



Universidad Autónoma de Guerrero

Facultad de Matemáticas

Maestría en Matemáticas Aplicadas

Identificación de Parámetros en el Controlador PID Fraccionario

T E S I S

PARA OBTENER EL GRADO DE:

Maestro en Matemáticas Aplicadas

PRESENTA:

Alexeis Figueroa Ferrer

DIRECTORES DE TESIS:

Dr. Francisco Julián Ariza Hernández

Dr. Martín Patricio Arciga Alejandre

Abril del 2019

Dedicatoria

Dedico esta tesis a toda mi familia, porque cada uno de ellos contribuyó en mi formación, no solo como profesional, sino también como persona. Cada uno de ustedes me ayudó de alguna manera, así que este resultado es también suyo.

A todos los que creyeron en mí y me dieron la oportunidad de compartir con ustedes, también les dedico esta tesis.

Agradecimientos

A mi madre, porque siempre luchó para que no me faltara lo necesario para convertirme en un hombre de bien. Por sus sacrificios y su amor, gracias. Esta también es tu tesis.

A mi abuelo Ángel, por siempre ser un apoyo incondicional y estar a mi lado. Papá, aun cuando no pude darte el último adiós, siempre estás conmigo.

A mis hermanos y a mi papá, porque con su amor y sacrificios han sido pilares fundamentales de lo que soy.

A mis amigos, los de Cuba y de México, por todos sus consejos y su apoyo incondicional.

A mis asesores, por creer en mí y permitirme trabajar a su lado.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por haberme brindado el apoyo económico necesario para realizar mis estudios de posgrado.

A todos los que de una forma u otra me apoyaron.

Siento una enorme gratitud por los que en algún momento no creyeron en mí y me dijeron “No”, gracias a ustedes lo hice yo mismo.

Resumen

Los controladores Proporcional-Integral-Derivativo (PID) son una herramienta muy usada en todas las esferas de la industria y la tecnología, donde se precise controlar el comportamiento de los estados de un sistema automático. Los controladores PID de orden fraccionario han demostrado tener un mejor desempeño en sistemas más complejos. El buen funcionamiento del sistema depende en gran medida de caracterizar los parámetros del controlador asociado a él. Por tanto, es de vital importancia la correcta estimación de estos parámetros. Una técnica estadística usada para este fin es la estimación Bayesiana. Este trabajo tiene como objetivo estimar los parámetros del controlador PID de orden fraccionario, formulando el problema de inferencia estadística como un problema inverso desde la perspectiva Bayesiana. Se presenta un ejemplo sobre el control del nivel en un tanque cónico, que es un proceso muy usado en las industrias. A través de este ejemplo, se muestra la metodología propuesta y se realiza un estudio de simulación para evaluar el desempeño de las estimaciones obtenidas en términos del sesgo y del error cuadrático medio.

Abstract

The Proportional-Integral-Derivative (PID) controllers are a broadly used tool in all areas of industry and technology, where it is necessary to control the behavior of the states of an automatic system. PID controllers of fractional order have been shown to perform better in more complex systems. The proper functioning of the system depends to a great extent on characterizing the parameters of the controller associated to it. Therefore, the correct estimation of these parameters is of vital importance. A statistical technique used for this purpose is Bayesian estimation. The objective of this work is to estimate the parameters of the fractional order PID controller, formulating the problem of statistical inference as an inverse problem, from the Bayesian perspective. An example about the level control is presented in a conical tank, which is a widely used process in industries. Through this example, the proposed methodology is shown and a simulation study is carried out, in order to evaluate the performance of the estimates obtained in terms of the bias and the mean square error.

Índice general

| | |
|--------------------------------------------------------------------------------------------------------------------------|-----------|
| 1. Introducción | 1 |
| 1.1. Antecedentes | 3 |
| 1.2. Problemática | 3 |
| 1.3. Objetivos | 4 |
| 2. Marco Teórico | 5 |
| 2.1. Cálculo fraccionario | 5 |
| 2.1.1. Integral de Riemann-Liouville de orden fraccionario | 6 |
| 2.1.2. Derivadas de orden fraccionario | 6 |
| 2.1.3. Transformada de Laplace | 7 |
| 2.2. Controladores PID de orden fraccionario | 8 |
| 2.2.1. Impedancia eléctrica | 9 |
| 2.3. Inferencia Bayesiana | 10 |
| 2.3.1. Cadenas de Markov Monte Carlo | 11 |
| 2.3.2. Diagnósticos de convergencia | 12 |
| 2.3.3. Estimadores Bayesianos | 15 |
| 2.3.4. Intervalos de credibilidad | 16 |
| 3. Metodología | 17 |
| 3.1. Datos | 18 |
| 3.2. Los Parámetros | 18 |
| 4. Resultados | 20 |
| 4.1. Estimación de los parámetros del controlador PID fraccionario para controlar el nivel en un tanque cónico | 20 |

| | | |
|-----------|------------------------------------------------------------------------------|-----------|
| 4.1.1. | Distribución <i>a priori</i> de los parámetros. | 21 |
| 4.1.2. | Estimaciones obtenidas | 23 |
| 4.1.3. | Diagnósticos de convergencia | 24 |
| 4.1.4. | Evaluación | 29 |
| 4.2. | Estimación de los parámetros del controlador PID fraccionario. Otro ejemplo. | 30 |
| 4.2.1. | Distribución <i>a priori</i> de los parámetros. | 31 |
| 4.2.2. | Estimaciones obtenidas | 32 |
| 4.2.3. | Diagnósticos de convergencia | 33 |
| 4.2.4. | Evaluación | 38 |
| 5. | Conclusiones | 40 |

Índice de figuras

| | |
|-----------------------------------------------------------------------------------------------------------------------|----|
| 1.1. Diagrama básico del proceso de control. | 2 |
| 2.1. Controlador PID (a) y PID fraccionario (b). | 8 |
| 4.1. Desviación de los datos respecto a la media. Magnitud (a) y Fase (b). | 21 |
| 4.2. Gráficas de las densidades <i>a priori</i> de los parámetros. | 23 |
| 4.3. Ajuste a los datos. Línea negra representa el modelo ajustado. | 24 |
| 4.4. Convergencia de las cadenas MCMC. | 25 |
| 4.5. Gráfica de las densidades de las distribuciones <i>a posteriori</i> | 26 |
| 4.6. Trazas de las MCMC y densidades <i>a posteriori</i> de los parámetros para el ejemplo del tanque cónico. | 27 |
| 4.7. Trazas de las MCMC y densidades <i>a posteriori</i> de los parámetros para el ejemplo del tanque cónico. | 28 |
| 4.8. Diagnóstico de convergencia de Gelman-Rubin. | 29 |
| 4.9. Desviación de los datos respecto a la media. Magnitud (a) y Fase (b). | 31 |
| 4.10. Gráficas de las densidades <i>a priori</i> de los parámetros. | 32 |
| 4.11. Ajuste a los datos. Línea negra representa el modelo ajustado. | 33 |
| 4.12. Convergencia de las cadenas MCMC. | 34 |
| 4.13. Gráfica de las densidades de las distribuciones <i>a posteriori</i> | 35 |
| 4.14. Trazas de las MCMC y densidades para el ejemplo tomado de Monje et al. (2010). | 36 |
| 4.15. Trazas de las MCMC y densidades para el ejemplo tomado de Monje et al. (2010). | 37 |
| 4.16. Diagnóstico de convergencia de Gelman-Rubin. | 38 |

Índice de tablas

| | |
|------------------------------------------------------------------------------------|----|
| 4.1. Estimaciones obtenidas por la técnica PSO. | 21 |
| 4.2. Estimaciones Bayesianas e intervalos de credibilidad de los parámetros. . . . | 24 |
| 4.3. Diagnóstico de Gelman-Rubin. | 28 |
| 4.4. Resultados de sesgo y ECM. | 30 |
| 4.5. Valores de los parámetros tomados de Monje et al. (2010). | 31 |
| 4.6. Estimaciones Bayesianas e intervalos de credibilidad de los parámetros. . . . | 33 |
| 4.7. Diagnóstico de Gelman-Rubin. | 37 |
| 4.8. Resultados de sesgo y ECM. | 39 |

Introducción

En la actualidad, los sistemas de control son utilizados prácticamente en todas las disciplinas científicas y tecnológicas, desde sistemas muy simples como control de temperatura o llenado de un tanque, hasta pilotos automáticos o robots que controlan los procesos industriales en fábricas. Estos sistemas son estudiados por una rama interdisciplinaria de la ingeniería y las matemáticas, conocida como Teoría de Control, que se ocupa de la modificación de sistemas dinámicos para obtener un comportamiento deseado en términos de un conjunto de especificaciones. Para obtener dicho comportamiento, un controlador detecta el funcionamiento del sistema, lo compara con el comportamiento deseado, y en base a esto realiza acciones correctivas y activa el sistema para obtener el cambio deseado (ver Figura 1.1). Un controlador fraccionario es una herramienta con buenos resultados para este fin. En este trabajo estaremos interesados en el controlador Proporcional-Integral-Derivativo de orden fraccionario (PID fraccionario o $PI^\lambda D^\mu$). En muchos fenómenos de la Teoría de Control, es común que los sensores del sistema proporcionen una serie de mediciones o datos de algunas de las variables que forman parte de un modelo, el cual involucra ciertos parámetros desconocidos que representan características importantes del fenómeno bajo estudio. Si contamos con observaciones de esas variables, es posible plantear el problema inverso como un problema de inferencia estadística.

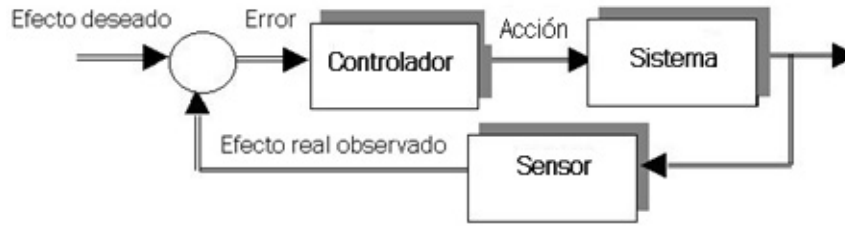


Figura 1.1: Diagrama básico del proceso de control.

El cálculo fraccionario es una extensión del cálculo que permite considerar la integración y la derivación de orden fraccionario, o sea, permite definir operadores como $\frac{d^\alpha f}{dt^\alpha}$, cuando $\alpha \in \mathbb{R}^+$. El desarrollo de los conceptos y definiciones del cálculo fraccionario se remontan al siglo XVII, con indicios en la correspondencia entre Leibniz y L'Hôpital. En las últimas décadas, el cálculo fraccionario se ha utilizado en varias ramas como son la ciencia de materiales, la mecánica, la Teoría de Control, entre otras. Un amplio campo de aplicaciones se pueden encontrar en Podlubny (1999).

Las primeras aplicaciones del cálculo fraccionario en la Teoría de Control aparecieron en la década del 60, aunque sin usar el término “fraccionario”. Un problema clave en el diseño de un amplificador de realimentación fue diseñar un bucle de realimentación para que el rendimiento del bucle cerrado sea invariante a los cambios en la ganancia del amplificador (A). Bode (1940) y Bode (1945) presentó una solución elegante a este robusto problema de diseño, que denominó la característica de corte ideal, hoy en día conocida como la función de transferencia de bucle ideal de Bode, cuya forma era $F(s) = \frac{A}{s^\alpha}$, $\alpha \in \mathbb{R}$, es decir, era un integrador fraccionario. Desde entonces, se han realizado trabajos relacionados con la Teoría de Control como son el diseño y realización de filtros, tratamientos de señales y estimación de parámetros. Precisamente, esta última tarea es de vital importancia en el funcionamiento de un controlador fraccionario. Es imprescindible determinar qué valores de los parámetros proporcionan el mejor funcionamiento del controlador. Una técnica usada para este fin es la estimación mediante el Teorema de Bayes.

1.1. Antecedentes

Varios son los autores que han desarrollado trabajos relacionados con problemas inversos, aplicaciones de derivadas fraccionarias y estimación de parámetros en la resolución de problemas prácticos.

Bates (2009) en su trabajo da un enfoque general de modelado inverso para mecánica pulmonar. Freeborn (2013) presenta un panorama general de las aplicaciones de los circuitos fraccionarios en biología. De gran interés es el trabajo de Ionescu y De Keyser (2008), quienes validan un modelo mecánico análogo a un circuito con componentes: una resistencia, un capacitor y una bobina (RCL, por sus siglas en inglés) fraccionario, para describir el proceso respiratorio. Otro ejemplo de un modelo fraccionario es el presentado por Moreles, Peña, Botello y Iturriaga (2013), el cual tiene su aplicación en el flujo en medios porosos heterogéneos. Relacionado con los problemas inversos que aparecen en geología, se tiene el trabajo de Xiang, Cheng, Schlindwein y Jones (2003), quienes se enfocan en desarrollar un método para estimar los parámetros del modelo. Monje, Chen, Vinagre, Xue y Feliu-Batlle (2010) realizan un amplio análisis de los sistemas fraccionarios aplicados a sistemas de la Teoría de Control. Hacen un interesante recorrido por toda la teoría del cálculo fraccionario y la aplican a ejemplos experimentales de controladores. Lainez (2015) hace una identificación y modelación de sistemas fraccionarios y propone modelos para el sistema respiratorio y el estudio de los componentes del suelo.

Hasta el momento, en la bibliografía revisada, en estos trabajos no se aborda la estimación de parámetros desde el enfoque Bayesiano en problemas en el contexto de la Teoría de Control, específicamente para el controlador PID fraccionario.

1.2. Problemática

Dado el amplio uso de los controladores fraccionarios en la industria y otras ramas, se hace imprescindible determinar qué valores de los parámetros proporcionan el mejor funcionamiento del controlador. El ajuste o estimación (también conocido como *tunning*) puede ser manual, a través de un método matemático o usando un programa de computadora. El ajuste manual puede ser un proceso lento, empírico, en ocasiones ineficiente y no exento de errores.

Entre los métodos matemáticos, el más conocido es el de Ziegler-Nichols. El problema con este método es que incluye una parte manual (con los problemas mencionados arriba) y no plantea nada para obtener los órdenes de la derivada y la integral fraccionarias. Es un método básicamente para el controlador PID o alguno de sus derivados (Proporcional-Integral PI, Proporcional-Derivativo PD). Por último, es posible usar un software, pero se necesita de equipos y programas posiblemente privativos, que son costosos. Dadas estas problemáticas, en este trabajo **estimaremos los parámetros del controlador PID fraccionario desde el paradigma Bayesiano.**

1.3. Objetivos

Objetivo general

Estimar los parámetros del controlador PID fraccionario desde el paradigma Bayesiano.

Objetivos específicos

- Simular datos a partir de valores conocidos de los parámetros del controlador PID fraccionario.
- Estimar los parámetros del controlador PID fraccionario y la varianza en un ejemplo real: control de nivel en un tanque cónico.
- Obtener los intervalos de credibilidad para los parámetros de interés.
- Evaluar el método de estimación en términos de sesgo y error cuadrático medio.

Para dar cumplimiento a estos objetivos y explicar la manera en que lo haremos, se ha estructurado el trabajo en varias secciones. En el Capítulo 2, abordaremos la teoría necesaria sobre cálculo fraccionario, se describirá el controlador PID fraccionario y explicaremos la técnica de estimación Bayesiana. En el Capítulo 3, se plantea la metodología usada para realizar la estimación de los parámetros de interés. Por último, en el Capítulo 4, se realiza dicha estimación, se verifican los diagnósticos de convergencia y se discuten los resultados obtenidos. En este capítulo presentamos un ejemplo de estimación del nivel de un fluido en un tanque cónico, un proceso muy importante en las industrias.

Marco Teórico

En este Capítulo se enuncia toda la teoría necesaria sobre la que se soporta esta investigación. Se introduce el marco teórico sobre cálculo fraccionario, controladores PID fraccionarios e inferencia Bayesiana.

2.1. Cálculo fraccionario

Explorando más allá del enfoque tradicional, consideraremos las funciones de transferencia de orden real arbitrario. Por tanto, estamos en presencia de sistemas que se describen mejor mediante modelos de orden fraccionario, que incluyen los sistemas de orden entero.

Una función sumamente importante en el cálculo fraccionario es la función Gamma de Euler $\Gamma(0, \infty) \rightarrow \mathbb{R}$, que se define como

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt.$$

Otra función que desempeña un papel muy importante en el cálculo fraccionario, es la función de dos parámetros del tipo Mittag-Leffler. Esta fue introducida por Agarwal (1953) y está definida como:

$$E_{\alpha, \beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}, \quad \alpha > 0, \quad \beta > 0.$$

Esta función se usa durante la aplicación de la Transformada de Laplace a las derivadas de orden fraccionario.

2.1.1. Integral de Riemann-Liouville de orden fraccionario

De acuerdo con la concepción de Riemann-Liouville, la integral de orden fraccionario puede expresarse como

$$\mathcal{I}_c^\alpha f(t) \triangleq \mathcal{D}_c^{-\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \int_c^t (t - \tau)^{\alpha-1} f(\tau) d\tau, \quad t > c, \quad \alpha \in \mathbb{R}^+.$$

Cuando tratamos con sistemas dinámicos, es usual que $f(t)$ sea una función que depende del tiempo t , y por lo tanto, en lo que sigue, la definición de la integral de orden fraccional a usar es

$$\mathcal{I}^\alpha f(t) \triangleq \mathcal{D}^{-\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t - \tau)^{\alpha-1} f(\tau) d\tau, \quad t > 0, \quad \alpha \in \mathbb{R}^+. \quad (2.1)$$

2.1.2. Derivadas de orden fraccionario

Para las derivadas de orden fraccionario existen varias definiciones, dos ampliamente usadas son la definición de Riemann-Liouville y la definición de Caputo.

Según Riemann-Liouville, la derivada fraccionaria de orden α se puede expresar como

$${}_R\mathcal{D}^\alpha f(t) \triangleq \mathcal{D}^m \mathcal{I}^{m-\alpha} f(t) = \frac{d^m}{dt^m} \left[\frac{1}{\Gamma(m-\alpha)} \int_0^t \frac{f(\tau)}{(t-\tau)^{\alpha-m+1}} d\tau \right],$$

donde $m-1 < \alpha < m$, $m \in \mathbb{N}$. En la literatura, suele usarse $0 < \alpha < 1$, por tanto, tomaremos $m = 1$ y la definición queda

$${}_R\mathcal{D}^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t \frac{f(\tau)}{(t-\tau)^\alpha} d\tau. \quad (2.2)$$

Por su parte, Caputo propuso la definición de la derivada de orden fraccionario como

$${}_C\mathcal{D}^\alpha f(t) \triangleq \mathcal{I}^{m-\alpha} \mathcal{D}^m f(t) = \frac{1}{\Gamma(m-\alpha)} \int_0^t \frac{f^{(m)}(\tau)}{(t-\tau)^{\alpha-m+1}} d\tau,$$

donde $m-1 < \alpha < m$, $m \in \mathbb{N}$. De igual manera que en la definición anterior, si $m = 1$, entonces la definición nos queda

$${}_C\mathcal{D}^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{f'(\tau)}{(t-\tau)^\alpha} d\tau, \quad (2.3)$$

2.1.3. Transformada de Laplace

Una herramienta importante y que, junto a la función de Mittag-Leffler es muy usada en la solución de ecuaciones con derivadas e integrales de orden fraccionario, es la Transformada de Laplace. La misma está definida como

$$F(s) = \mathcal{L}\{f(t)\}(s) = \int_0^{\infty} e^{-st} f(t) dt \quad (2.4)$$

donde s es una variable compleja y $f(t)$ es la función original, esta debe ser tal que no crezca más rápido que una cierta función exponencial cuando $t \rightarrow \infty$. Se suele denotar la transformada con letras mayúsculas y la función original con letras minúsculas. Por la importancia de esta herramienta, a continuación damos las transformadas de Laplace de los operadores de orden fraccionario que definimos.

$$\mathcal{L}\{\mathcal{I}^\alpha f(t)\} = s^{-\alpha} F(s), \quad (2.5)$$

$$\mathcal{L}\{{}_R\mathcal{D}^\alpha f(t)\} = s^\alpha F(s) - \sum_{k=0}^{m-1} s^k [{}_R\mathcal{D}^{\alpha-k-1} f(t)]_{t=0}, \quad (2.6)$$

$$\mathcal{L}\{{}_C\mathcal{D}^\alpha f(t)\} = s^\alpha F(s) - \sum_{k=0}^{m-1} s^{\alpha-k-1} f^{(k)}(0). \quad (2.7)$$

La función original $f(t)$ puede obtenerse a partir de la transformada de Laplace $F(s)$ con la ayuda de la transformada inversa de Laplace

$$f(t) = \mathcal{L}^{-1}\{F(s)\} = \int_{c-j\infty}^{c+j\infty} e^{st} F(s) ds, \quad c = \text{Re}(s) > c_0, \quad (2.8)$$

donde c_0 se encuentra en el semiplano derecho de la convergencia absoluta de la integral de Laplace (2.4).

La transformada de Laplace de la convolución

$$f(t) * g(t) = \int_0^t f(t-\tau)g(\tau)d\tau = \int_0^t f(\tau)g(t-\tau)d\tau \quad (2.9)$$

de dos funciones $f(t)$ y $g(t)$, ambas iguales a 0 para $t < 0$, es igual al producto de las transformadas de Laplace de estas funciones:

$$\mathcal{L}\{f(t) * g(t)\}(s) = F(s)G(s) \quad (2.10)$$

suponiendo que ambas funciones $F(s)$ y $G(s)$ existen. La propiedad (2.10) es usada en la evaluación de la transformada de Laplace de la integral fraccionaria de Riemann-Liouville (Petráš, 2011).

2.2. Controladores PID de orden fraccionario

Un controlador PID (Proporcional-Integral-Derivativo) es un mecanismo de control por realimentación ampliamente usado en sistemas de control industrial. Consta de tres partes diferentes: el Proporcional, el Integral, y el Derivativo. El valor Proporcional depende del error actual y consiste en el producto entre la señal de error y la constante proporcional para lograr que el error en estado estacionario se aproxime a cero. El Integral depende de los errores pasados y tiene como propósito disminuir y eliminar el error en estado estacionario, provocado por perturbaciones exteriores y el cual no puede ser corregido por el control proporcional. Por último, el Derivativo es una predicción de los errores futuros, se manifiesta cuando hay un cambio en el valor absoluto del error. Su función es mantener el error al mínimo corrigiéndolo proporcionalmente con la misma velocidad que se produce; de esta manera evita que el error se incremente.

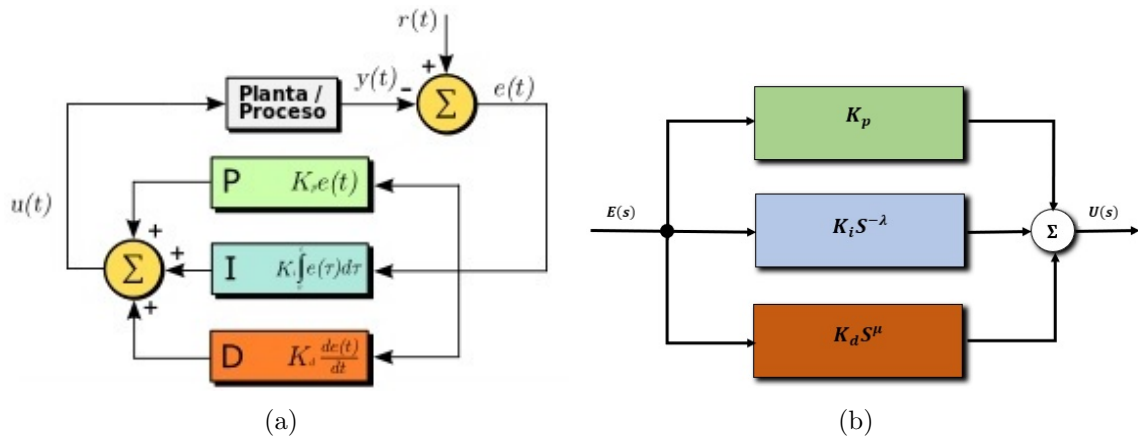


Figura 2.1: Controlador PID (a) y PID fraccionario (b).

El controlador de orden fraccionario $PI^\lambda D^\mu$ se propuso en Podlubny (1999) como una generalización del controlador PID con integrador λ y diferenciador μ , ambos de orden real. La ecuación íntegro-diferencial que define la acción de control del PID fraccionario es

$$u(t) = K_p e(t) + K_i \mathcal{D}^{-\lambda} e(t) + K_d \mathcal{D}^\mu e(t)$$

En general, pueden tomarse cualquiera de las dos definiciones de la derivada de orden fraccionario. En este trabajo usaremos la definición de Caputo. Esta decisión responde al hecho de que esta definición permite hacer interpretaciones físicas de la derivada. Aplicando la transformada de Laplace y las propiedades (2.5) y (2.7) al problema de valor inicial siguiente, se obtiene la función de transferencia del controlador (Podlubny, 1999):

$$\begin{cases} K_p e(t) + K_i \mathcal{D}^{-\lambda} e(t) + K_d \mathcal{D}^\mu e(t) \\ e(t_0) = e_0 \end{cases}$$

$$\begin{aligned} \mathcal{L}\{u(t)\} &= \mathcal{L}\{K_p e(t) + K_i \mathcal{D}^{-\lambda} e(t) + K_d \mathcal{D}^\mu e(t)\} \\ U(s) &= K_p E(s) + K_i s^{-\lambda} E(s) + K_d [s^\mu E(s) - s^{\mu-1} e_0] \\ E(s) &= \frac{1}{K_p + K_i s^{-\lambda} + K_d s^\mu} (K_d s^{\mu-1} e_0 + U(s)) \end{aligned}$$

La ecuación de transferencia es el denominador del primer factor

$$C(s) = K_p + K_i s^{-\lambda} + K_d s^\mu, \quad (\lambda, \mu > 0) \quad (2.11)$$

donde K_p es la constante proporcional, K_i es la constante integral y K_d es la constante derivativa. En la Figura 2.1 se pueden observar, de forma general, ambos controladores¹.

2.2.1. Impedancia eléctrica

Los controladores PID fraccionarios trabajan en el dominio de la frecuencia. Específicamente, miden la impedancia eléctrica como una señal del proceso que desean controlar.

Definición 2.2.1 La impedancia (Z) es una medida de oposición que presenta un circuito a una corriente cuando se aplica una tensión. La impedancia extiende el concepto de resistencia

¹Imagen (a) fue tomada de https://es.wikipedia.org/wiki/Controlador_PID

a los circuitos de corriente alterna (CA), y posee tanto magnitud como fase, a diferencia de la resistencia, que sólo tiene magnitud.

La impedancia es un número complejo $Z = R + jX$, donde R es la resistencia o parte real y X es la reactancia o parte imaginaria. Muchas veces, es mejor trabajar con la forma polar. En lugar de usar una parte imaginaria y otra real, se trabaja con la magnitud y la fase.

2.3. Inferencia Bayesiana

El proceso de inferencia Bayesiana consiste en analizar cómo la información muestral modifica la distribución probabilística del parámetro de interés. Para su realización práctica, se recurre a la aplicación del Teorema de Bayes, que combina la información muestral suministrada por la función de verosimilitud y el conocimiento que se tenga de los parámetros recogida en la distribución *a priori*. Ante un problema de inferencia sobre un parámetro o vector de parámetros θ de una población, el enfoque Bayesiano se puede estructurar en las siguientes etapas:

1. Obtención de una distribución *a priori*, representada por $\pi(\theta)$, que resume la información disponible sobre el parámetro.
2. Obtención de la función de verosimilitud, expresada como $L(y|\theta)$, de los datos y dado el valor del parámetro. Esta función recoge toda la información sobre la relación entre las observaciones muestrales (los datos) y el parámetro.
3. Calcular la distribución *a posteriori*, $\pi(\theta|y)$. Esta función, proporcional al producto de la distribución *a priori* y la verosimilitud, resume toda la información actualizada sobre el parámetro.
4. Establecer, a partir de la distribución *a posteriori*, los estimadores e intervalos de credibilidad.

Así, para realizar inferencias sobre un parámetro, el enfoque Bayesiano se basa en la distribución *a posteriori* $\pi(\theta|y)$, que es la distribución del parámetro dado los valores de la muestra. Se produce así un enfoque más natural y lógicamente coherente, pero que necesita de la buena elección de una distribución extramuestral o *a priori* para θ .

En el caso de que $\boldsymbol{\theta}$ sea un vector de parámetros, la distribución *a priori* es una distribución conjunta. En muchos casos (como el nuestro) se considera la distribución *a priori* como el producto de las distribuciones marginales, ya que se supone independencia entre los parámetros. Cuando el conocimiento inicial del parámetro $\boldsymbol{\theta}$ es poco, da lugar a distribuciones *a priori* no informativas.

Cuando se cuenta con pocas muestras, o sea, pocos datos, la mayor información la inyectan las distribuciones *a priori* de los parámetros. Por tanto, es de vital importancia en un enfoque Bayesiano, determinar cuáles son las *a priori* correctas. Cuando no se tiene información previa sobre los parámetros, la mayor información recae en la función de verosimilitud, o sea, en los datos. Es este caso, se usan *a prioris* no informativas y el análisis Bayesiano es equivalente al análisis clásico.

Definición 2.3.1 Una distribución *a priori* se dice no informativa si no contiene información sobre el parámetro que establezca si unos valores son más probables que otros.

Una vez conocida la verosimilitud de los datos $L(\mathbf{y}|\boldsymbol{\theta})$ y la distribución *a priori* de los parámetros $\pi(\boldsymbol{\theta})$, la distribución posterior de los parámetros de interés está dada por

$$\pi(\boldsymbol{\theta}|\mathbf{y}) = \frac{L(\mathbf{y}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\int_{\Theta} L(\mathbf{y}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}},$$

donde Θ denota el espacio de parámetros de $\boldsymbol{\theta}$. Dado que el denominador es una constante de normalización y no depende de $\boldsymbol{\theta}$, se suele usar la proporción

$$\pi(\boldsymbol{\theta}|\mathbf{y}) \propto L(\mathbf{y}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}). \tag{2.12}$$

Esta distribución *a posteriori* conjunta puede ser bien complicada, puede incluir integrales múltiples y en ocasiones llega a ser analíticamente intratable. Es por esta razón que se acude a algoritmos computacionales como el método de cadenas de Markov Monte Carlo (MCMC).

2.3.1. Cadenas de Markov Monte Carlo

Las cadenas de Markov Monte Carlo (MCMC) son métodos de simulación para generar muestras de las distribuciones *a posteriori* y estimar cantidades de interés a posteriori. En

los métodos MCMC se simulan valores sucesivamente de una densidad propuesta, que no tiene que ser necesariamente parecida a la densidad *a posteriori*. (Robert & Casella, 2010). Una cadena de Markov es un proceso estocástico² $\theta^{(0)}, \theta^{(1)}, \dots$ tal que la distribución de $\theta^{(t)}$ dados los valores previos $\theta^{(0)}, \theta^{(1)}, \dots, \theta^{(t-1)}$ solo depende de $\theta^{(t-1)}$, o sea

$$f(\theta^{(t+1)}|\theta^{(t)}, \dots, \theta^{(1)}) = f(\theta^{(t+1)}|\theta^{(t)})$$

Para generar una muestra de $f(\theta)$ es deseable que la cadena de Markov construida cumpla con dos propiedades, la primera es que $f(\phi^{(t+1)}|\phi^{(t)})$ debe ser fácil de generar y la segunda es que la distribución estacionaria debe ser la distribución *a posteriori* de interés.

Bajo estos requisitos, podemos aplicar varios algoritmos para construir la cadena de Markov, los más comunes son Gibbs Samplig (Gelfand & Smith, 1990) y Metropolis-Hastings (Chib & Greenberg, 1995). Aquí, expondremos un algoritmo general:

1. Seleccionar un valor inicial.
2. Generar T observaciones MCMC hasta obtener la distribución estacionaria.
3. Monitorear la convergencia del algoritmo usando diagnósticos de convergencia. Si el diagnóstico detecta que no hay convergencia, se generan más observaciones.
4. Desechar las primeras B observaciones.
5. Considerar $\{\theta^{(B+1)}, \theta^{(B+2)}, \dots, \theta^{(T)}\}$
6. Graficar la distribución *a posteriori* (generalmente se pone atención en las distribuciones marginales univariadas)
7. Resumir la distribución *a posteriori* (media, mediana, cuantiles)

2.3.2. Diagnósticos de convergencia

Para comprobar si realmente las cadenas MCMC alcanzan la convergencia hacia la distribución estacionaria que es la distribución *a posteriori* de interés, se necesitan aplicar pruebas

²Un proceso estocástico es una familia de variables aleatorias $\{X(t) : t \in T\}$ definidas sobre un mismo espacio de probabilidad (Ω, μ, P) y con valores en un mismo espacio, llamado espacio de estados, S .

estadísticas. Dos muy usadas son el Diagnóstico de Gelman-Rubin y el Diagnóstico de Geweke.

Diagnóstico de Gelman-Rubin

Este diagnóstico propone hacer una prueba de convergencia basada en 2 o más cadenas paralelas, cada una partiendo de diferentes valores iniciales que están sobre-dispersos con respecto a la verdadera distribución posterior. Su método se basa en una comparación entre las varianzas dentro y entre cadenas para cada variable (esencialmente un análisis clásico de varianza).

Particularmente, el diagnóstico de Gelman-Rubin implementado en la biblioteca *coda* del software R, calcula el “factor de reducción de escala potencial” para cada variable, junto con los límites de confianza superior e inferior (Plummer, Best & Vines, 2006). La convergencia aproximada se diagnostica cuando el límite superior está cerca de 1. Para las cadenas multivariadas, se calcula un valor multivariado que se extiende por encima del factor de reducción de escala potencial para cualquier combinación lineal de las variables.

El factor de reducción de escala potencial (Estadístico de Gelman-Rubin) se estima de la siguiente forma:

$$\sqrt{\widehat{R}} = \sqrt{\frac{\widehat{Var}(\theta)}{W}},$$

donde W es la varianza en las cadenas

$$W = \frac{1}{m(n-1)} \sum_{j=1}^m \sum_{i=1}^n (\theta_j^i - \bar{\theta}_j)^2,$$

\widehat{Var} es la varianza estimada

$$\widehat{Var}(\theta) = \left(1 - \frac{1}{n}\right) W + \frac{1}{n} B$$

y B es la varianza entre cadenas

$$B = \frac{n}{(m-1)} \sum_{j=1}^m (\theta_j - \bar{\theta})^2.$$

El factor \hat{R} disminuye a 1 cuando la longitud de la cadena tiende al infinito, por lo que \hat{R} debería estar cerca de 1 si las cadenas han convergido aproximadamente a la distribución objetivo. Gelman sugiere que \hat{R} debe ser menor que 1.1 ó 1.2 (Rizzo, 2007).

Cuando declaramos la convergencia aproximada, en realidad estamos concluyendo que cada secuencia individual parece estacionaria y que las secuencias observadas se han mezclado bien entre sí. Por esta razón, en la práctica, es común graficar las secuencias de las cadenas y comprobar visualmente si se mezclaron bien.

Diagnóstico de Geweke

Geweke (1992) sugiere un criterio para evaluar la convergencia en base a la comparación de medias, en intervalos diferentes, después de una fase de calentamiento del algoritmo. Si la convergencia fue alcanzada, los comportamientos en esos intervalos deben ser semejantes.

En este diagnóstico se aplica una prueba de Z para comprobar si las medias estimadas a partir de dos submuestras diferentes de la muestra total, son iguales. Usualmente se compara el 10% inicial y el 50% final de la muestra total. Se construyen las medias $\bar{\theta}^A$ y $\bar{\theta}^B$ en base a los grupos de iteraciones T_A y T_B . Entonces el error estándar de la media está dado por $\sqrt{S_\theta(0)/T}$ y por lo tanto el estadístico Z se define como

$$Z = \frac{\bar{\theta}^A - \bar{\theta}^B}{\sqrt{\frac{S_\theta^A(0)}{T_A} + \frac{S_\theta^B(0)}{T_B}}}$$

Bajo la hipótesis de convergencia, asintóticamente Z tiene distribución normal estándar, donde $\bar{\theta}^A, \bar{\theta}^B$ son las medias muestrales de las dos submuestras mencionadas, T_A y T_B son los tamaños, y $S_\theta^A(0)$ y $S_\theta^B(0)$ son las varianzas.

Los parámetros con $|Z| \geq 2$ tienen medias diferentes de las primeras y últimas iteraciones, y por lo tanto indican no convergencia de la cadena MCMC (Ntzoufras, 2009). Sin embargo, valores pequeños de esta estadística no significan que existe convergencia. Luego, la decisión debe ser tomada conjuntamente con otros criterios y con la inspección visual. Para implementar el criterio de Geweke, basta con efectuar solo una cadena larga del algoritmo MCMC.

2.3.3. Estimadores Bayesianos

Aunque la información sobre θ está en su distribución *a posteriori*, es conveniente resumirla en un valor representativo de éste. Una opción puede ser utilizar su esperanza matemática, como indicador del valor medio del parámetro, conocida como estimador a posteriori, y que responde a la expresión $E[\pi(\theta|y)]$.

Otra opción, basada en la teoría de la decisión, consiste en la especificación de una función de pérdida $f_p(\theta, a)$, que cuantifique la penalización por decidir que el parámetro tiene el valor a , cuando de hecho es igual a θ . Con ella, se puede calcular el llamado riesgo de Bayes, $R_a(\theta) = E[f_p(\theta, a)]$, como la esperanza de la pérdida dada la distribución *a posteriori* del parámetro. Entonces se puede utilizar el denominado estimador de Bayes, consistente en elegir la estimación que minimiza el riesgo de Bayes.

Tenga en cuenta que especificar la función de pérdida es difícil, ya que la pérdida no siempre es fácilmente medible. Algunas de las funciones de pérdida más utilizadas son:

- La función de pérdida cuadrática:

$$f_p(\theta, a) = (\theta - a)^2$$

produce como estimador Bayesiano la media de la distribución *a posteriori*.

- La función de pérdida lineal absoluta:

$$f_p(\theta, a) = |\theta - a|$$

produce como estimador Bayesiano la mediana de la distribución *a posteriori*.

- La función de pérdida todo/nada:

$$f_p(\theta, a) = \begin{cases} 0 & \text{si } \theta = a \\ 1 & \text{si } \theta \neq a \end{cases}$$

produce como el estimador Bayesiano la moda de la distribución *a posteriori* (Gelman, 2014).

2.3.4. Intervalos de credibilidad

En el enfoque Bayesiano, se define una región creíble a un cierto nivel de probabilidad, $1 - \alpha$, como un conjunto de valores que contenga al parámetro con una probabilidad de, al menos, $1 - \alpha$, es decir C_α tal que $P(\theta \in C_\alpha) \geq (1 - \alpha)$

Definición 2.3.2 Un intervalo de credibilidad al $100(1 - \alpha)\%$ para θ es un subconjunto C_α de Θ tal que

$$1 - \alpha \leq P(C_\alpha|y) = \begin{cases} \int_{C_\alpha} f(\theta|y)d\theta & \text{caso continuo} \\ \sum_{\theta \in C_\alpha} f(\theta|y) & \text{caso discreto} \end{cases}$$

Dado que es frecuente que se esté interesado en los valores más probables del parámetro, es deseable que la longitud del intervalo sea la mínima posible (Cepeda, Aguilar & Cervantes, 2008).

Metodología

Utilizamos un enfoque Bayesiano para estimar los parámetros del modelo (2.11). En esta configuración, los parámetros $K_p, K_d, K_i, \mu, \lambda$ se consideran como variables aleatorias. Las mismas tienen distribuciones *a priori* que reflejan el conocimiento sobre sus valores antes de observar los datos. El procedimiento consiste en determinar la distribución *a posteriori* de los parámetros.

Para poder aplicar la función de transferencia (2.11) del controlador PID fraccionario a los datos de impedancia eléctrica, necesitamos hacer la transformación $s = j\omega$. Luego, utilizando propiedades de números complejos:

$$\begin{aligned}
 C(j\omega) &= K_p + K_d(j\omega)^\mu + \frac{K_i}{(j\omega)^\lambda} \\
 &= K_p + K_d\omega^\mu \exp\left(\frac{\mu\pi}{2}j\right) + \frac{K_i}{\omega^\lambda} \exp\left(-\frac{\lambda\pi}{2}j\right) \\
 &= K_p + K_d\omega^\mu \left[\cos\left(\frac{\mu\pi}{2}\right) + j \sin\left(\frac{\mu\pi}{2}\right) \right] + \frac{K_i}{\omega^\lambda} \left[\cos\left(\frac{\lambda\pi}{2}\right) - j \sin\left(\frac{\lambda\pi}{2}\right) \right] \\
 &= K_p + K_d\omega^\mu \cos\left(\frac{\mu\pi}{2}\right) + \frac{K_i}{\omega^\lambda} \cos\left(\frac{\lambda\pi}{2}\right) + j \left[K_d\omega^\mu \sin\left(\frac{\mu\pi}{2}\right) - \frac{K_i}{\omega^\lambda} \sin\left(\frac{\lambda\pi}{2}\right) \right]
 \end{aligned}$$

$$z_r = K_p + K_d\omega^\mu \cos\left(\frac{\mu\pi}{2}\right) + \frac{K_i}{\omega^\lambda} \cos\left(\frac{\lambda\pi}{2}\right) \quad z_j = K_d\omega^\mu \sin\left(\frac{\mu\pi}{2}\right) - \frac{K_i}{\omega^\lambda} \sin\left(\frac{\lambda\pi}{2}\right)$$

3.1. Datos

Suponemos la siguiente ecuación de observación para los datos de impedancia eléctrica

$$\mathbf{y}_i = \mathbf{z}_i + \mathbf{e}_i; \quad i = 1, \dots, n$$

donde $\mathbf{z}_i = \begin{pmatrix} z_r \\ z_j \end{pmatrix}_i$, $\mathbf{y}_i = \begin{pmatrix} y_r \\ y_j \end{pmatrix}_i$, $\mathbf{e}_i \sim N_2(\mathbf{0}, \Sigma)$, donde $\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$. y_r, y_j representan la parte real y la parte imaginaria de la impedancia eléctrica observada.

Por lo tanto, la función de verosimilitud queda expresada como:

$$f(\mathbf{y}_i | \mathbf{z}_i) = \left(\frac{1}{2\pi} \right) |\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{y}_i - \mathbf{z}_i)' \Sigma^{-1} (\mathbf{y}_i - \mathbf{z}_i) \right\}.$$

3.2. Los Parámetros

Además de los parámetros del modelo (2.11) estimamos la desviación estándar σ , entonces los parámetros de interés son $\boldsymbol{\theta} = (K_p, K_d, K_i, \mu, \lambda, \sigma)$. Para construir la distribución *a priori* de $\boldsymbol{\theta}$ se asume la independencia entre los parámetros, por lo tanto:

$$\pi(\boldsymbol{\theta}) = \pi_p(K_p) \times \pi_i(K_i) \times \pi_d(K_d) \times \pi_\mu(\mu) \times \pi_\lambda(\lambda) \times \pi_\sigma(\sigma)$$

Todos los parámetros involucrados en el controlador son positivos, por tal razón se eligieron distribuciones cuyo soporte fuera positivo. Para K_p, K_d y K_i se eligieron distribuciones Gammas. En el caso del orden de la derivada y de la integral, solo nos interesan los valores en el intervalo $(0, 1)$, por lo que una distribución Beta es una buena candidata. En el caso de la desviación estándar, asignamos una distribución *a priori* no informativa Uniforme. Veamos:

- $K_p \sim \Gamma(\alpha_1, \beta_1) \quad \pi_p(K_p) = \frac{1}{\beta_1^{\alpha_1} \Gamma(\alpha_1)} K_p^{\alpha_1-1} e^{-\frac{K_p}{\beta_1}} \quad \alpha_1 > 0, \beta_1 > 0$
- $K_i \sim \Gamma(\alpha_2, \beta_2) \quad \pi_i(K_i) = \frac{1}{\beta_2^{\alpha_2} \Gamma(\alpha_2)} K_i^{\alpha_2-1} e^{-\frac{K_i}{\beta_2}} \quad \alpha_2 > 0, \beta_2 > 0$
- $K_d \sim \Gamma(\alpha_3, \beta_3) \quad \pi_d(K_d) = \frac{1}{\beta_3^{\alpha_3} \Gamma(\alpha_3)} K_d^{\alpha_3-1} e^{-\frac{K_d}{\beta_3}} \quad \alpha_3 > 0, \beta_3 > 0$

- $\mu \sim B(\alpha_4, \beta_4)$ $\pi_\mu(\mu) = \frac{\Gamma(\alpha_4 + \beta_4)}{\Gamma(\alpha_4)\Gamma(\beta_4)} \mu^{\alpha_4-1} (1 - \mu)^{\beta_4-1}$ $\alpha_4 > 0, \beta_4 > 0$
- $\lambda \sim B(\alpha_5, \beta_5)$ $\pi_\lambda(\lambda) = \frac{\Gamma(\alpha_5 + \beta_5)}{\Gamma(\alpha_5)\Gamma(\beta_5)} \lambda^{\alpha_5-1} (1 - \lambda)^{\beta_5-1}$ $\alpha_5 > 0, \beta_5 > 0$
- $\sigma \sim U(a, b)$ $\pi_\sigma(\sigma) = \frac{1}{b - a}$ $a < \sigma < b$

Según lo analizado en la Sección 2.3, la distribución *a posteriori* viene dada por la ecuación (2.12), por tanto,

$$\begin{aligned}
\pi(\boldsymbol{\theta}|\mathbf{y}) &\propto L(\mathbf{y}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}) \\
&= \left(\frac{1}{2\pi}\right)^{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{y}_i - \mathbf{z}_i)' \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \mathbf{z}_i)\right\} \\
&\times \pi_p(K_p) \times \pi_i(K_i) \times \pi_d(K_d) \times \pi_\mu(\mu) \times \pi_\lambda(\lambda) \times \pi_\sigma(\sigma)
\end{aligned} \tag{3.1}$$

Usando la función de pérdida cuadrática, el estimador puntual Bayesiano es la **Media condicional (CM)**

$$\hat{\boldsymbol{\theta}}_B = E(\boldsymbol{\theta}|\mathbf{y}).$$

Tenga en cuenta que la distribución posterior conjunta (3.1) es analíticamente intratable y las distribuciones posteriores marginales de los parámetros son complicadas. Por tanto, usaremos la técnica de Cadenas de Markov Monte Carlo para aproximarla. La estimación se hace usando las herramientas JAGS y R. En JAGS se escribe el modelo con los parámetros a estimar y este software conjuntamente con R construyen las cadenas de Markov.

Resultados

En este capítulo se muestran dos ejemplos de estimación de los parámetros del controlador PID fraccionario ($PI^\lambda D^\mu$). Para ambos casos se detallan las *a priori* con los hiperparámetros, las estimaciones Bayesianas de los parámetros y los diagnósticos de convergencia de las cadenas de Markov.

4.1. Estimación de los parámetros del controlador PID fraccionario para controlar el nivel en un tanque cónico

Uno de los tipos de estanques más utilizados en la industria son los tanques cónicos, ya que garantizan un drenaje completo de los fluidos que contienen. Una tarea importante es controlar el nivel del tanque para mantenerlo en las condiciones de llenado que se precisen. Varios son los estudios que se han realizado para tratar esta tarea. Se han sintonizado controladores PID, tanto enteros como fraccionarios, mediante reglas heurísticas u optimización no lineal, respectivamente, de acuerdo a especificaciones de control en el dominio de la frecuencia. Aquí se han empleado técnicas de computación evolutiva, utilizando algoritmos genéticos para determinar los valores que minimicen una cierta función objetivo. Jauregui et al. (2016) realizan un experimento para sintonizar los parámetros de un controlador PID fraccionario para controlar el nivel de un tanque cónico que tiene un sensor de presión para obtener los

datos. La técnica empleada fue la Optimización por Enjambre de Partículas (PSO, por sus siglas en inglés), que no explicaremos aquí y que puede encontrar en (Kennedy & Eberhart, 1995). Ahora realizaremos la estimación pero desde el enfoque Bayesiano, y compararemos nuestros resultados con los obtenidos por Jauregui et al. (2016).

El primer paso consiste en la simulación de los datos. A partir del modelo (2.11) y con las estimaciones de los parámetros obtenidas mediante PSO por Jauregui et al. (2016), se generó una muestra sintética de 50 observaciones. La Tabla 4.1 muestra los valores de los parámetros. La varianza la tomamos con valor 5.

| K_p | K_d | K_i | μ | λ |
|--------|-------|-------|-------|-----------|
| 122.46 | 3.62 | 100 | 0.18 | 0.86 |

Tabla 4.1: Estimaciones obtenidas por la técnica PSO.

La Figura 4.1 muestra el gráfico de los datos de impedancia simulados. Se observan la magnitud y la fase que representan a la impedancia.

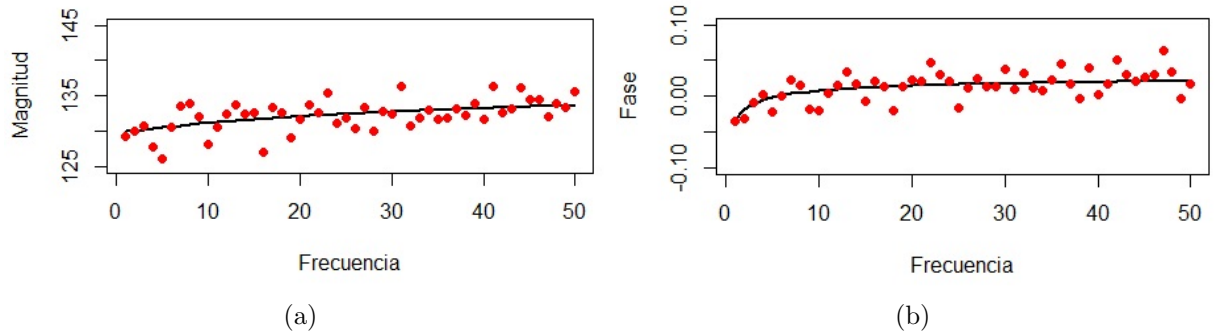


Figura 4.1: Desviación de los datos respecto a la media. Magnitud (a) y Fase (b).

4.1.1. Distribución *a priori* de los parámetros.

Los valores de los parámetros del controlador PID fraccionario, son dependientes del fenómeno en el que se esté usando el controlador. Por esta razón, es difícil encontrar un consenso para establecer las *a priori*. Como se explicó en la Sección 3.2, lo que sí suele observarse en la bibliografía, es que generalmente los parámetros se toman positivos. Siguiendo este principio, optamos por asignar distribuciones *a priori* Gammas para los parámetros K_p , K_d y K_i ,

ya que su soporte es positivo. Para el caso de los órdenes de las derivada y la integral fraccionarias, en la bibliografía Jauregui et al. (2016) reportan estos valores entre 0 y 1, por lo que decidimos usar distribuciones *a priori* Beta. Además de los parámetros anteriores, estimamos la varianza, para ella usamos una *a priori* no informativa Uniforme en el intervalo de 0 a 100.

- $\pi_p(K_p) = \Gamma(1500, 12.246)$
- $\pi_d(K_d) = \Gamma(6.5522, 1.81)$
- $\pi_i(K_i) = \Gamma(1000, 10)$
- $\pi_\mu(\mu) = \mathcal{B}(1, 1)$
- $\pi_\lambda(\lambda) = \mathcal{B}(1, 1)$
- $\pi_\sigma(\sigma) = \mathcal{U}(0, 100)$

Los hiperparámetros de las distribuciones garantizan que las zonas de mayor probabilidad estén cerca de sus medias. Las gráficas pueden verse en la Figura 4.2.

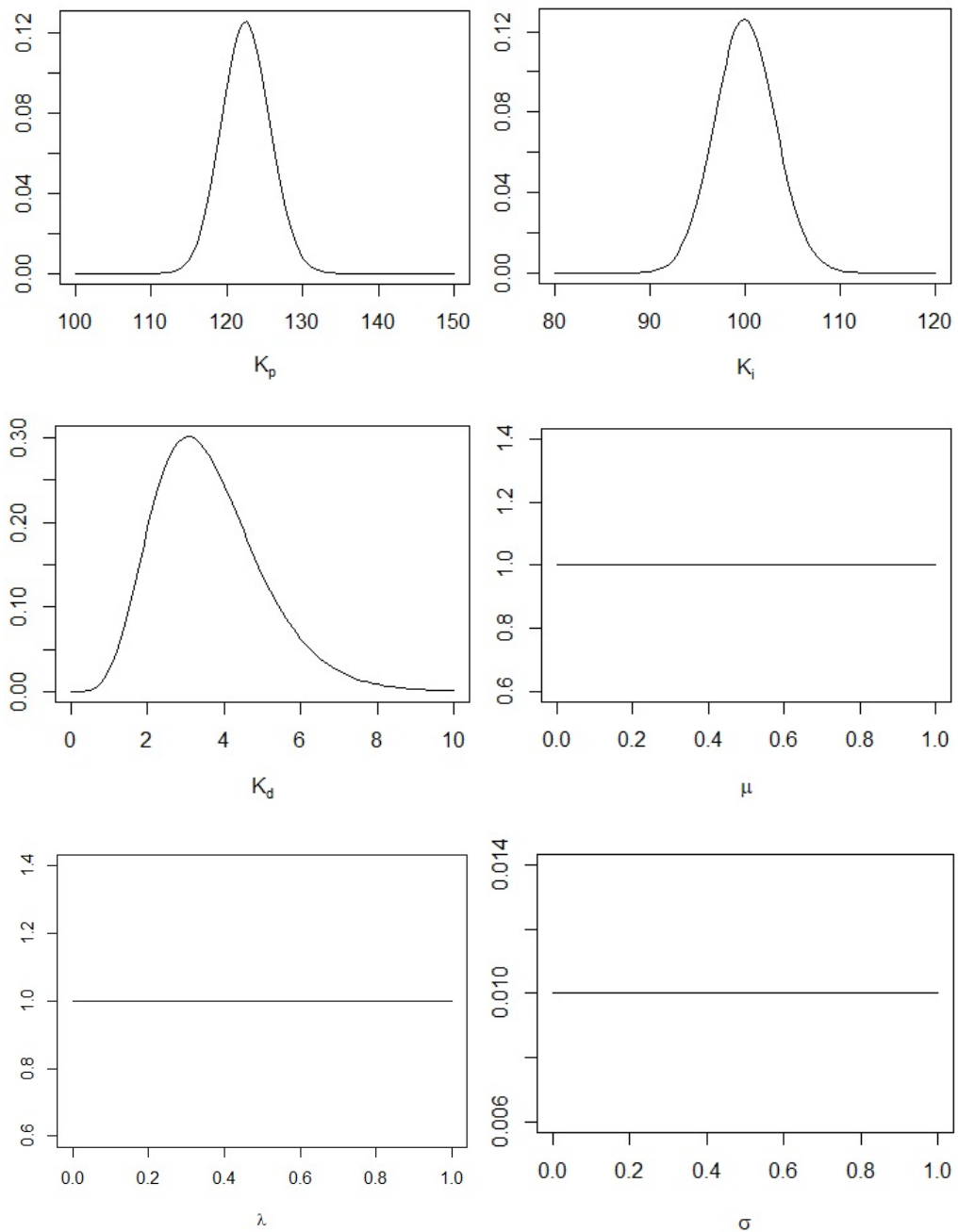


Figura 4.2: Gráficas de las densidades *a priori* de los parámetros.

4.1.2. Estimaciones obtenidas

Establecidas las distribuciones *a priori*, procedemos a hacer la estimación usando la técnica Bayesiana. Para aproximar la distribución posterior se usan 3 cadenas MCMC y 70000 iteraciones y descartamos las primeras 30000. Las estimaciones obtenidas se muestran en la Tabla

4.2, la que también incluye los valores obtenidos por Jauregui et al. (2016) y los intervalos de credibilidad.

| Parámetro | Estimación | Valor referencia | 2.5 % | 97.5 % |
|------------|------------|------------------|---------|---------|
| K_p | 122.814 | 122.460 | 118.998 | 126.299 |
| K_d | 3.643 | 3.620 | 1.587 | 6.287 |
| K_i | 100.111 | 100.000 | 94.029 | 106.362 |
| μ | 0.179 | 0.180 | 0.121 | 0.258 |
| λ | 0.859 | 0.860 | 0.739 | 0.982 |
| σ^2 | 4.920 | 5.000 | 3.713 | 6.564 |

Tabla 4.2: Estimaciones Bayesianas e intervalos de credibilidad de los parámetros.

La Figura 4.3 muestra el ajuste a los datos con las estimaciones obtenidas. Se puede ver que hay un buen ajuste.

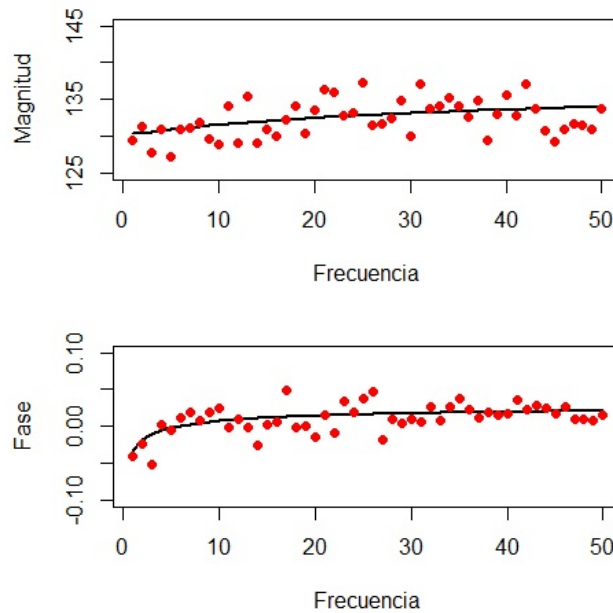


Figura 4.3: Ajuste a los datos. Línea negra representa el modelo ajustado.

4.1.3. Diagnósticos de convergencia

Para verificar si las tres cadenas de Markov Monte Carlo convergieron a la distribución estacionaria que aproxima la *a posteriori*, se realizaron pruebas visuales y los diagnósticos

de Gelman-Rubin y de Geweke. En la Figura 4.4, puede observarse que las tres cadenas se mezclaron y podemos decir, que visualmente, las cadenas convergieron.

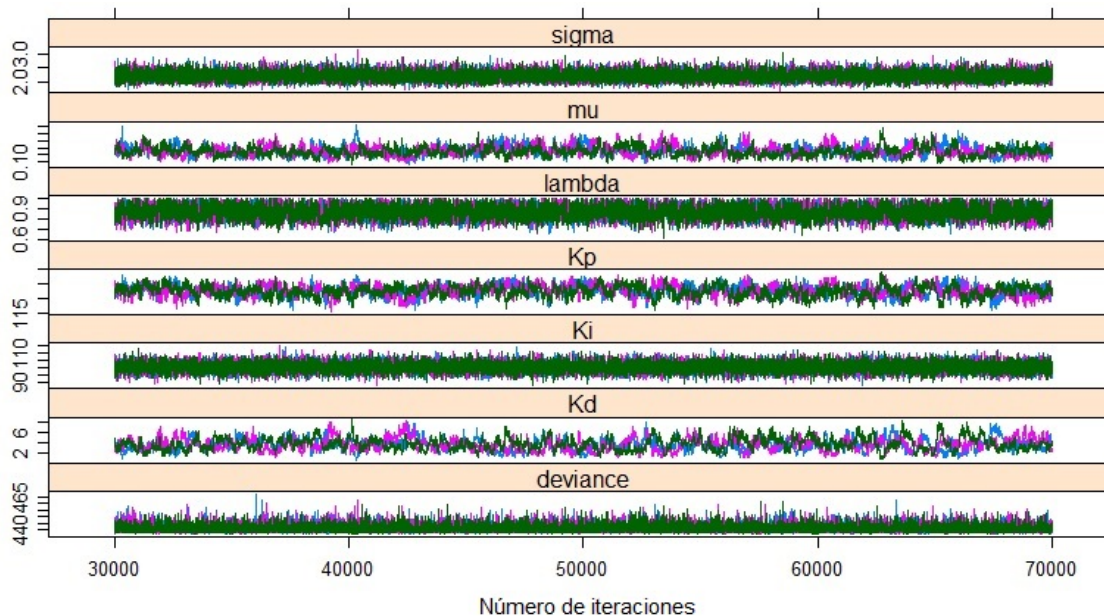


Figura 4.4: Convergencia de las cadenas MCMC.

La Figura 4.5 muestra las gráficas de las densidades de las distribuciones *a posteriori* marginales de cada parámetro. Podemos ver que las tres cadenas tienen densidades muy parecidas, otra razón para asegurar que las cadenas convergieron. También podemos percatarnos de que las medias de estas densidades están cerca de los valores de referencia de los parámetros, o sea, los obtenidos por Jauregui et al. (2016).

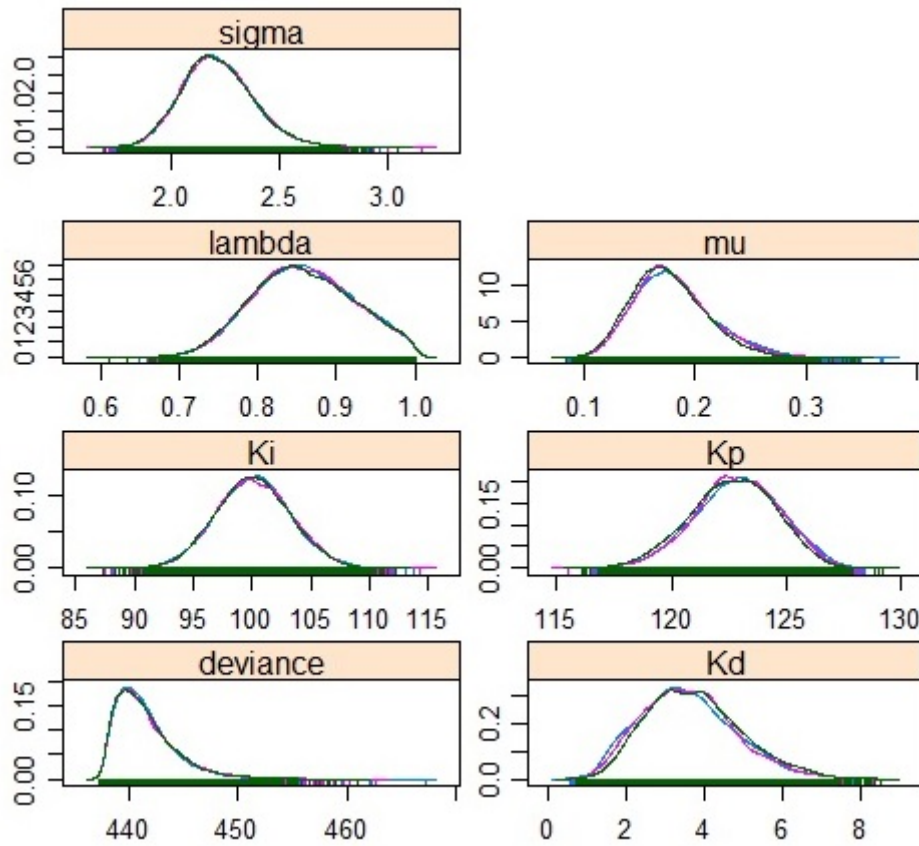


Figura 4.5: Gráfica de las densidades de las distribuciones *a posteriori*.

La Figura 4.6 y la Figura 4.7 también muestran otros gráficos de las trazas de las cadenas y las densidades de las distribuciones *a posteriori*. Se puede observar que las cadenas se mezclan bien y las densidades son cerradas.

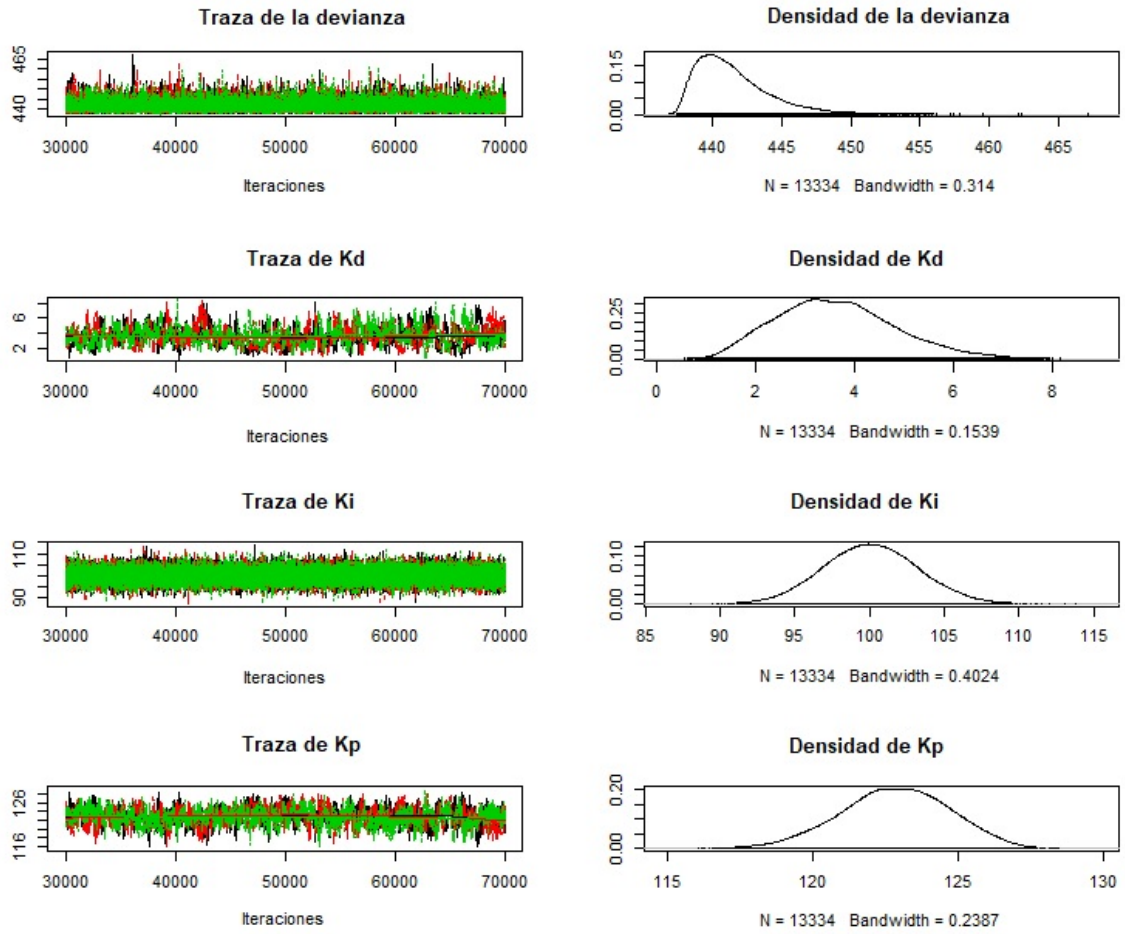


Figura 4.6: Trazas de las MCMC y densidades *a posteriori* de los parámetros para el ejemplo del tanque cónico.

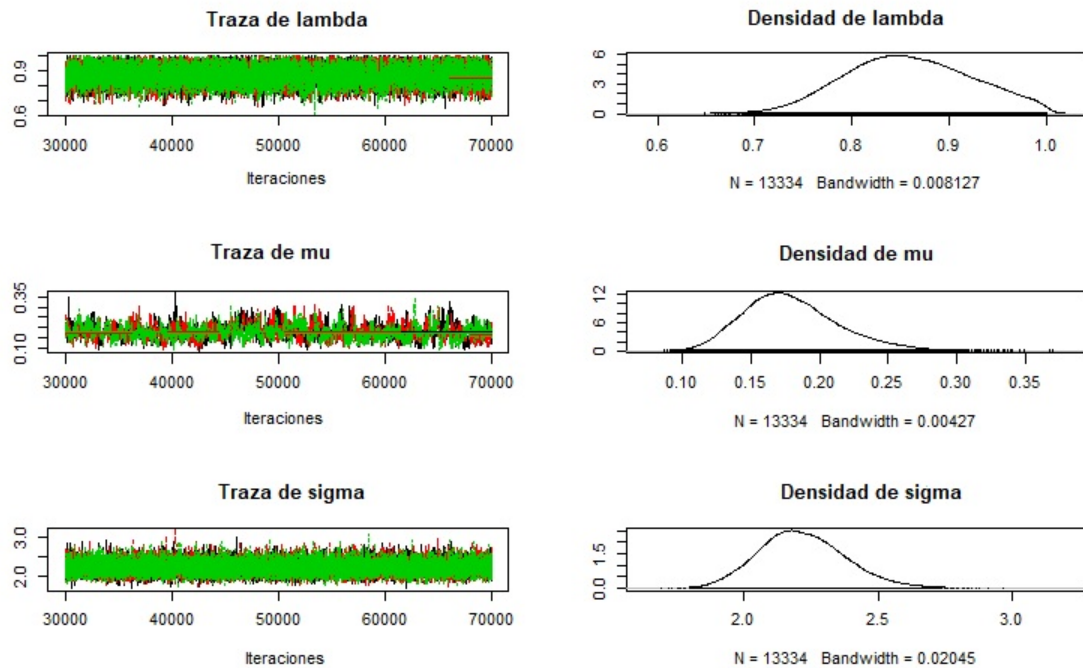


Figura 4.7: Trazas de las MCMC y densidades *a posteriori* de los parámetros para el ejemplo del tanque cónico.

Para tener una mayor seguridad y corroborar los resultados de los exámenes visuales, se realizaron los diagnósticos de convergencia de Gelman-Rubin y de Geweke. La Tabla 4.3 contiene el resultado de la prueba de Gelman-Rubin, note que todos los valores son menores que 1.2. Además, en la Figura 4.8 se puede observar que el comportamiento de las gráficas es cercano a 1. Tomando en cuenta ambas razones, podemos decir, que según la prueba de Gelman-Rubin, las cadenas alcanzaron la convergencia.

| Parámetro | Estimación Puntual | IC Superior |
|-----------|--------------------|-------------|
| K_p | 1.01 | 1.02 |
| K_d | 1.01 | 1.02 |
| K_i | 1.00 | 1.00 |
| μ | 1.00 | 1.02 |
| λ | 1.00 | 1.00 |
| σ | 1.00 | 1.00 |

Tabla 4.3: Diagnóstico de Gelman-Rubin.

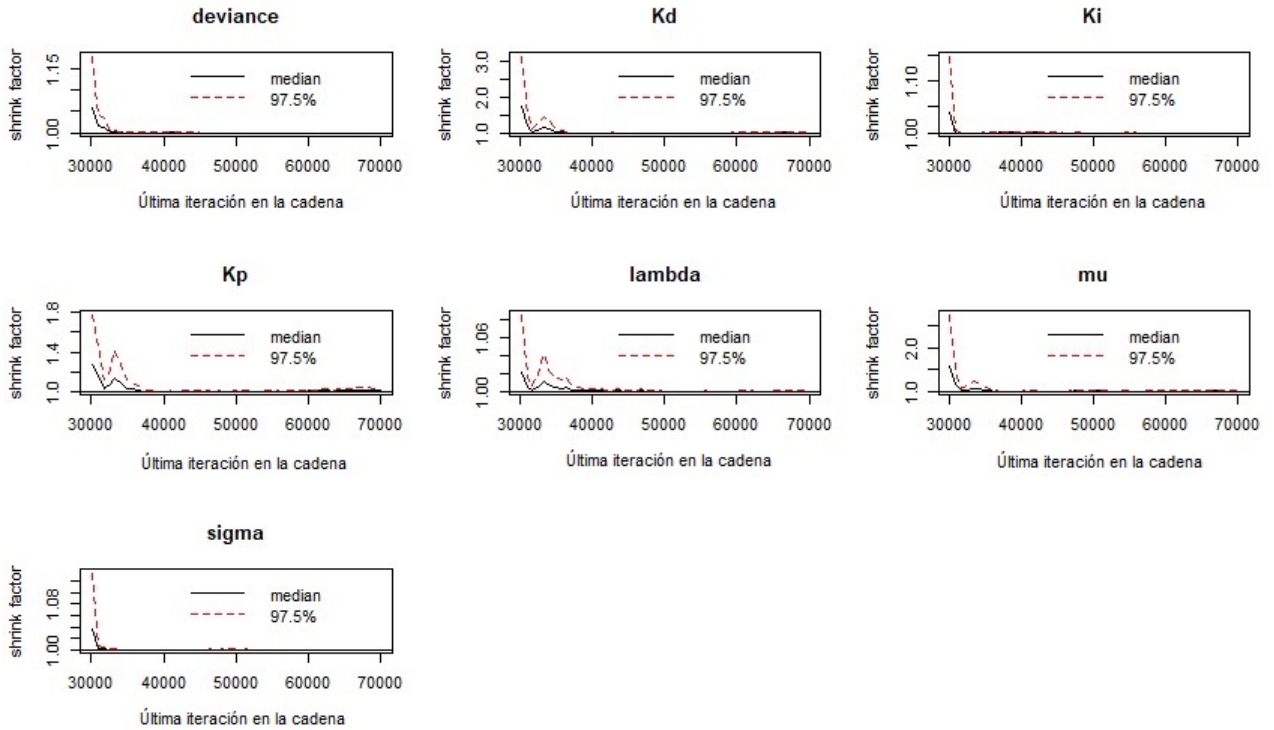


Figura 4.8: Diagnóstico de convergencia de Gelman-Rubin.

Para cada una de las cadenas y en cada parámetro, el diagnóstico de convergencia de Geweke dio como resultado $|Z| < 2$. Esto indica que las medias de las primeras y últimas iteraciones son iguales y se alcanza la convergencia de las cadenas.

4.1.4. Evaluación

El comportamiento del método de estimación se evaluó en términos de sesgo y Error Cuadrático Medio (ECM) siguiendo el Algoritmo 1:

1. Generar 100 muestras de tamaño 50 usando los valores de referencia de los parámetros
2. Para cada muestra de 1 a 100
 - Estimar los parámetros usando 3 cadenas MCMC
3. Para cada parámetro p en $\theta = (K_p, K_d, K_i, \mu, \lambda, \sigma)$
 - Obtener \hat{p} que es la media de las 100 estimaciones
 - Calcular el sesgo por la fórmula $sesgo(\hat{p}, p_R) = \hat{p} - p_R$, donde p_R es el valor real del parámetro p

- Calcular el ECM por la fórmula $ECM(\hat{p}) = Var(\hat{p}) + sesgo(\hat{p}, p_R)^2$

Los resultados se muestran en la Tabla 4.4. Podemos observar que los valores, tanto del sesgo como del error cuadrático medio, son pequeños. Dado que el sesgo es una diferencia del valor de referencia con el valor estimado, podemos decir, que como es pequeño, las estimaciones obtenidas están cerca de los valores reales. Además, en el cálculo del error cuadrático medio interviene la varianza, si el ECM es pequeño, significa que la varianza no es grande. Luego, tenemos estimadores con valores cercanos a los de referencia y con varianzas pequeñas. Podemos decir que son buenos estimadores.

| Parámetros | Sesgo | ECM |
|------------|---------|--------|
| K_p | 0.2356 | 0.3151 |
| K_d | -0.0980 | 0.0523 |
| K_i | -0.0982 | 0.0435 |
| μ | 0.0070 | 0.0003 |
| λ | 0.0187 | 0.0028 |
| σ | 0.0261 | 0.0378 |

Tabla 4.4: Resultados de sesgo y ECM.

4.2. Estimación de los parámetros del controlador PID fraccionario. Otro ejemplo.

Dado que no contamos con datos reales de impedancia eléctrica, para realizar la estimación de los parámetros $\theta = (K_p, K_d, K_i, \mu, \lambda, \sigma)$, primero se generó una muestra sintética a partir del modelo (2.11). Se obtuvo una muestra con 50 observaciones con los valores de los parámetros obtenidos de Monje et al. (2010), mostrados en la Tabla 4.5. La varianza σ^2 la tomamos igual a 5. Los datos simulados de impedancia se muestran en la Figura 4.9 en puntos. La línea negra representa la media de la distribución normal, a partir de la que se generan los datos.

| K_p | K_d | K_i | μ | λ |
|--------|--------|--------|--------|-----------|
| 0.6152 | 4.3867 | 0.0100 | 0.4773 | 0.8968 |

Tabla 4.5: Valores de los parámetros tomados de Monje et al. (2010).

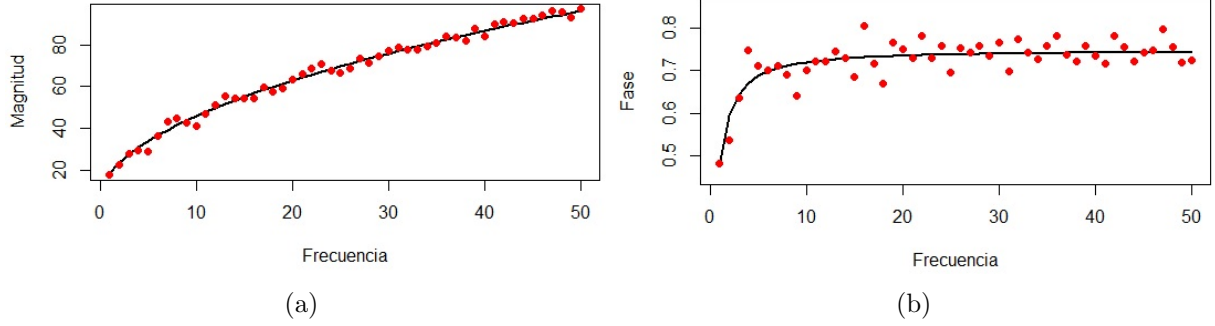


Figura 4.9: Desviación de los datos respecto a la media. Magnitud (a) y Fase (b).

4.2.1. Distribución *a priori* de los parámetros.

En la bibliografía consultada no encontramos referentes para asignar *a priori*s específicas a los parámetros. Por tanto, las distribuciones *a priori* se tomaron con hiperparámetros de manera tal que sus medias estuvieran cerca de los valores de los parámetros tomados de Monje et al. (2010) y que su soporte fuera consecuente con el soporte del parámetro. De acuerdo a lo analizado en la Sección 3.2, las *a priori*s son:

- $\pi_p(K_p) = \Gamma(1.2304, 2)$
- $\pi_d(K_d) = \Gamma(3.8486, 0.8773)$
- $\pi_i(K_i) = \Gamma(0.033, 3.3333)$
- $\pi_\mu(\mu) = \mathcal{B}(1, 1)$
- $\pi_\lambda(\lambda) = \mathcal{B}(1, 1)$
- $\pi_\sigma(\sigma) = \mathcal{U}(0, 100)$

La distribución uniforme de σ es una distribución no informativa, ya que le da la misma probabilidad a todos los valores del intervalo.

En la Figura 4.10 se pueden ver las gráficas de las densidades de las distribuciones *a priori*. Note como las zonas de mayor probabilidad están cerca de la media.

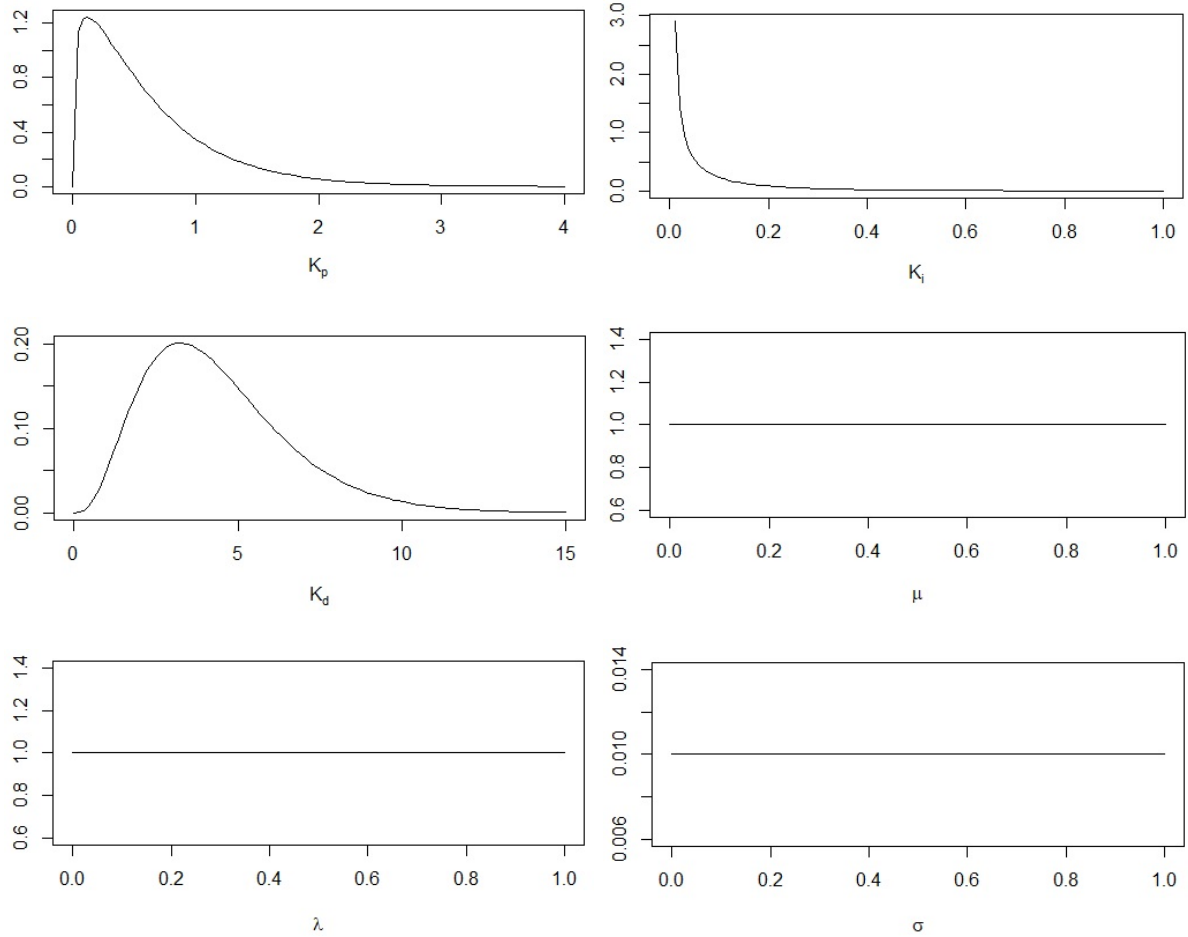


Figura 4.10: Gráficas de las densidades *a priori* de los parámetros.

4.2.2. Estimaciones obtenidas

Una vez establecidas las distribuciones *a priori* de los parámetros de interés, se procedió a estimar. Las estimaciones se obtuvieron a través de 3 cadenas MCMC, empleando 70000 iteraciones y descartando las primeras 30000. Los resultados se muestran en la Tabla 4.6.

| Parámetro | Estimación | Valor real | 2.5 % | 97.5 % |
|------------|------------|------------|--------|--------|
| K_p | 0.6388 | 0.6152 | 0.0310 | 2.0800 |
| K_d | 4.3875 | 4.3867 | 4.0400 | 4.6320 |
| K_i | 0.0165 | 0.0100 | 0.0050 | 0.0620 |
| μ | 0.4775 | 0.4773 | 0.4690 | 0.4890 |
| λ | 0.8232 | 0.8968 | 0.4600 | 0.9930 |
| σ^2 | 4.6565 | 5.0000 | 3.5044 | 6.2450 |

Tabla 4.6: Estimaciones Bayesianas e intervalos de credibilidad de los parámetros.

Para todos los parámetros, el estimador puntual usado es la media de la distribución *a posteriori* marginal del parámetro.

La Figura 4.11 muestra el ajuste a los datos con las estimaciones obtenidas. Se puede ver que hay un buen ajuste.

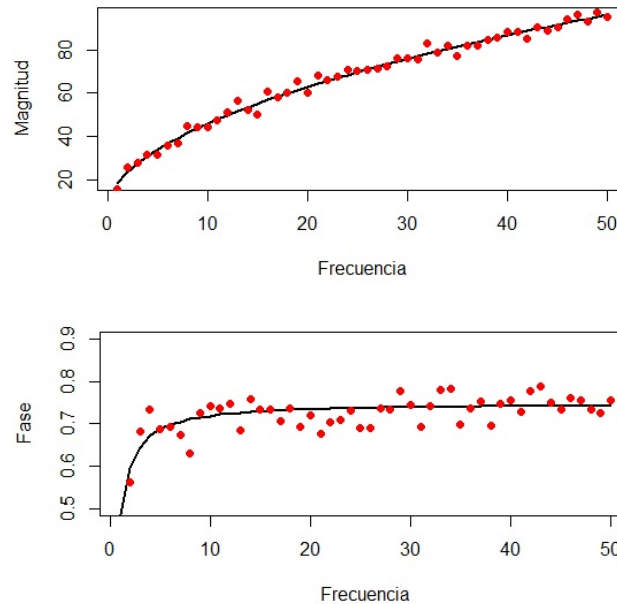


Figura 4.11: Ajuste a los datos. Línea negra representa el modelo ajustado.

4.2.3. Diagnósticos de convergencia

Para validar los resultados de las estimaciones anteriores y constatar que las cadenas de Markov convergieron, se graficaron las trazas de las cadenas y se les hizo un examen visual.

En la Figura 4.12, se puede observar que las 3 cadenas se mezclaron bien y que se alcanza la convergencia.

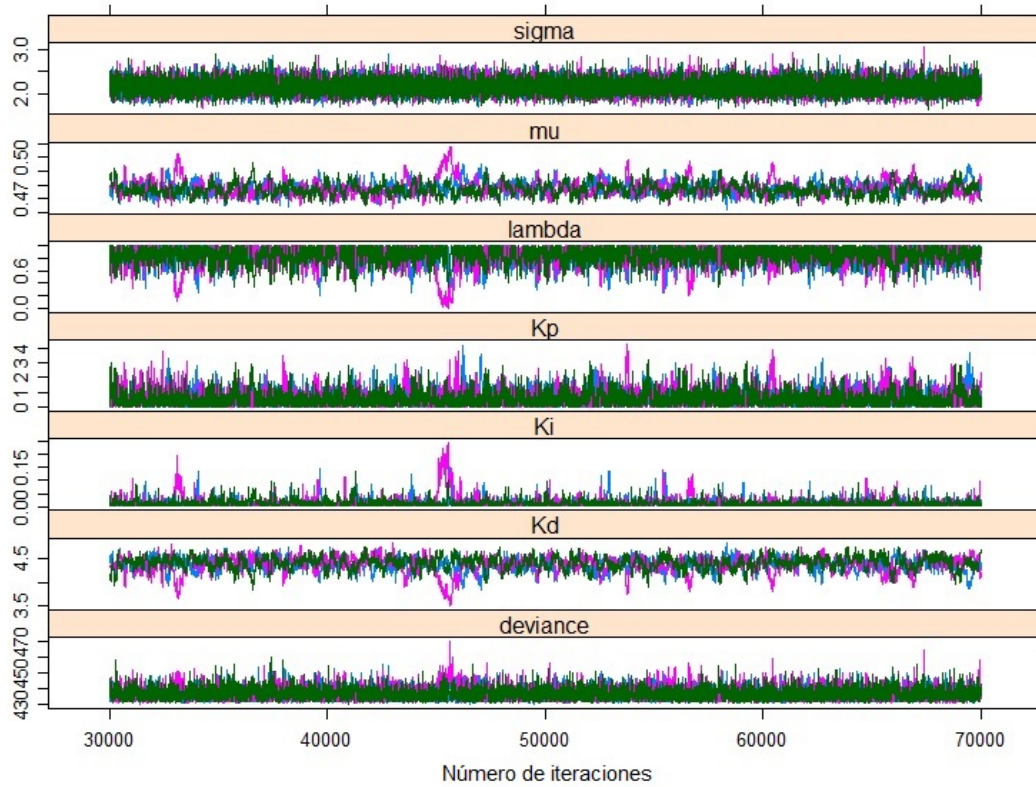


Figura 4.12: Convergencia de las cadenas MCMC.

En la Figura 4.13, se pueden observar las distribuciones de las densidades *a posteriori* de los parámetros. Vemos que las medias de las distribuciones *a posteriori* de cada parámetro, están muy cerca del valor de referencia tomado de Monje et al. (2010). La distribución posterior de σ tiene una media cercana a 2.236, que sería la raíz cuadrada de 5 que es el valor real de σ^2 . Además, las densidades de cada cadena están muy cercanas, de manera que se concluye que se obtuvo una buena mezcla y las cadenas convergieron.

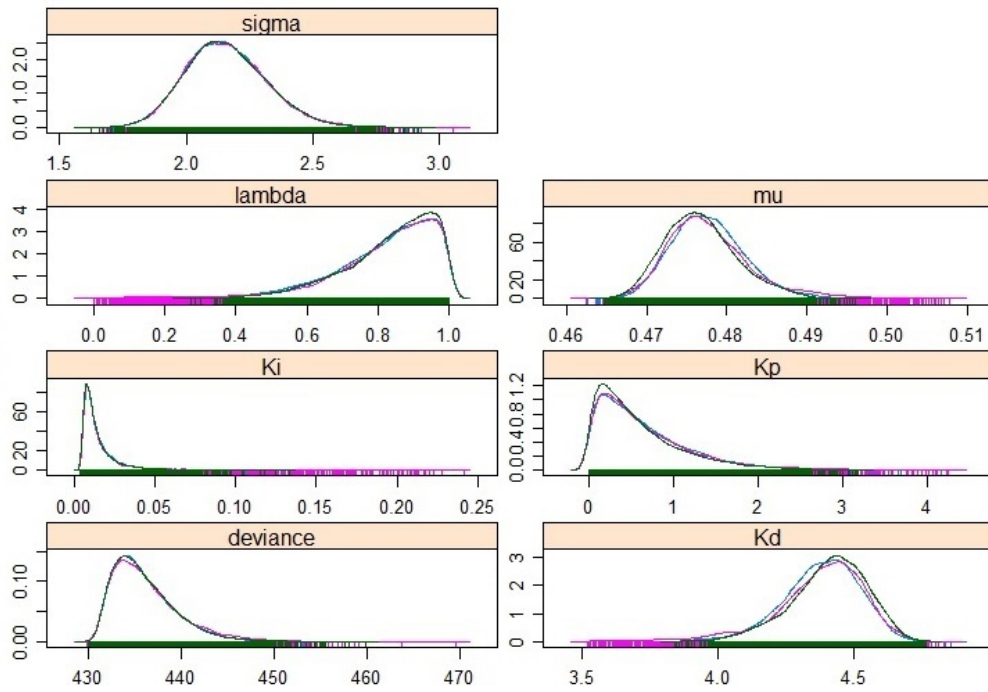


Figura 4.13: Gráfica de las densidades de las distribuciones *a posteriori*.

En la Figura 4.14 y en la Figura 4.15, se muestra otra vista de los gráficos de las trazas de las cadenas MCMC y de las densidades de las distribuciones *a posteriori* de los parámetros estimados del controlador PID fraccionario. La gráfica muestra que las cadenas se mezclan bien, las densidades son cerradas y sus medias están cerca de los valores reales de los parámetros.

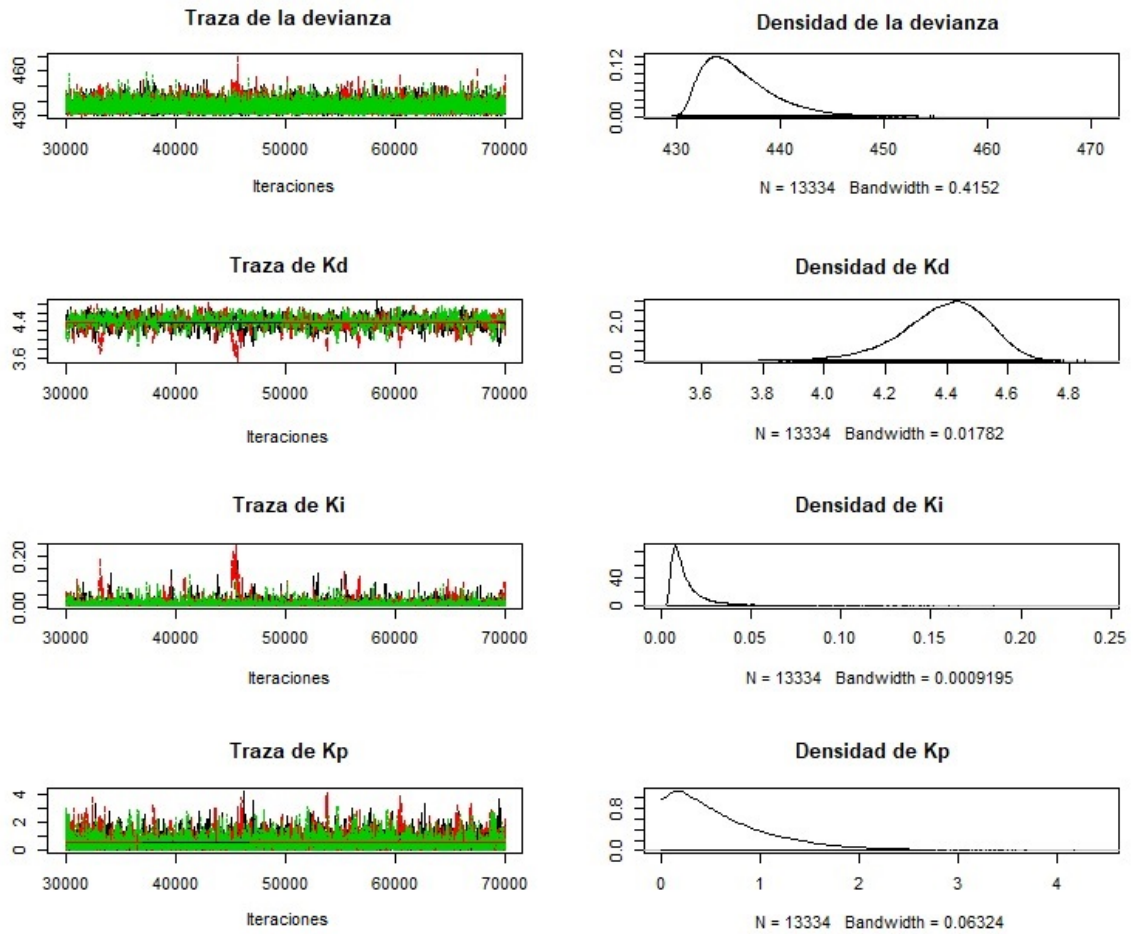


Figura 4.14: Trazas de las MCMC y densidades para el ejemplo tomado de Monje et al. (2010).

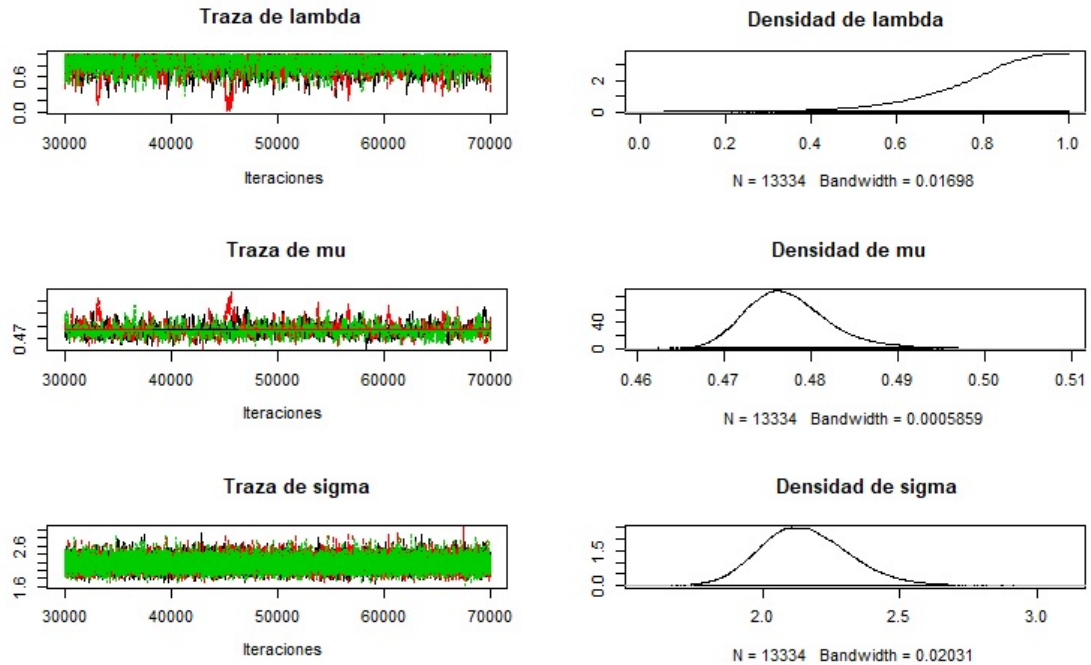


Figura 4.15: Trazas de las MCMC y densidades para el ejemplo tomado de Monje et al. (2010).

Para tener mayor confianza en la convergencia, además de los diagnósticos visuales, se realizaron las pruebas de convergencia de Geweke y Gelman-Rubin. En la Tabla 4.7 se muestran los resultados del diagnóstico de Gelman-Rubin y todos los valores son menores que 1.2. En la Figura 4.16, se puede apreciar que a partir de la segunda mitad el comportamiento de cada una de las gráficas es cercano a 1. Por ambas razones, con el diagnóstico de Gelman-Rubin se puede afirmar que las cadenas alcanzaron la convergencia.

| Parámetro | Estimación Puntual | IC Superior |
|------------|--------------------|-------------|
| K_p | 1.00 | 1.01 |
| K_d | 1.02 | 1.04 |
| K_i | 1.07 | 1.09 |
| μ | 1.02 | 1.04 |
| λ | 1.01 | 1.02 |
| σ^2 | 1.00 | 1.00 |

Tabla 4.7: Diagnóstico de Gelman-Rubin.

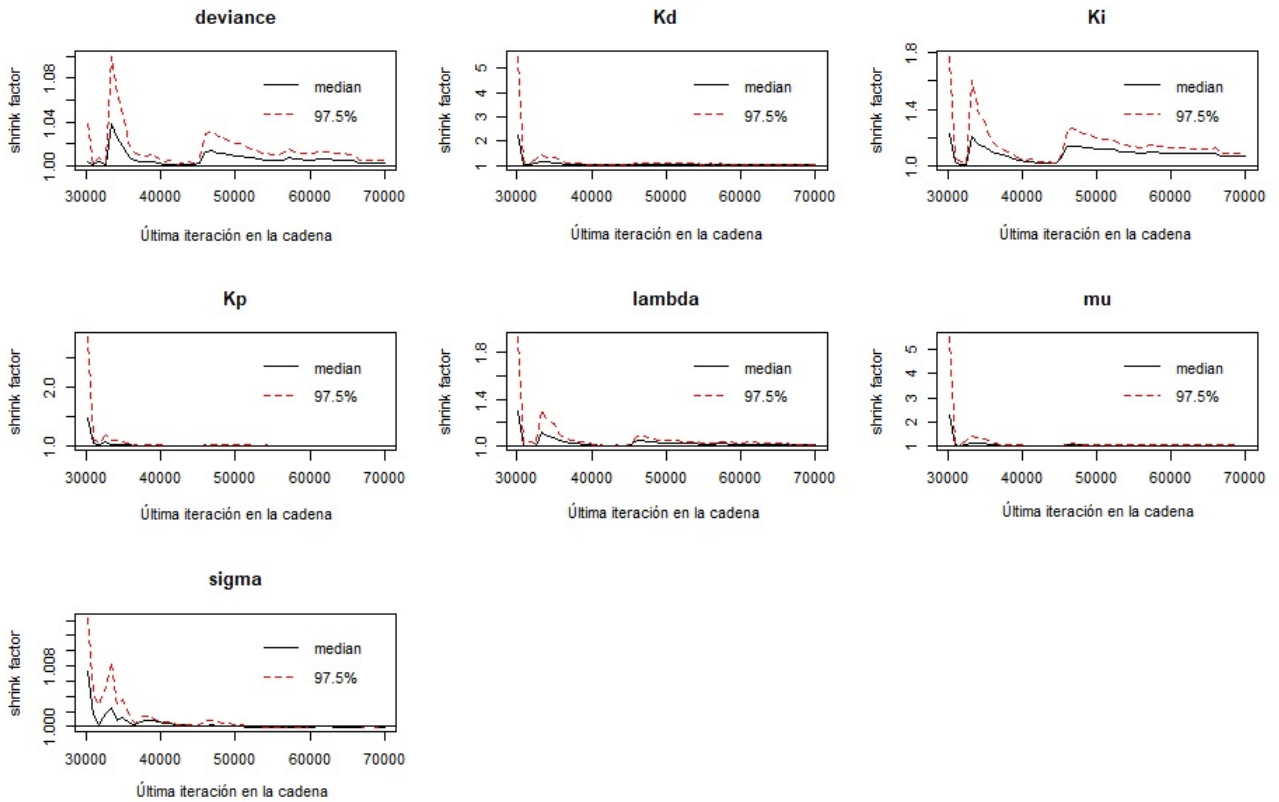


Figura 4.16: Diagnóstico de convergencia de Gelman-Rubin.

El diagnóstico de convergencia de Geweke dio como resultado $|Z| < 2$ para cada una de las cadenas con cada uno de los parámetros. Esto indica que las medias de las primeras y últimas iteraciones son iguales y se alcanza la convergencia de las cadenas.

4.2.4. Evaluación

Para la evaluación del método de estimación en este ejemplo, usamos el Algoritmo 1 definido en la Sección 4.1.4. En este caso, los valores de referencia de los parámetros son los obtenidos de Monje et al. (2010). Los resultados se muestran en la Tabla 4.8. Vemos que los valores de sesgo y ECM son pequeños. Como el sesgo es la diferencia del valor de referencia con el valor estimado, que esta diferencia sea pequeña es favorable para decir que se obtuvo una buena estimación. Por otro lado, en el ECM, además del sesgo, interviene la varianza del estimador, por tanto, que sea pequeño también sugiere que las estimaciones obtenidas son buenas.

| Parámetros | Sesgo | ECM |
|------------|---------|---------|
| K_p | -0.0406 | 0.0090 |
| K_d | -0.0445 | 0.0122 |
| K_i | 0.0235 | 0.0012 |
| μ | 0.0017 | 1.6e-05 |
| λ | -0.1312 | 0.0303 |
| σ | -0.0019 | 0.0182 |

Tabla 4.8: Resultados de sesgo y ECM.

Conclusiones

En este trabajo de tesis se realizó de manera exitosa la estimación de los parámetros que caracterizan a un controlador PID de orden fraccionario, desde el punto de vista de la Estadística Bayesiana. Adicionalmente, podemos concluir de manera particular en los siguientes puntos:

- Presentamos dos ejemplos donde se aplica el controlador PID fraccionario. El caso del tanque cónico es un caso muy común, donde estimar los parámetros del controlador es esencial. Para cada caso se analizaron las distribuciones *a priori* adecuadas y sus respectivos hiperparámetros. Además, se realizó la estimación de los parámetros, se verificó la convergencia de las cadenas MCMC y se evaluó el método de estimación, arrojando buenos resultados.
- Las distribuciones finales de los parámetros fueron obtenidas usando algoritmos Monte Carlo vía Cadenas de Markov. Estos algoritmos están implementados en JAGS dentro del paquete estadístico R, lo que facilitó enormemente el trabajo de cómputo, optimizando el tiempo de ejecución del algoritmo.
- Dado el buen ajuste a los datos con las estimaciones obtenidas, podemos afirmar la eficacia de la estimación Bayesiana de los parámetros del controlador PID fraccionario. Además, tanto el sesgo como el error cuadrático medio mostraron estar en una vecindad cercana al cero, lo cual muestra una buena evaluación de las estimaciones obtenidas a partir de la metodología empleada.

- Ante problemas concretos en los que se aplique el controlador, solo será necesario ajustar las distribuciones *a priori*, de acuerdo al conocimiento que ya se tiene de los parámetros.
- Como trabajo futuro se recomienda realizar comparaciones entre el método de estimación Bayesiana utilizado en este trabajo y otros métodos como algoritmos evolutivos, lógica difusa, ect. Los resultados obtenidos por cada método se deben evaluar y comparar con los valores reales conocidos de los parámetros, para determinar qué método obtiene los mejores resultados.

Bibliografía

- Agarwal, R. (1953). A propos d'une note de m. pierre humbert. *CR Acad. Sci. Paris*, 236(21):2031–2032.
- Bates, J. H. (2009). *Lung mechanics: an inverse modeling approach*. Cambridge University Press.
- Bode, H. W. (1940). Relations between attenuation and phase in feedback amplifier design. *Bell System Technical Journal*, 19(3):421–454.
- Bode, H. W. et al. (1945). Network analysis and feedback amplifier design.
- Chen, M.-H., Shao, Q.-M., and Ibrahim, J. G. (2012). *Monte Carlo methods in Bayesian computation*. Springer Science & Business Media.
- Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335.
- Freeborn, T. J. (2013). A survey of fractional-order circuit models for biology and biomedicine. *IEEE Journal on emerging and selected topics in circuits and systems*, 3(3):416–424.
- Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409.
- Gelman, A. et al. (2014). Bayesian data analysis (crc, boca raton, fl).
- Ionescu, C. M. and De Keyser, R. (2008). Time domain validation of a fractional order model for human respiratory system. In *MELECON 2008-The 14th IEEE Mediterranean Electrotechnical Conference*, pages 89–95. IEEE.

- Jauregui, C., Mermoud, M. D., Lefranc, G., Orostica, R., Torres, J. C. T., and Beytia, O. (2016). Conical tank level control with fractional pid. *IEEE Latin America Transactions*, 14(6):2598–2604.
- Kennedy, J. (1995). Eberhart, r.: Particle swarm optimization. In *Proceedings of IEEE international conference on neural networks*, volume 4, pages 1942–1948. IEEE Press.
- Kimeldorf, G. S. and Wahba, G. (1970). A correspondence between bayesian estimation on stochastic processes and smoothing by splines. *The Annals of Mathematical Statistics*, 41(2):495–502.
- Lainez, R. A. (2015). Modelación e identificación de sistemas fraccionarios. Master’s thesis, Centro de Investigación en Matemáticas, A.C.
- Monje, C. A., Chen, Y., Vinagre, B. M., Xue, D., and Feliu-Batlle, V. (2010). *Fractional-order systems and controls: fundamentals and applications*. Springer Science & Business Media.
- Moreles, M. A., Peña, J., Botello, S., and Iturriaga, R. (2013). On modeling flow in fractal media from fractional continuum mechanics and fractal geometry. *Transport in porous media*, 99(1):161–174.
- Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS*, volume 698. John Wiley & Sons.
- Petráš, I. (2011). *Fractional-order nonlinear systems: modeling, analysis and simulation*. Springer Science & Business Media.
- Plummer, M. (2013). rjags: Bayesian graphical models using mcmc. *R package version*, 3(10).
- Podlubny, I. (1999). Fractional-order systems and pi/spl lambda//d/spl mu//-controllers. *IEEE Transactions on automatic control*, 44(1):208–214.
- Surguladze, T. (2002). On certain applications of fractional calculus to viscoelasticity. *Journal of Mathematical sciences*, 112(5):4517–4557.
- Team, R. C. (2014). A language and environment for statistical computing. r foundation for statistical computing, vienna, austria2014. *URL:(<https://www.R-project.org>)*.

- Westerlund, S. and Ekstam, L. (1994). Capacitor theory. *IEEE Transactions on Dielectrics and Electrical Insulation*, 1(5):826–839.
- Xiang, J., Cheng, D., Schlindwein, F., and Jones, N. (2003). On the adequacy of identified cole–cole models. *Computers & geosciences*, 29(5):647–654.