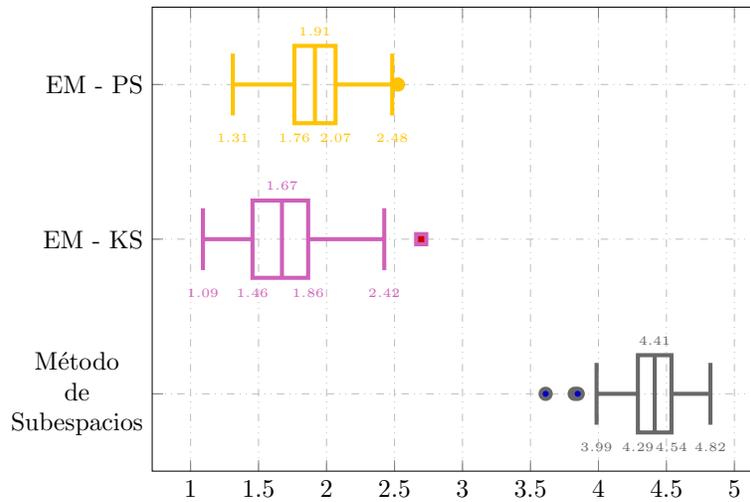


Figura 6.10: Boxplot de la estimación de la covarianza del ruido de medición R , obtenida para las 200 realizaciones, en sistema de segundo orden.

Al tener un sistema de segundo orden no es posible analizar la estimación de parámetros, por lo que se analiza la respuesta en frecuencia, en la Figura 6.9.

A través de la Figura 6.9, se puede ver como el mejor resultado se obtiene a través del método desarrollado. Sin embargo, el método de Subespacios logra un resultado bastante similar, con la media de las 200 realizaciones superpuesta al sistema real y presentando apenas un aumento en la varianza con respecto al resultado obtenido con el método aquí desarrollado. Por otro lado, el algoritmo EM usando Suavizador de Kalman presenta errores evidentes en algunas frecuencias, con sesgos que producen que se convierta en el peor resultado de los tres algoritmos utilizados.

La estimación de la covarianza del ruido de medición R , se presenta en el *Box Plot* obtenido en la Figura 6.10, donde se aprecia claramente que el resultado más cercano al valor real corresponde al aquí presentado.

Una resultado destacable es que el algoritmo EM usando Suavizador de Kalman estima una covarianza \hat{R} menor al valor real. En contraste a lo que se pensaría, ya que al no disponer de información del proceso de cuantización, esta podría ser interpretada como un aumento en el ruido de medición, tal como sucedió en el Ejemplo 5 y con el resultado que muestra en este caso el método de Subespacios.

Por otro lado, aunque no es posible analizar el sistema de forma paramétrica debido a las infinitas representaciones existentes, es posible analizar los polos del sistema, los que, como se presenta en el Apéndice B, son invariantes a transformaciones de similaridad. En la Figura 6.11, se muestra la estimación de ambos polos obtenida a través de los tres métodos y en las 200 realizaciones ejecutadas. Cada método presenta su respectivo color, mientras la media obtenida por cada método se muestra con un marcador de mayor tamaño. Nuevamente es evidente en que el Algoritmo EM usando Suavizador de Kalman presenta el mayor error en la estimación de los polos, presentando un sesgo evidente, el que es mayor para el polo 2. Luego, el resultado obtenido por el método de Subespacios es similar al del método propuesto, entregando ambos métodos una estimación bastante cercana a los polos verdaderos, presentando una varianza

levemente menor para el Algoritmo 5.1.

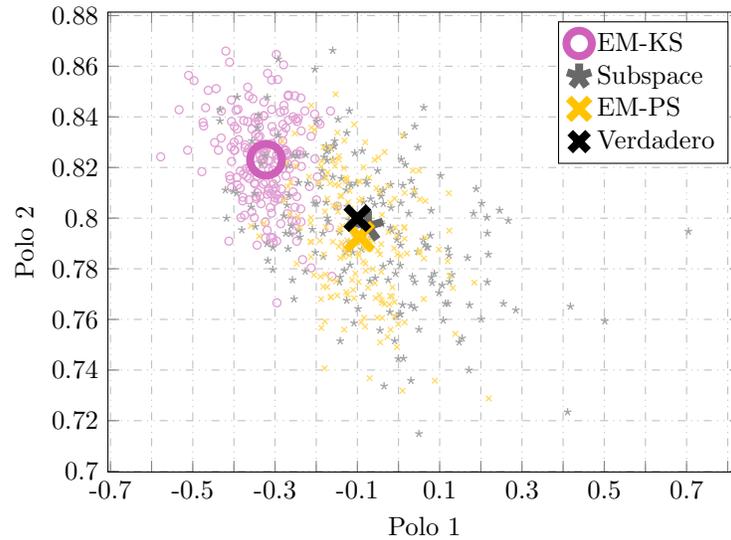


Figura 6.11: Estimación de ambos polos para 200 realizaciones, con su respectiva media.

Por último, al igual que que en el ejemplo anterior, en la Figura 6.12 se presenta el error relativo calculado a través de (6.2) con el método propuesto versus el obtenido con los otros dos métodos. La Figura 6.12, muestra a la izquierda como el error es bastante similar usando el Algoritmo 5.1 y el método de Subespacios, presentando una pequeña tendencia a errores mayores para éste último. A la derecha por el contrario, se aprecia una clara tendencia a mayores errores para el método EM usando Suavizador de Kalman. Por esto, la Figura 6.12 evidencia un mejor desempeño para el método desarrollado, entre los tres utilizados.

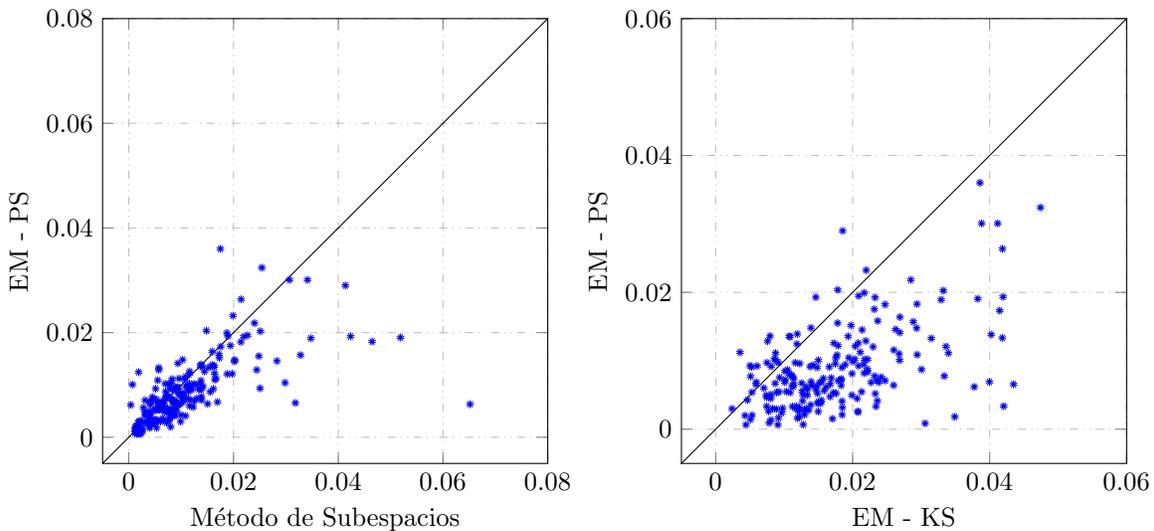


Figura 6.12: Error de estimación relativo, en sistema de segundo orden.

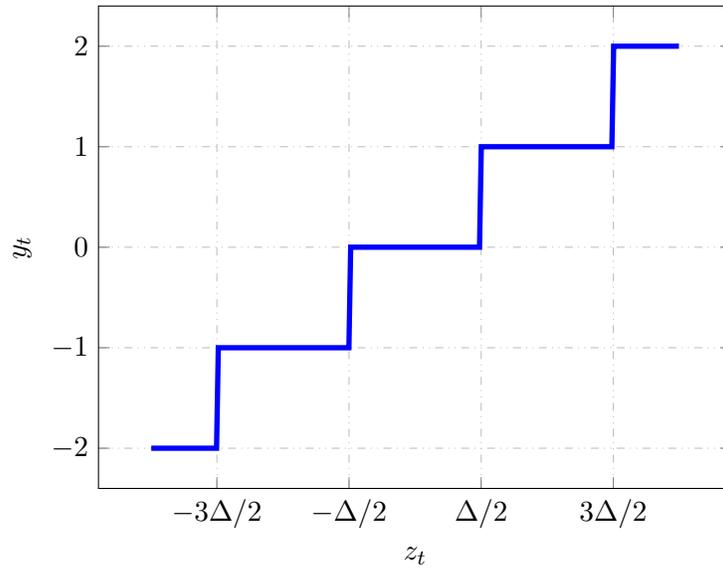


Figura 6.13: Cuantizador con pendiente, utilizado en el Ejemplo 7.

Ejemplo 7: Cuantizador con Pendiente distinta de 1

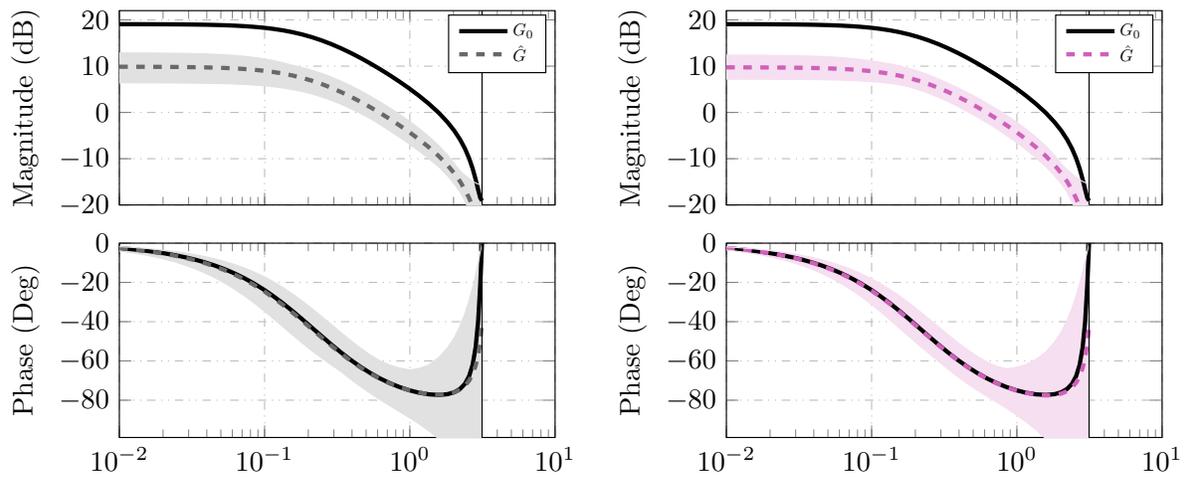
Considere el sistema dado por

$$\begin{aligned}
 x_{t+1} &= 0.8x_t + 1.6u_t + w_t \\
 z_t &= x_t + u_t + v_t \\
 y_t &= \mathfrak{q}\{z_t\},
 \end{aligned} \tag{6.4}$$

donde $\mathfrak{q}\{\cdot\}$ corresponde al cuantizador representado en la Figura 6.13, $w_t \sim \mathcal{N}(0, 2)$ y $v_t \sim \mathcal{N}(0, 4)$. Nótese que el cuantizador de la Figura 6.13 genera un mapa entre los intervalos Ω_ℓ y los números enteros. Por el contrario, el cuantizador de la Figura 3.3, genera un mapa que tiene como conjunto de llegada el centro de cada intervalo. Esta diferencia en el cuantizador, podría interpretarse como un cuantizador con pendiente distinta de 1 al comparar ambas gráficas.

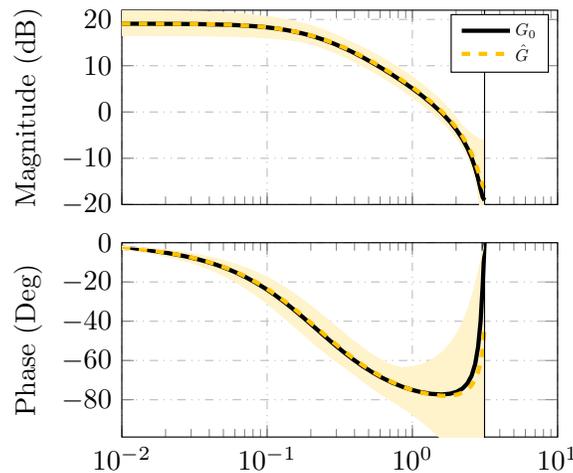
Para los efectos de simulación se ha utilizado $N = 1000$, 300 iteraciones del algoritmo EM y secuencia $u_{1:N}$ como ruido blanco de media 0 y covarianza 0.3. Además, la estimación inicial θ_0 se ha generado de la misma forma que en los ejemplos anteriores.

En este se muestran las gráficas usuales y se busca mostrar el rendimiento ante un cuantizador distinto. Es evidente que si la señal cuantizada es cercana o del orden de la señal no cuantizada el resultado a través de un método que no considere la cuantización no es necesariamente malo. Sin embargo, si la forma de cuantización es distinta, es decir, se escogen los valores ν_ℓ en (1.1c) de una forma no homogénea, el resultado podría empeorar de forma evidente. Por el contrario, el método desarrollado no debería presentar una gran degeneración ya que utiliza el mapa entre Ω_ℓ y ν_ℓ .



(a) Método de Subespacios.

(b) Algoritmo EM usando Kalman Smoother.



(c) Algoritmo EM usando Particle Smoother.

Figura 6.14: Diagrama de Bode de 200 realizaciones, donde se muestra la respectiva media con línea segmentada.

La Figura 6.14 muestra los diagramas de Bode obtenidos por los tres métodos, donde se puede observar resultados en general similares a los anteriores con la gran diferencia de la magnitud en los métodos que no consideran la cuantización. Esto viene directamente del tipo de cuantizador utilizado, y que el espacio de llegada del cuantizador no es comparable en magnitud con el que tiene la señal no cuantizada. Así, uno esperaría que la magnitud sea mayor o menor, dependiendo de si la señal cuantizada es mayor o menor respecto a la no cuantizada.

La Figura 6.15 muestra el *Box Plot* de la covarianza R , donde se aprecia que los métodos usados para comparar entregan una estimación totalmente sesgada y menor al valor real, lo que se justifica por la contracción que sufrió la señal de salida al pasar por el cuantizador.

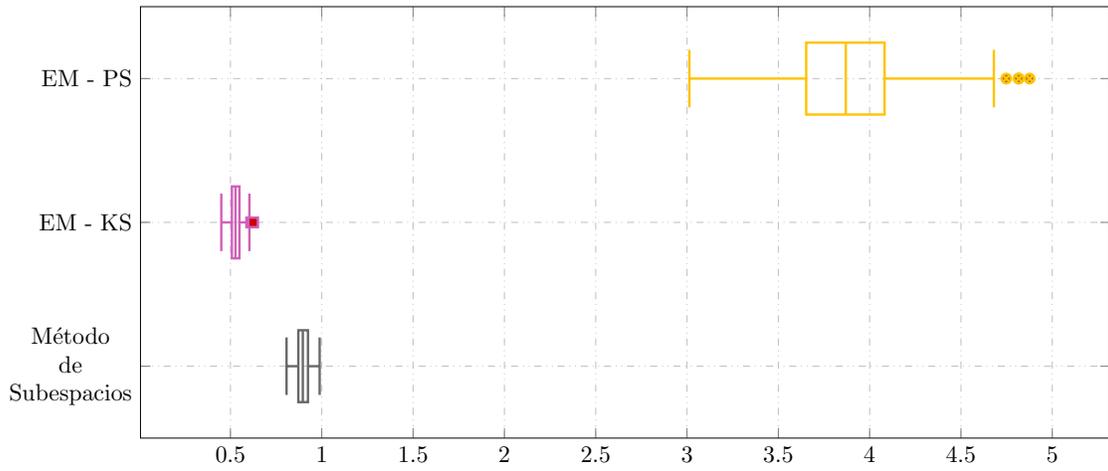


Figura 6.15: Boxplot de la estimación de la covarianza del ruido de medición R , obtenida para las 200 realizaciones.

Por último, la Figura 6.16 muestra la comparación de errores relativos con respecto al sistema verdadero, donde el resultado es claro. Debido al sesgo obtenido en la magnitud del sistema. Como se vio en el diagrama de Bode, el error calculado de la forma (6.2) es ampliamente mayor para los métodos de Subespacios y EM usando Suavizador de Kalman.

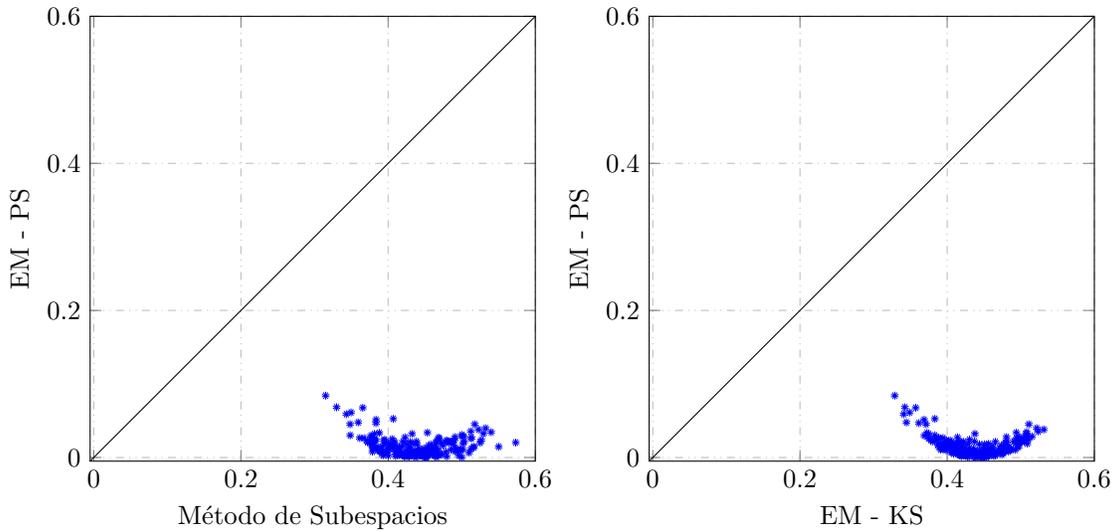


Figura 6.16: Error de estimación relativo.

7

Conclusiones

En este trabajo de Tesis se planteó un problema común en los tiempos actuales debido a las restricciones de comunicación. Éste es el de identificación de un sistema lineal e invariante en el tiempo, representado en espacio de estados y donde la señal de salida experimenta un proceso de cuantización, limitando así la información disponible. Dicho problema de identificación se abordó con el objetivo de obtener el estimador de Máxima Verosimilitud, a través del algoritmo EM.

Como objetivo preliminar se planteó un problema de estimación Bayesiana, con el fin de obtener una estimación de la función de densidad que resuelve los problemas de Filtraje y Suavizado de un sistema en espacio de estados. Dicho problema requiere resolver las respectivas ecuaciones Bayesianas de Filtraje y Suavizado no lineal, pues la salida del sistema lineal no está disponible como medición, sino que atraviesa un proceso de cuantización, lo que introduce una no linealidad en las ecuaciones.

Para el desafío de obtener un algoritmo de Filtraje se caracterizó la función de masa de probabilidad de las mediciones dado el estado, lo que permitió en primer lugar, implementar un algoritmo basado en partículas. Y además, se logró una aproximación natural de dicha función, para desarrollar un algoritmo basado en suma de Gaussianas, que permite caracterizar cualquier función de densidad de probabilidad.

Para complementar el problema de estimación Bayesiana se implementó un algoritmo de Suavizado basado en partículas, utilizando la caracterización de las funciones de densidad de probabilidad del sistema de interés.

La solución e implementación de estimación Bayesiana del sistema a través de Filtraje y Suavizado se puso a prueba a través de simulaciones que demostraron que el disponer de algoritmos que consideren la pérdida de información introducida por la cuantización consiguen mejorar la estimación de estados de forma significativa, especialmente cuando el nivel de cuantización es grande. Además, es posible caracterizar la varianza de la estimación en forma adecuada, lo que no se consigue cuando se considera un algoritmo para sistemas lineales estándar, como el Filtro de Kalman.

Para el caso de los algoritmos de Filtrajes se pudo evaluar el impacto en la calidad de la

estimación tanto para diferente número de Gaussianas, como número de partículas. Luego, para el caso del algoritmo de Suavizado se observó como éste contribuye con una mejoría en la estimación y una reducción en la varianza de la estimación con respecto a los algoritmos de Filtraje.

Posteriormente se realizó un estudio de las bondades del estimador de Máxima Verosimilitud y de la teoría del algoritmo EM para obtener el estimador señalado, lo que sirvió como motivación del por qué abordar el problema desde allí.

Luego, se desarrolló el algoritmo EM para obtener el estimador de Máxima Verosimilitud que responde al problema fundamental de este trabajo de Tesis. El algoritmo planteado, hace uso de estadísticas obtenidas a través del Suavizador implementado, y además, utiliza dichos resultados para obtener momentos asociados a la variable de salida no cuantizada, la que no es medible en el presente problema.

Finalmente, el algoritmo desarrollado se evaluó a través de simulaciones y los resultados se compararon, tanto con el estimador de Máxima Verosimilitud obtenido a través del algoritmo EM para un sistema lineal en variables de estados sin cuantización, como con el método de Subespacios. Los resultados mostraron que para niveles de cuantización pequeña, los tres algoritmos presentan resultados similares, aunque el método presentado estima de mejor forma las características del ruido de medición. Al aumentar el nivel de cuantización se presentaron problemas con la estimación de la covarianza, sin embargo, al suponer las covarianzas de los ruidos conocidas, se mostró claramente como el método desarrollado consigue mejores resultados en la estimación del sistema. Además, entre los métodos usados para comparar no se evidenció uno que lograra mejores resultados de forma constante. Intercambiando el logro de mejor desempeño dependiendo el caso de simulación. Esto indica una clara ventaja del método presentado, el que consiguió resultados al menos iguales (o mejores) que el mejor de los métodos usados para comparar, en todos los casos simulados.

Adicionalmente, se consideró un cuantizador con pendiente, el que responde a la suposición de un tipo de cuantización no usual y por ende, podría ser reemplazado por algún otro. En este caso, a pesar de usar un nivel de cuantización pequeña, el método propuesto consigue una estimación evidentemente mejor. Esto se debe al conocimiento del cuantizador, lo que permite utilizar la información del intervalo al que pertenece la medición sin cuantizar..

7.1 Trabajo Futuro

Sin duda un trabajo de Tesis culmina con una infinidad de puertas que muestran caminos de trabajos futuros. Aquí nombramos sólo algunas ideas de las extensiones que se podrían realizar de esta Tesis.

Una extensión natural es resolver el problema de Suavizado a través de una Suma de Gaussianas, usando el desarrollo obtenido para el Filtro por Suma de Gaussianas. Esto debería permitir obtener las estadísticas de Suavizado necesarias para el algoritmo EM a través de una sumatoria de pocas componentes, comparada al número de partículas, tal como sucedió con la implementación del filtro. Para esto, se puede abordar el problema usando la *Generalized Two-Filter Formula*, presentada en la Sección 2.2.2 y basados en la teoría en [77, 78].

En [59] se presenta un algoritmo de Filtraje por Suma de Gaussianas acompañado de una implementación numéricamente estable basada en las mismas ideas utilizadas para la implementación robusta del Filtro de Kalman estándar [76], por lo que se podría extender el algoritmo desarrollado en esta Tesis para obtener una implementación del mismo estilo. Además, una

extensión de Suavizado también podría buscar una implementación numéricamente estable.

Al obtener los trabajos futuros mencionados, y utilizarlos para el algoritmo EM presentado en este trabajo, se podría estudiar la implicancia de implementaciones numéricamente estables en el problema encontrado en este trabajo al estimar la covarianza del ruido de proceso con niveles de cuantización grandes.

Por otro lado, otra extensión del trabajo viene de considerar perturbaciones no Gaussianas. Para el caso de que dichas perturbaciones sean conocidas y por lo tanto estén caracterizadas, las soluciones de Filtraje y Suavizado basadas en partículas son naturales para calcular estadísticas requeridas. Si por el contrario, la caracterización de las densidades de los ruidos es desconocida, se puede suponer una densidad de suma de Gaussianas, y estimar los parámetros de la misma, dentro del problema de identificación. Dicho enfoque se puede realizar de igual forma a través de partículas, para lo que se debe suponer la estructura del ruido, por ejemplo, como suma de Gaussianas y muestrear estas distribuciones para propagar las partículas. En dicho caso, disponer de un algoritmo directamente basado en suma de Gaussianas es la solución ideal.

Por último, es evidente que las bases de los enfoques utilizados en este trabajo vienen de la teoría de Filtraje y Suavizado de sistemas no lineales, por lo que, una extensión a sistemas no lineales de características similares está dentro de un alcance futuro. Para esto, se debe estudiar la derivación del algoritmo EM del sistema en particular.

A

Propiedades de Sumatorias y Distribuciones Gaussianas

Lema A.1 Dos sumatorias anidadas, se pueden escribir como una sola sumatoria con un nuevo índice de la forma

$$\sum_{i=1}^I \sum_{j=1}^J g_{ij} = \sum_{r=1}^R h_r \quad (\text{A.1})$$

donde

$$g_{ij} = h_r, \quad R = IJ \quad \text{y} \quad r = (i-1)J + j \quad (\text{A.2})$$

Demostración. Expandiendo la doble sumatoria se tienen los siguientes índices

$$\begin{aligned} i = 1 \quad j = 1, \dots, J &\rightarrow r = 1, \dots, J \\ i = 2 \quad j = 1, \dots, J &\rightarrow r = J + 1, \dots, 2J \\ &\vdots \\ i = I \quad j = 1, \dots, J &\rightarrow r = (I-1)J + 1, \dots, IJ \end{aligned}$$

lo que muestra la correspondencia entre índices y coeficientes. □

Lema A.2 Si $p(x) = \mathcal{N}_x(\mu_1, P_1)$ y $p(y|x) = \mathcal{N}_y(Cx + \mu_2, P_2)$, entonces

- $p(x, y) = \mathcal{N}_y(\mu_3, P_3)\mathcal{N}_x(\mu_4, P_4)$

- $\int p(x, y)dx = \mathcal{N}_y(\mu_3, P_3)$

donde

$$\begin{aligned}\mu_3 &= C\mu_1 + \mu_2 \\ P_3 &= P_2 + CP_1C^T\end{aligned}$$

y

$$\begin{aligned}K &= P_1C^TP_3^{-1} \\ \mu_4 &= \mu_1 + K(y - \mu_3) \\ P_4 &= (I - KC)P_1.\end{aligned}$$

Demostración. Comenzamos encontrando la densidad conjunta,

$$\begin{aligned}p(x, y) &= \mathcal{N}_x(\mu_1, P_1)\mathcal{N}_y(Cx + \mu_2, P_2) \\ &= \frac{1}{\sqrt{\det\{2\pi P_1 P_2\}}} \exp\left\{-\frac{1}{2}L\right\},\end{aligned}$$

donde hemos llamado L a la suma de las formas cuadráticas. Y nos enfocamos en desarrollarlas como sigue

$$\begin{aligned}L &= (x - \mu_1)^T P_1^{-1}(x - \mu_1) + (y - Cx - \mu_2)^T P_2^{-1}(y - Cx - \mu_2) \\ L &= (x - \mu_1)^T P_1^{-1}(x - \mu_1) + x^T C^T P_2^{-1} Cx - x^T C P_2^{-1}(y - \mu_2) - (y - \mu_2)^T P_2^{-1} Cx \\ &\quad + (y - \mu_2)^T P_2^{-1}(y - \mu_2),\end{aligned}$$

factorizando algunos terminamos cuadráticos

$$\begin{aligned}L &= (x - \mu_1)^T (P_1^{-1} + C^T P_2^{-1} C) (x - \mu_1) + x^T C^T P_2^{-1} C\mu_1 + \mu_1^T C^T P_2^{-1} Cx - \mu_1^T C^T P_2^{-1} C\mu_1 \\ &\quad - x^T C^T P_2^{-1}(y - \mu_2) - (y - \mu_2)^T P_2^{-1} Cx + (y - \mu_2)^T P_2^{-1}(y - \mu_2),\end{aligned}$$

sumando y restando términos convenientes

$$\begin{aligned}L &= (x - \mu_1)^T (P_1^{-1} + C^T P_2^{-1} C) (x - \mu_1) - (x - \mu_1)^T C^T P_2^{-1}(y - C\mu_1 - \mu_2) \\ &\quad - \mu_1^T C^T P_2^{-1}(y - \mu_2) + \mu_1^T C^T P_2^{-1} Cx - (y - \mu_2)^T P_2^{-1} Cx - (y - \mu_2)^T P_2^{-1}(y - \mu_2) \\ L &= (x - \mu_1)^T (P_1^{-1} + C^T P_2^{-1} C) (x - \mu_1) - (x - \mu_1)^T C^T P_2^{-1}(y - C\mu_1 - \mu_2) \\ &\quad + (y - C\mu_1 - \mu_2)^T P_2^{-1}(y - \mu_2) + \mu_1^T C^T P_2^{-1} Cx - (y - C\mu_1 - \mu_2)^T P_2^{-1} C\mu_1 \\ &\quad + (y - C\mu_1 - \mu_2)^T P_2^{-1} C\mu_1,\end{aligned}$$

para finalmente volver a factorizar todos los términos cuadráticos

$$\begin{aligned}L &= (x - \mu_1)^T (P_1^{-1} + C^T P_2^{-1} C) (x - \mu_1) - (x - \mu_1)^T C^T P_2^{-1}(y - C\mu_1 - \mu_2) \\ &\quad - (y - C\mu_1 - \mu_2)^T P_2^{-1} C(x - \mu_1) + (y - C\mu_1 - \mu_2)^T P_2^{-1}(y - C\mu_1 - \mu_2).\end{aligned}$$

Por otro lado, podemos escribir

$$\begin{aligned}
P_1^{-1} + C^T P_2^{-1} C &= (P_1 - P_1 C^T (P_2 + C P_1 C^T)^{-1} C P_1)^{-1} \\
P_2^{-1} &= (P_2 + C P_1 C^T)^{-1} \\
&\quad + (P_2 + C P_1 C^T)^{-1} C P_1 \\
&\quad \times (P_1 - P_1 C^T (P_2 + C P_1 C^T)^{-1} C P_1)^{-1} \\
&\quad \times P_1 C^T (P_2 + C P_1 C^T)^{-1} \\
C^T P_2^{-1} &= (P_1 - P_1 C^T (P_2 + C P_1 C^T)^{-1} C P_1)^{-1} \\
&\quad \times P_1 C^T (P_2 + C P_1 C^T)^{-1}
\end{aligned}$$

donde la primera igualdad se puede comprobar utilizando la identidad de Woodbury o lema de inversión matricial al lado derecho, con lo que se obtiene directamente la expresión al lado izquierdo. En la segunda igualdad podemos escribir el término en el lado izquierdo como $(P_2 + C P_1 C^T - C P_1 P_1^{-1} P_1 C^T)^{-1}$ y utilizar la identidad de Woodbury para obtener la expresión en el lado derecho. Finalmente, la última igualdad viene de utilizar la identidad $A^{-1} B (D - C A^{-1} B)^{-1} = (A - B D^{-1} C)^{-1} B D^{-1}$ en el término del lado izquierdo, que se puede escribir como $P_1^{-1} P_1 C^T (P_2 + C P_1^{-1} C^T - C P_1 P_1^{-1} P_1 C^T)^{-1}$.

Luego, definimos

$$\mu_3 = C \mu_1 + \mu_2 \tag{A.3}$$

$$P_3 = P_2 + C P_1 C^T \tag{A.4}$$

$$K = P_1 C^T (P_2 + C P_1 C^T)^{-1} = P_1 C^T P_3^{-1} \tag{A.5}$$

$$\begin{aligned}
P_4 &= P_1 - P_1 C^T (P_2 + C P_1 C^T)^{-1} C P_1 \\
&= (I - K C) P_1
\end{aligned} \tag{A.6}$$

con lo que es posible reducir la notación escribir y L como sigue

$$\begin{aligned}
L &= (x - \mu_1)^T P_4^{-1} (x - \mu_1) - (x - \mu_1)^T P_4^{-1} K (y - \mu_3) - (y - \mu_3)^T K^T P_4^{-1} (x - \mu_1) \\
&\quad + (y - \mu_3)^T (P_3^{-1} + K^T P_4^{-1} K) (y - \mu_3)
\end{aligned}$$

desarrollando el último sumando y factorizando algunos términos se tiene

$$L = (x - \mu_1 - K(y - \mu_3))^T P_4^{-1} (x - \mu_1 - K(y - \mu_3)) + (y - \mu_3)^T P_3^{-1} (y - \mu_3).$$

A continuación definimos

$$\mu_4 = \mu_1 + K(y - \mu_3), \tag{A.7}$$

por lo que se puede escribir la densidad conjunta como

$$\begin{aligned}
p(x, y) &= \frac{1}{\sqrt{\det(P_1 P_2)}} \exp \left\{ -\frac{1}{2} (y - \mu_3)^T P_3^{-1} (y - \mu_3) \right\} \\
&\quad \times \exp \left\{ -\frac{1}{2} (x - \mu_4)^T P_4 (x - \mu_4) \right\}.
\end{aligned}$$

Además, para la constante de normalización podemos ver que

$$\det(P_3 P_4) = \det(P_2 + C P_1 C^T) \det(P_1 - P_1 C^T (P_2 + C P_1 C^T)^{-1} C P_1),$$

y utilizando la identidad $\det(A) \det(D - C^T A^{-1} B) = \det(D) \det(A - B D^{-1} C)$

$$\begin{aligned} \det(P_3 P_4) &= \det(P_1) \det(P_2 + C P_1 C^T - C P_1 P_1^{-1} P_1 C^T) \\ &= \det(P_1) \det(P_2). \end{aligned}$$

Entonces, se puede concluir que

$$p(x, y) = \mathcal{N}_y(\mu_3, P_3) \mathcal{N}_x(\mu_4, P_4) \tag{A.8}$$

y al integrar con respecto a x la distribución marginal es

$$\int p(x, y) dx = \mathcal{N}_y(\mu_3, P_3). \tag{A.9}$$

Así, las ecuaciones (A.3) a (A.9) completan la prueba. □

B

Transformaciones de Similaridad

En este apéndice se muestra que la representación en variable de estados de un sistema no es única, contenido basado en [79].

Es posible tener un sistema con entrada u_t , salida y_t y dos elecciones distintas para el vector de estados: x_t y $\tilde{x}_t \in \mathbb{R}^n$, con sus matrices y covarianzas de ruidos asociados $\{A, B, C, D, Q, R\}$ y $\{\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}, \tilde{Q}, \tilde{R}, \}$, respectivamente. El siguiente lema establece la relación entre ambos modelos.

Lema B.1 Suponga un sistema estocástico cuyo vector de estado es x_t y con representación en espacio de estados definido por $\{A, B, C, D, Q, R\}$. Suponga además que se elige una matriz de transformación $T \in \mathbb{R}^{n \times n}$ no singular, que define un nuevo vector de estado:

$$\tilde{x}_t = T \cdot x_t. \quad (\text{B.1})$$

Entonces, el modelo en espacio de estados para la nueva representación, queda determinada por las matrices:

$$\begin{aligned} \tilde{A} &= TAT^{-1}, & \tilde{B} &= TB, & \tilde{C} &= CT^{-1} \\ \tilde{D} &= D, & \tilde{Q} &= TQT^T, & \tilde{R} &= R \end{aligned} \quad (\text{B.2})$$

Demostración. Dada la transformación (B.1), en que T es no singular y, por tanto, invertible, tenemos que:

$$x_t = T^{-1} \cdot \tilde{x}_t. \quad (\text{B.3})$$

Si reemplazamos entonces el vector de estado x_t en las ecuaciones de estado (1.1a)-(1.1b), obtenemos

$$T^{-1}\tilde{x}_{t+1} = AT^{-1}\tilde{x}_t + Bu_t + w_t \quad (\text{B.4})$$

$$y_t = CT^{-1}\tilde{x}_t + Du_t + v_t \quad (\text{B.5})$$

donde, al multiplicar la primera ecuación por T por la izquierda, se obtiene el modelo

$$\tilde{x}_{t+1} = TAT^{-1}x_t + TBu_t + Tw_t \quad (\text{B.6})$$

$$y_t = CT^{-1}\tilde{x}_t + Du_t + v_t \quad (\text{B.7})$$

y definiendo nuevos ruidos como

$$\tilde{w}_t = Tw_t \quad (\text{B.8})$$

$$\tilde{v}_t = v_t \quad (\text{B.9})$$

donde $w_t \sim \mathcal{N}(0, Q)$ y $v_t \sim \mathcal{N}(0, R)$, por lo que $\tilde{w}_t \sim \mathcal{N}(0, TQT^T)$ y $\tilde{v}_t \sim \mathcal{N}(0, R)$. □

A pesar de las diferentes posibles representaciones, es natural que ciertas importantes características y propiedades del sistema permanezcan invariantes, cualquiera sea la representación elegida. Si, por ejemplo, consideramos los autovalores del sistema, tenemos que:

$$\det(\lambda I - \tilde{A}) = \det(\lambda TT^{-1} - TAT^{-1}) \quad (\text{B.10})$$

$$= \det(T(\lambda I - A)T^{-1}) \quad (\text{B.11})$$

$$= \det(T) \det(\lambda I - A) \det(T^{-1}) \quad (\text{B.12})$$

$$= \det(\lambda I - A) \quad (\text{B.13})$$

lo que demuestra que los autovalores, polos o frecuencias naturales del sistema, son invariantes respecto a transformaciones de similaridad.

Bibliografía

- [1] G. Goodwin, S. Graebe, and M. Salgado, *Control System Design*. Prentice Hall, 2001.
- [2] K. D. Kim and P. R. Kumar, “Cyber-physical systems: A perspective at the centennial,” *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1287–1308, 2012.
- [3] D. Hristu-Varsakelis and W. Levine, *Handbook of Networked and Embedded Control Systems*, ser. Control Engineering - Birkhäuser. Birkhäuser Boston, 2005.
- [4] E. Silva, J. Aguero, G. Goodwin, K. Lau, and M. Wang, “The snr approach to networked control,” in *The Control Handbook, Second Edition: Control System Applications*, 2nd ed., W. Levine, Ed. CRC Press, 2010, ch. 25.
- [5] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, “Feedback control under data rate constraints: An overview,” *Proceedings of the IEEE*, vol. 95, pp. 108–137, 2007.
- [6] S. Ding, *Model-based fault diagnosis techniques: Design schemes, algorithms, and tools*. Springer, 01 2008.
- [7] F. Gustafsson and F. Gustafsson, *Adaptive filtering and change detection*. Citeseer, 2000, vol. 1.
- [8] E. L. Russell, L. H. Chiang, and R. D. Braatz, *Data-driven Methods for Fault Detection and Diagnosis in Chemical Processes*. Springer, 2000.
- [9] L. Ljung, *System Identification: Theory for the User*, ser. Prentice Hall information and system sciences series. Prentice Hall PTR, 1999.
- [10] T. Soderstrom and P. Stoica, *System Identification*, ser. Prentice Hall International Series In Systems And Control Engineering. Prentice Hall, 1989.
- [11] L. Y. Wang, G. G. Yin, J. F. Zhang, and Y. Zhao, *System Identification with Quantized Observations*. Springer, 2010.
- [12] T. Wigren, “Approximate gradients, convergence and positive realness in recursive identification of a class of non-linear systems,” *International Journal of Adaptive Control and Signal Processing*, vol. 9, no. 4, pp. 325–354, 1995.
- [13] —, “Adaptive filtering using quantized output measurements,” *IEEE transactions on signal processing*, vol. 46, no. 12, pp. 3423–3426, 1998.
- [14] L. Y. Wang, G. G. Yin, and J.-F. Zhang, “System identification using quantized data,” *IFAC Proceedings Volumes*, vol. 39, no. 1, pp. 255–260, 2006.

-
- [15] H. Suzuki and T. Sugie, "System identification based on quantized i/o data corrupted with noises," in *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, 2006.
- [16] —, "System identification based on quantized i/o data corrupted with noises and its performance improvement," in *Proceedings of the 45th IEEE Conference on Decision and Control*. IEEE, 2006, pp. 3684–3689.
- [17] J. C. Agüero, G. C. Goodwin, and J. I. Yuz, "System identification using quantized data," in *2007 46th IEEE Conference on Decision and Control*. IEEE, 2007, pp. 4263–4268.
- [18] Y. Zhao, L. Y. Wang, G. G. Yin, and J.-F. Zhang, "Identification of wiener systems with binary-valued output observations," *Automatica*, vol. 43, no. 10, pp. 1752–1765, 2007.
- [19] J. C. Agüero, B. I. Godoy, G. C. Goodwin, and T. Wigren, "Scenario-based em identification for fir systems having quantized output data," *IFAC Proceedings Volumes*, vol. 42, no. 10, pp. 66–71, 2009.
- [20] E. Weyer, S. Ko, and M. C. Campi, "Finite sample properties of system identification with quantized output data," in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*. IEEE, 2009, pp. 1532–1537.
- [21] F. Gustafsson and R. Karlsson, "Statistical results for system identification based on quantized observations," *Automatica*, vol. 45, no. 12, pp. 2794–2801, 2009.
- [22] E. Colinet and J. Juillard, "A weighted least-squares approach to parameter estimation problems based on binary measurements," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 148–152, 2009.
- [23] B. I. Godoy, G. C. Goodwin, J. C. Agüero, D. Marelli, and T. Wigren, "On identification of fir systems having quantized output data," *Automatica*, vol. 47, no. 9, pp. 1905–1915, 2011.
- [24] R. Carvajal, J. C. Agüero, B. I. Godoy, G. C. Goodwin, and J. I. Yuz, "Em-based identification of sparse fir systems having quantized data," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 553–558, 2012.
- [25] T. Chen, Y. Zhao, and L. Ljung, "Impulse response estimation with binary measurements: a regularized fir model approach," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 113–118, 2012.
- [26] D. Marelli, K. You, and M. Fu, "Identification of arma models using intermittent and quantized output observations," *Automatica*, vol. 49, no. 2, pp. 360–369, 2013.
- [27] J. Guo and Y. Zhao, "Recursive projection algorithm on fir system identification with binary-valued observations," *Automatica*, vol. 49, no. 11, pp. 3396–3401, 2013.
- [28] B. I. Godoy, J. C. Agüero, R. Carvajal, G. C. Goodwin, and J. I. Yuz, "Identification of sparse fir systems using a general quantisation scheme," *International Journal of Control*, vol. 87, no. 4, pp. 874–886, 2014.

- [29] K. You, "Recursive algorithms for parameter estimation with adaptive quantizer," *Automatica*, vol. 52, pp. 192–201, 2015.
- [30] A. Moschitta, J. Schoukens, and P. Carbone, "Parametric system identification using quantized data," *IEEE transactions on instrumentation and measurement*, vol. 64, no. 8, pp. 2312–2322, 2015.
- [31] J. Guo, L. Y. Wang, G. Yin, Y. Zhao, and J.-F. Zhang, "Asymptotically efficient identification of fir systems with quantized observations and general quantized inputs," *Automatica*, vol. 57, pp. 113–122, 2015.
- [32] C. Yu, C. Zhang, and L. Xie, "Blind system identification using precise and quantized observations," *Automatica*, vol. 49, no. 9, pp. 2822–2830, 2013.
- [33] C. Yu, K. You, and L. Xie, "Quantized identification of arma systems with colored measurement noise," *Automatica*, vol. 66, pp. 101–108, 2016.
- [34] J. C. Agüero, K. González, and R. Carvajal, "Em-based identification of arx systems having quantized output data," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8367–8372, 2017.
- [35] D. E. Marelli, B. I. Godoy, and G. C. Goodwin, "A scenario-based approach to parameter estimation in state-space models having quantized output data," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 2011–2016.
- [36] T. Katayama, *Subspace methods for system identification*. Springer Science & Business Media, 2006.
- [37] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge University Press, 2013, vol. 3.
- [38] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [39] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of basic engineering*, vol. 83, no. 1, pp. 95–108, 1961.
- [40] Y.-C. Ho and R. C. K. Lee, "A bayesian approach to problems in stochastic estimation and control," *IEEE Transactions on Automatic Control*, vol. 9, no. 4, pp. 333–339, October 1964.
- [41] R. C. K. Lee, *Optimal Estimation, Identification and Control*. MIT Press, 1964.
- [42] A. Jazwinski, "Filtering for nonlinear dynamical systems," *IEEE Transactions on Automatic Control*, vol. 11, no. 4, pp. 765–766, 1966.
- [43] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970, vol. 64.
- [44] R. L. Stratonovich, *Conditional Markov processes and their application to the theory of optimal control*. American Elsevier, 1968.
- [45] H. Rauch, "Solutions to the linear smoothing problem," *IEEE Transactions on Automatic Control*, vol. 8, no. 4, pp. 371–372, 1963.

-
- [46] H. E. Rauch, F. Tung, and C. T. Striebel, “Maximum likelihood estimates of linear dynamic systems,” *AIAA journal*, vol. 3, no. 8, pp. 1445–1450, 1965.
- [47] M. H. DeGroot, *Optimal statistical decisions*. John Wiley & Sons, 2005, vol. 82.
- [48] C. Robert and G. Casella, *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.
- [49] J. Geweke, “Bayesian inference in econometric models using monte carlo integration,” *Econometrica: Journal of the Econometric Society*, pp. 1317–1339, 1989.
- [50] A. Doucet, S. Godsill, and C. Andrieu, “On sequential monte carlo sampling methods for bayesian filtering,” *Statistics and computing*, vol. 10, no. 3, pp. 197–208, 2000.
- [51] N. Gordon, *Bayesian methods for tracking*. Imperial College London (University of London), 1993.
- [52] B. Ninness, “Strong laws of large numbers under weak assumptions with application,” *IEEE Transactions on Automatic Control*, vol. 45, no. 11, pp. 2117–2122, 2000.
- [53] B. Anderson and J. B. Moore, *Optimal filtering*. Prentice-Hall, 1979.
- [54] I. Arasaratnam, S. Haykin, and R. J. Elliott, “Discrete-time nonlinear filtering algorithms using gauss–hermite quadrature,” *Proceedings of the IEEE*, vol. 95, no. 5, pp. 953–977, 2007.
- [55] T. Söderström, *Discrete-time Stochastic Systems: Estimation and Control*. Springer Science & Business Media, 2002.
- [56] R. Albornoz, R. Carvajal, and J. C. Agüero, “A novel bayesian filtering method for systems with quantized output data,” in *2019 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*. IEEE, 2019, pp. 1–7.
- [57] H. W. Sorenson and D. L. Alspach, “Recursive bayesian estimation using gaussian sums,” *Automatica*, vol. 7, no. 4, pp. 465–479, 1971.
- [58] A. R. Runnalls, “Kullback-leibler approach to gaussian mixture reduction,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 989–999, 2007.
- [59] A. G. Wills, J. Hendriks, C. Renton, and B. Ninness, “A bayesian filtering algorithm for gaussian mixture models,” *arXiv preprint arXiv:1705.05495*, 2017.
- [60] T. B. Schön, A. Wills, and B. Ninness, “System identification of nonlinear state-space models,” *Automatica*, vol. 47, no. 1, pp. 39–49, 2011.
- [61] Y. Bresler, “Two-filter formulae for discrete-time non-linear bayesian smoothing,” *International Journal of Control*, vol. 43, no. 2, pp. 629–641, 1986.
- [62] M. Briers, A. Doucet, and S. Maskell, “Smoothing algorithms for state–space models,” *Annals of the Institute of Statistical Mathematics*, vol. 62, no. 1, p. 61, 2010.

- [63] R. Douc, A. Garivier, E. Moulines, J. Olsson *et al.*, “Sequential monte carlo smoothing for general state space hidden markov models,” *The Annals of Applied Probability*, vol. 21, no. 6, pp. 2109–2145, 2011.
- [64] P. Fearnhead, D. Wyncoll, and J. Tawn, “A sequential smoothing algorithm with linear computational cost,” *Biometrika*, vol. 97, no. 2, pp. 447–464, 2010.
- [65] S. J. Godsill, A. Doucet, and M. West, “Monte carlo smoothing for nonlinear time series,” *Journal of the american statistical association*, vol. 99, no. 465, pp. 156–168, 2004.
- [66] G. Pillonetto and B. M. Bell, “Optimal smoothing of non-linear dynamic systems via monte carlo markov chains,” *Automatica*, vol. 44, no. 7, pp. 1676–1685, 2008.
- [67] R. A. Fisher, “On the mathematical foundations of theoretical statistics,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 222, no. 594-604, pp. 309–368, 1922.
- [68] G. C. Goodwin and R. L. Payne, *Dynamic system identification: experiment design and data analysis*. Academic press, 1977.
- [69] X.-L. Meng and D. Van Dyk, “The em algorithm—an old folk-song sung to a fast new tune,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 59, no. 3, pp. 511–567, 1997.
- [70] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the em algorithm,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.
- [71] C. J. Wu *et al.*, “On the convergence properties of the em algorithm,” *The Annals of statistics*, vol. 11, no. 1, pp. 95–103, 1983.
- [72] G. J. McLachlan and T. Krishnan, *The EM algorithm and extensions*. Wiley, 1997.
- [73] K. Lange, D. R. Hunter, and I. Yang, “Optimization transfer using surrogate objective functions,” *Journal of computational and graphical statistics*, vol. 9, no. 1, pp. 1–20, 2000.
- [74] D. R. Hunter and K. Lange, “Quantile regression via an mm algorithm,” *Journal of Computational and Graphical Statistics*, vol. 9, no. 1, pp. 60–77, 2000.
- [75] —, “A tutorial on mm algorithms,” *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.
- [76] S. Gibson and B. Ninness, “Robust maximum-likelihood estimation of multivariable dynamic systems,” *Automatica*, vol. 41, no. 10, pp. 1667–1682, 2005.
- [77] G. Kitagawa, “The two-filter formula for smoothing and an implementation of the gaussian-sum smoother,” *Annals of the Institute of Statistical Mathematics*, vol. 46, no. 4, pp. 605–623, 1994.
- [78] G. Kitagawa and W. Gersch, *Smoothness priors analysis of time series*. Springer Science & Business Media, 1996, vol. 116.
- [79] M. E. Salgado, J. I. Yuz, R. A. Rojas *et al.*, *Análisis de sistemas lineales*. Pearson Educación, 2005.