UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICA Y CIENCIAS DE LA COMPUTACIÓN
MAESTRÍA EN MATEMÁTICA

Modelos de espacio de estados con filtración de partículas aplicado a un modelo DSGE

Tesis presentada para optar al Título de máster en matemática con orientación en Estadística Matemática

Tesis presentada por: Pedro José Molina Morales

Asesor tesis: Cristian Andrés Cruz Torres
Fecha de defensa: 24 de julio del 2026

Tegucigalpa, Honduras 2026.

Modelos de espacio de estados con filtración de partículas aplicado a un modelo DSGE

Presentada por: Pedro José Molina Morales

Tesis presentada para optar al Título de máster en matemática con orientación en Estadística Matemática

Asesor: Cristian Andrés Cruz Torres
Universidad Nacional Autónoma de Honduras

TERNA EXAMINADORA

Ph.D Javier Enrique Aguilar Romero
TU Dortmund University

M.Sc Darwin Omar Gutierrez
Universidad Nacional Autónoma de Honduras

M.Sc Marvin Levi Villafranca Rivera
Universidad Nacional Autónoma de Honduras

Tegucigalpa M.D.C 25 de agosto del 2026.

Agradecimientos

La culminación de esta tesis ha sido posible gracias al apoyo y acompañamiento de muchas personas a lo largo de este camino académico, a quienes deseo expresar mi más sincero agradecimiento.

En primer lugar, agradezco a Dios por darme la fortaleza, la salud y la perseverancia necesaria para completar este trabajo.

A mi familia, en especial a mis padres Pedro Joaquín Molina y Xiomara Angelina Morales, por su amor incondicional, su apoyo constante y por enseñarme, con su ejemplo, el valor del esfuerzo y la dedicación.

También a mi esposa Eliany Fabiola Enamorado Tinoco por su amor incondicional, quien me ha inspirado y motivado en la culminación de la misma.

A mi asesor Cristian Andrés Cruz Torres, por su guía constante y sus valiosas observaciones. Su experiencia y disposición han sido fundamentales para el desarrollo de esta investigación.

A mis profesores del programa de Maestría en Matemáticas, quienes a lo largo de los cursos impartieron conocimientos que fueron esenciales para este proyecto.

Pedro José Molina Morales

Índice general

Agradecimientos II

Índice de figuras VI

Índice de tablas IX

Resumen X

Abstract XI

1. Introducción 1

1.1. Organización de la Tesis 3

2. Conceptos Preliminares 4

2.1. Estimación Bayesiana 4

2.1.1. Regla de Bayes 4

2.1.2. Predicción Bayesiana 5

2.2. Series de Tiempo 6

2.2.1. Funciones de Autocorrelación 9

2.3. Métodos de MonteCarlo 10

2.3.1. Extensiones 12

2.4. Estimación Robusta 14

2.4.1. Medidas de Robustez 14

2.5. Validación Cruzada Bayesiana: ELPD 16

3. Modelo en Espacio de Estado 18

3.1. Modelos Lineales Dinámicos 19

3.1.1. Filtración 21

3.1.2. Suavizamiento 23

3.1.3. Pronóstico 23

3.2. Modelos de Espacios de Estados con Filtro de Partículas 24

3.2.1. Limitaciones del filtro de Kalman 27
3.2.2. Filtro de Partículas Regularizado 29
3.3. Predicción a Partir de Filtros de Partículas 33
3.3.1. Predicción Mediante la Media 33
3.3.2. Predicción Robusta Mediante la Mediana 34
3.4. Validación del método de Filtro de Partículas 35
3.4.1. Consideraciones Finales 37

4. Modelos DSGE **39**

4.1. Estructura General de un Modelo DSGE 40
4.1.1. Linealización y Métodos Clásicos de Solución 42
4.2. El Problema del Consumidor 44
4.2.1. Variables de Estado y de Control 45
4.3. Condiciones de Equilibrio 47
4.4. Estado Estacionario 48
4.5. Calibración y estimación de Parámetros 50
4.6. Aplicación 51
4.6.1. Modelo DSGE con Familia Representativa, Firmas y Banco Central 52
4.6.1.1. Problema del Hogar Representativo 53
4.6.2. Empresas Productoras y Fijación de Precios 54
4.6.2.1. Política Monetaria 60
4.6.3. Ecuaciones de Equilibrio 61
4.6.4. Representación en Espacio de Estado y Resolución Numérica 63
4.6.4.1. Representación en Espacio de Estado 64
4.6.4.2. Vector de Estados 64
4.6.4.3. Ecuaciones de Transición 65
4.6.4.4. Ecuación de Observación 66
4.6.4.5. Implementación del Filtro de Partículas 67
4.6.5. Seudocódigo del Algoritmo de Filtro de Partículas 67

5. Simulaciones y análisis de datos **69**

5.1. Análisis Exploratorio de las Series de Tiempo 70

5.2. Parámetros Iniciales 72

5.3. Resultados con Estimación por la Media 73

5.3.1. Análisis de Errores 77

5.4. Resultados con Estimación por la Mediana 79

5.4.1. Análisis de Errores 85

5.5. Consideraciones computacionales 87

6. Conclusiones y trabajos futuros **90**

6.1. Trabajos futuros 91

Bibliografía **93**

A. La función de Lagrange en el Problema del Consumidor **100**

B. Cuadratura de Gauss-Hermite y Cálculo de Esperanzas **103**

C. Interpolación por Elementos Finitos en Modelos DSGE **107**

Índice de figuras

2.1. Proceso autorregresivo de primer orden $AR(1)$ definido mediante $x_t = 0.8x_{t-1} + \epsilon_t$, donde $\epsilon_t \sim N(0,1)$ 10

2.2. Flujo básico de un método de Monte Carlo. 11

3.1. Independencia condicional en un modelo de espacio de estados. 19

3.2. Relación temporal entre filtración, suavizamiento y pronóstico en modelos de espacio de estados. 21

3.3. Ecuación original vs aproximación lineal 28

3.4. Simulación de Filtro de Kalman vs Filtro de Partículas 29

3.5. Comparación de variables observadas vs. estimadas con $N = 50$. 36

3.6. Comparación de variables observadas vs. estimadas con $N = 10000$. 36

4.1. Relación entre los conceptos teóricos y la aplicación en modelos DSGE. . . 42

4.2. Estructura general de un modelo DSGE. 48

5.1. Series observadas de crecimiento del PIB real, inflación y tasa de interés nominal (2000-2024). 71

5.2. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 100$ partículas, en el periodo del 2000 al 2024 de forma trimestral. 73

5.3. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 200$ partículas, en el periodo del 2000 al 2024 de forma trimestral. 74

5.4. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 500$ partículas, en el periodo del 2000 al 2024 de forma trimestral. 74

5.5. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 1000$ partículas, en el periodo del 2000 al 2024 de forma trimestral. 75

5.6. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 2000$ partículas, en el periodo del 2000 al 2024 de forma trimestral. 75

5.7. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 6000$ partículas, en el periodo del 2000 al 2024 de forma trimestral. 76

5.8. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 8000$ partículas, en el periodo del 2000 al 2024 de forma trimestral. 76

5.9. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 16000$ partículas, en el periodo del 2000 al 2024 de forma trimestral. 77

5.10. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 100$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral. 80

5.11. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 200$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral. 80

5.12. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 500$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral. 81

5.13. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 1000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral. 81

5.14. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 2000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral. 82

5.15. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 6000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral. 82

5.16. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 8000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral. 83

5.17. Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 16000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral. 83

5.18. Distribución de $N = 16,000$ partículas de las variables g_t, π_t, r_t en los periodos de predicción Q1 2024, Q2 2024, Q3 2024 y Q4 2024. 85

C.1. Ejemplo de interpolación bilineal sobre una malla rectangular en (a, z). . . 110

C.2. Funciones base para la interpolación bilineal en las dimensiones a y z. . . . 110

Índice de tablas

3.1. Raiz de error cuadrático medio (RMSE) para las estimaciones y pronósticos de los filtros de Kalman y filtros de partículas. 29
3.2. RMSE de las variables estimadas 36
3.3. Expected Log Predictive Density (ELPD) para estimación y predicción fuera de muestra según el número de partículas utilizadas. 37

5.1. Parámetros calibrados utilizados en la estimación del modelo DSGE. 72
5.2. Errores absolutos medios (MAE) fuera de muestra según el número de partículas utilizadas. 77
5.3. Raíz del error cuadrático medio (RMSE) fuera de muestra según el número de partículas utilizadas. 78
5.4. Expected Log Predictive Density (ELPD) para estimación y predicción fuera de muestra según el número de partículas utilizadas. 78
5.5. Resumen numérico de media, mediana, percentiles 5-95 y real en $Q1 - 2024$ 84
5.6. Resumen numérico de media, mediana, percentiles 5-95 y real en $Q2 - 2024$ 84
5.7. Resumen numérico de media, mediana, percentiles 5-95 y real en $Q3 - 2024$ 84
5.8. Resumen numérico de media, mediana, percentiles 5-95 y real en $Q4 - 2024$ 84
5.9. MAE de las variables estimadas con mediana para distintos tamaños de partículas 85
5.10. RMSE de las variables estimadas con mediana para distintos tamaños de partículas 86
5.11. Tiempo de ejecución en horas para cada tamaño de N partículas 88

Modelos de espacio de estados con filtración de partículas aplicado a un modelo DSGE

Resumen

En esta tesis se desarrolla un modelo Dinámico Estocástico de Equilibrio General (DSGE, por sus siglas en inglés) el cual surge como una evolución natural de los modelos de crecimiento neoclásico. Estos modelos combinan el rigor macroeconómico con estructuras dinámicas y estocásticas, permitiendo evaluar cuantitativamente las políticas económicas. El modelo propuesto por las ecuaciones de equilibrio, se representó en su forma de espacio de estados multivariado, lo que facilita la aplicación de técnicas secuenciales de Monte Carlo (SMC) de un marco Bayesiano, calibrando los parámetros estructurales del modelo en base a la literatura y el filtro de partículas regularizado para la estimación de los estados latentes y generar pronósticos fuera de la muestra. Durante el proceso de filtrado, se emplearon técnicas numéricas como la cuadratura de Gauss-Hermite y la interpolación mediante elementos finitos para aproximar esperanzas racionales que no tienen solución analítica. Asimismo, se incorporaron técnicas de estimación robusta como la mediana para resumir la distribución de partículas. La base de datos proviene del sistema FRED (Federal Reserve Economic Data, EE. UU), enfocándose en la estimación y la predicción de tres variables macroeconómicas fundamentales: g_t (PIB), π_t (Inflación) y r_t (Tasa de interés) de manera trimestral, realizando estimaciones de los años $2000 - 2023$ y predicciones para el año 2024. Los resultados obtenidos demostraron que al utilizar técnicas Secuenciales de Monte Carlo y estimación robusta en modelos DSGE, se tienen estimaciones y predicciones aceptables para la inferencia.

Palabras clave: Modelos DSGE, modelos de espacio de estados, filtros de partículas, métodos numéricos, estimación bayesiana.

State-space models with particle filtering applied to a DSGE model

Abstract

This research develops a Dynamic Stochastic General Equilibrium (DSGE) model, which emerges as a natural evolution of neoclassical growth models. These models combine macroeconomic rigor with dynamic and stochastic structures, enabling a quantitative evaluation of economic policies. The proposed model, consisting of the equilibrium equations, is represented in its multivariate state-space form, which facilitates the application of Sequential Monte Carlo (SMC) techniques within a Bayesian framework, calibrating the model's structural parameters based on the literature, and utilizing the regularized particle filter for the estimation of latent states and the generation of out-of-sample forecasts. During the filtering process, numerical techniques such as Gauss-Hermite quadrature and finite element interpolation were employed to approximate rational expectations that lack an analytical solution. Furthermore, robust estimation techniques, such as the median, were incorporated to summarize the particle distribution. The dataset is sourced from the FRED (Federal Reserve Economic Data, USA) system, focusing on the estimation and forecasting of three fundamental macroeconomic variables: g_t (GDP), π_t (Inflation), and r_t (Interest rate) on a quarterly basis, performing estimations for the years $2000 - 2023$ and forecasts for the year 2024. The results obtained demonstrated that utilizing Sequential Monte Carlo techniques and robust estimation in DSGE models yields acceptable estimations and forecasts for inference.

Keywords: DSGE models, state-space models, particle filters, numerical methods, bayesian estimation.

Capítulo 1

Introducción

Los modelos dinámicos estocásticos de equilibrio general (DSGE) se han consolidado como una de las principales herramientas en el análisis macroeconómico moderno, al permitir representar de forma estructural la interacción entre agentes económicos bajo incertidumbre y esperanzas racionales. Estos modelos son ampliamente utilizados en el análisis de política económica, particularmente en bancos centrales, debido a su capacidad para incorporar fundamentos microeconómicos en la dinámica agregada. [1].

No obstante, una de las principales dificultades en su implementación empírica radica en su naturaleza altamente no lineal. En la práctica, esta dificultad suele abordarse mediante técnicas de log-linealización alrededor del estado estacionario, lo que permite reformular el modelo en un sistema lineal en espacio de estados y aplicar el filtro de Kalman bajo supuestos de normalidad. Sin embargo, estas aproximaciones pueden limitar la capacidad del modelo para capturar dinámicas no lineales y distribuciones no gaussianas presentes en los datos.

En este contexto, surge la necesidad de explorar metodologías que permitan trabajar directamente con la estructura no lineal del modelo. Los métodos de Monte Carlo secuenciales (SMC), particularmente los filtros de partículas, constituyen una alternativa adecuada para realizar inferencia en modelos de espacio de estados no lineales y no gaussianos, al aproximar la distribución posterior de los estados mediante un conjunto de partículas ponderadas [2].

La implementación de filtros de partículas en modelos DSGE presenta desafíos relevantes, tales como la aproximación numérica de esperanzas racionales, la degeneración de partículas y la evaluación de su desempeño predictivo.

El objetivo de esta tesis es integrar y analizar el desempeño de un filtro de partículas regularizado aplicado a un modelo DSGE no lineal basado en [3] para resolver el problema de la degeneración de partículas, formulado en representación de espacio de estados

multivariado con parámetros estructurales calibrados en base a literatura. A diferencia de los enfoques tradicionales, en este trabajo no se recurre a la log-linealización del modelo, sino que se preserva su estructura no lineal en la etapa de inferencia de los estados.

También, uno de los desafíos es la resolución de las esperanzas racionales, la cual se encuentran en las ecuaciones de equilibrio resultantes del modelo DSGE. Estas expresiones, al no tener solución analítica, requieren de métodos numéricos de aproximación. En esta tesis dichas esperanzas se resolvieron mediante dos técnicas complementarias, la cuadratura de Gauss-Hermite y la interpolación mediante elementos finitos. La cuadratura de Gauss-Hermite permite aproximar integrales de funciones no lineales, utilizando una combinación ponderada de nodos [4], y la interpolación mediante elementos finitos, se empleó para construir funciones aproximadas continuas en espacios de estados discretizados para aproximación de valores futuros [5]. Ambos métodos en conjunto permiten una resolución de las esperanzas racionales, preservando la estructura no lineal del modelo.

La evaluación empírica se lleva a cabo utilizando datos macroeconómicos de Estados Unidos obtenidos del sistema FRED, considerando tres variables observadas: crecimiento económico (g_t), inflación (π_t) y tasa de interés nominal (r_t). El análisis incluye tanto la estimación dentro de muestra como la evaluación predictiva fuera de muestra, utilizando métricas de error para el desempeño del modelo, como el error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE) para las estimaciones y predicciones, asimismo se utilizó el Expected Log Predictive Density (ELPD) para medir la capacidad predictiva del modelo de las distribuciones a posterioris.

Finalmente, este trabajo aporta evidencia empírica sobre el desempeño de los filtros de partículas en un modelo DSGE no lineal, destacando tanto sus ventajas como sus limitaciones, particularmente en la capacidad de predicción de ciertas variables. Asimismo, se analiza el papel de estimadores robustos, como la mediana, en la representación de distribuciones posteriores en entornos no gaussianos.

1.1. Organización de la Tesis

Esta tesis contiene seis capítulos incluyendo la introducción, un capítulo de conceptos preliminares, cuatro capítulos teóricos para la implementación del modelos DSGE, incluyendo el de las simulaciones y análisis de datos y finalmente las conclusiones y trabajos futuros.

El Capítulo 2 presenta los conceptos fundamentales y generales, necesarios que servirán como base inicial para el comprendimiento de la presente tesis.

En el Capítulo 3 se presentan los conceptos teóricos fundamentales del modelo en espacio de estado, donde una de sus características fundamentales es la capacidad de representar sistemas donde las variables de observación tienen dependencia de variables latentes que evolucionan en el tiempo bajo variaciones. Primeramente, se habla de los Modelos Lineales Dinámicos (DLM, por sus siglas en inglés), posteriormente de modelos de espacio de estados generales, los cuales permiten incorporar relaciones no lineales.

En el Capítulo 4 se presenta la formulación teórica de los modelos DSGE, partiendo desde el problema del consumidor y construyendo las ecuaciones de equilibrio que conforman el modelo estructural general para su posterior resolución, luego se realiza la formulación del modelo DSGE en un modelo de espacio de estados, paso esencial para aplicar técnicas de filtrado secuencial. Posteriormente se describe el funcionamiento del filtro de partículas y su implementación numérica.

El Capítulo 5 expone los resultados obtenidos a partir de la estimación y predicción del modelo, comparando distintos tamaños de partículas y evaluando el desempeño de la media y la mediana como estimadores.

Finalmente, en el Capítulo 6, se resumen los hallazgos principales del trabajo y se plantean posibles direcciones para investigaciones futuras.

Capítulo 2

Conceptos Preliminares

Este capítulo presenta los fundamentos teóricos necesarios para la formulación y estimación del modelo DSGE no lineal desarrollado en esta tesis. Se introducen los conceptos de inferencia bayesiana, series de tiempo, métodos de Monte Carlo secuenciales (SMC), estimación robusta y validación cruzada bayesiana.

2.1. Estimación Bayesiana

La estadística bayesiana se caracteriza por realizar inferencia en presencia de incertidumbre, donde tanto los parámetros de un modelo como las predicciones futuras se tratan como variables aleatorias. Su base es la regla de Bayes, que permite combinar informaciòn previa (priori) con evidencia empírica (datos) para obtener una distribución posterior que resume todo el conocimiento disponible [6].

2.1.1. Regla de Bayes

Sea $\boldsymbol{\theta}$ un vector de parámetros desconocidos y $\mathbf{y}_{1:T} = \{y_1, y_2, \ldots, y_T\}$ un conjunto de observaciones. Uno de los objetivos de la inferencia bayesiana consiste en caracterizar la distribución posterior $p(\boldsymbol{\theta} \mid \mathbf{y}_{1:T})$, la cual describe la incertidumbre sobre los parámetros una vez observados los datos.

La distribución a priori de los parámetros se denota por $p(\boldsymbol{\theta})$, y representa la información disponible sobre $\boldsymbol{\theta}$ antes de observar los datos. En general, esta distribución se especifica a partir de resultados previos de la literatura, restricciones teóricas del modelo o consideraciones específicas.

La verosimilitud se define como la función de probabilidad de los datos condicionada en los parámetros, es decir,

$$p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta}).$$

Desde un punto de vista probabilístico, la verosimilitud mide qué tan probable es observar la muestra $\mathbf{y}_{1:T}$ bajo un valor específico de $\boldsymbol{\theta}$. Es importante notar que, aunque tiene la forma de una función de probabilidad, en el contexto inferencial se interpreta como una función de los parámetros, manteniendo los datos fijos.

Bajo el supuesto de independencia condicional de las observaciones dado $\boldsymbol{\theta}$, la verosimilitud puede factorizarse como

$$p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta}) = \prod_{t=1}^{T} p(y_t \mid \boldsymbol{\theta}).$$

En contextos donde las observaciones presentan dependencia temporal, la verosimilitud puede expresarse de manera más general mediante la regla del producto condicional [7]:

$$p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta}) = \prod_{t=1}^{T} p(y_t \mid \mathbf{y}_{1:t-1}, \boldsymbol{\theta}).$$

Aplicando el teorema de Bayes, la distribución posterior de los parámetros viene dada por

$$p(\boldsymbol{\theta} \mid \mathbf{y}_{1:T}) = \frac{p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\mathbf{y}_{1:T})},$$

donde

$$p(\mathbf{y}_{1:T}) = \int p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})\, d\boldsymbol{\theta}$$

es la distribución marginal de los datos o evidencia bayesiana.

Dado que $p(\mathbf{y}_{1:T})$ no depende de $\boldsymbol{\theta}$, en la práctica se trabaja con la forma proporcional

$$p(\boldsymbol{\theta} \mid \mathbf{y}_{1:T}) \propto p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta}). \tag{2.1}$$

Por tanto, la inferencia bayesiana consiste en actualizar la información previa sobre $\boldsymbol{\theta}$ mediante la información aportada por los datos, donde la verosimilitud desempeña el papel central de enlazar el modelo probabilístico con la evidencia empírica.

2.1.2. Predicción Bayesiana

La estadística bayesiana no solo permite estimar parámetros, también genera predicciones futuras integrando la incertidumbre. Si y_{T+1} es una observación nueva, su distribución predictiva es:

$$p(y_{T+1} \mid \mathbf{y}_{1:T}) = \int p(y_{T+1} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta} \mid \mathbf{y}_{1:T})\, d\boldsymbol{\theta}. \tag{2.2}$$

Según [8] en muestras pequeñas, la elección de la priori influye fuertemente en la distribución a posterior; mientras que en muestras grandes converge a la función de máxima verosimilitud. Este balance es una característica clave, ya que la posterior refleja tanto el conocimiento previo como la evidencia empírica.

En modelos complejos, la distribución posterior no puede calcularse de forma analítica. Se utilizan métodos numéricos ampliamente conocidos de Markov Chain Monte Carlo (MCMC) que se detallan en la sección de Monte Carlo de este capítulo.

2.2. Series de Tiempo

El análisis de series de tiempo constituye un componente esencial para comprender fenómenos dinámicos en economía, finanzas, ingeniería y ciencias sociales. En el contexto de esta tesis, las series de tiempo son la base para la estimación de modelos DSGE no lineales en su forma de espacio de estado (se introduce en detalle a partir del capítulo 3) mediante métodos bayesianos y Filtros de Partículas.

Una serie de tiempo o serie cronológica es una realización muestral observable de un proceso estocástico. Se define como una colección indexada de variables aleatorias $\{Y_t\}_{t=1}^{T}$, donde cada Y_t es una función medible definida sobre un espacio de probabilidad común $(\Omega, \mathcal{F}, P)$ y mapeada al espacio de los números reales $\mathbb{R}^n$. Así, para un estado de la naturaleza específico $\omega \in \Omega$, la secuencia de valores observados $\{Y_t(\omega)\}_{t=1}^{T}$ constituye la trayectoria temporal o serie cronológica de la variable macroeconómica como PIB, inflación o tasas de interés. Una de las características principales de estas series es la dependencia en el tiempo, lo que significa que las observaciones no son independientes entre sí, sino que dependen una de otra a lo largo del tiempo [9].

Uno de los conceptos fundamentales es la estacionariedad estricta y la débil, que según [10] se definen de la siguiente manera:

Definición 2.1. *Se dice que un proceso estocástico es estrictamente estacionario si sus propiedades no se ven afectadas por un cambio en el origen del tiempo, es decir, si*

la distribución de probabilidad conjunta asociada con m observaciones $y_{t_1}, y_{t_2}, \ldots, y_{t_m}$*, realizadas en cualquier conjunto de tiempos* $t_1, t_2, \ldots, t_m$*, es la misma que la asociada con las m observaciones* $y_{t_1+k}, y_{t_2+k}, \ldots, y_{t_m+k}$*, realizadas en los tiempos* $t_1+k, t_2+k, \ldots, t_m+k$*. Por lo tanto, para que un proceso discreto sea estrictamente estacionario, la distribución conjunta de cualquier conjunto de observaciones debe permanecer inalterada al desplazar todos los tiempos de observación hacia adelante o hacia atrás en cualquier cantidad entera* k*.*

Definición 2.2. *Media y Varianza de un Proceso Estacionario. Cuando* $m = 1$*, el supuesto de estacionariedad implica que la distribución de probabilidad* $p(y_t)$ *es la misma para todos los tiempos t y puede escribirse simplemente como* $p(y)$*. Por lo tanto, el proceso estocástico tiene una media constante:*

$$\mu = \mathbb{E}[y_t] = \int_{-\infty}^{\infty} yp(y)dy \tag{2.3}$$

la cual define el nivel alrededor del cual fluctúa, y una varianza constante:

$$\sigma_y^2 = \mathbb{E}[(y_t - \mu)^2] = \int_{-\infty}^{\infty} (y-\mu)^2 p(y)dy \tag{2.4}$$

la cual mide su dispersión alrededor de dicho nivel. Dado que la distribución de probabilidad $p(y)$ *es la misma para todos los tiempos t, su forma puede inferirse construyendo el histograma de las observaciones* $y_1, y_2, \ldots, y_N$*, que conforman la serie de tiempo observada. Además, la media* μ *del proceso estocástico puede estimarse mediante la media muestral:*

$$\bar{y} = \frac{1}{N}\sum_{t=1}^{N} y_t \tag{2.5}$$

de la serie de tiempo, y la varianza σ_z^2 *del proceso estocástico puede estimarse mediante la varianza muestral:*

$$\hat{\sigma}_y^2 = \frac{1}{N}\sum_{t=1}^{N} (y_t - \bar{y})^2 \tag{2.6}$$

de la serie de tiempo.

Definición 2.3. *Coeficientes de Autocovarianza. El supuesto de estacionariedad también implica que la distribución de probabilidad conjunta* $p(y_{t_1}, y_{t_2})$ *es la misma para todos los*

tiempos t_1, t_2, que se encuentren a un intervalo constante de distancia. En particular, se deduce que la covarianza entre los valores y_t y y_{t+k}, separados por k intervalos de tiempo, o por un rezago (lag) k, debe ser la misma para todo t bajo el supuesto de estacionariedad. Esta covarianza se denomina la autocovarianza en el rezago k y se define por:

$$\gamma_k = \text{cov}[y_t, y_{t+k}] = \mathbb{E}[(y_t - \mu)(y_{t+k} - \mu)] \tag{2.7}$$

Hemos visto que para que un proceso sea estrictamente estacionario, toda la estructura de probabilidad debe depender únicamente de las diferencias de tiempo. Un requisito menos restrictivo, llamado estacionariedad débil de orden f se define de la siguiente forma.

Definición 2.4. *La estacionariedad débil de orden f, es que los momentos hasta cierto orden f dependan únicamente de las diferencias de tiempo.*

Por ejemplo, la existencia de una media fija μ y una matriz de autocovarianza $\mathbf{\Gamma}_n$ es suficiente para garantizar la estacionariedad hasta el segundo orden. Es decir, un proceso $\{y_t\}$ es débilmente estacionario (de orden 2), o estacionario de segundo orden, si la media $\mathbb{E}[y_t] = \mu$ es una constante fija para todo t y las autocovarianzas $\text{cov}[y_t, y_{t+k}] = \gamma_k$ dependen únicamente de la diferencia de tiempo o rezago temporal (lag) k para todo t. Por lo tanto, la estacionariedad de segundo orden y el supuesto de normalidad son suficientes para producir una estacionariedad estricta.

En modelos macroeconómicos, pueden existir variables que no son estacionarias, por lo que se puede aplicar transformaciones como tasas de crecimiento o logaritmos para cumplir con dicha propiedad [9, 11].

Las series de tiempo mas comunes son los procesos autorregresivos (AR), de media móvil (MA) y su combinación (ARMA). Se describen las ecuaciones a continuación:

1. **AR(p)**: $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$,
2. **MA(q)**: $Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$,
3. **ARMA(p,q)**: $Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$.

donde ε_t es un *ruido blanco* (comúnmente con distribución normal estándar) con media cero y varianza σ^2.

2.2.1. Funciones de Autocorrelación

La función de autocorrelación (ACF) mide la relación entre las observaciones pasadas y presentes:

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad k = 1, 2, \ldots$$

El comportamiento de las autocorrelaciones es una herramienta útil para identificar la estructura de la serie de tiempo.

Muchas series económicas presentan tendencias o raíces unitarias ($\phi = 1$ en un AR(1)). Para tratar esto:

1. Se aplican **diferencias**: $\Delta y_t = y_t - y_{t-1}$,
2. Se transforman en **tasas de crecimiento logarítmico**: $g_t = \log(y_t) - \log(y_{t-1})$.
3. Se transforman en **tasas de crecimiento clásico**: $g_t = y_t / y_{t-1}$.

En la práctica, la autocorrelación se estima a partir de los datos mediante el estimador muestral:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2},$$

donde $\bar{y}$ representa la media muestral.

El gráfico 2.1 de la función de autocorrelación (ACF) exhibe el comportamiento teórico típico de un proceso autorregresivo de primer orden $AR(1)$ definido mediante $x_t = 0.8x_{t-1} + \epsilon_t$, donde $\epsilon_t \sim N(0, 1)$.

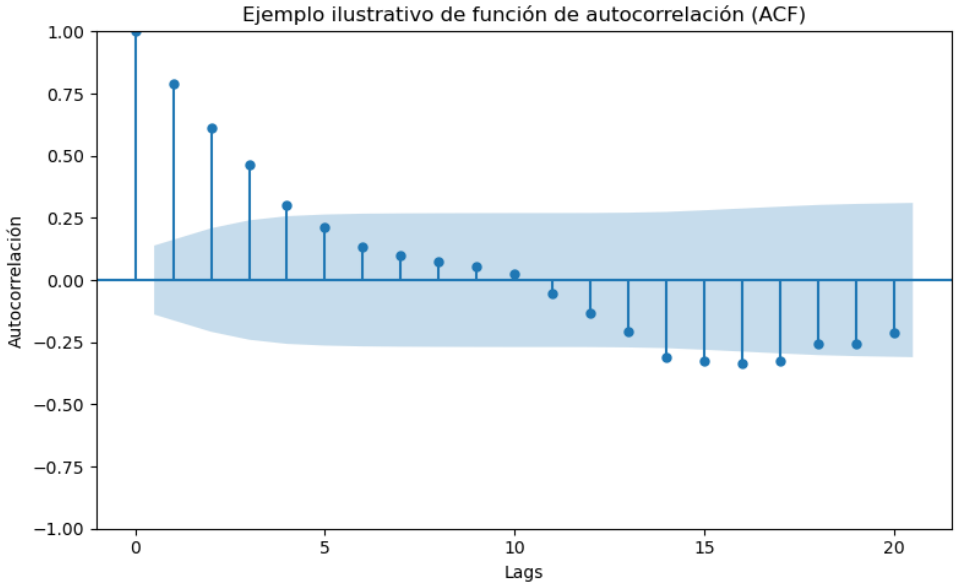


Figura 2.1: Proceso autorregresivo de primer orden $AR(1)$ definido mediante $x_t = 0.8x_{t-1} + \epsilon_t$, donde $\epsilon_t \sim N(0, 1)$

La función de autocorrelación proporciona información sobre la persistencia de las variables en el tiempo. En particular, procesos con alta autocorrelación presentan una fuerte dependencia temporal, lo cual es característico de muchos modelos macroeconómicos.

2.3. Métodos de MonteCarlo

Los métodos de Monte Carlo constituyen una familia de técnicas numéricas diseñadas para aproximar integrales mediante muestreo aleatorio. La idea general es reemplazar cálculos analíticos complejos por promedios (u otros estimadores) de simulaciones, lo que resulta especialmente útil cuando el espacio de estados es de alta dimensión o la integral carece de solución analítica [12].

Sea una integral de la forma:

$$I = \int h(x)p(x)\,dx, \tag{2.8}$$

donde $p(x)$ es una densidad de probabilidad y $h(x)$ una función de interés. Si se tiene N

muestras independientes $x^{(1)}, \ldots, x^{(N)} \sim p(x)$, el estimador de Monte Carlo es:

$$\hat{I}_N = \frac{1}{N} \sum_{i=1}^{N} h(x^{(i)}). \tag{2.9}$$

La veracidad de este estimador se sustenta en la Ley débil de los Grandes Números [13], que garantiza que:

$$\hat{I}_N \xrightarrow[N\to\infty]{p} I,$$

y en el Teorema Central del Límite [13], que establece que:

$$\sqrt{N}(\hat{I}_N - I) \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

donde $\sigma^2 = \mathrm{Var}[h(X)]$ para $X \sim p(x)$. Esto permite cuantificar la precisión del estimador mediante intervalos de confianza.

En la figura 2.2, visualizamos la idea general de los métodos de Monte Carlo

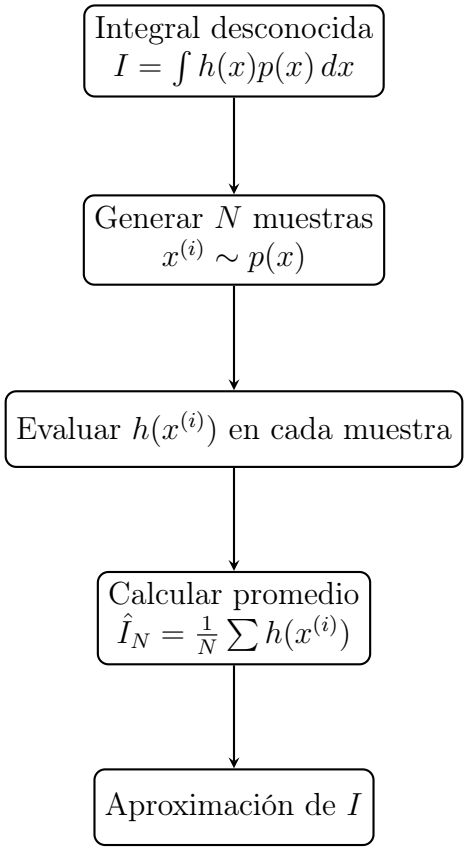


Figura 2.2: Flujo básico de un método de Monte Carlo.

2.3.1. Extensiones

En aplicaciones reales, no siempre es posible muestrear directamente de $p(x)$. Esto motiva variantes (durante el contenido de esta tesis, se revisan mas a detalle) como:

1. **Muestreo por importancia**: es una técnica de simulación que permite aproximar cantidades de interés mediante muestras generadas a partir de una distribución auxiliar $q(x)$, elegida de manera que facilite el muestreo. En vista que esta distribución puede diferir a la distribución $p(x)$, cada muestra recibe un peso proporcional de la siguiente manera $\omega(x) = p(x)/q(x)$. De esta manera, los pesos corrigen la diferencia entre las distribuciones y permiten obtener aproximaciones ponderadas de las cantidades de interés.

2. **Métodos de Monte Carlo secuenciales (SMC)**: ajustan el muestreo durante el tiempo, base fundamental de los filtros de partículas.

   En el contexto de la inferencia bayesiana, estas técnicas se utilizan para aproximar cantidades como la distribución posterior o la distribución predictiva. Sin embargo, en modelos dinámicos, la presencia de dependencia temporal y variables latentes introduce dificultades adicionales.

   En particular, la verosimilitud de los datos puede escribirse de forma recursiva como

   $$p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta}) = \prod_{t=1}^{T} p(\boldsymbol{y_t} \mid \mathbf{y}_{1:t-1}, \boldsymbol{\theta}),$$

   donde cada término requiere integrar sobre variables no observables. Estas integrales no suelen tener forma cerrada en modelos no lineales, lo que impide aplicar métodos tradicionales de inferencia.

   Los métodos de Monte Carlo Secuenciales (SMC), proporcionan una solución a este problema al aproximar de forma iterativa las distribuciones involucradas. [14]

   La principal ventaja de estos métodos radica en su capacidad para tratar modelos no lineales y no gaussianos, donde los métodos analíticos no son aplicables. Por esta razón, el SMC, sobre todo el filtro de partículas constituyen una herramienta fundamental en la estimación de modelos DSGE en su forma no lineal.

3. **Markov Chain Monte Carlo (MCMC)**: permiten generar muestras de distribuciones complejas mediante la construcción de cadenas de Markov cuya distribución estacionaria coincide con la distribución objetivo, en este caso la distribución posterior $p(\boldsymbol{\theta} \mid \mathbf{y}_{1:T})$.

   El algoritmo de Gibbs Sampling constituye un caso particular de MCMC en el que se muestrea directamente de las distribuciones condicionales completas. Si el vector de parámetros se descompone como $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$, el algoritmo consiste en generar iterativamente:

$$\theta_1^{(j+1)} \sim p(\theta_1 \mid \theta_2^{(j)}, \ldots, \theta_k^{(j)}, \mathbf{y}_{1:T}),$$

$$\theta_2^{(j+1)} \sim p(\theta_2 \mid \theta_1^{(j+1)}, \theta_3^{(j)}, \ldots, \theta_k^{(j)}, \mathbf{y}_{1:T}),$$

   y así sucesivamente hasta completar todos los parámetros $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$.

   Este procedimiento requiere que las distribuciones condicionales completas tengan forma conocida y sean fáciles de muestrear, lo cual no siempre ocurre en modelos complejos.

   Por su parte, el algoritmo Metropolis - Hasting construye una cadena de Markov mediante el uso de una distribución propuesta, denotada por $q(\boldsymbol{\theta}' \mid \boldsymbol{\theta})$, la cual define un mecanismo para generar candidatos $\boldsymbol{\theta}'$ a partir del estado actual $\boldsymbol{\theta}$.

   El objetivo de la distribución propuesta $q(\cdot)$ es explorar el espacio de parámetros. Su elección es crucial, ya que determina la eficiencia del algoritmo.

   Dado un estado actual $\boldsymbol{\theta}$, el algoritmo genera un candidato $\boldsymbol{\theta}' \sim q(\boldsymbol{\theta}' \mid \boldsymbol{\theta})$, el cual se acepta con probabilidad:

$$\alpha = \text{mín}\left(1, \frac{p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta}')\, p(\boldsymbol{\theta}')\, q(\boldsymbol{\theta} \mid \boldsymbol{\theta}')}{p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})\, q(\boldsymbol{\theta}' \mid \boldsymbol{\theta})}\right).$$

   Si el candidato es aceptado, se establece $\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}'$; en caso contrario, se mantiene el valor actual.

La principal ventaja del algoritmo de Metropolis-Hastings es que permite generar muestras de la distribución posterior sin necesidad de conocer la constante de normalización $p(\mathbf{y}_{1:T})$, lo cual resulta especialmente útil en modelos donde la verosimilitud es difícil de evaluar.

2.4. Estimación Robusta

En contextos bayesianos, se han desarrollado enfoques robustos para definir las prioris que reduzcan sensibilidad a datos extremos, utilizar distribuciones como t-Student que tienen colas pesadas para modelar la verosimilitud, construir estimadores de tipo mediana a partir de la distribución posterior.

En el caso de métodos de simulación como MCMC o Filtros de Partículas, la distribución aproximada de los estados o parámetros puede presentar *colas pesadas* o *asimetrías*. En tales contextos, la media de las simulaciones puede ser inestable, mientras que considerar otros estimadores robustos pueden proporcionar un resumen más robusto de la distribución posterior [15].

En inferencia estadística, un estimador se considera robusto cuando su desempeño no se ve afectado de manera significativa por pequeñas desviaciones en los supuestos del modelo o por la presencia de *valores atípicos* (*outliers*). A diferencia de los estimadores clásicos como la media que son altamente sensibles a observaciones extremas, los estimadores robustos buscan mantener estabilidad en escenarios donde los datos no siguen exactamente la distribución teórica asumida [16].

También el método secuencial de Monte Carlo, el Filtro de Partículas crea aproximaciones de la distribución posterior mediante un conjunto finito de partículas. Estas partículas pueden concentrarse de manera desproporcional o también verse afectadas por el remuestreo, produciendo valores extremos. Además, en las predicciones, la dispersión aumenta y la presencia de trayectorias atípicas es más probable.

2.4.1. Medidas de Robustez

Existen dos conceptos fundamentales en la teoría de estimación robusta:

1. **Breakdown Point**: Proporción máxima de observaciones contaminadas que un estimador puede tolerar antes de producir resultados arbitrariamente erróneos. Por ejemplo, la media tiene un breakdown point de 0 %, mientras que la mediana alcanza el 50 %.

2. **Función de Influencia (Influence Function)**: Se define como el efecto infinitesimal de una observación sobre el estimador. Un estimador robusto tiene una función de influencia acotada, es decir, ningún valor individual puede tener un impacto desproporcionado.

Estos conceptos formalizan la idea de estabilidad ante datos adversos y permiten comparar el desempeño de diferentes estimadores en situaciones no ideales. La definición formal es

Definición 2.5 (Estimador robusto, [16])**.** *Un estimador se considera robusto si su desempeño no se ve significativamente afectado por pequeñas desviaciones respecto a los supuestos del modelo, como la presencia de valores atípicos o leves cambios estructurales.*

Además de la mediana, existen otros estimadores robustos ampliamente utilizados en estadística, especialmente cuando se busca mitigar el efecto de valores atípicos o de distribuciones sesgadas. Entre ellos se encuentran los cuantiles, los cuales tienen propiedades robustas bajo ciertos criterios.

Definición 2.6 (Cuantil τ-ésimo)**.** *Dada una distribución de probabilidad F, el cuantil de orden $\tau \in (0,1)$ se define como:*

$$q_\tau = \text{ínf}\{x \in \mathbb{R} : F(x) \geq \tau\}.$$

El cuantil τ-ésimo es un estimador robusto con función de influencia acotada. El caso particular $\tau = 0.5$ corresponde a la mediana. Los cuartiles serían $\tau = 0.25, 0.5, 0.75$ y los percentiles $\tau = 0.01, 0.02, ..., 0.99$.

Aunque nos centramos en la mediana como medida robusta, es relevante mencionar que otros cuantiles (como los percentiles 25 y 75, o bien, cuartil 1 y 3) también pueden utilizarse para describir la dispersión de las partículas.

2.5. Validación Cruzada Bayesiana: ELPD

En el marco de la inferencia bayesiana, la evaluación de modelos se fundamenta en su capacidad para generar predicciones probabilísticas coherentes con los datos observados. Esta evaluación se realiza a través de la densidad predictiva, la cual incorpora de manera natural la incertidumbre sobre los parámetros del modelo.

Sea y_t una observación y $\boldsymbol{\theta}$ el vector de parámetros, la densidad predictiva se define como

$$p(y_t \mid \mathbf{y}_{1:t-1}) = \int p(y_t \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta} \mid \mathbf{y}_{1:t-1})\, d\boldsymbol{\theta}.$$

A partir de esta expresión, una medida central de desempeño es la *Expected Log Predictive Density* (ELPD) [17], definida como

$$\text{ELPD} = \sum_{t=1}^{T} \log p(y_t \mid \mathbf{y}_{1:t-1}),$$

donde $p(y_t \mid \mathbf{y}_{1:t-1})$ representa la densidad predictiva fuera de muestra, y al $\log p(y_t \mid \mathbf{y}_{1:t-1})$ se le conoce como el *log score*. Esta cantidad resume la capacidad del modelo para asignar alta probabilidad a datos no observados, constituyendo un criterio fundamental de comparación entre modelos bayesianos.

Dado que esta integral no posee una solución analítica cerrada en la mayoría de los escenarios reales, se implementa una aproximación numérica basada en los Métodos de Monte Carlo. Extrayendo N extracciones aleatorias independientes $\{\theta^{(i)}\}_{i=1}^{N}$ provenientes de la distribución posterior, la densidad predictiva estimada se aproxima mediante el promedio muestral discreto:

$$p(y_t \mid y_{1:t-1}) \approx \frac{1}{N} \sum_{i=1}^{N} p(y_t \mid y_{1:t-1}, \theta^{(i)}) \tag{2.10}$$

De esta manera, tanto la estimación del modelo como su evaluación predictiva se fundamentan en aproximaciones consistentes dentro del marco bayesiano, lo que garantiza coherencia entre inferencia y validación.

El ELPD entre valores más altos indican una mejor capacidad predictiva del modelo,

ya que implican que el sistema asigna mayor probabilidad a los datos efectivamente observados.

Capítulo 3

Modelo en Espacio de Estado

En este capítulo se presenta una introducción general a los modelos en espacio de estados, los cuales nos ayudan a modelar fenómenos dinámicos y estocásticos en una amplia variedad de campos, incluyendo economía, ingeniería, biología y finanzas. Su principal característica es la capacidad de representar sistemas donde las variables observadas dependen de variables latentes que evolucionan en el tiempo bajo variaciones.

Comenzaremos definiendo uno de los modelos más elementales de los Modelos de espacio de estado: los Modelos Lineales Dinámicos (DLM, por sus siglas en inglés), los cuales definen series temporales con supuestos de linealidad y normalidad. Estos modelos permiten la aplicación directa del filtro de Kalman, una técnica de inferencia secuencial óptima.

Posteriormente, ampliaremos la discusión hacia modelos en espacio de estado no lineales y con ruido no gaussiano, los cuales capturan de manera más realista muchos procesos económicos y financieros, aunque se requiere técnicas computacionales más avanzadas para su estimación, como los Filtros de Partículas.

Según [14], un modelo de espacio de estados consiste en dos procesos de tiempo discretos $\{X_t\} := (X_t)_{t>0}$, $\{Y_t\} := (Y_t)_{t>0}$ tomando valores en los espacios χ y $\mathbb{Y}$ respectivamente. Estos espacios pueden ser espacios euclideanos multidimensionales, espacios discretos, pero suelen ser menos estandar.

Una especificación simplificada del modelo se realiza, primero definiendo un vector de parámetros $\boldsymbol{\theta} \in \Theta$ y un conjunto de densidades que definen la densidad conjunta de los procesos

$$p_x(\boldsymbol{x}_t|\boldsymbol{x_{0:t-1}}) = p_x(\boldsymbol{x}_t|\boldsymbol{x_{t-1}}) \tag{3.1}$$

$$p_x(\boldsymbol{y}_t|\boldsymbol{x_{0:t}}, \boldsymbol{y_{0:t-1}}) = f_x(\boldsymbol{y}_t|\boldsymbol{x_t}) \tag{3.2}$$

donde $p_x(\boldsymbol{x_0})$ es la densidad inicial (prior), $p_\theta(\boldsymbol{x}_t|\boldsymbol{x_{t-1}})$ y $f_\theta(\boldsymbol{y}_t|\boldsymbol{x_t})$ son las densidades de $\boldsymbol{x_t}$ y $\boldsymbol{y_t}$ respectivamente, y pertenecen a la familia indexada por dos conjuntos de parámetros $(\boldsymbol{\theta}, \boldsymbol{x_t})$ y $(\boldsymbol{\theta}, \boldsymbol{y_t})$ respectivamente. La ecuación 3.1 es conocida como la ecuación de estados, mientras que la ecuación 3.2 es la ecuación de observación, donde las distribuciones de ambas ecuaciones pueden ser no lineales y no gaussianas.

Se visualiza también a través de la figura 3.1, la independencia condicional de las variables aleatorias del modelo, que Y_t es una variable independiente en el tiempo, pero que depende del estado actual del modelo, mientras que los estados dependen del estado anterior.

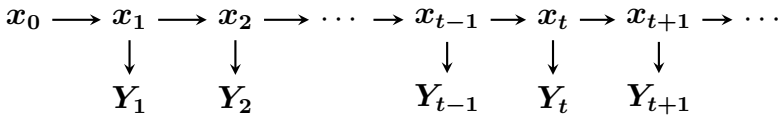


Figura 3.1: Independencia condicional en un modelo de espacio de estados.

3.1. Modelos Lineales Dinámicos

Los modelos lineales dinámicos con un enfoque bayesiano fueron introducidos por [18], donde son un caso especial de los modelos de espacio de estados cuando se consideran relaciones lineales, que es definido como

Definición 3.1. *El modelo lineal dinámico (DLM) general está caracterizado por el conjunto de cuadruplas*

$$\{\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{V}, \boldsymbol{W}\}_t = \{\boldsymbol{F_t}, \boldsymbol{G_t}, \boldsymbol{V_t}, \boldsymbol{W_t}\}$$

donde para cada t

1. $\boldsymbol{F_t}(n \times r)$ *es una matriz conocida*
2. $\boldsymbol{G_t}(n \times n)$ *es una matriz conocida*
3. $\boldsymbol{V_t}(r \times r)$ *es una matriz de covarianza conocida*
4. $\boldsymbol{W_t}(n \times n)$ *es una matriz de covarianza conocida*

La cuadrupla $\{\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{V}, \boldsymbol{W}\}_t$ define la relación de las observaciones $\boldsymbol{Y_t}$ y el vector de estado $\boldsymbol{X_t}(n \times 1)$ en el tiempo t, donde durante el tiempo tiene distribución normal

$$(\boldsymbol{Y_t}|\boldsymbol{X_t}) \sim N(\boldsymbol{F_t'X_t}, \boldsymbol{V}_t),\ (\boldsymbol{X_t}|\boldsymbol{X_{t-1}}) \sim N(\boldsymbol{G_tX_{t-1}}, \boldsymbol{W_t}) \tag{3.3}$$

La ecuación anterior, implicítamente ya está condicionado en el conjunto de información a priori $\boldsymbol{D_{t-1} = \{D_0, Y_1, Y_2, ..., Y_{t-1}\}}$, donde $\boldsymbol{D_0}$ es el conjunto de información inicial.

Con lo visto anteriormente, de manera general se define como

$$\boldsymbol{Y_t = F_t'X_t + v_t},\quad \boldsymbol{v_t} \sim N(\boldsymbol{0}, \boldsymbol{V_t}) \tag{3.4}$$

$$\boldsymbol{X_t = G_tX_{t-1} + w_t},\quad \boldsymbol{w_t} \sim N(\boldsymbol{0}, \boldsymbol{W_t}) \tag{3.5}$$

$$(\boldsymbol{X_0}|\boldsymbol{D_0}) \sim N(\boldsymbol{m_0}, \boldsymbol{C_0}) \tag{3.6}$$

$\boldsymbol{v_t}, \boldsymbol{w_t}$ son independientes entre si y se les conoce como los errores de la ecuación. La ecuación 3.4 se le conoce como la ecuación de observación y a la ecuación 3.5 como la ecuación de estado, y la ecuación 3.6 es la información inicial para algunos valores a prioris $\boldsymbol{m_0}$ y $\boldsymbol{C_0}$.

Al uso de los datos recientes para inferir en los estados actuales con respecto a la información actual, se conoce como filtración. Entonces la distribución de interés es $(\boldsymbol{X_t}|\boldsymbol{D_t})$, utilizando las ecuaciones de actualización presentados en el teorema 3.1.

Un enfoque más detallado sobre los filtrajes en particular el filtro de Kalman lo da [19], en la cual detalla ciertas caracteríticas de este filtraje para la solución del modelo, donde el problema a resolver es estimar

$$\hat{\boldsymbol{X}}_{t|t-1} = E[\boldsymbol{X_t}|\boldsymbol{Y_{t-1}}],\quad \hat{\boldsymbol{X}}_{t|t} = E[\boldsymbol{X_t}|\boldsymbol{Y_t}] \tag{3.7}$$

mediante las ecuaciones de actualización, donde se explican detalladamente en el teorema 3.1 de la siguiente sección.

3.1.1. Filtración

Cuando se tiene construido el modelo, uno de los principales objetivos es realizar inferencia sobre los estados utilizando la información disponible hasta el momento. Supongamos que tenemos datos observados hasta el tiempo t, con notación $\boldsymbol{Y_{1:t}}$, y el objetivo es inferir en el vector de estados. En este contexto, la distribución de interés es la densidad condicional $p(\boldsymbol{X_s}|\boldsymbol{Y_{1:t}})$.

Dependiendo de la relación entre s y t, se distinguen tres casos:

1. **Filtración:** cuando $s = t$, se estima el estado actual dado los datos observados hasta el mismo periodo.

2. **Pronóstico:** cuando $s > t$, se proyectan estados futuros a partir de la información disponible.

3. **Suavizamiento (análisis retrospectivo):** cuando $s < t$, se reevalúan estados pasados incorporando información posterior.

La Figura 3.2 resume la diferencia entre filtración, suavizamiento y pronóstico.

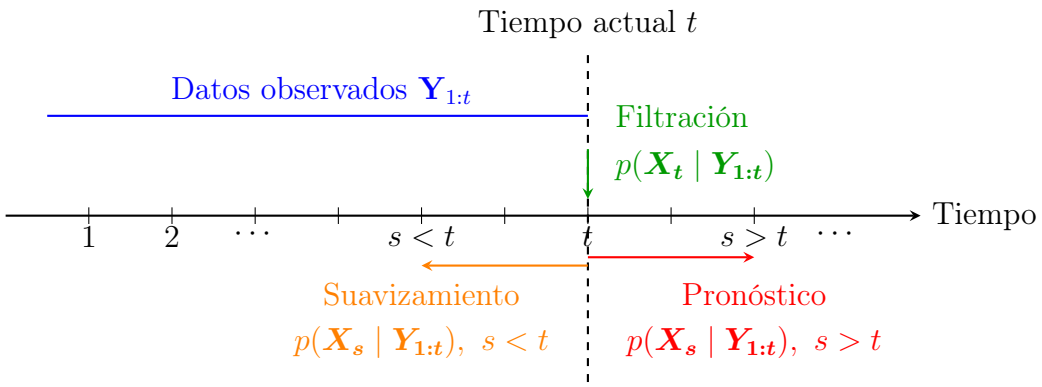


Figura 3.2: Relación temporal entre filtración, suavizamiento y pronóstico en modelos de espacio de estados.

En la filtración, se considera que las observaciones se pueden ir actualizando de manera continua en el tiempo mediante las ecuaciones de actualización que se definen en el teorema 3.1, obteniendo la densidad condicional $p(\boldsymbol{X_t}|\boldsymbol{Y_{1:t}})$, lo que permite resolver esto se le conoce como el filtro de Kalman.

La idea general para resolver el problema de filtrado de una manera general es

1. Se obtiene la distribución predictiva del estado $\boldsymbol{X_t}$, dadas las observaciones $\boldsymbol{Y_{1:t-1}}$, usando la densidad filtrada $p(\boldsymbol{X_{t-1}}|\boldsymbol{Y_{1:t-1}})$ y la distribución condicional $p(\boldsymbol{X_t}|\boldsymbol{X_{t-1}})$.

2. Se obtiene la distribución predictiva de la siguiente observación $\boldsymbol{Y_t}$ dadas las observaciones anteriores.

3. Se obtiene la distribución filtrada $p(\boldsymbol{X_t}|\boldsymbol{Y_{1:t}})$ usando la regla de Bayes con $p(\boldsymbol{X_t}|\boldsymbol{Y_{1:t-1}})$ como la distribución a priori y $p(\boldsymbol{Y_t}|\boldsymbol{X_t})$ como la verosimilitud.

Ahora bien, para los DLM, la filtración se resuelve mediante el filtro de Kalman, que se enuncia a continuación [18]:

Teorema 3.1 (Filtro de Kalman)**.** *Consideremos el DLM enunciado anteriormente, y supongamos que*

$$\boldsymbol{X_{t-1}}|\boldsymbol{Y_{1:t-1}} \sim N(\boldsymbol{m_{t-1}}, \boldsymbol{C_{t-1}})$$

Entonces los siguientes resultados se cumplen

1. *La distribución predictiva (del primer paso) de $\boldsymbol{X_t}$ dado $\boldsymbol{Y_{1:t-1}}$ es Gaussiana con parámetros*

$$\boldsymbol{a_t} = E(\boldsymbol{X_t}|\boldsymbol{Y_{1:t-1}}) = \boldsymbol{G_t m_{t-1}}, \quad \boldsymbol{R_t} = Var(\boldsymbol{X_t}|\boldsymbol{Y_{1:t-1}}) = \boldsymbol{G_t C_{t-1} G_t'} + \boldsymbol{W_t} \tag{3.8}$$

2. *La distribución predictiva (del primer paso) de $\boldsymbol{Y_t}$ dado $\boldsymbol{Y_{1:t-1}}$ es Gaussiana con parámetros*

$$\boldsymbol{f_t} = E(\boldsymbol{Y_t}|\boldsymbol{Y_{1:t-1}}) = \boldsymbol{F_t' a_t}, \quad \boldsymbol{Q_t} = Var(\boldsymbol{Y_t}|\boldsymbol{Y_{1:t-1}}) = \boldsymbol{F_t' R_t F_t} + \boldsymbol{V_t} \tag{3.9}$$

3. *La distribución filtrada $\boldsymbol{X_t}$ de $\boldsymbol{Y_t}$ dado $\boldsymbol{Y_{1:t}}$ es Gaussiana con parámetros*

$$\boldsymbol{m_t} = E(\boldsymbol{X_t}|\boldsymbol{Y_{1:t}}) = \boldsymbol{a_t} + \boldsymbol{R_t F_t Q_t^{-1} e_t}, \quad \boldsymbol{C_t} = Var(\boldsymbol{X_t}|\boldsymbol{Y_{1:t}}) = \boldsymbol{R_t} - \boldsymbol{R_t F_t Q_t^{-1} F_t' R_t} \tag{3.10}$$

donde $\boldsymbol{e_t} = \boldsymbol{Y_t} - \boldsymbol{f_t}$

Las ecuaciones 3.8, 3.9 y 3.10 las conocemos como las *ecuaciones de actualización.*

3.1.2. Suavizamiento

En ciertos problemas de las series de tiempo se requiere conocer datos pasados, es decir, el comportamiento del sistema a partir de los datos disponibles. Para ello se utiliza el algoritmo recursivo regresivo, con el que se obtiene la distribución condicional $p(\boldsymbol{X_t}|\boldsymbol{Y_{1:T}})$, para cualquier $t < T$, empezando con la distribución filtrada $p(\boldsymbol{X_T}, \boldsymbol{Y_{1:T}})$ y estimando de forma retrospectiva todos los estados. Se enuncia el algoritmo a continuación:

Teorema 3.2 (Suavizamiento de Kalman)**.** *Consideremos el DLM enunciado anteriormente, si*

$\boldsymbol{X_{t+1}}|\boldsymbol{Y_{1:T}} \sim N(\boldsymbol{s_{t+1}}, \boldsymbol{S_{t+1}})$*, entonces* $\boldsymbol{X_t}|\boldsymbol{Y_{1:T}} \sim N(\boldsymbol{s_t}, \boldsymbol{S_t})$*, donde*

$$\boldsymbol{s_t} = \boldsymbol{m_t} + \boldsymbol{C_t G'_{t+1} R^{-1}_{t+1}}(\boldsymbol{s_{t+1}} - \boldsymbol{a_{t+1}})$$

$$\boldsymbol{S_t} = \boldsymbol{C_t} - \boldsymbol{C_t G'_{t+1} R^{-1}_{t+1}}(\boldsymbol{R_{t+1}} - \boldsymbol{S_{t+1}})\boldsymbol{R^{-1}_{t+1} G_{t+1} C_t}$$

Observemos que para poder realizar el algoritmo de suavizamiento, también hay que ir realizando el algoritmo del filtro de Kalman.

3.1.3. Pronóstico

Ahora bien, una de las características más importantes a considerar para la utilización de un modelo aplicado a la vida real es predecir observaciones futuras para la toma de decisiones, es decir estimar $\boldsymbol{Y_{t+1}}$ a través de observaciones ya disponibles $\boldsymbol{Y_{1:t}}$, lo cual conlleva a encontrar la distribución $p(\boldsymbol{Y_{t+1}}|\boldsymbol{Y_{1:t}})$, que se le conoce como distribución de predicción a primer paso. El objetivo es estimar el estado siguiente $\boldsymbol{X_{t+1}}$, obteniendo $\boldsymbol{Y_{t+1}}$, lo cual es válido debido a las dependencias del modelo.

Dada la naturaleza Markoviana del modelo, la distribución filtrada en el tiempo t, sirve como la distribución inicial para la futura evolución del modelo, se presenta el siguiente algoritmo que sirve para obtener la distribución de los pronósticos de los estados y observaciones:

Teorema 3.3. *Consideremos el DLM enunciado anteriormente, y supongamos que* $\boldsymbol{a_t}(0) = \boldsymbol{m_t}$ *y* $\boldsymbol{R_t(0)} = \boldsymbol{C_t}$*, entonces los siguientes enunciados se cumplen.*

1. *La distribución predictiva de* $\boldsymbol{X_{t+k}}$ *dado* $\boldsymbol{Y_{1:t}}$ *es Gaussiana con parámetros*

$$\boldsymbol{a_t}(k) = \boldsymbol{G_{t+k}}\boldsymbol{a_t}(k-1), \quad \boldsymbol{R_t}(k) = \boldsymbol{G_{t+k}}\boldsymbol{R_t}(k-1)\boldsymbol{G'_{t+k}} + \boldsymbol{W_{t+k}} \tag{3.11}$$

2. *La distribución predictiva de* $\boldsymbol{Y_{t+k}}$ *dado* $\boldsymbol{Y_{1:t}}$ *es Gaussiano con parámetros*

$$\boldsymbol{f_t}(k) = \boldsymbol{F'_{t+k}}\boldsymbol{a_t}(k), \quad \boldsymbol{Q_t}(k) = \boldsymbol{F'_{t+k}}\boldsymbol{R_t}(k)\boldsymbol{F_{t+k}} + \boldsymbol{V_{t+k}} \tag{3.12}$$

3.2. Modelos de Espacios de Estados con Filtro de Partículas

Se visualizó en la sección anterior, un modelo de espacio de estado lineal con ruido gaussiano, y como ajustarlo a una serie de tiempo a través del filtro de Kalman. Ahora bien, en general un modelo de espacio de estados puede ser no lineal y al mismo tiempo tener ruido no gaussiano, por ende, para poder resolver estos tipos de modelos debemos de recurrir a otros métodos, dado que el filtraje de Kalman, es útil para los modelos de espacio de estados lineales.

Las técnicas de filtraje a utilizar para abordar este tipo de modelos son los filtros de partículas, que son Métodos Secuenciales de Monte Carlo (SMC siglas en inglés) [20], donde el método ha demostrado ser muy exitoso en cualquier tipo de aplicacion de filtrado y no tiene limitante a diferentes modelos de series temporales, sin embargo, la desventaja de esta técnica es el alto costo computacional [21].

Se considera el siguiente modelo de espacio de estado general

$$\boldsymbol{X_t} \sim p(\boldsymbol{X_t}|\boldsymbol{X_{t-1}}), \quad \boldsymbol{Y_t} \sim p(\boldsymbol{Y_t}|\boldsymbol{X}_t) \tag{3.13}$$

donde $\boldsymbol{X_t} \in \mathbb{R}^d$ lo definimos como los estados latentes y $\boldsymbol{Y_t}$ la observación al tiempo t. Como se visualizó en los modelos DLM, el objetivo del filtraje es obtener la distribución a

posteriori secuencial $p(\boldsymbol{X_t}|\boldsymbol{Y_{1:t}})$ para cada t, que no requiere asumir linealidades ni ruidos gaussianos. El objetivo principal del filtro de partículas es representar la distribución $p(\boldsymbol{X_t}|\boldsymbol{Y_{1:t}})$ a través de un conjunto de N partículas $\{\boldsymbol{X_t}^{(i)}\}_{i=1}^{N}$, cada uno con un peso asociado $\omega_t^{(i)}$, tal que

$$p(\boldsymbol{X_t}|\boldsymbol{Y_{1:t}}) \approx \sum_{i=1}^{N} \omega_t^{(i)} \boldsymbol{X_t}^{(i)}$$

donde $\omega_t^{(i)}$ son pesos normalizados, es decir, se cumple que $\sum \omega_t^{(i)} = 1$. El conjunto de partículas obtenidas del filtrado evolucionan durante el tiempo en tres pasos en cada iteración t:

1. **Propagación:** para cada partícula $\boldsymbol{X_{t-1}}^{(i)}$, se genera una muestra $\boldsymbol{X_t}^{(i)} \sim p(\boldsymbol{X_t}|\boldsymbol{X_{t-1}}^{(i)})$.

2. **Actualización de pesos:** se actualiza el peso de cada partícula según su verosimilitud $\omega_t^{(i)} \propto \omega_{t-1}^{(i)} \cdot p(\boldsymbol{Y_t}|\boldsymbol{X_t}^{(i)})$, y luego se normalizan.

3. **Remuestreo:** si los pesos se degeneran (es decir, unos pocos dominan), se remuestrean las partículas con reemplazo según sus pesos para obtener una muestra no ponderada.

Esto surge del deseo de aproximar la densidad a posteriori. Supongamos que queremos aproximar la densidad $p(\boldsymbol{X_t})$, sin embargo puede ser dificil muestrearlo directamente [22]. Podemos muestrearlo de otra densidad $q(\boldsymbol{X_t})$ de la siguiente forma

$$p(\boldsymbol{X_t}) = \frac{p(\boldsymbol{X_t})}{q(\boldsymbol{X_t})} q(\boldsymbol{X_t})$$

y aproximar la esperanza por el principio de Monte Carlo

$$E_p(g(\boldsymbol{X_t})) = \int g(\boldsymbol{X_t}) p(\boldsymbol{X_t}) dx = \int g(\boldsymbol{X_t}) \frac{p(\boldsymbol{X_t})}{q(\boldsymbol{X_t})} q(\boldsymbol{X_t}) dx \approx \frac{1}{N} \sum_{i=1}^{N} g(\boldsymbol{X_t}^{(i)}) \omega_t^{(i)}$$

donde $\boldsymbol{X_t}^{(i)} \sim q(\boldsymbol{X_t})$ y $\omega^{(i)} \propto \frac{p(\boldsymbol{X_t}^{(i)})}{q(\boldsymbol{X_t}^{(i)})}$ que son los pesos.

Ahora bien, para la actualización de los pesos del paso $t-1$ al paso t se sigue de la factorización de la posteriori

$$p(\boldsymbol{X_t}|\boldsymbol{Y_t}) \propto p(\boldsymbol{Y_t}|\boldsymbol{X_t})p(\boldsymbol{X_t}|\boldsymbol{X_{t-1}})p(\boldsymbol{X_{t-1}}|\boldsymbol{Y_{t-1}})$$

Si nuestra distribución de propuesta se factoriza como $q(\boldsymbol{X_t}) = q(\boldsymbol{X_t}|\boldsymbol{X_{t-1}}, \boldsymbol{Y_t})q(\boldsymbol{X_{t-1}}|\boldsymbol{Y_{t-1}})$, los pesos se convierten

$$\omega_t^{(i)} \propto \omega_{t-1}^{(i)} \frac{p(\boldsymbol{Y_t}|\boldsymbol{X_t}^{(i)})p(\boldsymbol{X_t}|\boldsymbol{X_{t-1}}^{(i)})}{q(\boldsymbol{X_t}^{(i)}|\boldsymbol{X_{t-1}}^{(i)}, \boldsymbol{Y_t})} = \omega_{t-1}^{(i)} p(\boldsymbol{Y_t}|\boldsymbol{X_t^{(i)}})$$

donde $p(\boldsymbol{X_t}|\boldsymbol{X_{t-1}}^{(i)}) = q(\boldsymbol{X_t}^{(i)}|\boldsymbol{X_{t-1}}^{(i)}, \boldsymbol{Y_t})$, este método es conocido como el filtro de importancia con remuestreo (SIR), propuesto por [23]. A continuación se visualiza en forma de seudocódigo:

Algoritmo 1 Filtro de Partículas SIR

Inicializar: para $i = 1, ..., N$, muestrear $\boldsymbol{X}_0^{(i)} \sim p(\boldsymbol{X_0})$, establecer $\omega_0^{(i)} = \frac{1}{N}$
para $t = 1, 2, ..., T$ **hacer**
    para cada $i = 1, ..., N$ **hacer**
        Propagar $\boldsymbol{X_t}^{(i)} \sim p(\boldsymbol{X_t}|\boldsymbol{X_{t-1}}^{(i)})$
        Calcular peso no normalizado $\tilde{\omega}_t^{(i)} = p(\boldsymbol{Y_t}|\boldsymbol{X_t}^{(i)})$
    fin para
    Normalizar los pesos $\omega_t^{(i)} = \tilde{\omega}_t^{(i)} / \sum_{j=1}^{N} \tilde{\omega}_t^{(i)}$
    Evaluar el *número efectivo de particulas*(NEF): **NEF**$=\frac{1}{\sum_{i=1}^{N}(\omega_t^{(i)})^2}$
    si $NEF < umbral$ **entonces**
        Remuestrear N partículas según $\omega_t^{(i)}$, asignar $\omega_t^{(i)} = \frac{1}{N}$
    fin si
fin para

Este es el filtro de partículas que se considera como el clásico, pero según [24] la técnica SIR tiene una debilidad fundamental, que cuando existe valores atípicos, los pesos podrían ser muy desproporcionales, lo puede requerir una gran cantidad partículas. Esto fue una motivación para que a partir de esta técnica se creen diferentes variantes, donde se ha mejorado ya sea la propagación, actualización de pesos o remuestreo de las partículas, según las series de tiempo a estudiar.

Antes de presentar el filtro de particulas empleado, se mostrará primeramente la limitación del filtro de Kalman donde el modelo de espacio de estado es no lineal.

3.2.1. Limitaciones del filtro de Kalman

El filtro de Kalman proporciona una solución óptima para la estimación de estados en modelos de espacio de estados lineales y gaussianos. En particular, este método asume que el sistema puede escribirse como:

$$\begin{aligned} \boldsymbol{Y_t} &= \boldsymbol{F_t'} \boldsymbol{X_t} + \boldsymbol{v_t}, \\ \boldsymbol{X_t} &= \boldsymbol{G_t} \boldsymbol{X_{t-1}} + \boldsymbol{w_t} \end{aligned} \tag{3.14}$$

donde $\boldsymbol{v_t} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{V_t})$ y $\boldsymbol{w_t} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{W_t})$ son perturbaciones gaussianas.

Sin embargo, en muchos problemas prácticos, estas condiciones no se cumplen. A continuación se presenta un ejemplo de procesamiento de señales que ilustra la violación de la linealidad en la ecuación de observación.

Considérese un objeto en movimiento cuya posición en el tiempo t está dada por el estado latente $\boldsymbol{X_t} \in \mathbb{R}^2$. La dinámica del sistema puede modelarse como:

$$\boldsymbol{X_t} = \boldsymbol{X_{t-1}} + \boldsymbol{U_t}, \quad \boldsymbol{U_t} \sim \mathcal{N}_2(\boldsymbol{0}, \sigma^2 \boldsymbol{I_2}),$$

mientras que la observación corresponde a una medición angular obtenida, mediante una señal de radar:

$$Y_t = \arctan\left(\frac{\theta_t^{(2)}}{\theta_t^{(1)}}\right) + V_t, \quad V_t \sim \mathcal{N}(0, \sigma_Y^2).$$

En este caso, la ecuación de observación es claramente no lineal debido a la función $\arctan(\cdot)$. Como consecuencia, la distribución de Y_t condicionada en $\boldsymbol{X_t} = (\theta_t^{(1)}, \theta_t^{(2)})$ no es gaussiana en términos lineales del estado, lo que invalida los supuestos fundamentales del filtro de Kalman.

Ahora bien, la ecuación de observación se puede linealizar mediante serie de McLaurin,

$$\arctan x \approx x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots, \quad x \in [0, 1]$$

entonces para cumplir con la condición de la ecuación 3.14 lo hacemos de un término teniendo

$$Y_t = \frac{\theta_t^{(2)}}{\theta_t^{(1)}} + V_t, \quad V_t \sim \mathcal{N}(0, \sigma_Y^2), \quad \frac{\theta_t^{(2)}}{\theta_t^{(1)}} \in [0, 1] \tag{3.15}$$

Al graficar la aproximación lineal contra la ecuación original, visualmente se puede apreciar que hay errores entre las ecuación original y la aproximación lineal.

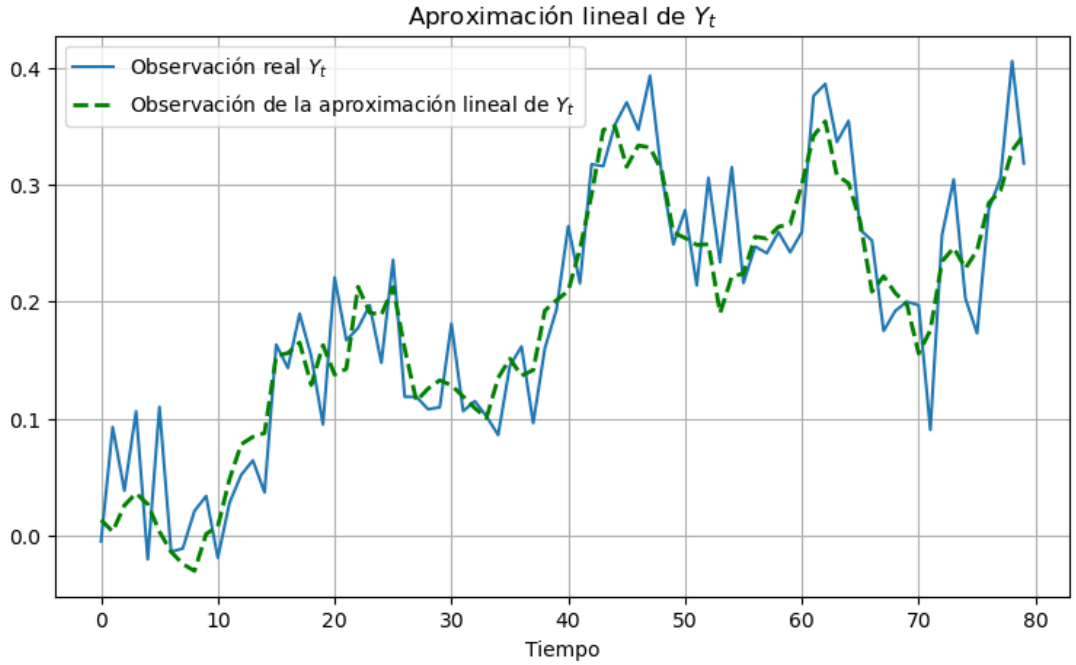


Figura 3.3: Ecuación original vs aproximación lineal

Al realizar una simulación aplicando el filtro de Kalman a la versión linealizada y el filtro de partículas directamente a la versión no linealizada de $T = 80$ observaciones donde $t \in [0, 70]$ se utilizó la filtración para estimación de los estados y en $t \in [71, 80]$ se aplica el pronóstico, obteniendo el gráfico 3.4, y obteniendo el RMSE (raiz del error cuadrático medio) para las estimaciones en $t \in [2, 70]$ y para los pronósticos $t \in [71, 80]$ en la tabla 3.1.

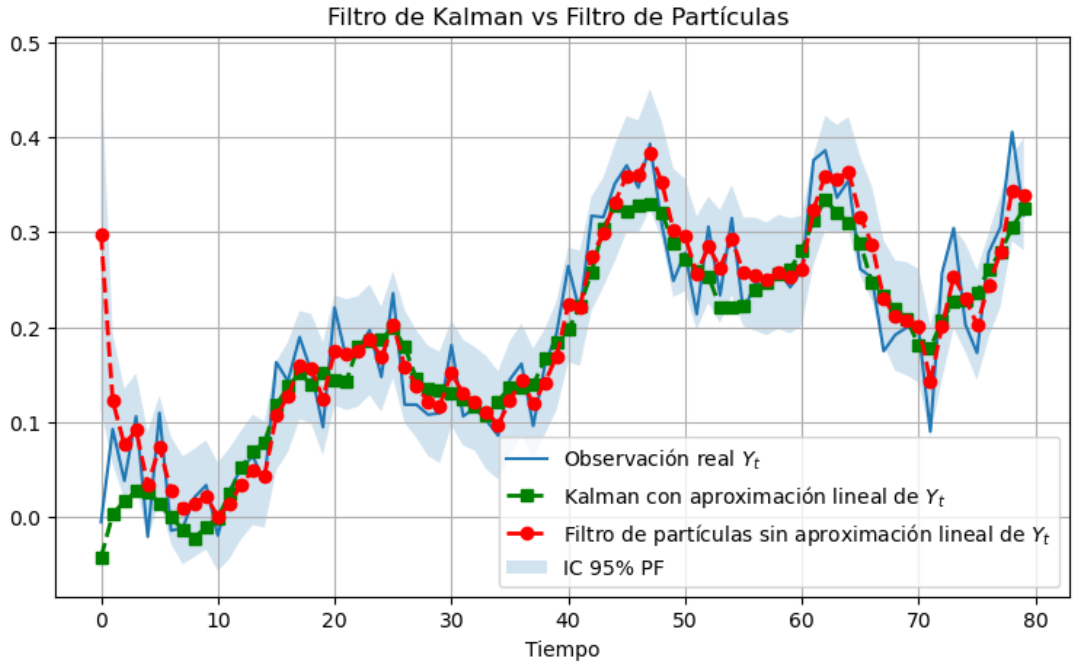


Figura 3.4: Simulación de Filtro de Kalman vs Filtro de Partículas

Filtro	**Estimación**	**Pronóstico**
Filtro de Kalman	0.031	0.051
Filtro de Partículas	0.027	0.038

Tabla 3.1: Raiz de error cuadrático medio (RMSE) para las estimaciones y pronósticos de los filtros de Kalman y filtros de partículas.

Tanto gráficamente como en el RMSE se visualiza una mejor aproximación en el filtro de partículas que el filtro de Kalman. En consecuencia, es necesario recurrir a métodos más generales, como los filtros de partículas, los cuales permiten aproximar distribuciones posteriores completas sin requerir supuestos de linealidad o normalidad. Además no siempre la aproximacion lineal es una buena aproximación.

Ahora bien, el filtro de particulas utilizado fue, el filtro de partículas regularizado ya que evita la degeneración de las partículas.

3.2.2. Filtro de Partículas Regularizado

Uno de los principales problemas que enfrentan los filtros de partículas clásicos es el fenómeno conocido como degeneración de partículas. Este problema ocurre debido a la actualización secuencial de los pesos asociados a cada partícula.

En un filtro de partículas estándar, después de varias iteraciones, una pequeña cantidad de partículas acumula prácticamente toda la probabilidad posterior, mientras que el resto recibe pesos cercanos a cero. Como consecuencia, aunque el número total de partículas permanezca constante, únicamente unas pocas contribuyen efectivamente a la aproximación de la distribución posterior.

Matemáticamente, la degeneración puede cuantificarse mediante el Número Efectivo de Partículas.

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N} (w_t^{(i)})^2}. \tag{3.16}$$

Cuando todos los pesos son iguales, $w_t^{(i)} = 1/N$, se obtiene

$$N_{\text{eff}} = N,$$

lo que indica máxima diversidad.
Por el contrario, cuando una única partícula concentra todo el peso,

$$w_t^{(1)} = 1,$$

y el resto posee peso nulo, se obtiene

$$N_{\text{eff}} = 1,$$

indicando degeneración completa.

Para mitigar este problema, los filtros de partículas incorporan una etapa de remuestreo, cuyo objetivo consiste en eliminar partículas con pesos bajos y replicar aquellas con pesos elevados.

Aunque el remuestreo reduce la degeneración de pesos, introduce un nuevo problema denominado empobrecimiento de partículas..

Después del remuestreo, múltiples copias idénticas de una misma partícula pueden aparecer en la población. En consecuencia, el conjunto de partículas pierde diversidad y la aproximación de la distribución posterior se vuelve cada vez más concentrada.

Entonces [25] proponen una solución basada en reemplazar la distribución empírica discreta generada por el remuestreo mediante una aproximación continua obtenida a través de técnicas de suavizado kernel, siendo un nuevo método llamado filtro de particulas regularizado.

Sea el conjunto de partículas remuestreadas

$$\left\{\tilde{\boldsymbol{x}}_t^{(1)}, \ldots, \tilde{\boldsymbol{x}}_t^{(N)}\right\},$$

la distribución empírica asociada puede escribirse como

$$\hat{p}(\boldsymbol{x_t}) = \frac{1}{N}\sum_{i=1}^{N} \delta\left(\boldsymbol{x_t} - \tilde{\boldsymbol{x}}_t^{(i)}\right), \tag{3.17}$$

donde $\delta(\cdot)$ representa la función delta de Dirac.
La idea fundamental del filtro regularizado consiste en sustituir esta distribución discreta por una versión suavizada

$$\hat{p}_h(\boldsymbol{x_t}) = \frac{1}{N}\sum_{i=1}^{N} K_h\left(\boldsymbol{x_t} - \tilde{\boldsymbol{x}}_t{}^{(i)}\right), \tag{3.18}$$

donde $K_h(\cdot)$ es una función kernel.
Asumiendo normalidad se puede utilizar un kernel Gaussiano multivariado,

$$K_h(\boldsymbol{x}) = \mathcal{N}\left(0, h^2\boldsymbol{\Sigma_t}\right), \tag{3.19}$$

donde Σ_t representa la matriz de covarianza empírica de las partículas y h corresponde al parámetro de suavizado o bandwidth, que puede ser calculado como

$$h = A(K)N^{-\frac{1}{n_x+4}} \quad con \quad [8c_{n_x}^{-1}(n_x+4)(2\pi)^{n_x}]^{\frac{1}{n_x+4}}$$

donde n_x representa la dimensión del vector de estado, c_{n_x} el volumen de la esfera unitaria de $\mathbb{R}^{n_x}$

Este remuestreo puede interpretarse como una perturbación aleatoria controlada alrededor de cada partícula remuestreada.

El objetivo no es modificar significativamente la posición de las partículas, sino restaurar

la diversidad perdida durante el remuestreo manteniendo la estructura general de la distribución posterior.

La regularización permite preservar la diversidad de partículas después de cada re-muestreo, reduciendo la degeneración de la muestra y mejorando la estabilidad de las estimaciones obtenidas para las variables macroeconómicas observadas. A continuación se presenta algoritmo del filtro de partículas.

Algoritmo 2 Filtro de Partículas Regularizado

Inicializar: para $i = 1, \ldots, N$, muestrear $\boldsymbol{X_0}^{(i)} \sim p(\boldsymbol{X_0})$, establecer $w_0^{(i)} = \frac{1}{N}$
para t = 1, 2, ..., T **hacer**
  para cada $i = 1, ..., N$ **hacer**
    Propagar $\boldsymbol{X_t}^{(i)} \sim p(\boldsymbol{X_t}|\boldsymbol{X_{t-1}}^{(i)})$
    Calcular peso no normalizado $\tilde{w}_t^{(i)} = p(\boldsymbol{Y_t}|\boldsymbol{X_t}^{(i)})$
  fin para
  Normalizar pesos: $w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^{N} \tilde{w}_t^{(j)}}$
  Evaluar NEF $= \frac{1}{\sum_{i=1}^{N} (w_t^{(i)})^2}$
  si NEF $<$ umbral **entonces**
    Re-muestrear $\tilde{\boldsymbol{X}}_t^{(1)}, ..., \tilde{\boldsymbol{X}}_t^{(N)} \sim$ Multinomial$(w_t^{(i)})$
    Estimar la matriz de covarianza empírica $\hat{\boldsymbol{\Sigma}}_t$ de $\{\tilde{\boldsymbol{X}}_t^{(i)}\}$
    para cada $i = 1, ..., N$ **hacer**
      Generar partícula regularizada:

$$\boldsymbol{X_t}^{(i)} \sim \mathcal{N}(\tilde{\boldsymbol{X_t}}^{(i)}, h^2 \hat{\boldsymbol{\Sigma}}_t), \ h \textit{ parmetro de suavizado}$$

    fin para
    Reiniciar pesos: $w_t^{(i)} = \frac{1}{N}$
  si no
    Conservar partículas originales $\boldsymbol{X_t}^{(i)}$
  fin si
fin para

Sin embargo, según [26] que en ciertas circunstancias el uso de kernels de suavizado estándar en filtros de partículas incrementa la varianza del sistema, degradando la precisión de estimación y predicción de los estados, esto puede suceder cuando los vectores de estados suelen ser de alta dimensión, por lo cual proponen técnicas de contracción donde el ancho de banda está determinado por un factor de descuento

$$h^2 = 1 - ((3\beta - 1)/2\beta)^2$$

Estos factores de descuentos suelen estar entre $(0.95, 0.99]$, lo cual implicaría que $h \in [0.1, 0.23)$.

3.3. Predicción a Partir de Filtros de Partículas

Una vez obtenida la distribución predictiva de una variable de interés mediante el uso de filtros de partículas, es posible derivar tanto predicciones puntuales como medidas de incertidumbre asociadas. En el enfoque bayesiano, la predicción se describe completamente a través de la distribución:

$$p(\boldsymbol{X_{t+1}} \mid \boldsymbol{X_{1:t}}),$$

la cual es aproximada mediante un conjunto de partículas $\{\boldsymbol{X_{t+1}}^{(i)}\}_{i=1}^{N}$.

En este contexto, dos opciones principales para resumir la distribución predictiva se presentan: el uso de la media (esperanza matemática) y el uso de la mediana, esta última asociada a la teoría de estimación robusta.

3.3.1. Predicción Mediante la Media

Supongamos que se generó un conjunto de partículas $\{\boldsymbol{x_{t+1}}^{(i)}\}_{i=1}^{N}$ mediante el filtro de partículas. La predicción puntual basada en la media se calcula como:

$$\hat{\boldsymbol{X}}_{t+1}^{\text{media}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x_{t+1}}^{(i)}.$$

La elección de este estimador está sustentada en su propiedad óptima bajo la función de pérdida cuadrática, ya que la media minimiza el error cuadrático medio esperado:

$$\hat{\boldsymbol{x}} = \arg\min_{a} \mathbb{E}[(\boldsymbol{X} - \boldsymbol{a})^2].$$

No obstante, la distribución predictiva generada por filtros de partículas puede presentar asimetrías, colas pesadas o estructuras multimodales, lo que puede afectar la estabilidad de la media como estimador puntual.

Adicionalmente, la misma muestra de partículas permite cuantificar la incertidumbre de la predicción mediante intervalos de credibilidad. Por ejemplo, un intervalo de credibilidad del 95 % puede obtenerse a partir de los cuantiles de las partículas:

$$[\hat{q}_{0.025},\ \hat{q}_{0.975}]\,,$$

donde $\hat{q}_\alpha$ representa el cuantil α de la distribución de $\{\boldsymbol{X_{t+1}}^{(i)}\}_{i=1}^{N}$.

Ahora bien, con las dificultades sobre estimar las predicciones con la media, según lo mencionado anteriormente, fue lo que motivó a la teoría de estimación robusta como se detalla en el capítulo 2, donde se explica la introducción de otros estimadores la cual nos ayuda a evitar a que el estimador sea influenciado por las distribuciones asimétricas, colas pesadas o valores atípicos, teniendo estimaciones más exactas a las reales.

3.3.2. Predicción Robusta Mediante la Mediana

Para mejorar la estimación en presencia de valores atípicos o distribuciones asimétricas, se propone utilizar la mediana como estimador robusto. La mediana minimiza la función de pérdida absoluta esperada:

$$\hat{\boldsymbol{X}}_{t+1}^{\text{mediana}} = \arg\min_{a}\ \mathbb{E}[|\boldsymbol{X} - \boldsymbol{a}|].$$

Esta propiedad convierte a la mediana en un estimador más estable frente a valores extremos, siendo especialmente útil en contextos no lineales y no gaussianos.
Aplicado al conjunto de partículas predictivas, la mediana se calcula como:

$$\hat{\boldsymbol{X}}_{t+1}^{\text{mediana}} = \text{mediana}\left(\{\boldsymbol{X_{t+1}}^{(i)}\}_{i=1}^{N}\right).$$

De manera análoga al caso de la media, la mediana puede acompañarse de intervalos de credibilidad construidos a partir de los cuantiles de la distribución empírica de las partículas, permitiendo así una descripción completa tanto del valor central como de la dispersión de la predicción.

En este capítulo se han presentado los fundamentos y conceptos teóricos necesarios para modelar fenómenos dinámicos y estocásticos mediante estructuras en espacio de estados.

La incorporación de inferencia bayesiana dentro de estos modelos es ventaja crucial. Esta inferencia se aplica en dos niveles fundamentales:

1. **Estados dinámicos:** aplicando filtros de partículas para aproximar la posterior de estados latentes $p(\boldsymbol{X_t} \mid \boldsymbol{Y_{1:t}})$.

2. **Parámetros fijos:** parámetros que han sido calibrados en base a la literatura, utilizando valores estándar ampliamente aceptados en estudios similares.

Ahora bien, primeramente es importante realizar una validación de los métodos a implementar para garantizar que el método funciona, lo cual veremos en la siguiente sección.

3.4. Validación del método de Filtro de Partículas

Antes de aplicar el Filtro de Partículas al modelo DSGE desarrollado, resulta conveniente ilustrar su funcionamiento mediante un modelo de espacio de estados no lineal. Este ejemplo permite evaluar la capacidad del algoritmo para estimar y predecir estados en presencia de no linealidades, así como analizar el efecto del número de partículas sobre la calidad de las estimaciones.

Sea el siguiente modelo de espacio de estados

$$y_{1,t} = \frac{\theta_t^2}{20} + \epsilon_{1,t}, \quad y_{2,t} = \sin(\theta_t) + \epsilon_{2,t} \quad \epsilon_{1,t}, \epsilon_{2,t} \sim N(0,1) \tag{3.20}$$

$$\theta_t = \frac{\theta_{t-1}}{2} + \frac{25\theta_{t-1}}{1+\theta_{t-1}^2} + 8\cos(1.2t) + \epsilon_t, \quad \epsilon_t \sim N(0,1) \tag{3.21}$$

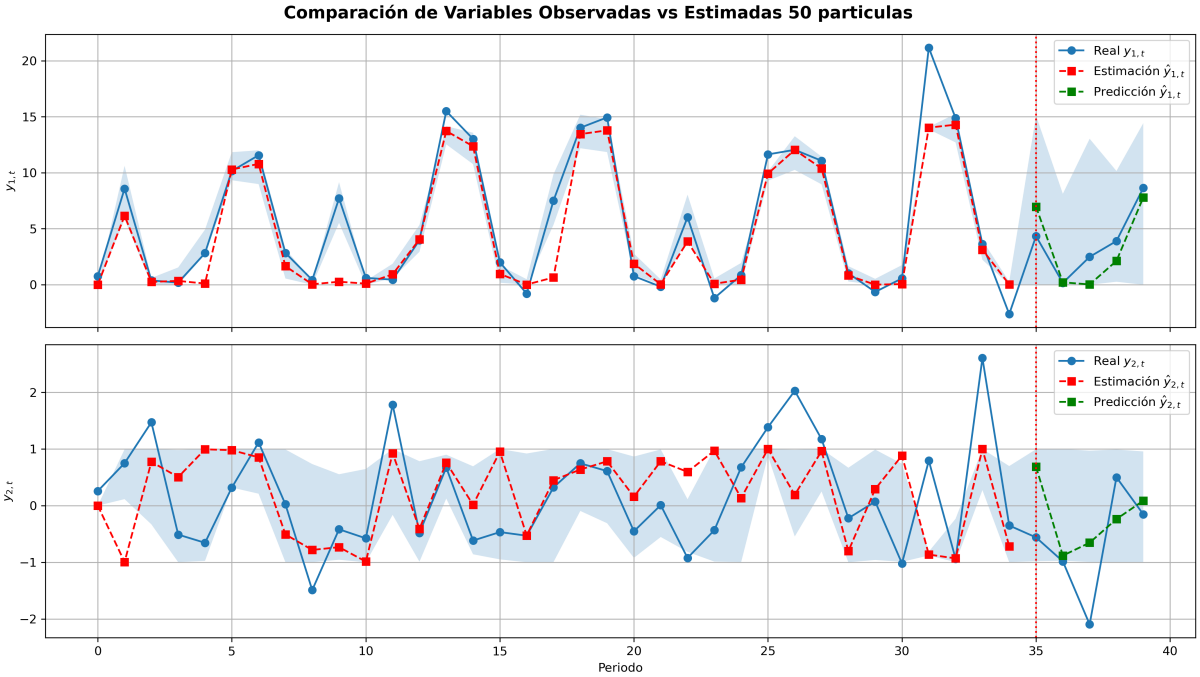


Figura 3.5: Comparación de variables observadas vs. estimadas con $N = 50$.

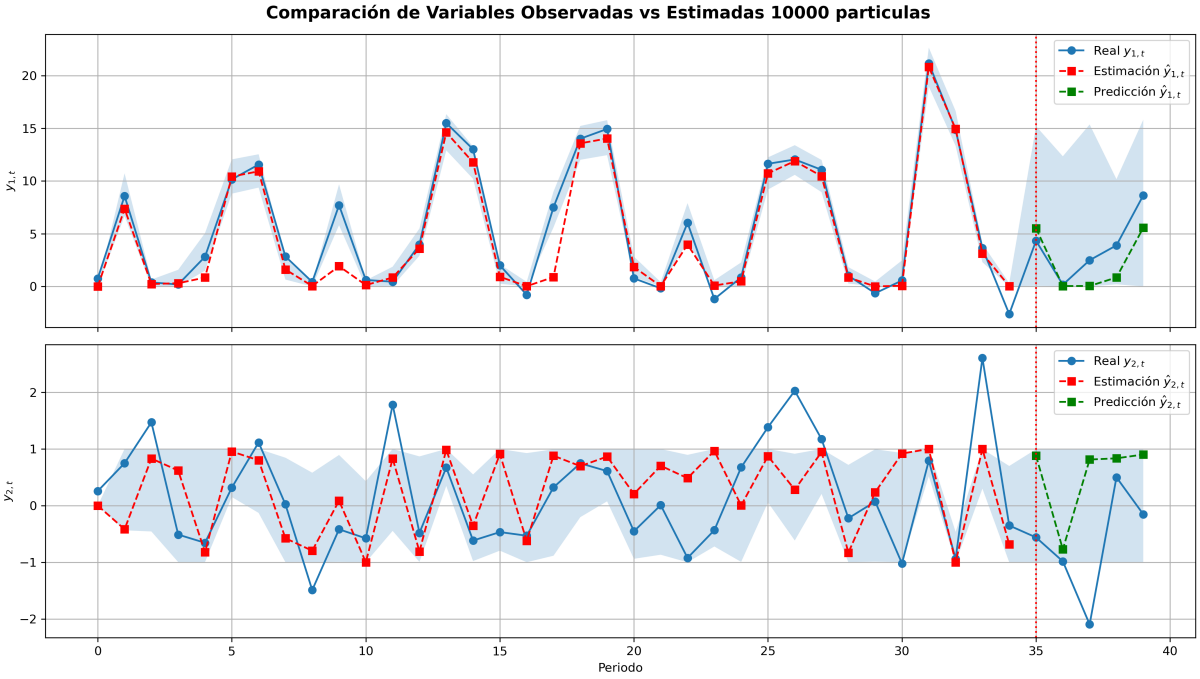


Figura 3.6: Comparación de variables observadas vs. estimadas con $N = 10000$.

N	50	100	2000	4000	6000	8000	10000
$y_{1,t}$	1.839	2.099	2.1919	2.2419	2.2655	2.236	2.286
$y_{2,t}$	0.9215	1.118	1.5255	1.545	1.5427	1.533	1.537

Tabla 3.2: RMSE de las variables estimadas

N	**50**	**100**	**2000**	**4000**	**6000**	**8000**	**10000**
elpd estimación	-119.472	-91.340	-91.518	-91.911	-91.669	-91.746	-92.218
elpd predicción	-20.615	-20.146	-20.467	-20.535	-20.590	-20.470	-20.490

Tabla 3.3: Expected Log Predictive Density (ELPD) para estimación y predicción fuera de muestra según el número de partículas utilizadas.

Los resultados obtenidos muestran que el Filtro de Partículas permite aproximar la dinámica del sistema no lineal propuesto. Sin embargo, los valores de RMSE presenta una tendencia diferente a la esperada respecto al incremento del número de partículas, pero a partir de cierto umbral se visualiza fluctuaciones relativamente pequeñas, lo cual sugiere cierta estabilidad en las estimaciones obtenidas, donde el error puede venir principalmente de la naturaleza estocástica y no lineal del sistema, por lo cual es necesario medir el desempeño del modelo mediante el ELPD.

Los resultados del Expected Log Predictive Density (ELPD) muestran una mejora significativa al pasar de un número reducido de partículas a configuraciones más grandes. Posteriormente, el ELPD se estabiliza, lo cual sugiere convergencia de la densidad predictiva aproximada. Este comportamiento confirma que el Filtro de Partículas logra construir distribuciones predictivas aceptables con los datos observados.

En términos predictivos, el ELPD fuera de muestra permanece relativamente estable para distintos tamaños de partículas, lo cual indica robustez en la capacidad predictiva del modelo.

En conjunto, estos resultados validan empíricamente el uso del Filtro de Partículas como herramienta adecuada para la estimación de modelos dinámicos no lineales y no gaussianos, especialmente en escenarios donde métodos lineales tradicionales, como el Filtro de Kalman estándar, no son apropiados.

3.4.1. Consideraciones Finales

Los modelos en espacio de estados proporcionan un marco general para representar sistemas dinámicos con variables latentes y observaciones ruidosas.

Sin embargo, cuando las ecuaciones que describen la dinámica del sistema y el proceso de observación son no lineales o no gaussianas, como ocurre en muchos modelos económicos

modernos, los métodos analíticos clásicos dejan de ser aplicables. En particular, el filtro de Kalman estándar no puede utilizarse directamente, lo que motiva el uso de técnicas numéricas más generales.

Esta situación es especialmente relevante en los modelos Dinámicos Estocásticos de Equilibrio General (siglas en inglés DSGE), donde las relaciones económicas derivadas de la teoría microfundamentada suelen implicar estructuras altamente no lineales y la presencia de múltiples fuentes de incertidumbre. En consecuencia, la estimación y predicción en estos modelos requieren herramientas capaces de aproximar distribuciones completas en contextos complejos.

En este sentido, los filtros de partículas se presentan como una alternativa, ya que permiten aproximar de forma flexible las distribuciones posteriores de los estados latentes sin imponer restricciones de linealidad o normalidad. Por lo tanto, el siguiente capítulo introduce los modelos DSGE y su formulación, sentando las bases para su posterior estimación mediante filtros de partículas.

Capítulo 4

Modelos DSGE

Este capítulo primeramente presenta la formulación teórica de los modelos dinámicos estocásticos de equilibrio general (DSGE) y desarrolla el modelo específico que servirá como base para la aplicación de filtros de partículas en capítulos posteriores, posteriormente se detalla la aplicación al modelo DSGE inspirado por [3] , representandolo como un modelo de espacio de estados para posteriormente aplicar el filtro de partículas.

Los modelos DSGE combinan tres elementos fundamentales: dinámica intertemporal, incertidumbre estocástica y equilibrio general. La dimensión dinámica surge porque las decisiones actuales de los agentes afectan variables futuras, como el capital, el consumo o la inflación. La dimensión estocástica aparece mediante shocks exógenos que perturban la economía. Finalmente, el equilibrio general exige que las decisiones de los agentes sean compatibles entre sí y con las restricciones agregadas del sistema

En esta tesis se utiliza una versión simplificada de un modelo DSGE no lineal, construido a partir del problema de optimización de un agente representativo (el problema del consumidor), una restricción de recursos y un proceso exógeno para la productividad. Esta formulación permite identificar con claridad las variables de estado, las variables de control, los parámetros estructurales y los shocks del modelo.

Según [27], los modelos DSGE han alcanzado gran popularidad tanto en la academia como en instituciones de política económica por su capacidad para ser estimados, comparados empíricamente y utilizados para pronóstico y simulación de políticas. Su representación como sistemas dinámicos en espacio de estado los hace adecuados para ser resueltos mediante técnicas de inferencia bayesiana, como el Filtro de Partículas, el cual fue introducido en el capítulo 3.

4.1. Estructura General de un Modelo DSGE

Los modelos DSGE se caracterizan por describir el comportamiento agregado de una economía mediante un conjunto de ecuaciones que representan las decisiones óptimas de los agentes económicos —hogares, empresas, gobierno y banco central— sujetos a restricciones presupuestarias, condiciones de mercado y evolución estocástica de los shocks. Estos modelos parten de fundamentos microeconómicos y son formulados en términos dinámicos, estocásticos y de equilibrio general.

Siguiendo la formulación de [28], un modelo DSGE puede escribirse en términos de tres componentes clave:

1. Un vector de variables de estado $\boldsymbol{x_t}$, que incluye tanto variables endógenas (como el capital acumulado) como exógenas (como shocks tecnológicos);

2. Un vector de variables de control $\boldsymbol{y_t}$, que representa las decisiones de los agentes (como consumo, inversión, trabajo);

3. Un conjunto de ecuaciones de equilibrio que relacionan la dinámica del modelo en estas variables:

$$E_t \left[f(\boldsymbol{x_{t+1}}, \boldsymbol{y_{t+1}}, \boldsymbol{x_t}, \boldsymbol{y_t}, \boldsymbol{\varepsilon_{t+1}} | \boldsymbol{\theta}) \right] = 0 \tag{4.1}$$

$$\boldsymbol{x_{t+1}} = g(\boldsymbol{x_t}, \boldsymbol{y_t}, \boldsymbol{\varepsilon_{t+1}} | \boldsymbol{\theta}) \tag{4.2}$$

donde:

1. El operador E_t denota la esperanza condicional dada la información disponible en el tiempo t, consistente con el enfoque de esperanzas racionales donde los agentes optimizan usando toda la información presente.

2. $\boldsymbol{\theta}$ representa el vector de parámetros estructurales del modelo. La dimensión de $\boldsymbol{\theta}$ depende del tipo de modelo DSGE, y también de los shocks estructurales del modelo.;

3. $\boldsymbol{\varepsilon_{t+1}}$ son los shocks estructurales, típicamente modelados como procesos estocásticos;

4. La ecuación (4.1) incorpora las condiciones de primer orden derivadas del problema de optimización de los agentes;

5. La ecuación (4.2) define la ley de evolución de las variables de estado.

donde las variables endógenas son aquellas determinadas dentro del modelo. Por ejemplo, el consumo, el capital, la producción o la inflación dependen de las decisiones de los agentes y de las condiciones de equilibrio. En cambio, las variables exógenas son aquellas cuya evolución no se determina dentro del sistema económico, sino que se especifica mediante un proceso estocástico externo, como por ejemplo un shock tecnológico o de productividad.

Dentro de las variables endógenas se distingue entre variables de estado y variables de control. Las variables de estado resumen la información relevante heredada del pasado y necesaria para describir la evolución futura del sistema. El capital acumulado es un ejemplo de variable de estado, ya que su valor actual depende de decisiones de inversión realizadas en periodos anteriores. Las variables de control, por su parte, son aquellas que los agentes eligen en cada periodo para optimizar su función objetivo. En el problema del consumidor, el consumo y el capital futuro pueden interpretarse como variables de control.

Los shocks exógenos en modelos macroeconómicos son perturbaciones inesperadas y externos al sistema modelado (como crisis financieras, desastres naturales, cambios bruscos en precios, cambios de tecnologías, preferencias, productividad) que alteran drásticamente variables endógenas como el PIB, la inflación y el empleo

Una vez teniendo la formulación del modelo, el próximo paso en la teoría de inferencia bayesiana consiste en definir una función de verosimilitud $p(\boldsymbol{y_{1:t}} \mid \boldsymbol{x_t})$, y combinarla con una distribución a priori $p(\boldsymbol{x_t})$, obteniendo la distribución a posteriori:

$$p(\boldsymbol{x_t} \mid \boldsymbol{y_{1:t}}) \propto p(\boldsymbol{y_{1:t}} \mid \boldsymbol{x_t}) p(\boldsymbol{x_t}) \tag{4.3}$$

Este enfoque permite aplicar técnicas como MCMC y Filtros de Partículas para la estimación y simulación de modelos DSGE, incluso en contextos no lineales o no gaussianos [29].

Una característica importante de esta formulación es que la solución del modelo implica encontrar una ley de movimiento para las variables endógenas que sea consistente con el

equilibrio intertemporal bajo esperanzas racionales. En contextos empíricos, esta forma facilita la conexión con los datos observados y podemos formular el modelo como un sistema de espacio de estado.

Esta formulación general proporciona un marco flexible y riguroso para estudiar el impacto de políticas económicas, shocks estructurales y dinámicas agregadas, y será el punto de partida conceptual para las aplicaciones desarrolladas en capítulos posteriores. La figura 4.1 sintetiza la convergencia de los principales conceptos teóricos desarrollados.

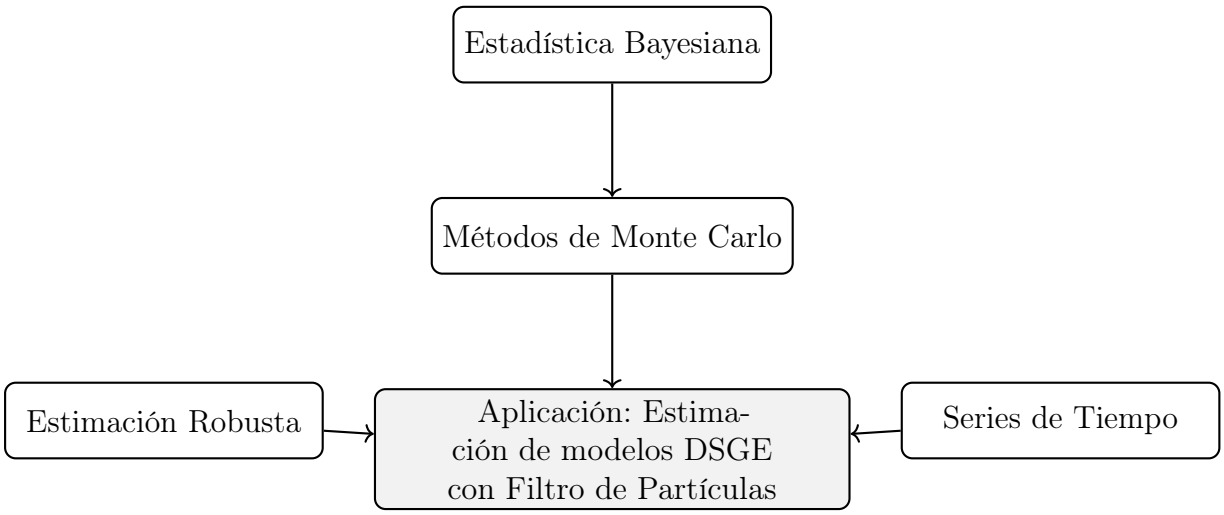


Figura 4.1: Relación entre los conceptos teóricos y la aplicación en modelos DSGE.

Es importante destacar que los modelos DSGE han logrado tener éxito en la macronomía moderna, como ser en los bancos centrales, donde en [30] han utilizado este modelo en el banco central europeo, que se logró analizar las crisis económica postpandemia, también en [31] que se utilizó el modelo para entender el comportamiento de inflación y el PIB durante el COVID-19. Asimismo, en [32] mediante los modelos DSGE se logró interpretar como los cambios físicos como los eventos climáticos críticos o agudos afectan a la oferta y a la demanda, también para modelar los impactos de transición, permitiendo evaluar la trayectoria de los impuestos al carbono mediante procesos estocásticos.

4.1.1. Linealización y Métodos Clásicos de Solución

La resolución de modelos DSGE en su forma normal podría ser compleja por la presencia de no linealidades en las funciones de producción, utilidad y restricciones dinámicas.

En contextos donde se busca una solución analítica o computacional más manejable, una práctica común es aplicar una aproximación de primer orden alrededor del estado estacionario, conocida como *log-linealización*. Este procedimiento transforma las ecuaciones no lineales en relaciones lineales entre desviaciones logarítmicas de las variables con respecto a sus valores estacionarios.

El valor a_s corresponde al estado estacionario de la variable a_t, definido como el nivel de equilibrio de largo plazo al cual converge en ausencia de shocks y donde la economía opera de forma constante.

Formalmente, si a_t es una variable con estado estacionario a_s (un estado estacionario de una variable se define como un valor constante que deja de cambiar en el tiempo $a_t = a_{t+1} = a_s$), se define $\hat{a}_t = \log a_t - \log a_s$. Bajo esta transformación, las funciones exponenciales y productos se aproximan utilizando series de Taylor, por ejemplo:

$$e^{\hat{a}_t + r\hat{b}_t} \approx 1 + \hat{a}_t + r\hat{b}_t, \quad \hat{a}_t\hat{b}_t \approx 0. \tag{4.4}$$

Con estas aproximaciones logramos representar todo el modelo como un sistema lineal en diferencias, lo cual permite la aplicación de métodos clásicos de solución.

Una vez linealizado, el sistema DSGE toma la forma de ecuaciones diferenciales estocásticas lineales con esperanzas racionales. Como lo menciona [33], existen tres enfoques ampliamente utilizados para resolver:

1. El método de Blanchard-Kahn: establece condiciones de existencia y unicidad de solución.

2. El método de Klein: utiliza la descomposición generalizada de Schur.

3. El método de los coeficientes indeterminados: determina sus coeficientes resolviendo el sistema algebraico resultante.

Cada uno de estos métodos impone requisitos computacionales particulares, pero comparten una limitación común: la necesidad de aproximar el sistema original, lo cual puede inducir errores relevantes si las no linealidades son sustanciales, así como se visualiza en el ejemplo presentado en el capítulo 3.

Estos métodos pueden ocultar propiedades clave del modelo al eliminar interacciones no lineales o efectos de segundo orden. Por esta razón, este trabajo adopta un enfoque completamente no lineal, evitando la linealización del modelo.

Una alternativa sólida es el uso del *Filtro de Partículas*, tal como se introdujo en el capítulo anterior (véase capítulo 3) es una técnica de inferencia bayesiana que permite estimar modelos no lineales y no gaussianos sin necesidad de aproximaciones locales.

4.2. El Problema del Consumidor

Con el objetivo de introducir la metodología de resolución de los modelos DSGE, partimos de un modelo clásico importante: el modelo de crecimiento neoclásico estocástico. Este modelo permite entender cómo las decisiones intertemporales de los agentes se ven afectadas por shocks agregados, y sirve como base para estructuras más complejas.

Se asume que existe un agente representativo en la economía, quien toma decisiones sobre consumo c_t y acumulación de capital k_{t+1} con el fin de maximizar su utilidad intertemporal esperada, de la siguiente forma:

$$\max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{4.5}$$

donde: $\beta \in (0,1)$ representa el factor de descuento, y $u(\cdot)$ es una función de utilidad que cumple con las propiedades estándar: es creciente, estrictamente cóncava y dos veces diferenciable.

El problema anterior está sujeto a las siguientes restricciones económicas que describen el entorno productivo y tecnológico:

1. **Función de Producción:** la economía produce productos mediante capital, según la tecnología $y_t = e^{z_t} f(k_t)$, donde z_t es un shock de productividad.

2. **Evolución del capital:** la acumulación de capital está dada por $k_{t+1} = i_t + (1-\delta)k_t$, con $\delta \in (0,1)$ como la tasa de depreciación e i_t la inversión neta.

3. **Shock tecnológico:** se modela como un proceso autorregresivo de primer orden (AR(1)) $z_t = \rho z_{t-1} + \varepsilon_t$, con $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, y ρ mide la persistencia del shock.

4. **Identidad de recursos:** la producción se destina al consumo y la inversión: $c_t + i_t = y_t$.

En este contexto, la función de producción $f(\cdot)$ se asume con propiedades neoclásicas estándar: es continua, creciente y cóncava en k_t.

El conjunto de parámetros que rige la dinámica del modelo se resume en el vector estructural $\theta = (\beta, \delta, \rho, \sigma_\varepsilon)'$, y se asume que las condiciones iniciales k_0 y z_0 son conocidas. Bajo este entorno, el agente toma decisiones racionales optimizando su bienestar ante los shocks estocásticos.

4.2.1. Variables de Estado y de Control

En el desarrollo de modelos DSGE, resulta útil clasificar las variables económicas en dos conjuntos principales: *variables de estado* y *variables de control*. Esta clasificación permite organizar la dinámica del sistema y facilita la solución analítica o computacional del modelo.

Las variables de estado son aquellos valores donde en el periodo actual depende de decisiones tomadas en el pasado o de procesos estocásticos que evolucionan exógenamente. Por el contrario, las variables de control son determinadas activamente por los agentes económicos en cada periodo, con el objetivo de maximizar alguna función objetivo (como utilidad o beneficio).

En el modelo planteado, el *capital* k_t es una variable de estado endógena, ya que su nivel en el tiempo actual está determinado por decisiones de inversión realizadas en el pasado. Asimismo, el *shock tecnológico* z_t se considera una variable de estado exógena, gobernada por un proceso autorregresivo. En conjunto, el vector de estado queda definido como $[k_t, z_t]'$.

Una de las opciones comunes con respecto a las variables de control es tomar el consumo c_t como la variable de control, lo que implica que la inversión i_t y el capital futuro k_{t+1} se determinan mediante las identidades del modelo.

Otra opción es escoger directamente k_{t+1} como variable de control. En ese caso, utilizando la restricción de acumulación de capital y la identidad de recursos, es posible

expresar el consumo como:

$$c_t = e^{z_t} f(k_t) + (1-\delta)k_t - k_{t+1} \tag{4.6}$$

donde el consumidor elige una secuencia no negativa $\{c_t\}$ que maximiza la utilidad esperada $c_t + i_t \leq y_t$, y el capital no puede ser menor que cero $k_t \geq 0$ [33].

Sustituyendo esta expresión en la función objetivo original, el problema de maximización puede reescribirse como:

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u\left(e^{z_t} f(k_t) + (1-\delta)k_t - k_{t+1}\right) \tag{4.7}$$

Una tercera alternativa consiste en considerar simultáneamente a c_t y k_{t+1} como variables de control, lo cual resulta útil en enfoques computacionales donde se trabaja con un vector de decisiones.

En resumen, en este modelo contamos con:

1. Dos variables de estado: k_t (endógena) y z_t (exógena),
2. Dos variables de control: c_t y k_{t+1},
3. Una variable observable: y_t,
4. Y un conjunto de parámetros estructurales: $\boldsymbol{\theta} = (\beta, \delta, \rho, \sigma_\varepsilon)'$.

Esta clasificación servirá como base para derivar las condiciones de primer orden del modelo, construir su solución dinámica, y formular su representación en espacio de estado en las secciones siguientes.

Desde el punto de vista computacional, identificar claramente el conjunto de variables de estado y de control es esencial para aplicar métodos de resolución numérica. Esta distinción no solo facilita la simulación del modelo, sino también su formulación en espacio de estado, lo cual será fundamental para la posterior aplicación de técnicas de inferencia como el Filtro de Partículas o el Filtro de Kalman, dependiendo del grado de linealidad asumido [34].

4.3. Condiciones de Equilibrio

Dado que el problema dinámico del consumidor fue formulado con dos variables de control, c_t y k_{t+1}, es necesario plantearlo como un problema de optimización con restricciones.

$$\max_{\{c_t,k_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{4.8}$$

s.a.

$$c_t + k_{t+1} = e^{z_t} f(k_t) + (1-\delta)k_t \tag{4.9}$$

Los multiplicadores de Lagrange junto a la función Lagrangiana es una técnica esencial para optimizar una funcion objetivo sujeta a una o mas restricciones (para verlo con mas profundidad ver el apéndice A).

La condición de equilibrio se obtiene al construir la función Lagrangiana del problema:

$$\mathcal{L} = E_0 \sum_{t=0}^{\infty} \beta^t \left[u(c_t) + \lambda_t \left(e^{z_t} f(k_t) + (1-\delta)k_t - c_t - k_{t+1}\right)\right] \tag{4.10}$$

Las condiciones de primer orden se obtienen derivando $\mathcal{L}$ respecto a las variables de control y al multiplicador:

1. Derivada respecto a λ_t (restricción de recursos):

$$\frac{\partial \mathcal{L}}{\partial \lambda_t} = \beta^t \left[e^{z_t} f(k_t) + (1-\delta)k_t - c_t - k_{t+1}\right] = 0 \tag{4.11}$$

2. Derivada respecto a c_t:

$$\frac{\partial \mathcal{L}}{\partial c_t} = \beta^t \left[u'(c_t) - \lambda_t\right] = 0 \quad \Rightarrow \quad \lambda_t = u'(c_t) \tag{4.12}$$

3. Derivada respecto a k_{t+1} (con esperanza condicional):

$$\frac{\partial \mathcal{L}}{\partial k_{t+1}} = -\beta^t \lambda_t + \beta^{t+1} \mathbb{E}_t \left[\lambda_{t+1} \left(e^{z_{t+1}} f'(k_{t+1}) + 1 - \delta\right)\right] = 0 \tag{4.13}$$

Sustituyendo la expresión de λ_t, se obtiene la condición de Euler:

$$u'(c_t) = \beta\, \mathbb{E}_t \left[u'(c_{t+1}) \left(e^{z_{t+1}} f'(k_{t+1}) + 1 - \delta\right)\right] \tag{4.14}$$

Esta última ecuación captura el comportamiento óptimo intertemporal del agente representativo y constituye una de las ecuaciones fundamentales del modelo. Junto con la restricción de recursos y la ley de evolución del shock exógeno, se conforma el sistema dinámico completo que será resuelto y estimado en secciones posteriores.

Estas condiciones de primer orden, derivadas del problema de optimización intertemporal del agente representativo, constituyen los microfundamentos del modelo. Además, cuando se incorporan fricciones nominales o reales, estas condiciones explican parte de la persistencia observada empíricamente en variables como el producto, el consumo y la inflación [35].

En la figura 4.2, se puede visualizar una idea general de un modelo DSGE.

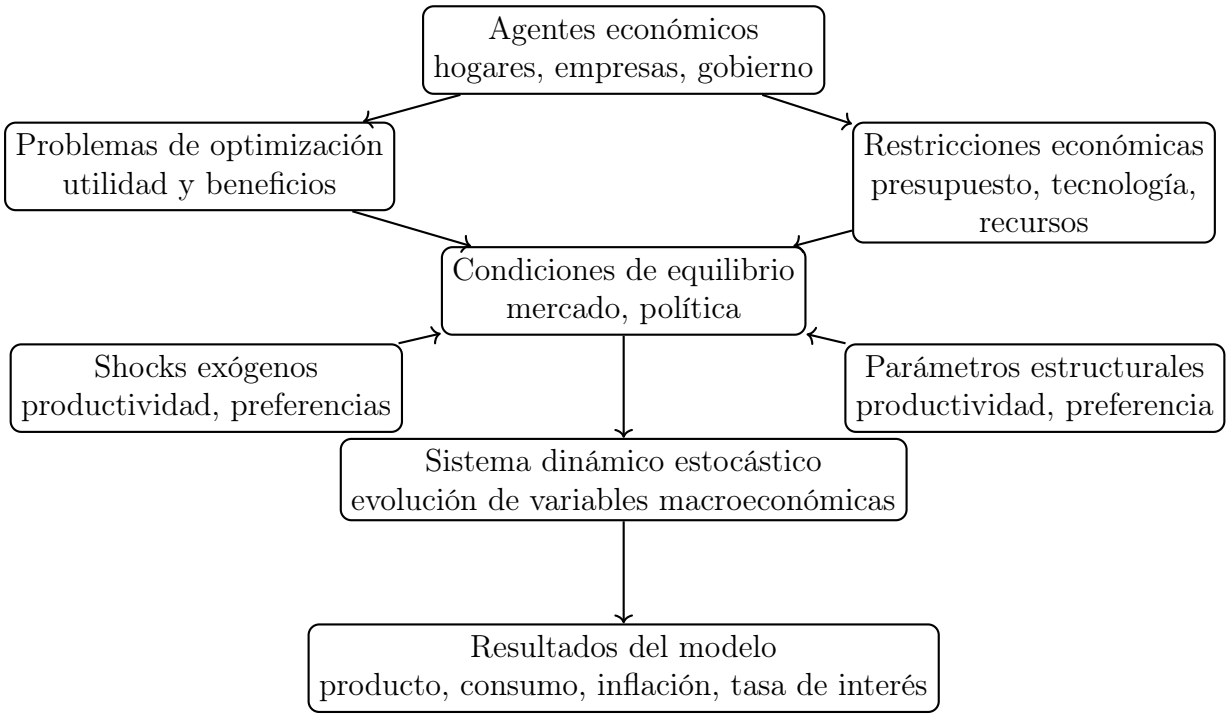


Figura 4.2: Estructura general de un modelo DSGE.

4.4. Estado Estacionario

El estado estacionario representa el equilibrio de largo plazo del sistema económico, donde las variables no cambian con el tiempo en ausencia de perturbaciones. Calcular este estado es fundamental, ya que permite caracterizar el comportamiento estructural del modelo y sirve como punto de partida para la simulación y la estimación de los parámetros.

Para determinar el estado estacionario, se asume que no hay shocks tecnológicos, es decir, $\sigma_\varepsilon = 0$. En este caso, el proceso que describe la evolución del shock se reduce a:

$$z_t = \rho z_{t-1} \Rightarrow z_t = \rho^t z_0$$

Como $|\rho| < 1$, se tiene que $\lim_{t\to\infty} z_t = 0$, y por lo tanto $e^{z_t} \to 1$. Sustituyendo esta condición en las ecuaciones del modelo, se obtiene el sistema en estado estacionario:

$$y_s = f(k_s) \tag{4.15}$$

$$c_s + k_s = f(k_s) + (1-\delta)k_s \tag{4.16}$$

$$c_s = f(k_s) - \delta k_s \tag{4.17}$$

Finalmente, reemplazando el valor de c_s en la condición de equilibrio intertemporal se obtiene:

$$1 = \beta\left(f'(k_s) + 1 - \delta\right)$$

Esta ecuación permite encontrar el valor del capital en estado estacionario k_s, que depende de los parámetros estructurales del modelo. Una vez determinado k_s, es posible obtener los valores correspondientes de y_s y c_s para describir el equilibrio de largo plazo de la economía.

El cálculo del estado estacionario no solo proporciona valores de referencia para las variables del modelo, sino que también es un punto central para entender la dinámica del sistema en ausencia de perturbaciones. Además, en enfoques computacionales, el estado estacionario sirve como insumo fundamental para algoritmos de simulación y estimación, incluyendo aquellos que no requieren linealización, como los métodos de Monte Carlo Secuencial. Aunque en este trabajo no se aplicarán aproximaciones log-lineales alrededor del estado estacionario, su determinación sigue siendo indispensable para evaluar trayectorias de equilibrio de largo plazo y comparar escenarios de política económica bajo distintos supuestos estructurales.

4.5. Calibración y estimación de Parámetros

En la literatura de modelos DSGE existen dos enfoques principales para asignar valores a los parámetros estructurales: la calibración y la estimación o ambas. La estimación bayesiana permite inferir parámetros a partir de los datos mediante distribuciones posteriores, la calibración consiste en fijar determinados parámetros utilizando evidencia previa proveniente de estudios empíricos relacionados.

De acuerdo con [36], la calibración constituye una estrategia ampliamente utilizada en modelos DSGE debido a las dificultades asociadas con la identificación conjunta de todos los parámetros estructurales. En particular, los modelos DSGE suelen involucrar sistemas altamente no lineales, esperanzas racionales y un gran número de parámetros, lo cual puede generar problemas de identificación y convergencia durante la estimación. Por esta razón, numerosos trabajos optan por calibrar ciertos parámetros y estimar únicamente aquellos relacionados con la dinámica estocástica del sistema.

En esta tesis no se realizó una estimación conjunta de parámetros. En consecuencia, los parámetros estructurales fueron calibrados utilizando información proveniente de la literatura especializada en modelos DSGE. Esta estrategia permitió reducir la complejidad computacional asociada con la estimación bayesiana completa y concentrar el análisis en la dinámica no lineal del sistema y en el desempeño predictivo del Filtro de Partículas Regularizado.

Siguiendo la clasificación propuesta por [36], los parámetros utilizados pueden dividirse en tres grupos principales:

1. **Parámetros tomados de la literatura:** corresponden a parámetros ampliamente documentados en estudios DSGE previos, especialmente en modelos tipo RBC y New Keynesian. Estos parámetros suelen asociarse con preferencias, descuento intertemporal y persistencia de choques.

2. **Parámetros calibrados mediante relaciones de estado estacionario:** algunos parámetros pueden ser determinados utilizando restricciones de equilibrio de largo plazo derivadas del propio modelo, garantizando coherencia macroeconómica entre

consumo, producción e inflación.

3. **Parámetros asociados a procesos estocásticos:** incluyen principalmente las desviaciones estándar y persistencias de los choques exógenos, las cuales pueden ser ajustadas buscando estabilidad numérica y capacidad predictiva adecuada.

La calibración presenta varias ventajas dentro del contexto DSGE. Primero, permite incorporar conocimiento económico previo obtenido de investigaciones anteriores. Segundo, evita estimaciones económicamente inconsistentes que pueden surgir cuando la función de verosimilitud presenta múltiples máximos locales o problemas de identificación. Finalmente, reduce considerablemente el costo computacional, especialmente en modelos no lineales cuya solución requiere métodos numéricos intensivos.

No obstante, la calibración también posee limitaciones importantes. Como señalan [36], los resultados obtenidos dependen directamente de los valores seleccionados para los parámetros, por lo que una mala calibración puede afectar significativamente la dinámica del modelo y sus capacidades predictivas. Además, este enfoque no incorpora explícitamente la incertidumbre asociada a los parámetros estructurales, a diferencia de la estimación bayesiana completa.

En esta tesis se adoptó una estrategia híbrida con la literatura DSGE: calibración de parámetros estructurales junto con estimación secuencial de estados mediante el Filtro de Partículas Regularizado. Esta combinación permitió preservar la no linealidad del modelo, reducir la carga computacional y mantener estabilidad en el proceso de inferencia.

4.6. Aplicación

En esta sección se desarrolla e implementa un modelo DSGE no lineal basado en la estructura propuesta por [3]. A diferencia de los enfoques tradicionales que recurren a aproximaciones log-lineales alrededor del estado estacionario [37, 38, 39], aqui se preserva la forma no lineal original del sistema con el objetivo de capturar de manera más precisa la dinámica estocástica de la economía. La decisión de mantener la no linealidad original responde al objetivo de capturar de forma más precisa los efectos no lineales del modelo. [2, 40].

El modelo incorpora hogares representativos, firmas con rigideces nominales y una autoridad de política monetaria. A partir de las condiciones de equilibrio obtenidas mediante optimización dinámica, se construye un sistema de ecuaciones no lineales que describe la evolución conjunta de variables macroeconómicas como es en este caso el PIB, la inflación y la tasa de interés.

Posteriormente, el modelo es reformulado en espacio de estado, permitiendo distinguir entre variables latentes y variables observadas. Debido a la presencia de esperanzas racionales y términos no lineales, las esperanzas condicionales son resueltas numéricamente mediante cuadratura de Gauss-Hermite e interpolación por elementos finitos. Se destaca que, a diferencia de enfoques como el Particle Marginal Metropolis-Hastings (PMMH) propuesto en [41], en el presente trabajo no se realiza estimación bayesiana de los parámetros sin dependencia del tiempo, sino que se emplean valores estándares de los parámetros que definen el modelo.

Finalmente, se implementa un filtro de partículas regularizado para aproximar la distribución posterior de los estados latentes y generar estimaciones y predicciones de las variables macroeconómicas observables.

4.6.1. Modelo DSGE con Familia Representativa, Firmas y Banco Central

En esta sección se presenta el sistema estructural específico utilizado, basado en el modelo propuesto por [3], partiendo desde optimizar la función objetivo de utilidad, para posteriormente encontrar ecuaciones que reflejan las condiciones de equilibrio bajo las diferentes restricciones tecnológicas y nominales.

Primeramente, se presentarán las ecuaciones que influyen en el modelo DSGE a través de los hogares representativos (hogares), las firmas representativas (empresas) y política monetaria (banco central), finalizando con las 8 ecuaciones de equilibrio que modelan el modelo DSGE.

4.6.1.1. Problema del Hogar Representativo

El hogar representativo elige secuencias de consumo C_t, trabajo h_t, bonos nominales B_t y dinero M_t para maximizar la utilidad intertemporal:

$$\max_{\{C_t,h_t\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left[a_t \ln(C_t) + \ln(\frac{M_t}{P_t}) - \frac{1}{\eta} h_t^{\eta} \right], \tag{4.18}$$

sujeta a la restricción presupuestaria real (simplificada):

$$\frac{M_{t-1} + B_{t-1} + T_t + W_t h_t + D_t}{P_t} \geq C_t + \frac{\frac{B_t}{r_t} + M_t}{P_t}, \tag{4.19}$$

donde a_t es el shock de preferencia que sigue un proceso autoregresivo de primer orden (véase ecuación 4.20, esta ecuación ya se considera para estudios de modelos DSGE en [42].), $0 < \beta < 1$ es el factor de descuento intertemporal, $\eta \geq 1$ es la elasticidad inversa de la oferta laboral, D_t son dividendos, B_t son bonos nominales, M_t dinero, W_t salario nominal, P_t nivel de precios, y T_t transferencias.

$$\ln(a_t) = (1 - \rho_a) \ln(a_s) + \rho_a \ln(a_{t-1}) + \epsilon_{a,t}. \tag{4.20}$$

Tal como se visualizó en el capítulo anterior, este tipo de problemas de optimización se resuelven mediante los multiplicadores de Lagrange junto a la función Lagrangiana (véase el apéndice A), teniendo lo siguiente:

Lagrangiano:

$$\begin{aligned} \mathcal{L} =& E_0 \sum_{t=0}^{\infty} \beta^t \left[a_t \ln(C_t) + \ln(\frac{M_t}{P_t}) - \frac{1}{\eta} h_t^{\eta} + \right. \\ & \left. \lambda_t \left(\frac{M_{t-1} + B_{t-1} + T_t + W_t h_t + D_t}{P_t} - C_t - \frac{\frac{B_t}{r_t} + M_t}{P_t} \right) \right]. \end{aligned} \tag{4.21}$$

Como se desea optimizar mediante las cuatro variables C_t, h_t, B_t y M_t, se obtienen las condiciones de primer orden derivando para cada una de estas variables, pero nótese que para B_t y M_t, existe derivada para $t+1$ por la existencia en la restricción de B_{t-1} y M_{t-1}. Además, considerar que los valores para el tiempo t se conocen, estas son constantes en la esperanza, mientras que para los valores futuros $t+1$, estas quedan dentro de la esperanza.

Condiciones de primer orden:

$$\frac{\partial \mathcal{L}}{\partial C_t}: \quad \frac{a_t}{C_t} - \lambda_t = 0 \quad \Rightarrow \quad \lambda_t = \frac{a_t}{C_t}, \tag{4.22}$$

$$\frac{\partial \mathcal{L}}{\partial h_t}: \quad -h_t^{\eta-1} + \lambda_t \frac{W_t}{P_t} = 0 \quad \Rightarrow \quad h_t^{\eta-1} = \frac{a_t}{C_t}\frac{W_t}{P_t}. \tag{4.23}$$

$$\frac{\partial \mathcal{L}}{\partial B_t}: \quad -\frac{\beta^t \lambda_t}{r_t P_t} + \beta^{t+1} E_t\left[\frac{\lambda_{t+1}}{P_{t+1}}\right] = 0 \quad \Rightarrow \quad \frac{a_t}{C_t} = \beta r_t E_t\left[\frac{a_{t+1}}{C_{t+1}}\frac{P_t}{P_{t+1}}\right] \tag{4.24}$$

Ahora bien, para obtener la última condición de primer orden, se requiere restar la primera parte de la condición 4.24, siendo lo siguiente

$$\frac{\partial \mathcal{L}}{\partial M_t}: \quad \beta^t\left(\frac{1}{M_t} - \frac{\lambda_t}{P_t}\right) + \beta^{t+1} E_t\left[\frac{\lambda_{t+1}}{P_{t+1}}\right] = 0 \tag{4.25}$$

Restando 4.24 - 4.25, se obtiene

$$\begin{aligned}
-\frac{\beta^t \lambda_t}{r_t P_t} + \beta^{t+1} E_t\left[\frac{\lambda_{t+1}}{P_{t+1}}\right] - \beta^t\left(\frac{1}{M_t} - \frac{\lambda_t}{P_t}\right) - \beta^{t+1} E_t\left[\frac{\lambda_{t+1}}{P_{t+1}}\right] &= 0 \\
-\frac{\lambda_t}{r_t P_t} - \frac{1}{M_t} + \frac{\lambda_t}{P_t} &= 0 \\
\frac{\lambda_t}{P_t}\left(-\frac{1}{r_t} + 1\right) &= \frac{1}{M_t} \\
\frac{M_t}{P_t}\left(\frac{-1 + r_t}{r_t}\right) &= \frac{1}{\lambda_t} \\
\frac{M_t}{P_t} &= \frac{C_t}{a_t}\left(\frac{r_t}{r_t - 1}\right)
\end{aligned} \tag{4.26}$$

Estas condiciones determinan la dinámica óptima del consumo y la oferta de trabajo, fundamentales para el equilibrio macroeconómico [43].

4.6.2. Empresas Productoras y Fijación de Precios

Las firmas representativas durante cada periodo $t = 0, 1, 2, 3...$, utilizan $Y_t(i)$ unidades de cada bien intermedio para $i \in [0, 1]$, comprados a un precio nominal P_t para realizar Y_t unidades del bien terminado de acuerdo con la tecnología de rendimientos constantes descrita por

$$\left[\int_0^1 Y_t(i)^{(\theta_t - 1)/\theta_t} di\right]^{\theta_t/(\theta_t - 1)} \geq Y_t$$

donde θ_t actúa como un shock de margen o costo, que desplaza el margen deseado entre precio y costo marginal. Lo modelamos como un proceso autorregresivo de primer orden alrededor de su valor de estado estacionario θ_s [1]:

$$\ln(\theta_t) \;=\; (1-\rho_\theta)\,\ln(\theta_s) \;+\; \rho_\theta\,\ln(\theta_{t-1}) \;+\; \varepsilon_{\theta,t}, \tag{4.27}$$

con $\rho_\theta \in (0,1)$ y $\varepsilon_{\theta,t} \sim \mathcal{N}(0,\sigma_\theta^2)$. El término ρ_θ captura la persistencia del shock y σ_θ^2 su varianza. En el estado estacionario, se cumple $E[\ln(\theta_t)] = \ln(\theta_s)$.

Las firmas productoras maximizan sus ganancias al elegir

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\theta} Y_t$$

para todo $i \in [0,1]$ y $t = 0,1,2,...$ el cual confirma que θ_t mide la elasticidad de la demanda respecto al tiempo para cada bien intermedio. La competencia reduce a cero las ganancias de la empresa productora de bienes terminados en el equilibrio, determinando P_t como

$$P_t = \left(\int_0^1 P_t(i)^{1-\theta_t} di\right)^{1/(1-\theta_t)}$$

Ahora bien, las firmas intermedias para cada tiempo $t = 0,1,2,...$ contratan $h_t(i)$ unidades de trabajo de una familia representativa para producir $Y_t(i)$ unidades del bien intermedio i de acuerdo con la tecnología descrita por

$$Z_t h_t(i) \geq Y_t(i) \tag{4.28}$$

donde Z_t es un choque tecnológico que sigue un proceso estocástico que representa innovaciones tecnológicas [44]:

$$\ln(Z_t) = \ln(z_s) + \ln(Z_{t-1}) + \epsilon_{Z,t}, \tag{4.29}$$

donde Z_s es el valor estacionario y $\epsilon_{Z,t}$ es un choque de innovación.

Además, la empresa productora de bienes intermedios enfrenta un costo explícito de ajuste de precios nominales, medido en términos del bien terminado dado por

$$\frac{\phi}{2}\left[\frac{P_t(i)}{\pi_s P_{t-1}(i)}-1\right]^2 Y_t,$$

donde $\phi \geq 0$ determina la magnitud del costo de ajuste de precios y $\pi_s \geq 1$ mide la tasa bruta de inflación en el estado estacionario.

La empresa debe elegir una secuencia para $P_t(i)$ para maximizar su valor total de mercado, dado por

$$\max_{P_t(i)} E_0 \sum_{0}^{\infty} \beta^t \frac{a_t}{C_t}\frac{D_t(i)}{P_t} \tag{4.30}$$

donde $\beta^t(a_t/C_t)$ mide el valor de la utilidad marginal para el hogar representativo de una unidad adicional de ganancias reales generadas durante el período t, y donde

$$\frac{D_t(i)}{P_t} = \left(\frac{P_t(i)}{P_t}\right)^{1-\theta_t} Y_t - \left(\frac{P_t(i)}{P_t}\right)^{-\theta_t} \frac{W_t}{P_t}\frac{Y_t}{Z_t} - \frac{\phi}{2}\left(\frac{P_t(i)}{\pi_s P_{t-1}(i)}-1\right)^2 Y_t \tag{4.31}$$

mide las ganancias reales, así como el requisito de que la empresa produzca y venda su producción a demanda al precio $P_t(i)$ elegido durante cada período $t = 0, 1, 2, \ldots$. Las condiciones de primer orden derivando con respecto a $P_t(i)$ (véase con detalle en apéndice A)

$$\begin{aligned} 0 =& (1-\theta_t)\frac{a_t}{C_t}\left(\frac{P_t(i)}{P_t}\right)^{-\theta_t}\frac{Y_t}{P_t} + \theta_t \frac{a_t}{C_t}\left(\frac{P_t(i)}{P_t}\right)^{-\theta_t-1}\frac{W_t}{P_t}\frac{Y_t}{Z_t}\frac{1}{P_t} \\ & - \phi\frac{a_t}{C_t}\left(\frac{P_t(i)}{\pi_s P_{t-1}(i)}-1\right)\left(\frac{Y_t}{\pi_s P_{t-1}}\right) + \\ & \beta\phi E_t\left[\left(\frac{a_{t+1}}{C_{t+1}}\right)\left(\frac{P_{t+1}(i)}{\pi_s P_t(i)}-1\right)\left(\frac{Y_{t+1}}{P_t(i)}\right)\left(\frac{P_{t+1}(i)}{\pi_s P_t(i)}\right)\right] \end{aligned} \tag{4.32}$$

En un equilibrio simétrico, todas las firmas hacen deciciones idénticas debido a que enfrentan las mismas condiciones, lo cual $Y_t(i) = Y_t$, $h_t(i) = h_t$, $P_t(i) = P_t$ y $D_t(i) = D_t$ para todo $i \in [0, 1]$ y $t = 0, 1, 2, ...$ Además, las condiciones de equilibrio de mercado son $M_t = M_{t-1} + T_t$ y $B_t = B_{t-1} = 0$ para todo $t = 0, 1, 2, ...$ Con las condiciones de equilibrio impuestas recientemente, podemos reescribir las siguientes ecuaciones.

La ecuación 4.28, la reescribimos

$$Z_t h_t = Y_t \Rightarrow h_t = \frac{Y_t}{Z_t} \tag{4.33}$$

La ecuación 4.31, la reescribimos considerando que $P_t(i)/P_t = 1$

$$\frac{D_t}{P_t} = Y_t - \frac{W_t}{P_t}\frac{Y_t}{Z_t} - \frac{\phi}{2}\left(\frac{P_t}{\pi_s P_{t-1}} - 1\right)^2 Y_t$$

Sustituyendo la ecuación 4.33, tenemos que

$$\frac{D_t}{P_t} = Y_t - \frac{W_t h_t}{P_t} - \frac{\phi}{2}\left(\frac{P_t}{\pi_s P_{t-1}} - 1\right)^2 Y_t \tag{4.34}$$

Ahora bien, para encontrar una de las condiciones de equilibrio, consideramos la restricción del hogar representativo 4.19, considerando las condiciones de mercado $M_t = M_{t-1} + T_t$, $B_t = B_{t-1} = 0$ y separando el denominador P_t

$$\frac{M_{t-1} + \cancel{B_{t-1}} + T_t + W_t h_t + D_t}{P_t} = C_t + \frac{\cancel{B_t} + M_t}{P_t}$$

Obtenemos

$$\frac{\cancel{M_t}}{\cancel{P_t}} + \frac{W_t h_t}{P_t} + \frac{D_t}{P_t} = C_t + \frac{\cancel{M_t}}{\cancel{P_t}}$$

Ahora bien, eliminando M_t/P_t y sustituyendo D_t/P_t por la ecuación 4.34, se tiene

$$\frac{\cancel{W_t h_t}}{\cancel{P_t}} + Y_t - \frac{\cancel{W_t h_t}}{\cancel{P_t}} - \frac{\phi}{2}\left(\frac{P_t}{\pi_s P_{t-1}} - 1\right)^2 Y_t = C_t$$

Cancelando y despejando, obtenemos la condición de restricción de recursos

$$Y_t = C_t + \frac{\phi}{2}\left(\frac{P_t}{\pi_s P_{t-1}} - 1\right)^2 Y_t \tag{4.35}$$

Ahora bien, para encontrar la condición de fijación de precios, partimos de la ecuación 4.32

$$\begin{aligned} 0 =& (1-\theta_t)\frac{a_t}{C_t}\frac{Y_t}{P_t} + \theta_t \frac{a_t}{C_t}\frac{W_t}{P_t}\frac{Y_t}{Z_t}\frac{1}{P_t} - \phi\frac{a_t}{C_t}\left(\frac{P_t}{\pi_s P_{t-1}} - 1\right)\left(\frac{Y_t}{\pi_s P_{t-1}}\right) + \\ & \beta\phi E_t\left[\left(\frac{a_{t+1}}{C_{t+1}}\right)\left(\frac{P_{t+1}}{\pi_s P_t} - 1\right)\left(\frac{Y_{t+1}}{P_t}\right)\left(\frac{P_{t+1}}{\pi_s P_t}\right)\right] \end{aligned}$$

Ahora bien, para tener una mejor forma de la condición de fijación de precios multiplicamos por $(C_t/a_t)(P_t/Y_t)$, se obtiene

$$0 = 1 - \theta_t + \theta_t \frac{W_t}{P_t}\frac{1}{Z_t} - \phi\left(\frac{P_t}{\pi_s P_{t-1}} - 1\right)\left(\frac{P_t}{\pi_s P_{t-1}}\right) + \beta\phi E_t\left[\left(\frac{a_{t+1}}{a_t}\right)\left(\frac{C_t}{C_{t+1}}\right)\left(\frac{P_{t+1}}{\pi_s P_t} - 1\right)\left(\frac{Y_{t+1}}{Y_t}\right)\left(\frac{P_{t+1}}{\pi_s P_t}\right)\right] \tag{4.36}$$

Ahora bien, considerando la segunda condición de primer orden al maximizar la utilidad del hogar representativo 4.23

$$h_t^{\eta-1} = \frac{a_t}{C_t}\frac{W_t}{P_t} \Rightarrow \frac{C_t h_t^{\eta-1}}{a_t} = \frac{W_t}{P_t}$$

y sustituyendo $h_t = Y_t/Z_t$, todo esto en la ecuación 4.36, finalmente obtenemos

$$0 = 1 - \theta_t + \theta_t \frac{1}{a_t}\frac{C_t}{Z_t}\left(\frac{Y_t}{Z_t}\right)^{\eta-1} - \phi\left(\frac{P_t}{\pi_s P_{t-1}} - 1\right)\left(\frac{P_t}{\pi_s P_{t-1}}\right) + \beta\phi E_t\left[\left(\frac{a_{t+1}}{a_t}\right)\left(\frac{C_t}{C_{t+1}}\right)\left(\frac{P_{t+1}}{\pi_s P_t} - 1\right)\left(\frac{Y_{t+1}}{Y_t}\right)\left(\frac{P_{t+1}}{\pi_s P_t}\right)\right] \tag{4.37}$$

para todo $t = 0, 1, 2, ...$ La esperanza racional es la dependencia intertemporal en la fijación de precios, y se resuelve numéricamente [42].

Asimismo un planificador social puede enfrentarse a un problema que puede superar las fricciones que hacen que los saldos monetarios reales aparezcan en la función de utilidad del hogar representativo y que dan origen al costo de ajuste de precios nominales que enfrenta la empresa representativa productora de bienes intermedios. Durante cada período $t = 0, 1, 2, \ldots$, este planificador social asigna $n_t(i)$ unidades de trabajo del hogar representativo para producir $Q_t(i)$ unidades de cada bien intermedio $i \in [0, 1]$, y luego utiliza esos diversos bienes intermedios para producir Q_t unidades del bien terminado, por lo cual el planificador social elige secuencias Q_t y $n_t(i)$ para todo $i \in [0, 1]$ para maximizar el bienestar del hogar dado por

$$\max_{Q_t, n_t(i)} E_0 \sum_{t=0}^{\infty} \beta^t \left[a_t \ln(Q_t) - \frac{1}{\eta}\left(\int_0^1 n_t(i)di\right)^{\eta}\right] \tag{4.38}$$

s.a

$$Z_t\left(\int_0^1 n_t(i)^{(\theta_t-1)/\theta_t} di\right)^{\theta_t/(\theta_t-1)} \geq Q_t \tag{4.39}$$

para todo $t = 0, 1, 2, ...$, aplicando los multiplicadores de Lagrange, tenemos que

$$\mathcal{L} = E_0 \sum_{t=0}^{\infty} \beta^t \left[a_t \ln(Q_t) - \frac{1}{\eta} \left(\int_0^1 n_t(i) di \right)^{\eta} + \lambda_t \left(Z_t \left(\int_0^1 n_t(i)^{(\theta_t - 1)/\theta_t} di \right)^{\theta_t/(\theta_t - 1)} - Q_t \right) \right]$$

entonces las condiciones de primer orden son

$$\frac{\partial \mathcal{L}}{\partial \lambda_t} : Z_t \left(\int_0^1 n_t(i)^{(\theta_t - 1)/\theta_t} di \right)^{\theta_t/(\theta_t - 1)} = Q_t \Rightarrow \left(\int_0^1 n_t(i)^{(\theta_t - 1)/\theta_t} di \right)^{\theta_t/(\theta_t - 1)} = \frac{Q_t}{Z_t} \quad (4.40)$$

$$\frac{\partial \mathcal{L}}{\partial Q_t} : \frac{a_t}{Q_t} - \lambda_t = 0 \Rightarrow \lambda_t = \frac{a_t}{Q_t} \quad (4.41)$$

Ahora bien, derivando con respecto a $n_t(i)$

$$- \left(\int_0^1 n_t(i) di \right)^{\eta - 1} + \lambda_t Z_t \left(\frac{\theta_t}{\theta_t - 1} \right) \left(\int_0^1 n_t(i)^{(\theta_t - 1)/\theta_t} di \right)^{1/(\theta_t - 1)} \left(\frac{\theta_t - 1}{\theta_t} \right) n_t(i)^{-1/\theta_t} = 0$$

Cancelando y sustituyendo λ_t

$$- \left(\int_0^1 n_t(i) di \right)^{\eta - 1} + \frac{a_t}{Q_t} Z_t \left(\int_0^1 n_t(i)^{(\theta_t - 1)/\theta_t} di \right)^{1/(\theta_t - 1)} n_t(i)^{-1/\theta_t} = 0 \quad (4.42)$$

Aplicando nuevamente el equilibrio simétrico por lo antes expuesto, se tiene que

$$n_t(i) = n_t \Rightarrow \int_0^1 n_t(i) di = \int_0^1 n_t di = n_t \int_0^1 di = n_t$$

La ecuación 4.42 se convierte en

$$-n_t^{\eta - 1} + \frac{a_t}{Q_t} Z_t \left(n_t^{(\cancel{\theta_t - 1})/\theta_t} \right)^{1/(\cancel{\theta_t - 1})} n_t^{-1/\theta_t} = 0$$

cancelando

$$-n_t^{\eta - 1} + \frac{a_t}{Q_t} Z_t \left(n_t^{\cancel{1/\theta_t}} \right) n_t^{\cancel{-1/\theta_t}} = 0$$

Simplificando y despejando

$$n_t^{\eta - 1} = \frac{a_t}{Q_t} Z_t \quad (4.43)$$

Ahora bien, para encontrar el valor de n_t, se obtiene de la condición de equilibrio 4.40

$$\left(\int_0^1 n_t(i)^{(\theta_t-1)/\theta_t} di\right)^{\theta_t/(\theta_t-1)} = \frac{Q_t}{Z_t} \Rightarrow n_t = \frac{Q_t}{Z_t}$$

Sustituyendo en la ecuación 4.43

$$\left(\frac{Q_t}{Z_t}\right)^{\eta} = \frac{a_t}{Q_t} Z_t \Rightarrow Q_t = a_t^{1/\eta} Z_t \tag{4.44}$$

para todo $t = 0, 1, 2, \ldots$. De acuerdo con esta definición, el nivel eficiente de producción aumenta después de un choque de preferencias favorable a_t o un choque tecnológico Z_t.

4.6.2.1. Política Monetaria

La tasa de política monetaria es implementada por un banco central que ajusta la tasa nominal de interés r_t, buscando estabilizar tanto la inflación como el producto. Para modelar esta decisión, se utiliza una regla tipo Taylor, ampliamente empleada en la literatura macroeconómica moderna [45].

La regla se expresa de forma general como:

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left[\phi_\pi(\pi_t - \pi^*) + \phi_x(x_t - x^*)\right] + v_{r,t}, \tag{4.45}$$

donde ρ_r es el grado de persistencia en la política monetaria, ϕ_π mide la respuesta del banco central a desviaciones de inflación, ϕ_y mide la respuesta ante desviaciones del producto, π^* y x^* son las metas de inflación y producto potencial, y $v_{r,t}$ es un choque monetario no anticipado. Además, sin pérdida de generalidad se puede considerar π^* y x^* iguales a cero, teniendo la ecuación resultante.

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left[\phi_\pi \pi_t + \phi_x x_t\right] + v_{r,t}, \tag{4.46}$$

El uso de esta regla nos ayuda a analizar cómo la política monetaria reacciona ante shocks y contribuye a estabilizar la economía, alineándose con prácticas empíricas observadas en bancos centrales modernos [43].

Cuando la inflación está por encima de la meta, el banco central eleva la tasa nominal para contener presiones inflacionarias. Asimismo, si la brecha del producto no es negativa,

la política se endurece. Este comportamiento es fundamental en la transmisión de la política monetaria en modelos DSGE [46].

A continuación se presentan las 8 ecuaciones incluyendo las 3 antes mencionadas que caracterizan el modelo DSGE.

4.6.3. Ecuaciones de Equilibrio

En esta sección sintetizamos las ecuaciones de equilibrio que componen el modelo DSGE, exceptuando la regla de política monetaria, con el fin de guiar accesiblemente al lector. Para la ecuación de restricción de recursos, partimos de la ecuación 4.35

$$Y_t = C_t + \frac{\phi}{2}\left(\frac{P_t}{\pi_s P_{t-1}} - 1\right)^2 Y_t.$$

Dividiendo ambos lados por Z_t, definimos $y_t = Y_t/Z_t$, $c_t = C_t/Z_t$ y $\pi_t = \frac{P_t}{P_{t-1}}$, de forma que la ecuación se reescribe como:

$$y_t = c_t + \frac{\phi}{2}\left(\frac{\pi_t}{\pi_s} - 1\right)^2 y_t. \tag{4.47}$$

El choque de preferencias a_t sigue un proceso autorregresivo de primer orden, donde el estado estacionario es $a_s > 0$, lo cual si lo reescribimos centrado en el estado estacionario, se tiene que

$$\ln(a_t) - \ln(a_s) = \rho_a(\ln(a_{t-1}) - \ln(a_s)) + \epsilon_{a,t}.$$

Siendo equivalente a

$$\ln(a_t) = (1 - \rho_a)\ln(a_s) + \rho_a \ln(a_{t-1}) + \epsilon_{a,t}. \tag{4.48}$$

Si se asume que el valor de estado estacionario es 0, es decir $E(\ln(a_t)) = 0$, se tendría la ecuación

$$\ln(a_t) = \rho_a \ln(a_{t-1}) + \epsilon_{a,t}$$

El choque tecnológico Z_t sigue:

$$\ln(Z_t) = \ln(z_s) + \ln(Z_{t-1}) + \epsilon_{z,t}. \tag{4.49}$$

Al definir $z_t = Z_t/Z_{t-1}$, esto implica que $\ln(z_t) = \ln(Z_t) - \ln(Z_{t-1})$, obteniendo la ecuación final:

$$\ln(z_t) = \ln(z_s) + \epsilon_{z,t}. \tag{4.50}$$

Para la ecuación de Euler del consumidor, partimos de la ecuación 4.24

$$\frac{a_t}{C_t} = \beta r_t E_t \left[\frac{a_{t+1}}{C_{t+1}} \frac{P_t}{P_{t+1}}\right]$$

definiendo $c_t = C_t/Z_t$, entonces $C_t = c_t Z_t$

$$\frac{a_t}{c_t Z_t} = \beta r_t E_t \left[\frac{a_{t+1}}{c_{t+1}} \frac{1}{\pi_{t+1}} \frac{1}{Z_{t+1}}\right] \Rightarrow \frac{a_t}{c_t} = \beta r_t E_t \left[\frac{a_{t+1}}{c_{t+1}} \frac{1}{\pi_{t+1}} \frac{Z_t}{Z_{t+1}}\right]$$

dado que $z_t = Z_t/Z_{t-1}$, finalizamos teniendo

$$\frac{a_t}{c_t} = \beta r_t E_t \left[\frac{a_{t+1}}{c_{t+1}} \frac{1}{z_{t+1}} \frac{1}{\pi_{t+1}}\right]. \tag{4.51}$$

Para la condición de fijación de precios, partimos de la ecuación 4.36

$$\begin{aligned} 0 =& 1 - \theta_t + \theta_t \frac{1}{a_t} \frac{C_t}{Z_t} \left(\frac{Y_t}{Z_t}\right)^{\eta-1} - \phi \left(\frac{P_t}{\pi_s P_{t-1}} - 1\right) \left(\frac{P_t}{\pi_s P_{t-1}}\right) + \\ & \beta\phi E_t \left[\left(\frac{a_{t+1}}{a_t}\right) \left(\frac{C_t}{C_{t+1}}\right) \left(\frac{P_{t+1}}{\pi_s P_t} - 1\right) \left(\frac{Y_{t+1}}{Y_t}\right) \left(\frac{P_{t+1}}{\pi_s P_t}\right)\right] \end{aligned}$$

Considerando $c_t = C_t/Z_t$, $y_t = Y_t/Z_t$ y $\pi_t = P_t/P_{t-1}$, obtenemos

$$\begin{aligned} 0 =& 1 - \theta_t + \theta_t \frac{c_t}{a_t} y^{\eta-1} - \phi \left(\frac{\pi_t}{\pi_s} - 1\right) \left(\frac{\pi_t}{\pi_s}\right) + \\ & \beta\phi E_t \left[\left(\frac{a_{t+1}}{a_t}\right) \left(\frac{c_t \cancel{Z_t}}{c_{t+1} \cancel{Z_{t+1}}}\right) \left(\frac{\pi_{t+1}}{\pi_s} - 1\right) \left(\frac{y_{t+1} \cancel{Z_{t+1}}}{y_t \cancel{Z_t}}\right) \left(\frac{\pi_{t+1}}{\pi_s}\right)\right] \end{aligned}$$

Cancelando los Z_t y Z_{t+1}, finalmente obtenemos la condición de fijación de precios

$$0 = 1 - \theta_t + \theta_t \left(\frac{c_t}{a_t}\right) y_t^{\eta-1} - \phi \left(\frac{\pi_t}{\pi_s} - 1\right)\left(\frac{\pi_t}{\pi_s}\right) + \beta\phi E_t \left[\left(\frac{a_{t+1}}{a_t}\right)\left(\frac{c_t}{c_{t+1}}\right)\left(\frac{\pi_{t+1}}{\pi_s} - 1\right)\left(\frac{\pi_{t+1}}{\pi_s}\right)\left(\frac{y_{t+1}}{y_t}\right)\right]. \tag{4.52}$$

En la condición de fijación de precios, θ_t actúa como un shock de margen o costo, que desplaza el margen deseado entre precio y costo marginal. Lo modelamos como un proceso autorregresivo de primer orden alrededor de su valor de estado estacionario θ_s:

$$\ln(\theta_t) \;=\; (1 - \rho_\theta)\,\ln(\theta_s) \;+\; \rho_\theta\,\ln(\theta_{t-1}) \;+\; \varepsilon_{\theta,t}, \tag{4.53}$$

con $\rho_\theta \in (0, 1)$ y $\varepsilon_{\theta,t} \sim \mathcal{N}(0, \sigma_\theta^2)$. El término ρ_θ captura la persistencia del shock y σ_θ^2 su varianza. En el estado estacionario, se cumple $\mathbb{E}[\ln(\theta_t)] = \ln(\theta_s)$.

Se define la tasa de crecimiento económico:

$$g_t = \frac{Y_t}{Y_{t-1}} = \frac{y_t}{y_{t-1}}\frac{Z_t}{Z_{t-1}} = \frac{y_t}{y_{t-1}} z_t. \tag{4.54}$$

Por último, partiendo de la ecuación 4.44, la brecha del producto se puede definir como

$$x_t = \frac{Y_t}{Q_t} = \frac{Y_t}{a_t^{1/\eta} Z_t}$$

y como $y_t = Y_t/Z_t$ finalmente se tiene la ecuación de equilibrio:

$$x_t = \frac{y_t}{a_t^{1/\eta}}. \tag{4.55}$$

Las ecuaciones finales son 4.46–4.55 donde capturan toda la información de un modelo DSGE: decisiones del hogar, fijación de precios, rigideces nominales y procesos de shocks. [27, 43, 46].

4.6.4. Representación en Espacio de Estado y Resolución Numérica

Una vez definidas las ecuaciones estructurales, el siguiente paso consiste en formular el sistema en forma de espacio de estado para permitir la implementación de métodos de

inferencia secuencial, como el filtro de partículas regularizado. Esta formulación facilita la separación explícita entre los estados latentes y las variables observadas, aspecto esencial para la estimación dinámica y la predicción.

4.6.4.1. Representación en Espacio de Estado

Una de las ventajas que tenemos con los modelos DSGE es que pueden ser representados como sistemas dinámicos en forma de espacio de estado. Esta característica es fundamental para aplicar métodos de inferencia y estimación como los filtros bayesianos, en particular el Filtro de Partículas, que será desarrollado en el siguiente capítulo [47].

En términos generales, un modelo de espacio de estado está compuesto por dos ecuaciones fundamentales. Como se definió en el Capítulo 3, esta estructura se basa en las propiedades de Markov y en la independencia condicional entre estados y observaciones, lo cual proporciona un marco general que ahora aplicamos en el contexto de los modelos DSGE.

Considerando las definiciones del capítulo 3, las ecuaciones 3.1 y 3.2 que conforman un modelo de espacio de estado, han sido utilizadas en la resolución de modelos DSGE, tanto en enfoques lineales como en contextos no lineales con métodos Monte Carlo secuenciales [48]. En particular, el uso del Filtro de Partículas ha demostrado ser una herramienta poderosa para tratar modelos no lineales y no gaussianos, sin necesidad de aproximaciones locales o linealización

[35] demostraron que es posible usar el filtro de partículas para estimar la verosimilitud de modelos DSGE totalmente no lineales, lo que abre la posibilidad de una inferencia más realista y flexible sin recurrir a aproximaciones locales.

Por tanto, esta representación no solo facilita la estimación y el pronóstico, sino que también conecta la teoría estructural del modelo con datos reales, constituyendo un puente clave entre teoría y evidencia empírica.

4.6.4.2. Vector de Estados

En esta tesis, el vector de estado S_t está compuesto por 12 variables principales, de acuerdo con la codificación implementada en el filtro de partículas:

$$\boldsymbol{X_t} = [y_t, r_t, \pi_t, g_t, x_t, c_t, \theta_t, a_t, z_t, \epsilon_{a,t}, \epsilon_{z,t}, v_{r,t}]\,. \tag{4.56}$$

Cada componente refleja variables fundamentales derivadas de las ecuaciones finales: y_t es el producto, r_t la tasa nominal, π_t la inflación, g_t el crecimiento, x_t la brecha, c_t el consumo, θ_t y a_t shocks de precios y preferencias, z_t el shock tecnológico, y $\epsilon_{a,t}$, $\epsilon_{z,t}$, $v_{r,t}$ son los errores de innovación.

4.6.4.3. Ecuaciones de Transición

Las ecuaciones de transición, al final dependen de las variables endógenas que se escojan en la ecuación de observación. Dado que se escoge la variable g_t y π_t, la ecuación 4.52 se excluye, ya que va implícito en los datos reales de estas variables. Entonces, la dinámica de los estados se determina a partir de las ecuaciones 4.46–4.55, resolviendo iterativamente para cada t:

$$\ln(a_t) = \rho_a \ln(a_{t-1}) + \epsilon_{a,t}, \tag{4.57}$$

$$\ln(z_t) = \ln(z_s) + \epsilon_{z,t}, \tag{4.58}$$

$$\ln(\theta_t) = (1-\rho_\theta)\ln(\theta_s) + \rho_\theta \ln(\theta_{t-1}) + \epsilon_{\theta,t}, \tag{4.59}$$

$$x_t = \frac{y_{t-1}}{a_t^{1/\eta}}, \tag{4.60}$$

$$\pi_t = \beta\pi_{t-1} + \kappa x_t + \epsilon_{\pi,t}, \tag{4.61}$$

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left(\phi_\pi \pi_t + \phi_x x_t\right) + v_{r,t} \tag{4.62}$$

$$g_t = \frac{y_t}{y_{t-1}} z_t. \tag{4.63}$$

donde la ecuación 4.61 se inspira por [49] y [50], y el consumo c_t se obtiene resolviendo numéricamente la condición de Euler:

$$c_t = \left(\frac{a_t}{\beta r_t E_t[\frac{a_{t+1}}{c_{t+1}} \frac{1}{z_{t+1}} \frac{1}{\pi_{t+1}}]} \right)^{-1}. \tag{4.64}$$

El producto y_t se deduce de la restricción de recursos:

$$y_t = \frac{c_t}{1 - \frac{\phi}{2}\left(\frac{\pi_t}{\pi_s} - 1\right)^2}. \tag{4.65}$$

La esperanza condicional

$$E_t\left[\frac{a_{t+1}}{c_{t+1}}\frac{1}{z_{t+1}}\frac{1}{\pi_{t+1}}\right] = \int\int \frac{a_{t+1}}{c_{t+1}z_{t+1}\pi_{t+1}}\, p(a_{t+1}, z_{t+1}|a_t, z_t)\, da_{t+1}dz_{t+1}$$

se resuelve mediante cuadratura de Gauss-Hermite en el apéndice B, lo cual permite aproximar integrales multidimensionales de manera eficiente. Ahora bien, se requiere aproximar valores futuros de c_{t+1} y π_{t+1}. Se utilizó interpolación por elementos finitos resuelto en el apéndice C.

4.6.4.4. Ecuación de Observación

En este trabajo, se consideran observables las siguientes variables macroeconómicas:

$$\boldsymbol{Y_t} = \begin{bmatrix} g_t \\ \pi_t \\ r_t \end{bmatrix}. \tag{4.66}$$

Teniendo la ecuación de observación

$$\boldsymbol{Y_t} = h(\boldsymbol{X_t}) + \epsilon_t \tag{4.67}$$

donde

$$h(\boldsymbol{X_t}) = \begin{bmatrix} g_t \\ \pi_t \\ r_t \end{bmatrix} \quad \boldsymbol{\epsilon_t} = \begin{bmatrix} \epsilon_{g,t} \\ \epsilon_{\pi,t} \\ \epsilon_{r,t} \end{bmatrix} \sim N(\mathbf{0}, \boldsymbol{\Sigma_y})$$

entonces la verosimilitud con distribución normal se puede escribir como

$$p(\boldsymbol{Y_t}|\boldsymbol{X_t^{(i)}}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma_y}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{Y_t} - h(\boldsymbol{X_t^{(i)}}))'\boldsymbol{\Sigma_y^{-1}}(\boldsymbol{Y_t} - h(\boldsymbol{X_t^{(i)}}))\right) \tag{4.68}$$

donde $n = 3$, porque tenemos tres variables de observación.

Estas corresponden directamente a las series de crecimiento económico, inflación y tasa nominal.

4.6.4.5. Implementación del Filtro de Partículas

En el código, se inicializan N partículas que representan posibles trayectorias de estados. Para cada t, se actualizan mediante la ecuación de transición, se evalúa la verosimilitud con base en Y_t, se ajustan los pesos y, si es necesario, se realiza el re-muestreo regularizado. Esto sigue la estructura general del filtro de partículas secuencial, adaptada al modelo DSGE no lineal propuesto.

4.6.5. Seudocódigo del Algoritmo de Filtro de Partículas

Se presenta a continuación un pseudocódigo resumido que describe el procedimiento seguido para la estimación de los estados latentes y la generación de predicciones a partir del modelo DSGE propuesto.

Algoritmo 3 Filtro de Partículas aplicado al modelo DSGE

Inicializar partículas $\{\boldsymbol{X_0}^{(i)}\}_{i=1}^{N}$
Asignar pesos iguales $w_0^{(i)} = \frac{1}{N}$
para $t = 1, \ldots, T$ **hacer**
    para $i = 1, \ldots, N$ **hacer**
        Actualizar $a_t^{(i)}$ y $z_t^{(i)}$ usando las ecuaciones de transición estocásticas.
        Calcular $x_t^{(i)}$ y $\pi_t^{(i)}$ con base en la brecha y la dinámica de precios.
        Calcular $r_t^{(i)}$ mediante la regla de política monetaria.
        Resolver numéricamente $c_t^{(i)}$ a partir de la condición de Euler..
        Obtener $y_t^{(i)}$ a partir de la restricción de recursos.
        Calcular $g_t^{(i)}$ como tasa de crecimiento.
    fin para
    Comparar $g_t^{(i)}$, $\pi_t^{(i)}$ y $r_t^{(i)}$ con las observaciones Y_t.
    Calcular la verosimilitud $p(Y_t|X_t^{(i)})$ y actualizar pesos $w_t^{(i)}$.
    Normalizar pesos.
    Evaluar verosimilitud
    Actualizar pesos
    Normalizar pesos
    si $N_{eff} < N/2$ **entonces**
        Realizar re-muestreo regularizado
    fin si
fin para
Calcular las medias (o medianas) ponderadas de las partículas para cada variable de interés.
Para prediccion, usar la última muestra de partículas $X_T^{(i)}$ como base. Propagar hacia adelante siguiendo la dinámica del modelo, sin actualizar con observaciones. Calcular estadísticas predictivas para cada $t = T + 1, \ldots, T + H$.
Generar gráficos de comparación entre observaciones y estimaciones.
Calcular métricas de error (RMSE, MAE, ELPD).
Exportar resultados numéricos a hojas de cálculo.

En el siguiente capítulo, se mostrarán los resultados de las simulaciones ejecutadas a través de este algoritmo, donde se consideraron datos de forma trimestral del PIB, inflación y tasas de interes.

Capítulo 5

Simulaciones y análisis de datos

En el presente capítulo se presentan los resultados obtenidos a partir de la implementación del modelo DSGE propuesto en su versión no lineal, estimado mediante el Filtro de Partículas. El objetivo principal es estimar el modelo DSGE mediante el método de filtro de partículas, asimismo analizar las predicciones realizadas en el modelo.

Para ello, primeramente se seleccionaron valores calibrados para los parámetros iniciales que no dependen del tiempo, con base en la literatura macroeconómica reciente y trabajos similares. Estos parámetros determinan el comportamiento del hogar representativo, la política monetaria y la rigidez nominal, y serán documentados en detalle en la primera sección.

A lo largo del capítulo se presentarán comparaciones entre las series observadas y las series estimadas del modelo, considerando tanto la media como la mediana de las partículas. El uso de la mediana busca estimaciones más robusta ante posibles valores atípicos o distribuciones sesgadas, lo cual es especialmente relevante en el contexto de modelos no lineales y no gaussianos.

Asimismo, se analizarán los errores de predicción mediante el Error Absoluto Medio (MAE), la Raíz del Error Cuadrático Medio (RMSE) y el Expected Log Predictive Density (ELPD). El MAE y el RMSE lo utilizaremos para comparar los valores puntuales de ambas medidas (media vs. mediana) para determinar cuál resulta más confiable en estos tipos de modelos, y el ELPD que nos ayuda a medir la capacidad predictiva del modelo a medida que crecen las partículas.

Los datos utilizados en esta investigación corresponden a series trimestrales de la economía de Estados Unidos, específicamente el crecimiento del PIB real, la inflación y la tasa de interés nominal. Las series abarcan desde el primer trimestre del año 2000 hasta el cuarto trimestre del 2023 para la estimación del modelo. A partir de esta estimación, se realizaron predicciones para los cuatro trimestres del año 2024 utilizando la estructura

dinámica del modelo sin incorporar nuevas observaciones. Los datos fueron extraídos de las siguientes fuentes oficiales:

1. **Crecimiento del PIB real:** Bureau of Economic Analysis (BEA), serie "Real Gross Domestic Product", disponible en: `https://fred.stlouisfed.org/series/GDPC1`.

2. **Inflación:** Bureau of Labor Statistics (BLS), serie "Consumer Price Index for All Urban Consumers (CPI-U)", disponible en: `https://fred.stlouisfed.org/series/CPIAUCSL`.

3. **Tasa de interés nominal:** Board of Governors of the Federal Reserve System, serie "Federal Funds Effective Rate", disponible en: `https://fred.stlouisfed.org/series/FEDFUNDS`.

5.1. Análisis Exploratorio de las Series de Tiempo

Las tres series se transformaron a tasas porcentuales trimestrales para adecuarse al diseño del modelo DSGE, como se puede ver a continuación (tasa de crecimiento clásico):

$$g_t = \frac{g_t}{g_{t-1}} - 1, \quad \pi_t = \frac{\pi_t}{\pi_{t-1}} - 1, \quad r_t = \frac{r_t}{r_{t-1}} - 1$$

Según capítulo 2, esto nos garantiza que pueda haber algo de estacionariedad, en la serie de tiempo.

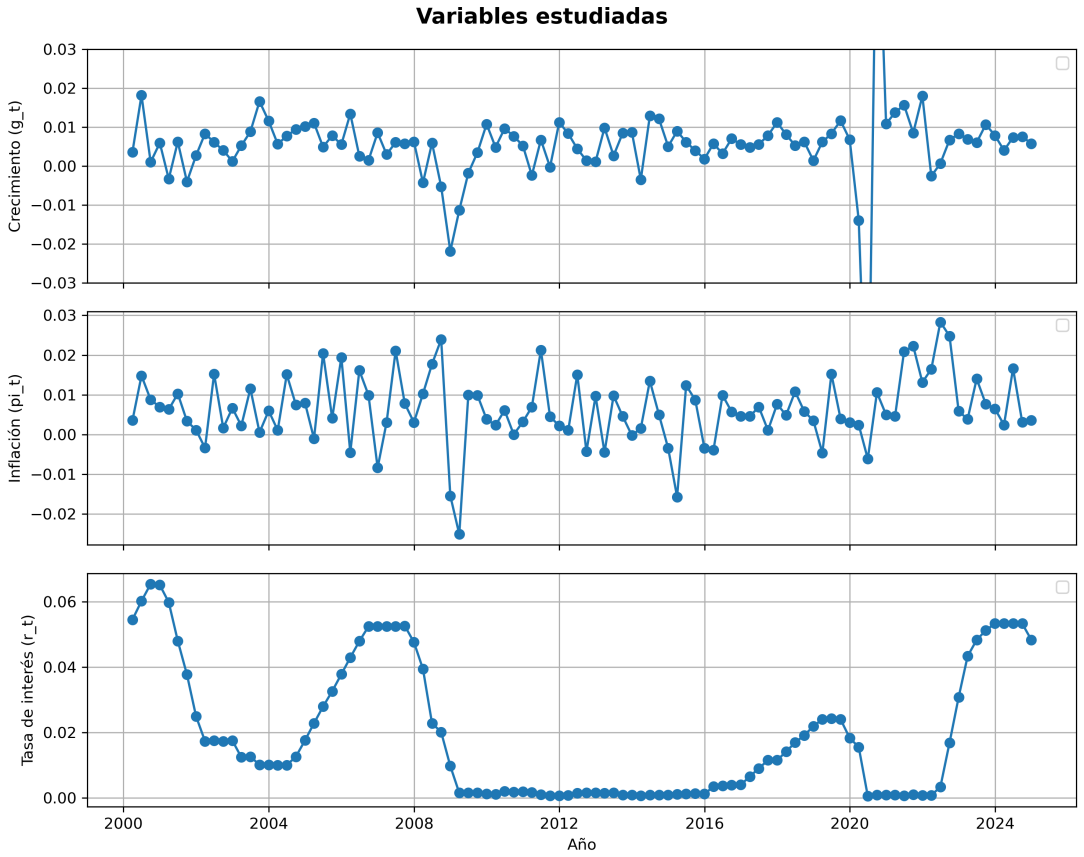


Figura 5.1: Series observadas de crecimiento del PIB real, inflación y tasa de interés nominal (2000-2024).

En la Figura 5.1, se observa:

1. **Crecimiento del PIB real:** Presenta una dinámica cíclica, con una fuerte caída durante la crisis financiera de 2008 y también durante el periodo del 2020 asociada a la pandemia de COVID-19, seguida por una recuperación pospandemia.

2. **Inflación (CPI):** Muestra una estabilidad entre el periodo 2000 y 2020, con un incremento significativo a partir de 2021, lo que coincide con los efectos pospandemia y presiones de oferta.

3. **Tasa de interés nominal:** Se mantuvo en niveles bajos durante la década posterior a 2008, con una política monetaria expansiva. A partir de 2022 se observa una tendencia ascendente en respuesta al aumento de la inflación.

5.2. Parámetros Iniciales

Antes de proceder al análisis de las estimaciones y predicciones, es fundamental detallar los valores iniciales de los parámetros empleados en la resolución del modelo DSGE. Estos parámetros no han sido estimados en esta tesis, sino que han sido calibrados en base a la literatura, utilizando valores estándar ampliamente aceptados en estudios similares como [43, 51, 42].

A continuación, se presenta una tabla con los parámetros empleados, su interpretación económica y la referencia bibliográfica principal que motiva su elección:

Parámetro	**Descripción**	**Valor**	**Fuente**
β	Factor de descuento intertemporal	0.99	[43]
η	Elasticidad inversa de la oferta laboral	1.5	[52]
ϕ	Parámetro de rigidez nominal	0.5	[42]
ρ_a	Persistencia del shock de preferencias	0.95	[51]
ρ_θ	Persistencia del shock de precio	0.8	[42]
ϕ_π	Sensibilidad a la inflación	1.2	[45]
ϕ_x	Sensibilidad al producto	0.3	[51]
ρ_r	Inercia de la tasa de interés nominal	0.85	[51]
π_s	Inflación objetivo de largo plazo	0.015	Calibración
θ_s	Valor estacionario del shock de precios	1.0	[42]
κ	Curva de Phillips (coef. rigidez nominal)	0.1	[43]
σ_ε	Desviación estándar shock observacional	0.05	Ajustado

Tabla 5.1: Parámetros calibrados utilizados en la estimación del modelo DSGE.

La elección de estos parámetros permite asegurar que el modelo tenga una dinámica razonable y acorde a estudios similares. Además, estos valores permiten que el sistema tenga un estado estacionario coherente y que las ecuaciones estructurales sean numéricamente estables para su implementación con el filtro de partículas regularizado.

También considerar tal como se mencionó en el capítulo 3, el vector de estado tiene dimensión 12, lo cual se optó por el método propuesto de [26] en contraer h, tomando un valor de 0.1 en vista que β se calibró con 0.99.

Cabe mencionar que el parámetro como π_s fueron calibrados tomando como referencia datos promedio de inflación en el periodo de estudio.

5.3. Resultados con Estimación por la Media

En esta sección se presentan los resultados obtenidos mediante la estimación de las variables macroeconómicas utilizando la media de partículas como medida de cada estado latente. Para evaluar la capacidad del modelo para replicar y predecir las series de crecimiento económico (g_t), inflación (π_t) y tasa de interés nominal (r_t), se realizaron estimaciones bajo diferentes valores del número de partículas N, explorando así la sensibilidad del filtro de partículas a este parámetro.

A continuación se muestran las figuras correspondientes a la comparación entre las variables observadas y estimadas para distintos valores de N (considerar que los últimos 4 puntos de datos son las predicciones del modelo). Cada gráfico presenta tres paneles: uno para el crecimiento económico, el siguiente para la inflación y por último la tasa de interés. Las series observadas se comparan con las estimadas por el modelo. De la figura 5.2 hasta la figura 5.17 siguen un formato similar: en el eje horizontal se muestra el tiempo (trimestres), y en el eje vertical, los valores porcentuales estimados u observados para cada variable. La línea azul corresponde a la serie observada y la línea roja a la serie estimada.

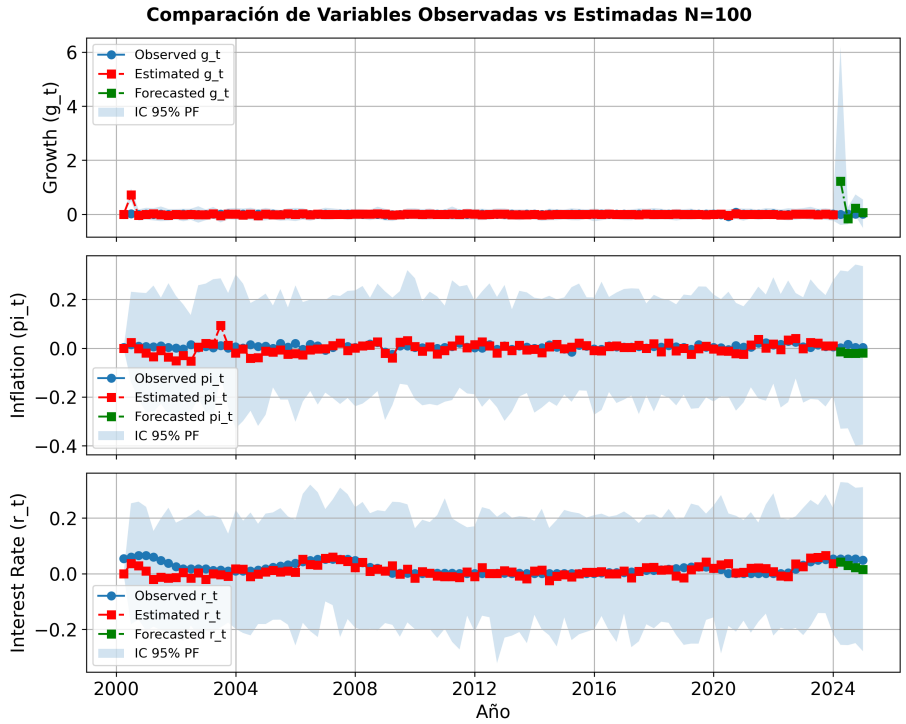


Figura 5.2: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 100$ partículas, en el periodo del 2000 al 2024 de forma trimestral.

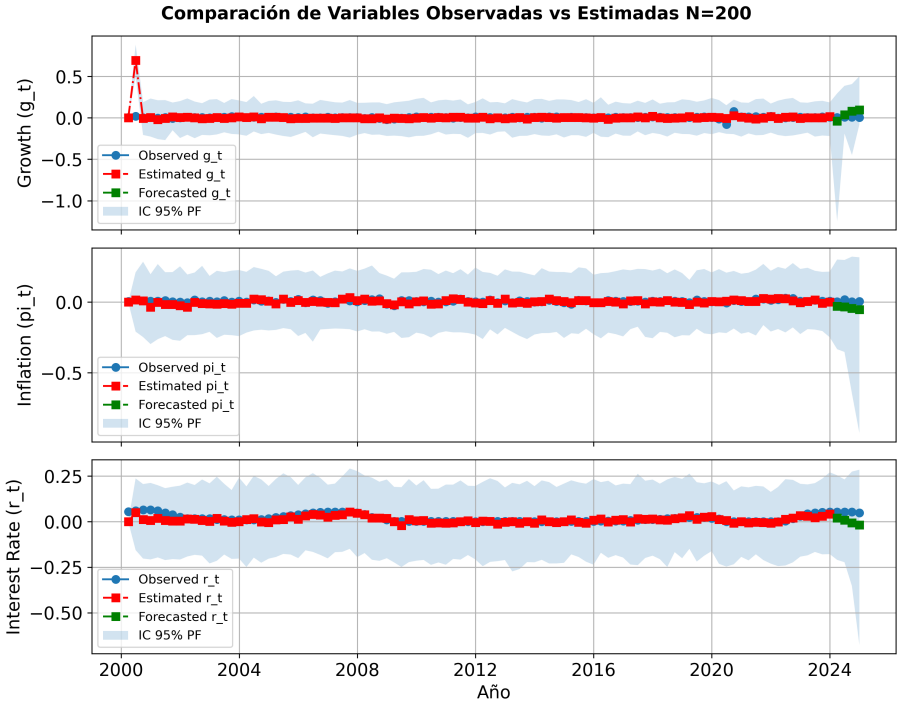


Figura 5.3: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 200$ partículas, en el periodo del 2000 al 2024 de forma trimestral.

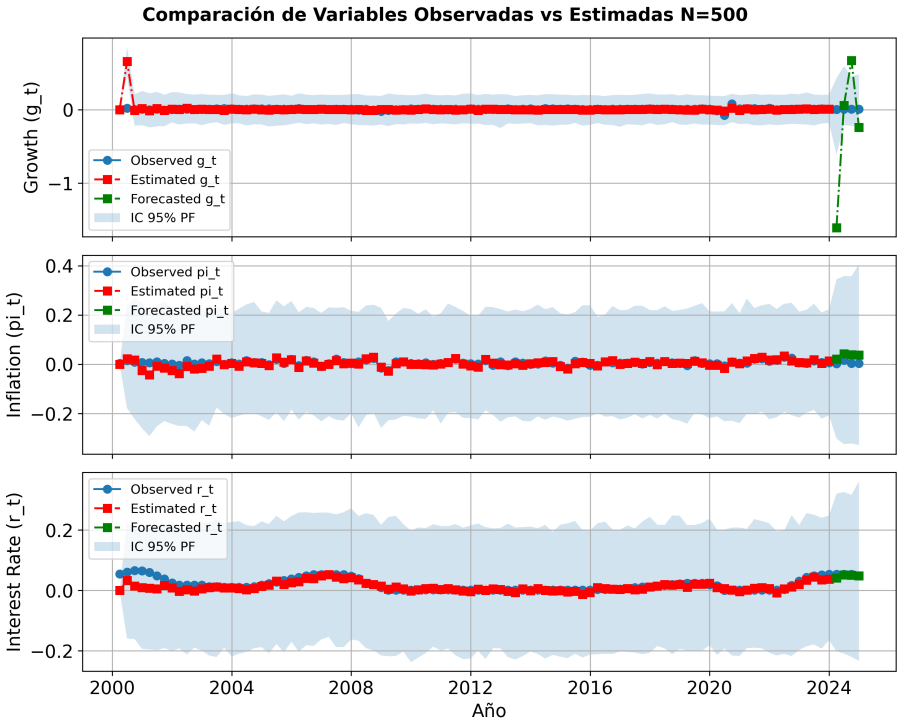


Figura 5.4: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 500$ partículas, en el periodo del 2000 al 2024 de forma trimestral.

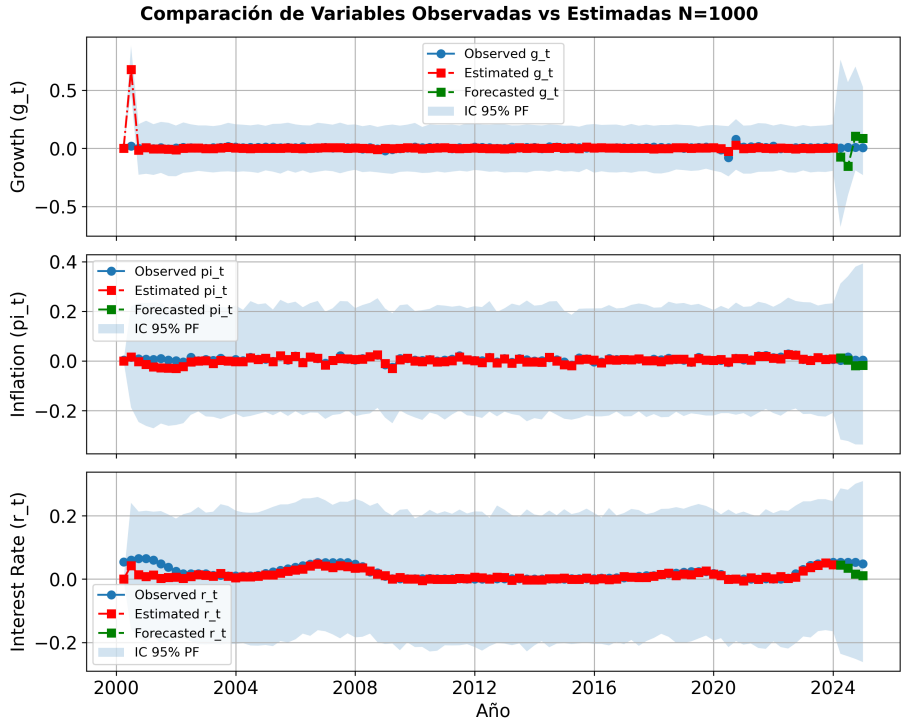


Figura 5.5: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 1000$ partículas, en el periodo del 2000 al 2024 de forma trimestral.

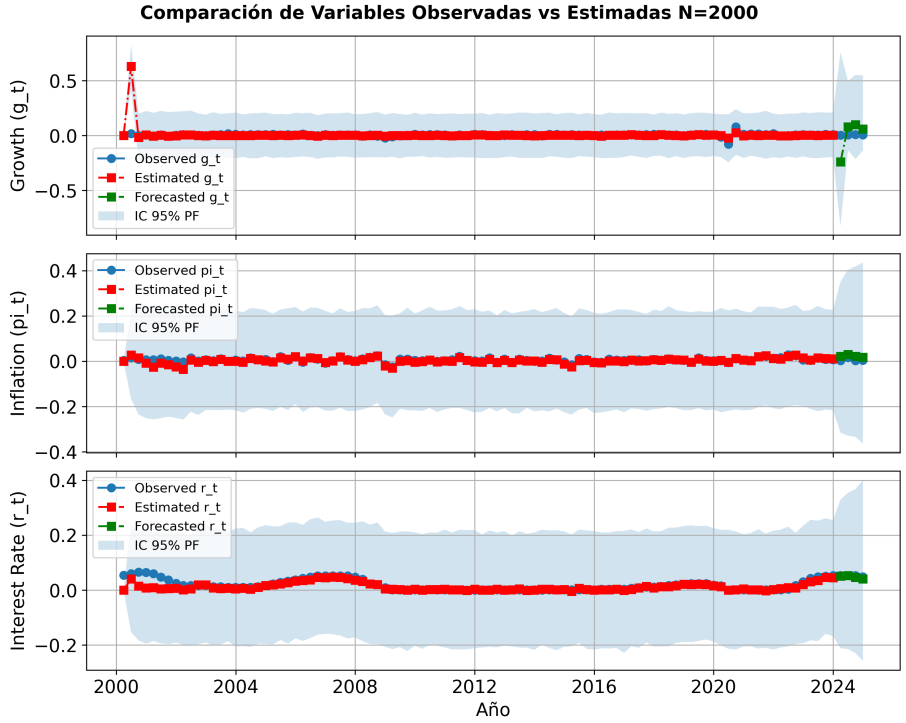


Figura 5.6: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 2000$ partículas, en el periodo del 2000 al 2024 de forma trimestral.

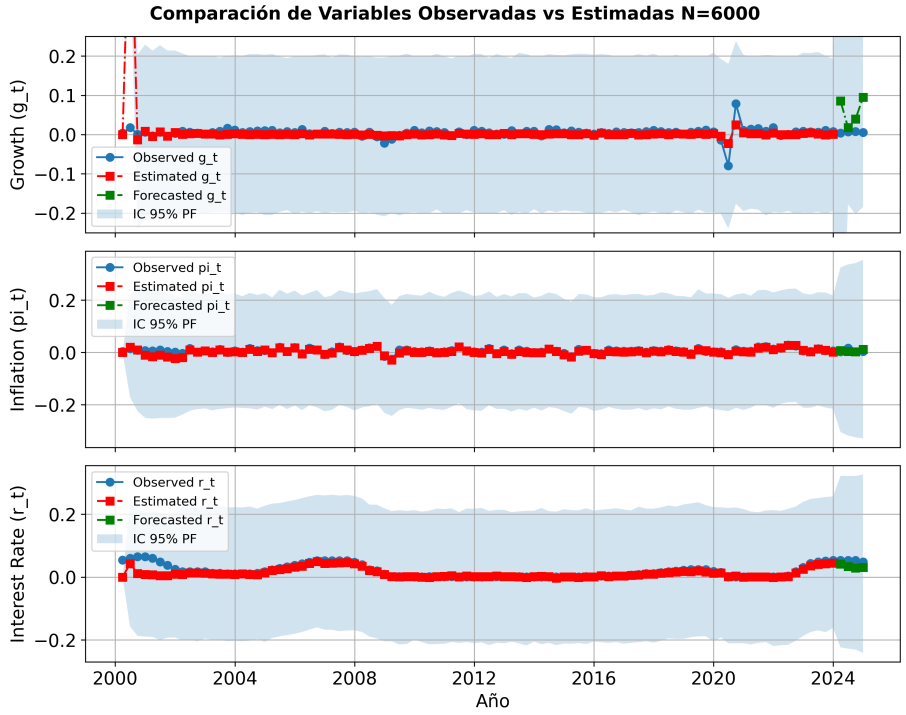


Figura 5.7: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 6000$ partículas, en el periodo del 2000 al 2024 de forma trimestral.

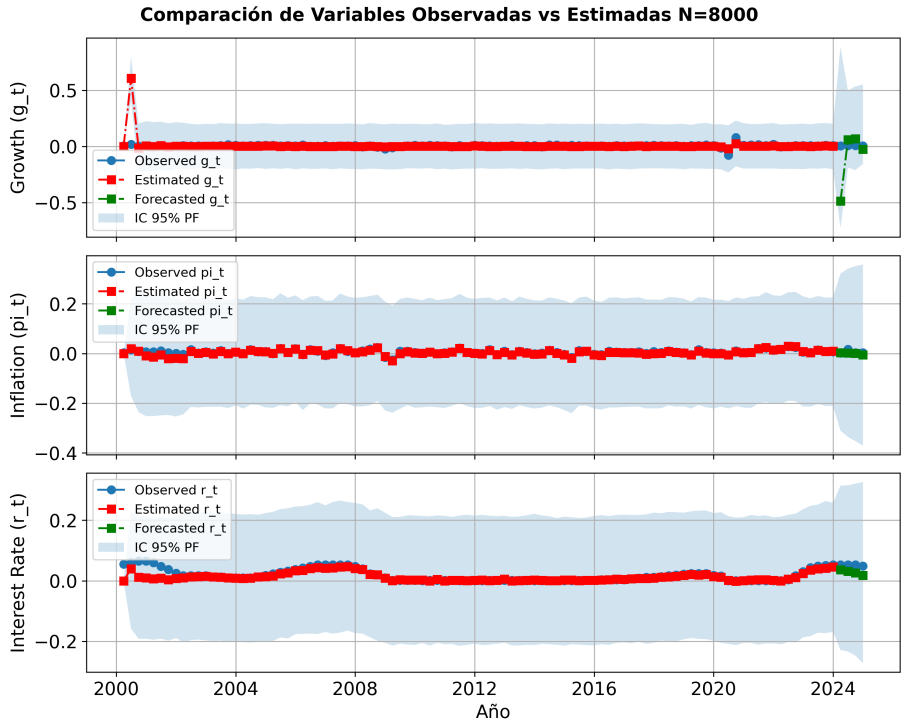


Figura 5.8: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 8000$ partículas, en el periodo del 2000 al 2024 de forma trimestral.

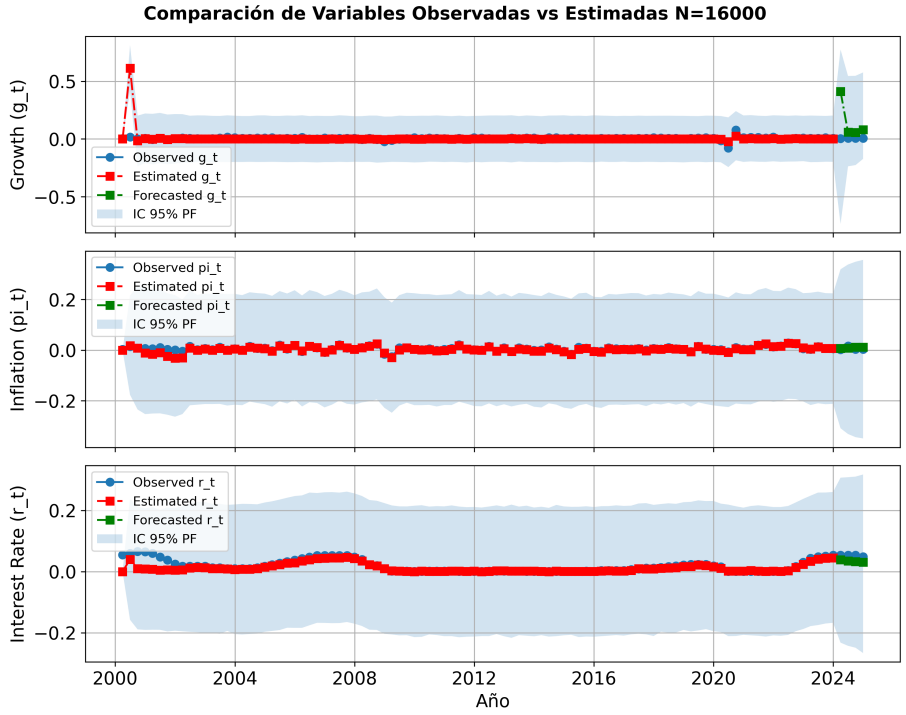


Figura 5.9: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 16000$ partículas, en el periodo del 2000 al 2024 de forma trimestral.

En las figuras, podemos observar que a medida que aumentamos las partículas; las estimaciones y predicciones mejoran considerablemente, lo cual demostraremos en la siguiente sección a través del análisis de errores.

5.3.1. Análisis de Errores

El desempeño de los valores puntuales en la predicción del modelo (correspondiente al año 2024) fueron evaluados por dos métricas estándar: el Error Absoluto Medio (MAE) y la Raíz del Error Cuadrático Medio (RMSE) para las tres variables de interés (g_t, π_t, r_t), bajo distintos tamaños del conjunto de partículas. Además, el ELPD para medir la capacidad predictiva en general del modelo a medida que se aumentan las partículas.

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
g_t	0.418	0.061	0.641	0.104	0.116	0.053	0.159	0.146
π_t	0.026	0.047	0.029	0.017	0.017	0.007	0.007	0.007
r_t	0.025	0.051	0.004	0.026	0.004	0.018	0.024	0.018

Tabla 5.2: Errores absolutos medios (MAE) fuera de muestra según el número de partículas utilizadas.

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
g_t	0.627	0.066	0.878	0.109	0.138	0.063	0.250	0.210
π_t	0.027	0.048	0.029	0.017	0.017	0.008	0.009	0.008
r_t	0.026	0.052	0.007	0.028	0.005	0.018	0.025	0.018

Tabla 5.3: Raíz del error cuadrático medio (RMSE) fuera de muestra según el número de partículas utilizadas.

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
ELPD Est	-1,799.94	-560.8	-59.45	53.24	116.14	123.62	190.51	191.43
ELPD Pred	-101.746	-9.546	-6.798	7.727	0.318	9.405	8.625	11.693

Tabla 5.4: Expected Log Predictive Density (ELPD) para estimación y predicción fuera de muestra según el número de partículas utilizadas.

Los resultados muestran que, en general, tanto el MAE y el RMSE disminuyen al incrementar el número de partículas, confirmando que un mayor número de muestras permite una mejor aproximación a la distribución posterior de los estados latentes. Esto es particularmente visible en las predicciones de la tasa de interés nominal (r_t), donde la reducción de errores es más marcada conforme aumenta N.

En el caso de la inflación (π_t), los errores se estabilizan desde valores intermedios ($N = 1000$ o $N = 2000$), lo que indica una menor sensibilidad a la cantidad de partículas. Sin embargo, se visualiza que para el crecimiento económico (g_t) pareciera que no hay disminución en los errores tan significativa.

Sin embargo, de forma integral, al visualizar los valores del ELPD va creciendo mientras aumentamos las particulas estabilizándose a partir de $N = 4000$, esto nos confirma que mientras aumenta la cantidad de partículas mayor es la capacidad predictiva del modelo. Además, los intervalos de credibilidad al 95 % contiene los valores reales de las tres variables.

Aunque, el uso de tamaños de muestra grandes de cierta manera minimiza errores de predicción, también es importante analizar los estimadores a utilizar, si bien, el uso de la media como estimador es una práctica común en métodos de filtrado, presenta ciertas limitaciones cuando se aplica a modelos no lineales y no gaussianos, ya que la distribución de las partículas puede ser asimétrica, con colas pesadas o influenciada por valores atípicos, lo cual sesga la media y reduce la precisión del estimador.

En particular, se observó que la media resultó sensible a valores extremos generadas por

partículas con pesos altos, sobre todo en el crecimiento económico (g_t). En consecuencia, se consideró oportuno explorar un estimador alternativo más robusto como la mediana. Este estimador tiene la ventaja de no verse afectado por valores extremos.

En la siguiente sección se examina el desempeño del modelo cuando se utiliza la mediana de partículas únicamente en la varible g_t como estimador principal, también exploramos la inclusión de otros estimadores como los percentiles 5-95.

5.4. Resultados con Estimación por la Mediana

En esta sección se presentan los resultados obtenidos al utilizar la mediana de las partículas como estimador mas robusto únicamente en la variable de crecimiento económico (g_t) para las variables no observadas del modelo DSGE, en lugar de la media. Esta elección busca eficientar la estimación puntual de la predicción frente a valores atípicos, asimetrías en la distribución de las partículas y valores inducidos por distribuciones sesgadas, lo cual es común en entornos no lineales y no gaussianos [15].

Se realizaron las estimaciones para distintos tamaños de partículas $N \in \{100, 200, 500, 1000, 2000, 6000, 8000, 16000\}$, presentando las comparaciones gráficas entre los valores observados y estimados de las tres variables clave del modelo: el crecimiento económico g_t, la inflación π_t y la tasa de interés nominal r_t.

A continuación, se presentan las figuras 5.10-5.17 para cada uno de los tamaños de las partículas.

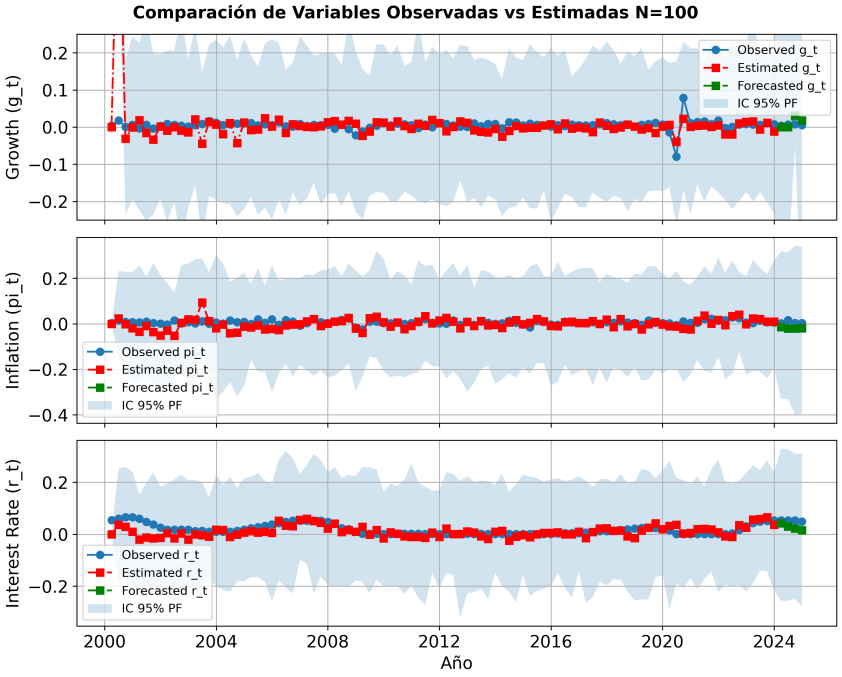


Figura 5.10: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 100$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral.

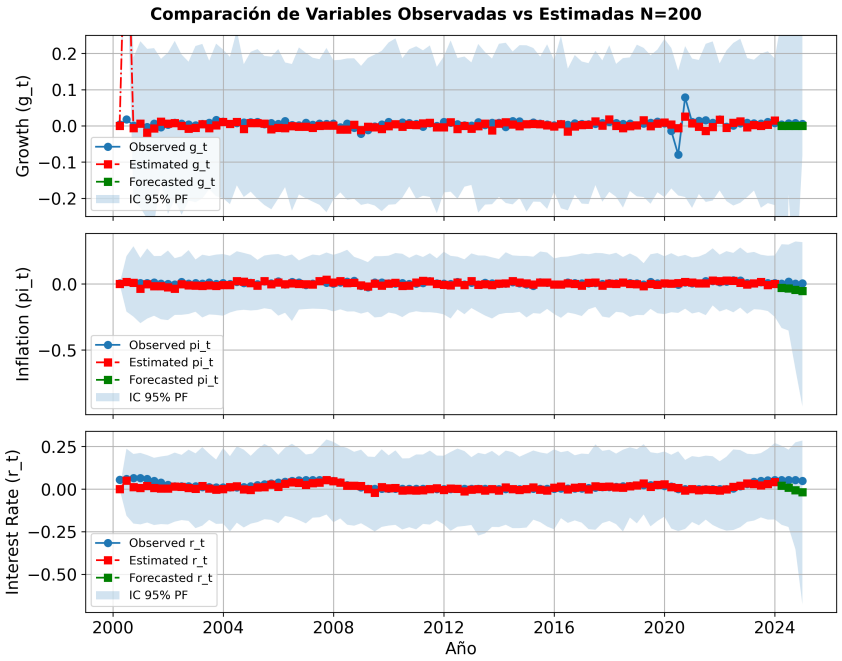


Figura 5.11: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 200$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral.

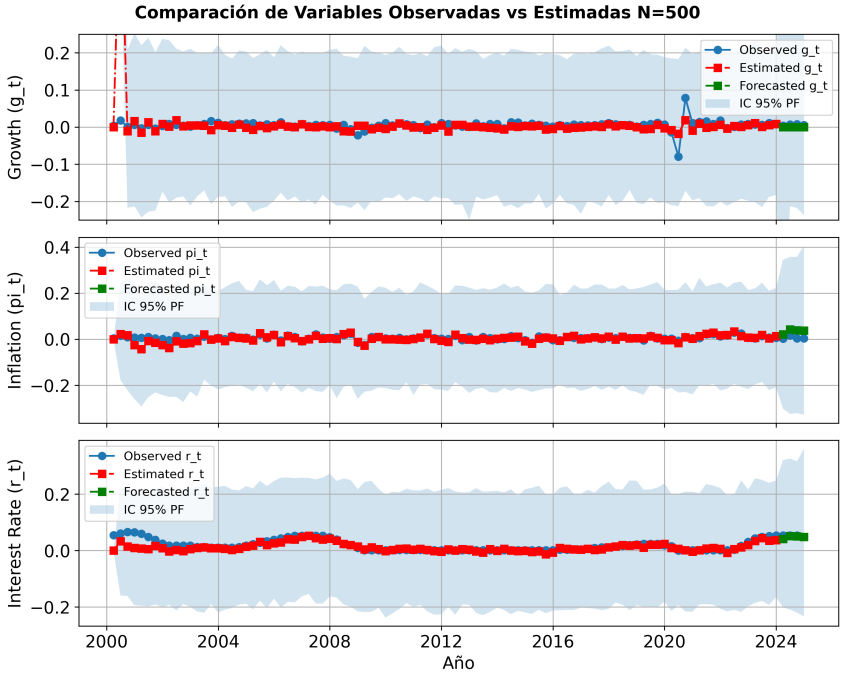


Figura 5.12: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 500$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral.

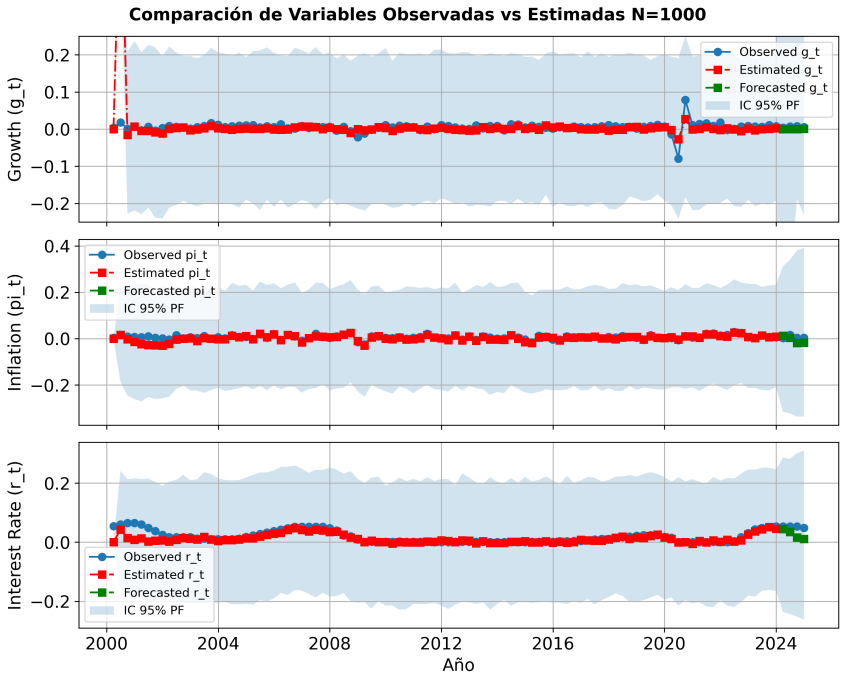


Figura 5.13: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 1000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral.

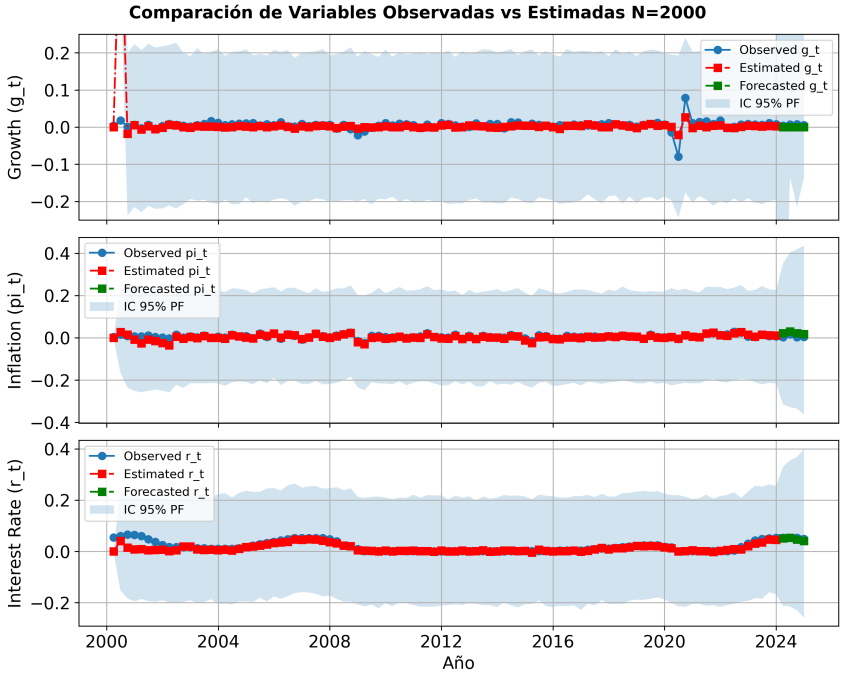


Figura 5.14: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 2000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral.

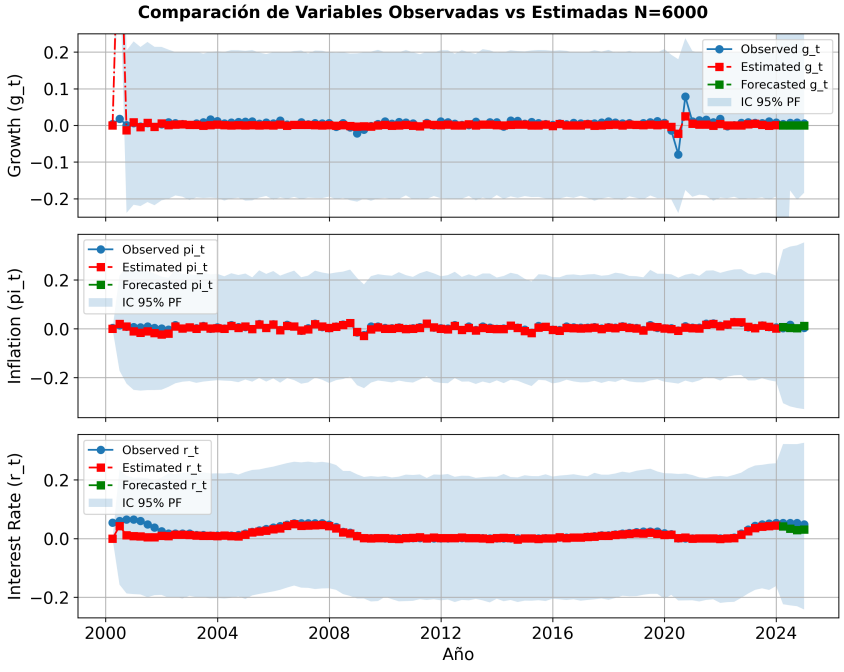


Figura 5.15: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 6000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral.

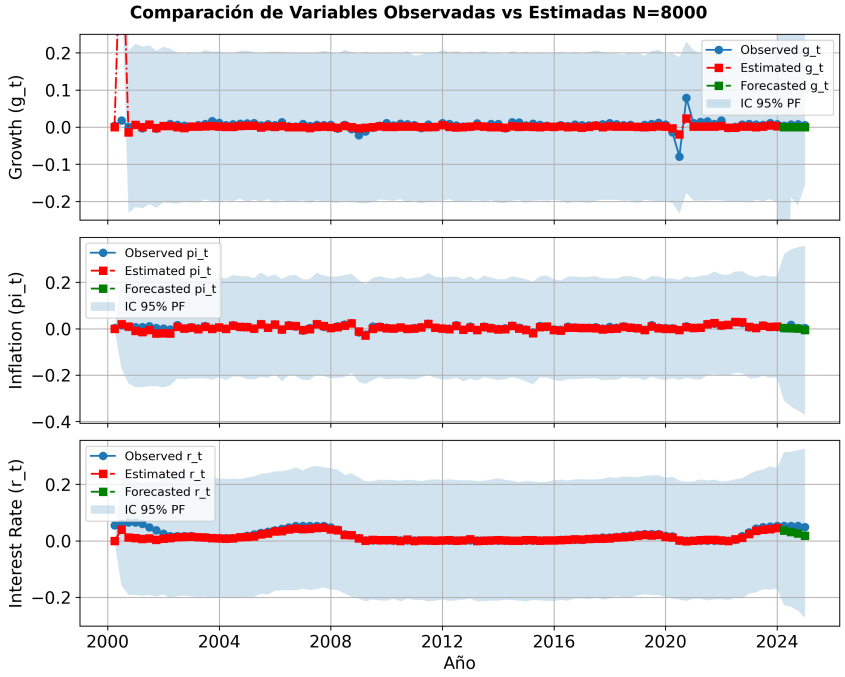


Figura 5.16: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 8000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral.

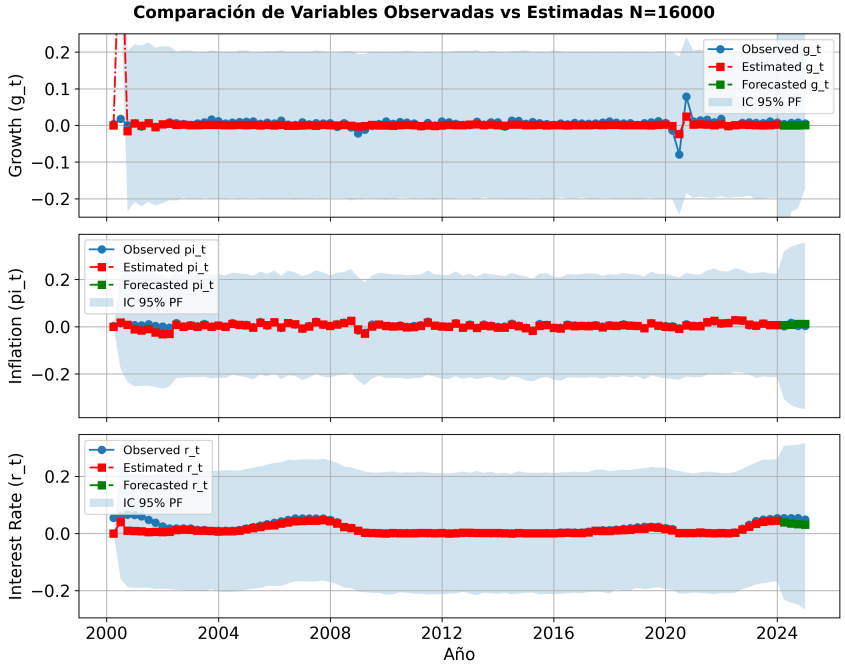


Figura 5.17: Comparación de las variables g_t, π_t y r_t vs los resultados estimados mediante el método de filtro de partículas con $N = 16000$ partículas (mediana para g_t), en el periodo del 2000 al 2024 de forma trimestral.

A continuación también se presenta resultados de la media, mediana, percentiles 5-95, en los tiempos de predicción que son el $Q1 - 2024$, $Q2 - 2024$, $Q3 - 2024$ y $Q4 - 2024$.

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
g_t real	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
g_t media	1.224	-0.040	-1.603	-0.075	-0.240	0.085	-0.488	0.412
g_t mediana	-1E^{-6}	-1E^{-6}	-2E^{-6}	-4E^{-7}	-2E^{-6}	-2E^{-6}	-2E^{-6}	-2E^{-6}
g_t p5	-0.93	-0.346	-0.167	-0.238	-0.262	-0.184	-0.223	-0.226
g_t p95	1.591	0.197	0.183	0.262	0.243	0.204	0.261	0.245

Tabla 5.5: Resumen numérico de media, mediana, percentiles 5-95 y real en $Q1 - 2024$

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
g_t real	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
g_t media	-0.161	0.040	0.057	-0.154	0.080	0.018	0.060	0.059
g_t mediana	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
g_t p5	-0.152	-0.061	-0.059	-0.123	-0.049	-0.042	-0.044	-0.057
g_t p95	0.184	0.285	0.301	0.369	0.330	0.363	0.346	0.364

Tabla 5.6: Resumen numérico de media, mediana, percentiles 5-95 y real en $Q2 - 2024$

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
g_t real	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
g_t media	0.230	0.083	0.667	0.103	0.100	0.039	0.067	0.058
g_t mediana	0.031	0.0001	0.0004	0.0003	0.0003	0.0002	0.0002	0.0003
g_t p5	-0.00003	-0.065	-0.043	-0.056	-0.045	-0.043	-0.054	-0.053
g_t p95	0.384	0.330	0.374	0.442	0.358	0.367	0.358	0.376

Tabla 5.7: Resumen numérico de media, mediana, percentiles 5-95 y real en $Q3 - 2024$

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
g_t real	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
g_t media	0.066	0.098	-0.243	0.085	0.061	0.095	-0.025	0.080
g_t mediana	0.018	0.0002	0.0002	0.0005	0.0003	0.0003	0.0003	0.0004
g_t p5	-0.144	-0.002	-0.073	-0.048	-0.028	-0.042	-0.032	-0.032
g_t p95	0.315	0.335	0.305	0.388	0.379	0.387	0.376	0.387

Tabla 5.8: Resumen numérico de media, mediana, percentiles 5-95 y real en $Q4 - 2024$

Los resultados obtenidos de las tablas 5.5-5.6 permiten observar una superioridad de la mediana como estimador puntual obtenidas mediante el filtro de partículas. En consecuencia, la utilización de la mediana constituye una alternativa adecuada para resumir la distribución posterior de los estados.

También es importante conocer la distribución de las partículas para ver si existe asimetrias entre las distintas variables $g_t.\pi_t$ y r_t. En el gráfico 5.18, se puede observar

valores extremos muy distantes entre el valor real en la variable g_t, comparado a las variables π_t y r_t.

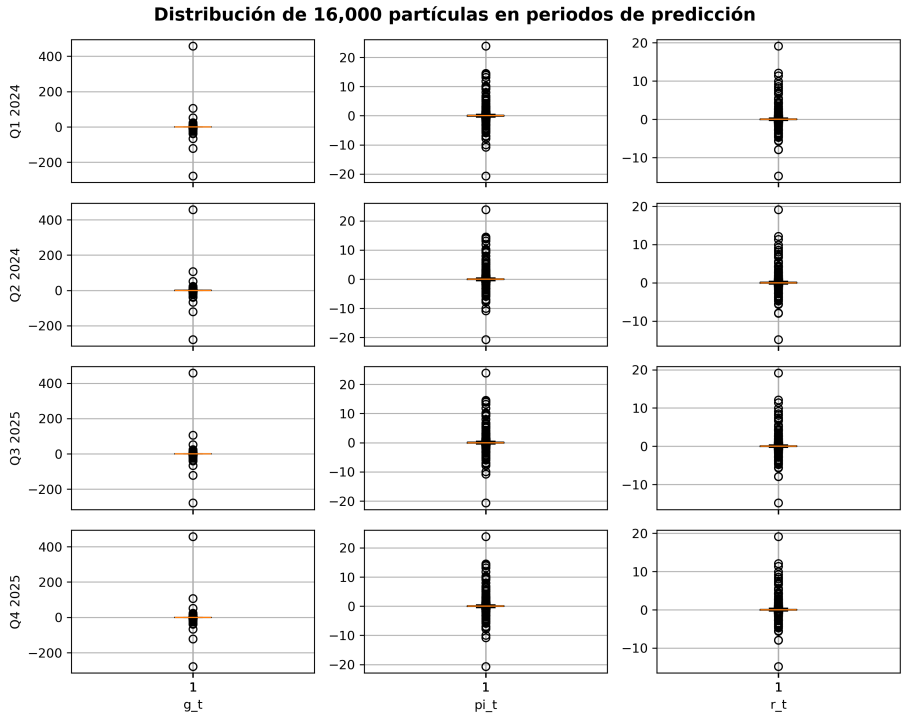


Figura 5.18: Distribución de $N = 16,000$ partículas de las variables g_t, π_t, r_t en los periodos de predicción Q1 2024, Q2 2024, Q3 2024 y Q4 2024.

5.4.1. Análisis de Errores

Para evaluar cuantitativamente el ajuste de las estimaciones puntuales se calcularon los errores MAE (Mean Absolute Error) y RMSE (Root Mean Squared Error). Las siguientes tablas muestran los resultados obtenidos

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
g_t	0.012	0.061	0.006	0.006	0.006	0.006	0.006	0.006
π_t	0.026	0.047	0.029	0.017	0.017	0.030	0.007	0.007
r_t	0.025	0.051	0.004	0.026	0.004	0.008	0.024	0.018

Tabla 5.9: MAE de las variables estimadas con mediana para distintos tamaños de partículas

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
g_t	0.014	0.066	0.006	0.006	0.006	0.006	0.006	0.006
π_t	0.027	0.048	0.029	0.017	0.017	0.032	0.009	0.008
r_t	0.026	0.052	0.007	0.028	0.005	0.008	0.025	0.018

Tabla 5.10: RMSE de las variables estimadas con mediana para distintos tamaños de partículas

A partir de los resultados gráficos y numéricos puede observarse que la utilización de la mediana genera estimaciones más estables y menos sensibles a valores extremos para el crecimiento económico g_t, donde se observa una notable mejora en los errores.

Ahora bien, en primer lugar, se demuestra que el filtro de partículas es un método para la estimación de modelos DSGE no lineales. El modelo fue capaz de replicar de manera consistente las dinámicas de crecimiento económico, inflación y tasa de interés nominal, y generar excelente predicciones, validando así la conveniencia del enfoque de filtrado secuencial en este tipo de modelos macroeconómicos.

En segundo lugar, los resultados obtenidos mediante el ELPD, demuestra que a medida aumentamos las partículas se obtiene una mejor capacidad predictiva del modelo. Además, el análisis comparativo entre los estimadores de la media y la mediana de partículas evidencia que la mediana ofrece una estimación más robusta, especialmente en variables sensibles a distribuciones sesgadas o colas pesadas. Este efecto es particularmente notable en la estimación del crecimiento económico g_t, donde los errores MAE y RMSE se reducen de manera considerable al utilizar la mediana como estimador central. Este hallazgo respalda la hipótesis de que, en entornos no lineales y no gaussianos, la mediana mitiga la influencia de series atípicas y mejora la capacidad predictiva del modelo.

Por último, desde una perspectiva económica, los resultados obtenidos muestran que el modelo DSGE no lineal logra capturar adecuadamente la interacción dinámica entre crecimiento económico, inflación y tasa de interés nominal, incluso bajo la presencia de shocks y rigideces nominales.

A pesar de los resultados obtenidos, esta investigación presenta algunas limitaciones que deben ser reconocidas. En primer lugar, aunque el modelo logró reproducir razonablemente las dinámicas observadas, existieron configuraciones del filtro de partículas donde ciertos resultados estimados mostraron sensibilidad al número de partículas utilizado. Particu-

larmente, con tamaños pequeños de partículas se observaron episodios de degeneración y pérdida de diversidad muestral, afectando la estabilidad de las estimaciones y la calidad de las predicciones.

Otra limitación importante es que los parámetros estructurales del modelo fueron calibrados utilizando valores estándares de la literatura, en lugar de estimarse completamente mediante un esquema bayesiano conjunto. Aunque esta decisión permitió concentrar el análisis en la dinámica de estados y en la implementación del filtro de partículas, implica que parte de la incertidumbre paramétrica no fue incorporada explícitamente en la inferencia.

Finalmente, debe reconocerse que la implementación computacional del filtro de partículas aplicada a modelos DSGE no lineales continúa siendo altamente demandante en términos de tiempo de procesamiento y estabilidad numérica. Aunque el presente trabajo demuestra la viabilidad del enfoque, todavía existen desafíos importantes relacionados con eficiencia computacional, degeneración de partículas y calibración óptima de hiperparámetros, los cuales representan oportunidades claras para futuras investigaciones.

Estos resultados refuerzan la idea de que, en contextos reales con alta volatilidad y distribuciones irregulares, el uso de filtros de partículas junto a estimadores robustos como la mediana resulta ser una estrategia sólida para la modelación macroeconómica moderna.

5.5. Consideraciones computacionales

Uno de los principales desafíos asociados a la implementación de modelos DSGE no lineales mediante métodos bayesianos corresponde al elevado costo computacional requerido para realizar inferencia conjunta sobre estados latentes y parámetros estructurales. A diferencia de los enfoques linealizados resueltos mediante Filtro de Kalman, los modelos DSGE no lineales requieren técnicas numéricas considerablemente más costosas, particularmente cuando se emplean métodos Secuenciales de Monte Carlo (SMC) o algoritmos MCMC acoplados con filtros de partículas.

Durante el desarrollo de esta investigación se exploró la posibilidad de extender el enfoque bayesiano no solo a la estimación de los estados latentes del modelo, sino también

a la estimación conjunta de los parámetros estructurales fijos del sistema DSGE, tales como el factor de descuento β, el parámetro de rigidez nominal ϕ, y las persistencias autorregresivas de los shocks exógenos ρ_a y ρ_z.

Con este propósito, se implementó experimentalmente el algoritmo *Particle Marginal Metropolis-Hastings* (PMMH), introducido formalmente por [41], el cual combina técnicas de Monte Carlo vía cadenas de Markov (MCMC) con filtros de partículas para aproximar simultáneamente la distribución posterior de estados y parámetros estructurales. Este enfoque pertenece a la familia de métodos *Particle Markov Chain Monte Carlo* (PMCMC), ampliamente utilizados en inferencia bayesiana para modelos no lineales y no gaussianos.

El procedimiento PMMH requiere que, en cada iteración de la cadena MCMC, se ejecute un filtro de partículas completo con el objetivo de obtener una aproximación insesgada de la verosimilitud marginal del modelo. En términos computacionales, esto implica que el costo total del algoritmo puede expresarse aproximadamente como:

$$\text{Costo Total} \approx N_{MCMC} \times N_{Partculas} \times T$$

donde N_{MCMC} representa el número de iteraciones de la cadena de Markov, $N_{Partculas}$ corresponde a la cantidad de partículas utilizadas en el filtro secuencial, y T denota el tamaño temporal de la muestra observada, considerando los tiempos reales de ejecución de esta tesis con N partículas y $T = 100$ observaciones, reflejados en la tabla 5.11.

N	**100**	**200**	**500**	**1000**	**2000**	**6000**	**8000**	**16000**
tiempo (horas)	0.068	0.171	0.328	0.986	1.265	4.879	5.689	9.915

Tabla 5.11: Tiempo de ejecución en horas para cada tamaño de N partículas

Por consiguiente, incluso configuraciones moderadas del algoritmo pueden generar tiempos de ejecución extremadamente elevados. Por ejemplo, una simulación con 50,000 iteraciones MCMC, 1,000 partículas y aproximadamente 100 observaciones implica la propagación y ponderación de millones de trayectorias estocásticas durante todo el proceso de ejecución.

El entorno de cómputo utilizado consistió en un servidor virtual de Google Cloud configurado con un procesador Intel Xeon de 2.60 GHz, 8 núcleos virtuales, 30 GB de

memoria RAM., Sistema operativo Linux Ubuntu, Entorno Python ejecutado mediante Anaconda.

El principal cuello de botella del algoritmo permaneció en la evaluación repetitiva de la verosimilitud mediante filtros de partículas secuenciales en cada iteración MCMC.

Estas dificultades computacionales coinciden con lo discutido por [53], quienes señalan que la estimación bayesiana de modelos DSGE no lineales representa uno de los problemas computacionalmente más exigentes dentro de la macroeconomía cuantitativa moderna.

Adicionalmente, [53] destacan que el desempeño de los algoritmos PMCMC depende críticamente de la varianza del estimador de verosimilitud generado por el filtro de partículas. Cuando el número de partículas es insuficiente, la cadena MCMC puede presentar problemas severos de convergencia; sin embargo, incrementar el número de partículas implica un aumento significativo del tiempo computacional total.

Debido a estas limitaciones prácticas, en esta tesis se decidió optar por una estrategia basada en calibración de parámetros estructurales tal como se detalló en el capítulo 4.

Capítulo 6

Conclusiones y trabajos futuros

La presente investigación desarrolló e implementó un esquema de estimación y predicción para modelos DSGE no lineales utilizando Filtros de Partículas, evitando la aproximación lineal y preservando la estructura no lineal original del sistema macroeconómico. La propuesta fue integrar una representación en espacio de estados en un modelo DSGE, junto con un método inferencia secuencial siendo el Filtro de Partículas Regularizado.

Uno de los principales aportes del trabajo consiste en demostrar que es posible estimar y predecir variables macroeconómicas relevantes sin recurrir a aproximaciones lineales. Esto permitió mantener explícitamente las relaciones no lineales derivadas de las ecuaciones estructurales del modelo DSGE.

La implementación realizada permitió modelar de manera conjunta la dinámica del crecimiento económico, la inflación y la tasa de interés nominal utilizando datos reales de la economía de Estados Unidos. Los resultados empíricos muestran que el modelo logra reproducir razonablemente la dinámica observada de la inflación y la tasa de interés nominal. Sin embargo, presenta mayores dificultades para capturar el comportamiento del crecimiento económico. Por tanto, el desempeño del modelo no es uniforme entre las variables analizadas, lo que evidencia ciertas limitaciones en su capacidad para representar las dinámicas macroecon ómicas observadas

Asimismo, se realizaron experimentos con distintos tamaños de partículas, observándose una mejora progresiva en la estabilidad de las trayectorias estimadas y en la aproximación de las distribuciones posteriores conforme aumentaba el número de partículas.

Otro hallazgo importante corresponde a la comparación entre los estimadores basados en la media y la mediana de partículas. Los resultados obtenidos mostraron que la mediana presentó un comportamiento más robusto frente a trayectorias extremas y distribuciones asimétricas generadas por la no linealidad del modelo. En términos estadísticos, los valores de MAE y RMSE fueron menores bajo el estimador basado en la mediana, sugiriendo

que los métodos robustos pueden ofrecer ventajas importantes en problemas de inferencia macroeconómica.

Adicionalmente, se incorporó métricas probabilísticas de evaluación predictiva, incluyendo densidad predictiva logarítmica y expected log predictive density (ELPD), permitiendo evaluar no solamente el error puntual de las predicciones, sino también la capacidad del modelo para asignar probabilidad a los valores realmente observados.

Desde el punto de vista metodológico, el trabajo evidencia que los Filtros de Partículas constituyen una alternativa viable frente a métodos basados en linealización y filtros gaussianos tradicionales, especialmente cuando el sistema económico presenta no linealidades importantes y distribuciones no gaussianas.

En términos generales, los resultados obtenidos respaldan la utilidad de combinar modelos DSGE no lineales con técnicas modernas de inferencia bayesiana y métodos secuenciales de Monte Carlo, proporcionando una herramienta flexible para el análisis macroeconómico dinámico y la generación de predicciones.

6.1. Trabajos futuros

Daremos continuación algunos trabajos futuros que pueden ser posibles tomando como base la presente investigación:

Estimación bayesiana de parámetros estructurales con independencia de tiempo: Actualmente, los parámetros del modelo fueron seleccionados con base en la literatura. La idea sería implementar un esquema completo de inferencia bayesiana tipo PMMH (Particle Marginal Metropolis-Hastings) para estimar los parámetros conjuntamente con los estados, aplicando medidas bayesianas como el WAIC o LOO para su validación.

Comparación con otros métodos: Sería de interés comparar el desempeño del Filtro de Partículas Regularizado con otros métodos avanzados de inferencia, como el Filtro de Partículas (SIR), el Filtro de Partículas (Resample-Move) y el Filtro de Partículas Auxiliar (APF).

Análisis económico: Realizar un análisis de impulse responses, es decir, como responden las variables macroeconómicas ante diferentes tipos de perturbaciones estructurales,

como los shocks de productividad, shocks de preferencia, etc.

Bibliografía

[1] Frank Smets and Raf Wouters. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5):1123–1175, 2003.

[2] Jesús Fernández-Villaverde. A baseline DSGE model. *Manuscript, University of Pennsylvania*, 2006.

[3] Peter N Ireland. Technology shocks in the new keynesian model. *The Review of Economics and Statistics*, 85(4):923–936, 2003.

[4] Kenneth L Judd. *Numerical Methods in Economics*. MIT Press, 1998.

[5] Alfio Quarteroni and Alberto Valli. *Numerical Approximation of Partial Differential Equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer, 2008.

[6] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton, FL, 3rd edition, 2013. ISBN 9781439840955.

[7] Satish Rao and David Tse. *Conditional Probability*. Fall 2009. Lecture notes.

[8] J.M. Bernardo and A.F.M. Smith. *Bayesian Theory*. John Wiley & Sons, 1994.

[9] James D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, NJ, 1994. ISBN 9780691042893.

[10] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time Series Analysis: Forecasting and Control*. Wiley Series in Probability and Statistics. John Wiley & Sons, 5th edition, 2015. ISBN 978-1-118-67502-1.

[11] J. Durbin and S. J. Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, 2nd edition, 2012.

[12] Christian P. Robert and George Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer, New York, 2nd edition, 2010.

[13] Dennis Wackerly, William Mendenhall, and Richard L. Scheaffer. *Mathematical Statistics with Applications*. Cengage Learning, 7th edition, 2014.

[14] Nicolas Chopin and Omiros Papaspiliopoulos. *An Introduction to Sequential Monte Carlo*. Springer Series in Statistics. Springer, 2020.

[15] Gary Koop. *Bayesian Econometrics*. Wiley Series in Probability and Statistics. Wiley, 2010. ISBN 978-0-470-74453-6.

[16] Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics*. Wiley Series in Probability and Statistics. Wiley, 2nd edition, 2009. ISBN 978-0-470-12990-6.

[17] Aki Vehtari, Daniel Simpson, Andrew Gelman, Yuling Yao, and Jonah Gabry. Approximate leave-future-out cross-validation for bayesian time series models. *Journal of Statistical Computation and Simulation*, 90(14):2499–2523, 2020.

[18] Mike West and Jeff Harrison. *Bayesian Forecasting and Dynamic Models (Springer Series in Statistics)*. Springer-Verlag, February 1997.

[19] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice-Hall, 1979.

[20] Arnaud Doucet, Nando de Freitas, and Neil Gordon. Sequential monte carlo methods in practice. In *Sequential Monte Carlo methods in practice*, pages 3–14. Springer, 2001.

[21] Giovanni Petris, Sonia Petrone, and Patrizia Campagnoli. *Dynamic Linear Models with R*. Springer, New York, 2009. ISBN 978-0-387-77237-0.

[22] Sahil Rajesh Dhayalkar. Particle filter made simple: A step-by-step beginner-friendly guide. *arXiv preprint arXiv:2511.01281*, 2025.

[23] N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEEE Proceedings F (Radar and Signal Processing)*, 140(2):107–113, April 1993. doi: 10.1049/ip-f-2.1993.0015.

[24] Michael K Pitt and Neil Shephard. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, 94(446):590–599, 1999.

[25] Cédric Musso, Nadia Oudjane, and François Le Gland. Improving regularized particle filters. *Sequential Monte Carlo Methods in Practice*, pages 247–271, 2001.

[26] Jane Liu and Mike West. Combined parameter estimation and state tracking in simulation-based filtering. In *Sequential Monte Carlo methods in practice*, pages 197–223. Springer, New York, NY, 2001.

[27] Jesús Fernández-Villaverde and Juan Rubio-Ramírez. Estimating macroeconomic models: A likelihood approach. *Review of Economic Studies*, 2007.

[28] Jesús Fernández-Villaverde and Juan F. Rubio-Ramírez. Solution and estimation methods for DSGE models. *NBER Working Paper No. 14677*, 2009.

[29] Fabio Canova. *Methods for Applied Macroeconomic Research*. Princeton University Press, 2007.

[30] Matthias Burgert, Oliver de Groot, Richard Harrison, et al. The european central bank's new macroeconomic model: New area-wide model iii. Occasional paper series, European Central Bank, 2024.

[31] Roberta Cardani, Olga Croitorov, Massimo Giovannini, Philipp Pfeiffer, Marco Ratto, and Lukas Vogel. The global macroeconomic effects of covid-19: A dsge approach. *Economic Modelling*, 108:105751, 2022.

[32] Network for Greening the Financial System. Climate macroeconomic modelling handbook. Technical document, NGFS, October 2024.

[33] Cristian Andres Cruz-Torres. *Modelos Dinâmicos Estocásticos de Equilíbrio Geral com Choques Heterocedásticos*. Tese de doutorado, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brasil, 2015.

[34] S. An and F. Schorfheide. Bayesian analysis of DSGE models. *Econometric Reviews*, 26(2-4):113–172, 2007.

[35] Jesús Fernández-Villaverde. The econometrics of DSGE models. *SERIEs*, 1(1-2):3–49, 2010. doi: 10.1007/s13209-009-0004-5.

[36] Anelí Bongers, Benedetto Molinari, and José L. Torres. Computers, programming and dynamic general equilibrium macroeconomic modeling. *MPRA Paper*, (112505), 2022. URL `https://mpra.ub.uni-muenchen.de/112505/`.

[37] Harald Uhlig. A toolkit for analyzing nonlinear dynamic stochastic models easily. *Computing in Economics and Finance*, 1999(213), 1999.

[38] Olivier Jean Blanchard and Charles M Kahn. The solution of linear difference models under rational expectations. *Econometrica*, 48(5):1305–1311, 1980.

[39] Paul Klein. Using the generalized schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control*, 24(10):1405–1423, 2000.

[40] Marco Del Negro and Frank Schorfheide. *Dynamic Macroeconomics*. Princeton University Press, 2015.

[41] Christophe Andrieu, Arnaud Doucet, and Roman Holenstein. Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):269–342, 2010.

[42] Peter N Ireland. A method for taking models to the data. *Journal of Economic Dynamics and Control*, 28(6):1205–1226, 2004.

[43] Jordi Galí. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications*. Princeton University Press, 2015.

[44] Finn E. Kydland and Edward C. Prescott. Time to build and aggregate fluctuations. *Econometrica*, 50(6):1345–1370, 1982.

[45] John B Taylor. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195–214, 1993.

[46] Michael Woodford. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, 2003.

[47] O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Springer, 2007.

[48] Jesús Fernández-Villaverde and Juan F. Rubio-Ramírez. Estimating dynamic equilibrium models with stochastic volatility. *Journal of Econometrics*, 142(2):755–789, 2008.

[49] Andreas Hornstein. Introduction to the new keynesian phillips curve. *Economic Quarterly-Federal Reserve Bank of Richmond*, 94(4):301–309, 2008.

[50] Jordi Galí and Mark Gertler. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44(2):195–222, 1999.

[51] Frank Smets and Raf Wouters. Shocks and frictions in us business cycles: A bayesian DSGE approach. *American Economic Review*, 97(3):586–606, 2007.

[52] Sebastián Garcés. Un modelo DSGE para la economía chilena: Calibración y simulación. Tesis de magíster en economía, Pontificia Universidad Católica de Chile, Santiago de Chile, 2013. URL `https://cliolab.economia.uc.cl/docs/tesis/tesis-sgarces-2013.pdf`.

[53] Edward P. Herbst and Frank Schorfheide. *Bayesian Estimation of DSGE Models*. Princeton University Press, Princeton, New Jersey, 2015.

[54] Alexandro Mayoral Terán. *Modelos Dinámicos Lineales Aplicados a Series de Tiempo*. 2013.

[55] Walter R. Gilks and Carlo Berzuini. Following a moving target—monte carlo inference for dynamic bayesian models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(1):127–146, 2001. doi: 10.1111/1467-9868.00280.

[56] Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez, Thomas J. Sargent, and Mark W. Watson. Abcs (and ds) of understanding vars. *American Economic Review*, 100(2): 1–20, 2010.

[57] Chang-Jin Kim and Charles R Nelson. Dynamic linear models with markov-switching. *Journal of Econometrics*, 60(1-2):1–22, 1994.

[58] Edward Herbst and Frank Schorfheide. Bayesian estimation of DSGE models. *Princeton University Press*, 2015.

[59] Frank Schorfheide. Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, 15(6):645–670, 2000.

[60] Lawrence J Christiano, Martin Eichenbaum, and Charles L Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45, 2005.

[61] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):721–741, 1984.

[62] Siddhartha Chib and Edward Greenberg. Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4):327–335, 1995.

[63] Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, 3rd edition, 2008.

[64] Carlos Márquez. *Solución Numérica de Modelos Dinámicos de Equilibrio General con Esperanzas Racionales: Un Enfoque mediante Elementos Finitos*. Tesis doctoral, Universidad de Chile, 2007.

[65] Richard L. Burden and J. Douglas Faires. *Numerical Analysis*. Brooks/Cole Cengage Learning, 9th edition, 2011.

[66] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27(5):1413–1432, 2017.

[67] Sumio Watanabe. Asymptotic equivalence of bayes cross validation and widely applicable information criterion. *Journal of Machine Learning Research*, 11:3571–3594, 2010.

[68] Andres Robayo Cesar Mora, Christian Holgun. Funcion de lagrange. 2023.

[69] Adrien Auclert, Matthew Rognlie, and Ludwig Straub. The microeconomic foundations of macroeconomic persistence. *The Quarterly Journal of Economics*, 138(4):2111–2160, 2023.

[70] Barbara Annicchiarico, Stefano Carattini, Carolyn Fischer, and Garth Heutel. Environmental macroeconomics: Environmental dsge models and climate policy. *Annual Review of Resource Economics*, 16, 2024.

[71] Edward Herbst and Frank Schorfheide. Sequential Monte Carlo sampling for DSGE models. *Journal of Applied Econometrics*, 34(1):1–23, 2019.

[72] Jesús Fernández-Villaverde, Albert Marbà-Sánchez, and Juan F Rubio-Ramírez. Estimating macroeconomic models with particle filters: An application to non-linear fiscal policy. Working Paper w31524, National Bureau of Economic Research, 2023.

Apéndice A

La función de Lagrange en el Problema del Consumidor

El objetivo de este apéndice es explicar con mayor detalle el uso de la función de Lagrange dentro del problema de optimización intertemporal del problema del consumidor. En los modelos DSGE, los agentes no eligen libremente sus decisiones económicas, sino que maximizan una función objetivo sujetos a restricciones presupuestarias, tecnológicas o acumulación de capital.

Para poder solucionar este tipo de problemas de optimización como lo es en los modelos DSGE, acudimos a métodos matemáticos para su resolución, como es en este caso las funciones de Lagrange. Consideremos las siguientes definiciones y teorema, según ([68]).

Definición A.1. *Un problema de optimización con restricciones de igualdad se define como*

$$\max_{\boldsymbol{x}\in S} f(\boldsymbol{x}) \tag{A.1}$$

donde $f : D \subset \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{g} = (g_1, ..., g_m)$, *con* $m < n$ *y* $S = \boldsymbol{x} \in D : g(\boldsymbol{x}) = \boldsymbol{b}$, *donde* $\boldsymbol{b} = (b_1, ..., b_m)$. *A* S *se le conoce como el conjunto de* m *restricciones.*

Definición A.2. *La función Lagrangiana de A.1 es*

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \lambda(b - g(\boldsymbol{x})) \tag{A.2}$$

donde $\boldsymbol{x} = (x_1, ..., x_n)$, *y* $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_m)$

Teorema A.1. *(Teorema de Lagrange) Si* $\boldsymbol{x_0} \in S$ *es solución de A.1, entonces existe*

un único vector $\boldsymbol{\lambda} = (\lambda_1^{(0)}, ..., \lambda_m^{(0)})$ *tal que* $(\boldsymbol{x_0}, \boldsymbol{\lambda_0})$ *es un punto crítico de la función Lagrangiana de A.1, es decir*

$$\begin{aligned} \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x_0}, \boldsymbol{\lambda_0}) = 0 \\ \nabla_{\boldsymbol{\lambda}} \mathcal{L}(\boldsymbol{x_0}, \boldsymbol{\lambda_0}) = 0 \end{aligned} \tag{A.3}$$

Definición A.3. *Si* $(\boldsymbol{x_0}, \boldsymbol{\lambda_0})$ *es un punto crítico de la función Lagrangiana, entonces se dice que* $\boldsymbol{x_0}$ *es un punto crítico de la función* f *en* S*, y que* $\lambda_1^{(0)}, ..., \lambda_m^{(0)}$ *son los multiplicadores de Lagrange asociados.*

Ahora bien, aplicando las definiciones y teorema anterior al problema de optimización asociado al modelo DSGE del problema del consumidor

$$\max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{A.4}$$

sujeto a la restricción de recursos

$$c_t + k_{t+1} = e^{z_t} f(k_t) + (1-\delta)k_t \Rightarrow 0 = e^{z_t} f(k_t) + (1-\delta)k_t - c_t - k_{t+1}$$

Tendriamos la función de Lagrange asociado a la ecuación A.4

$$\mathcal{L}(c_t, k_{t+1}, \lambda_t) = \sum_{t=0}^{\infty} \beta^t \left[u(c_t) + \lambda_t \left(e^{z_t} f(k_t) + (1-\delta)k_t - c_t - k_{t+1} \right) \right] \tag{A.5}$$

Aplicando las derivadas correspondientes para encontrar los puntos críticos que optimizan la función objetivo

1. Derivada respecto a λ_t (restricción de recursos):

$$\frac{\partial \mathcal{L}}{\partial \lambda_t} = \beta^t \left[e^{z_t} f(k_t) + (1-\delta)k_t - c_t - k_{t+1} \right] = 0 \tag{A.6}$$

2. Derivada respecto a c_t:

$$\frac{\partial \mathcal{L}}{\partial c_t} = \beta^t \left[u'(c_t) - \lambda_t \right] = 0 \quad \Rightarrow \quad \lambda_t = u'(c_t) \tag{A.7}$$

3. Derivada respecto a k_{t+1} (con esperanza condicional):

$$\frac{\partial \mathcal{L}}{\partial k_{t+1}} = -\beta^t \lambda_t + \beta^{t+1} \mathbb{E}_t \left[\lambda_{t+1} \left(e^{z_{t+1}} f'(k_{t+1}) + 1 - \delta \right) \right] = 0 \tag{A.8}$$

Sustituyendo la expresión de λ_t, se obtiene la condición de Euler:

$$u'(c_t) = \beta \, \mathbb{E}_t \left[u'(c_{t+1}) \left(e^{z_{t+1}} f'(k_{t+1}) + 1 - \delta \right) \right] \tag{A.9}$$

Esta ecuación resume la decisión intertemporal fundamental del consumidor. El lado izquierdo representa el beneficio marginal de consumir una unidad adicional hoy. El lado derecho representa el valor esperado descontado de ahorrar esa unidad, transformarla en capital futuro y consumir los recursos adicionales generados en el siguiente periodo.

Por tanto, la ecuación de Euler establece que, en equilibrio, el consumidor es indiferente entre consumir hoy o ahorrar para consumir mañana. Si el rendimiento esperado del capital aumenta, el consumidor tiene incentivos a reducir consumo presente y aumentar ahorro. Si la utilidad marginal futura esperada disminuye, el incentivo a ahorrar se reduce.

De esta manera, el método de Lagrange no solo permite obtener las condiciones matemáticas de optimalidad, sino que también proporciona una interpretación económica clara de las decisiones dinámicas del agente representativo dentro del modelo DSGE.

Apéndice B

Cuadratura de Gauss-Hermite y Cálculo de Esperanzas

En los modelos DSGE no lineales, se pueden encontrar ecuaciones de equilibrio que incorporan **esperanzas condicionales no lineales**. Estas esperanzas suelen tomar la forma:

$$\mathbb{E}_t \left[h(x_{t+1}, \varepsilon_{t+1}) \right], \tag{B.1}$$

donde ε_{t+1} representa un shock estocástico con distribución conocida. Lo más frecuente en la literatura es asumir que los choques son normales: $\varepsilon_{t+1} \sim \mathcal{N}(0, \sigma^2)$. La naturaleza no lineal de $h(\cdot)$ impide una solución analítica directa, por lo que se recurre a métodos numéricos para aproximar dichas esperanzas.

Uno de los métodos más utilizados en la literatura es la cuadratura de Gauss-Hermite [4], que permite evaluar con precisión integrales cuando los shocks son gaussianos.

Este método aproxima integrales del tipo:

$$\int_{-\infty}^{\infty} h(x)\, e^{-x^2} dx \approx \sum_{i=1}^{n} w_i\, h(x_i), \tag{B.2}$$

donde:

1. x_i son los **nodos** (raíces del polinomio de Hermite de grado n),
2. w_i son los **pesos** asociados a cada nodo,
3. n es el **número de puntos de cuadratura**, que define la precisión del método.

Para una densidad normal estándar $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, se aplica el cambio de variable $x = \sqrt{2}z$ y se obtiene:

$$\int_{-\infty}^{\infty} h(x)\phi(x)dx = \int_{-\infty}^{\infty} h(\sqrt{2}z)\, \frac{1}{\sqrt{\pi}} e^{-z^2} dz \approx \sum_{i=1}^{n} w_i\, h(\sqrt{2}z_i). \tag{B.3}$$

Este esquema transforma la integral en una suma ponderada que puede ser resuelta computacionalmente.

Aplicación en el Modelo DSGE

En nuestro modelo DSGE no lineal, una de las esperanzas clave surge de la condición de Euler para el consumo. Específicamente, la ecuación:

$$\mathbb{E}_t\left[\frac{a_{t+1}}{c_{t+1}z_{t+1}\pi_{t+1}}\right] = \int\int \frac{a_{t+1}}{c_{t+1}z_{t+1}\pi_{t+1}}\, p(a_{t+1}, z_{t+1}|a_t, z_t)\, da_{t+1}dz_{t+1}, \tag{B.4}$$

donde se considera que los shocks $\varepsilon_{a,t+1}$ y $\varepsilon_{z,t+1}$ son independientes y siguen una distribución normal estándar:

$$\varepsilon_{a,t+1} \sim \mathcal{N}(0,1), \quad \varepsilon_{z,t+1} \sim \mathcal{N}(0,1)$$

requiere resolver una esperanza condicional respecto a los shocks estocásticos z_{t+1} y π_{t+1}, lo cual se lleva a cabo mediante este método cuadrático.

Dado que se busca aplicar la regla de cuadratura de Gauss-Hermite para aproximar esta doble integral, realizamos el cambio de variable estándar:

$$x_a = \frac{\varepsilon_{a,t+1}}{\sqrt{2}}, \qquad x_z = \frac{\varepsilon_{z,t+1}}{\sqrt{2}}, \tag{B.5}$$

Esto nos permite reescribir las ecuaciones de los shocks como:

$$a_{t+1} = e^{(1-\rho_a)\ln a_s + \rho_a \ln a_t + \sqrt{2}x_a}, \tag{B.6}$$

$$z_{t+1} = e^{\ln z_s + \sqrt{2}x_z}. \tag{B.7}$$

Ahora podemos aplicar Gauss-Hermite para aproximar la integral:

$$E_t\left[\frac{a_{t+1}}{c_{t+1}}\frac{1}{z_{t+1}}\frac{1}{\pi_{t+1}}\right] \approx \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \frac{e^{(1-\rho_a)\ln a_s + \rho_a \ln a_t + \sqrt{2}x_i}}{c_{t+1}^{(i,j)}} \frac{1}{e^{\ln z_s + \sqrt{2}x_j}} \frac{1}{\pi_{t+1}^{(i,j)}}. \tag{B.8}$$

donde:

1. x_i, x_j: nodos de Gauss-Hermite.

2. w_i, w_j: pesos asociados a cada nodo.

3. $\pi_{t+1}^{(i,j)}$: Se interpola mediante elemento finico (ver C).

4. $c_{t+1}^{(i,j)}$: Se interpola mediante elemento finico (ver C).

Esta formulación permite resolver mediante métodos numéricos la esperanza racional mediante un conjunto finito de evaluaciones, haciendo viable la implementación del modelo sin necesidad de linearizar el sistema.

A continuación, se presentan algunos ejemplos de pares de nodos (x_i, x_j) y sus pesos conjuntos $w_i w_j$, utilizados en la aproximación de la integral bidimensional mediante la cuadratura de Gauss-Hermite con dos variables gaussianas independientes.

Para $n = 3$:

(x_i, x_j)	(w_i, w_j)	$w_i w_j$
$(-1.22474, -1.22474)$	$(0.29541, 0.29541)$	0.08720
$(-1.22474, 0.00000)$	$(0.29541, 1.18164)$	0.34880
$(-1.22474, 1.22474)$	$(0.29541, 0.29541)$	0.08720
$(0.00000, -1.22474)$	$(1.18164, 0.29541)$	0.34880
$(0.00000, 0.00000)$	$(1.18164, 1.18164)$	1.39626
$(0.00000, 1.22474)$	$(1.18164, 0.29541)$	0.34880
$(1.22474, -1.22474)$	$(0.29541, 0.29541)$	0.08720
$(1.22474, 0.00000)$	$(0.29541, 1.18164)$	0.34880
$(1.22474, 1.22474)$	$(0.29541, 0.29541)$	0.08720

Para $n = 4$:

(x_i, x_j)	(w_i, w_j)	$w_i w_j$
$(-1.65068, -1.65068)$	$(0.08131, 0.08131)$	0.00661
$(-1.65068, -0.52465)$	$(0.08131, 0.80491)$	0.06545
$(-1.65068, 0.52465)$	$(0.08131, 0.80491)$	0.06545
$(-1.65068, 1.65068)$	$(0.08131, 0.08131)$	0.00661
$\vdots$	$\vdots$	$\vdots$
$(1.65068, 1.65068)$	$(0.08131, 0.08131)$	0.00661

Para $n = 5$:

(x_i, x_j)	(w_i, w_j)	$w_i w_j$
$(-2.02018, -2.02018)$	$(0.01995, 0.01995)$	0.00040
$(-2.02018, -0.95857)$	$(0.01995, 0.39361)$	0.00785
$(-2.02018, 0.00000)$	$(0.01995, 0.94531)$	0.01885
$(-2.02018, 0.95857)$	$(0.01995, 0.39361)$	0.00785
$(-2.02018, 2.02018)$	$(0.01995, 0.01995)$	0.00040
$\vdots$	$\vdots$	$\vdots$
$(2.02018, 2.02018)$	$(0.01995, 0.01995)$	0.00040

Los valores de los nodos y pesos pueden obtenerse con la función `hermgauss` de `numpy`, y se construyen con:

```
from numpy.polynomial.hermite import hermgauss
import numpy as np


nodes, weights = hermgauss(n)
grid = np.array(np.meshgrid(nodes, nodes)).T.reshape(-1, 2)
W = np.outer(weights, weights).flatten()
```

En conclusión, el método de Gauss-Hermite constituye una herramienta fundamental para resolver de forma eficiente las esperanzas condicionales presentes en modelos DSGE no lineales. Su implementación, combinada con técnicas de interpolación como se describe en el Apéndice C, permite mantener la la estructura original del modelo, sin incurrir a linealizaciones.

Apéndice C

Interpolación por Elementos Finitos en Modelos DSGE

En el contexto de modelos DSGE resueltos numéricamente, la cuadratura de Gauss-Hermite permite aproximar las esperanzas condicionales, pero requiere evaluar funciones de variables endógenas c_{t+1}, π_{t+1}, etc., en puntos no observados directamente en la malla simulada. Para abordar esto, se recurre a técnicas de interpolación. [5]

Una de las más robustas es la interpolación por elementos finitos, también conocida como interpolación multilineal en mallas estructuradas. Esta técnica se basa en construir funciones aproximantes definidas localmente sobre una malla rectangular de nodos preestablecidos, y es especialmente útil en espacios de dimensión baja (como los dos estados utilizados en este modelo).

Sea $\mathcal{D} \subset \mathbb{R}^2$ un dominio rectangular cubierto por una malla regular de nodos (a_k, z_j). Supongamos que una función $f(a, z)$ está definida en los puntos de la malla. Queremos aproximar f en un punto (a^*, z^*) que no pertenece exactamente a la malla.

La interpolación lineal en dos dimensiones se define como:

$$f(a^*, z^*) \approx \sum_{k=1}^{K} \theta_k \, \Psi_k(a^*, z^*), \tag{C.1}$$

donde:

1. $\Psi_k(a, z)$ son funciones base (por ejemplo, funciones forma tipo bilineal),
2. θ_k son los coeficientes (valores de la función conocida en los nodos),
3. La suma se extiende sobre los nodos vecinos al punto (a^*, z^*).

Las funciones base más comunes para mallas estructuradas son productos de funciones lineales unidimensionales [65]. Por ejemplo, para un punto dentro de un rectángulo definido por los nodos $(a_i, z_j), (a_{i+1}, z_j), (a_i, z_{j+1}), (a_{i+1}, z_{j+1})$, se usa:

$$f(a^*, z^*) \approx \sum_{k=0}^{1} \sum_{l=0}^{1} \theta_{i+k,j+l} \cdot (1 - |a^* - a_{i+k}|/\Delta a) \cdot (1 - |z^* - z_{j+l}|/\Delta z),$$

donde Δa y Δz son las distancias entre nodos en la malla.

Aplicación en el Modelo DSGE

En nuestro modelo DSGE, la cuadratura de Gauss-Hermite genera valores de $a_{t+1}^{(i,j)}$ y $z_{t+1}^{(i,j)}$ en puntos fuera de la malla simulada. Para evaluar el consumo $c_{t+1}^{(i,j)}$ y la inflación $\pi_{t+1}^{(i,j)}$ en dichos puntos, usamos funciones de interpolación definidas sobre la malla (a, z), tal como:

$$c_{t+1}^{(i,j)} \approx \sum_{k} \theta_k \Psi_k(a_{t+1}^{(i,j)}, z_{t+1}^{(i,j)}), \tag{C.2}$$

$$\pi_{t+1}^{(i,j)} \approx \sum_{m} \lambda_m \Phi_m(a_{t+1}^{(i,j)}, z_{t+1}^{(i,j)}), \tag{C.3}$$

donde:

1. Ψ_k, Φ_m son funciones base correspondientes a la interpolación,

2. θ_k, λ_m son coeficientes estimados previamente durante la solución del modelo.

Estas expresiones permiten que c_{t+1} y π_{t+1} sean evaluadas en cualquier nodo generado por la cuadratura de Gauss-Hermite, facilitando así el cálculo de esperanzas racionales.

Ejemplo Ilustrativo: Interpolación de c_{t+1} y π_{t+1}

Para ilustrar la interpolación con elementos finitos, consideramos una mallado bidimensional sobre el espacio de estados (a, z), con los siguientes puntos de malla:

$$a \in \{0.9, 1.1\}, \quad z \in \{0.8, 1.0\}$$

y supongamos que se desea interpolar las funciones $c_{t+1}(a, z)$ y $\pi_{t+1}(a, z)$ en el punto:

$$(a^*, z^*) = (1.0, 0.9)$$

1. Interpolación de c_{t+1}: Supongamos que los valores de la función c_{t+1} en los nodos son:

	$z = 0.8$	$z = 1.0$
$a = 0.9$	0.92	1.02
$a = 1.1$	0.96	1.10

La interpolación se realiza con las funciones base:

$$\phi_1(a) = \frac{a_2 - a}{a_2 - a_1}, \quad \phi_2(a) = \frac{a - a_1}{a_2 - a_1}, \quad \psi_1(z) = \frac{z_2 - z}{z_2 - z_1}, \quad \psi_2(z) = \frac{z - z_1}{z_2 - z_1}$$

Evaluando:

$$\phi_1(1.0) = 0.5, \quad \phi_2(1.0) = 0.5, \quad \psi_1(0.9) = 0.5, \quad \psi_2(0.9) = 0.5$$

La interpolación es:

$$c_{t+1}(1.0, 0.9) \approx (0.92)(0.5)(0.5) + (1.02)(0.5)(0.5) + (0.96)(0.5)(0.5) + (1.10)(0.5)(0.5)$$

$$= 0.23 + 0.255 + 0.24 + 0.275 = \boxed{1.00}$$

2. Interpolación de π_{t+1}: Supongamos que los valores de π_{t+1} en los mismos nodos son:

	$z = 0.8$	$z = 1.0$
$a = 0.9$	1.03	1.10
$a = 1.1$	1.05	1.20

La interpolación es:

$$\pi_{t+1}(1.0, 0.9) \approx (1.03)(0.5)(0.5) + (1.10)(0.5)(0.5) + (1.05)(0.5)(0.5) + (1.20)(0.5)(0.5)$$

$$= 0.2575 + 0.275 + 0.2625 + 0.3 = \boxed{1.095}$$

Este procedimiento se aplica en cada paso temporal y para cada punto o partícula, permitiendo una evaluación eficiente de las funciones requeridas en la resolución del modelo DSGE no lineal mediante elementos finitos.

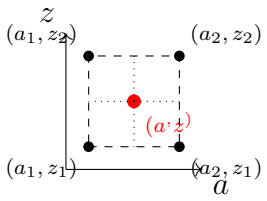


Figura C.1: Ejemplo de interpolación bilineal sobre una malla rectangular en (a, z).

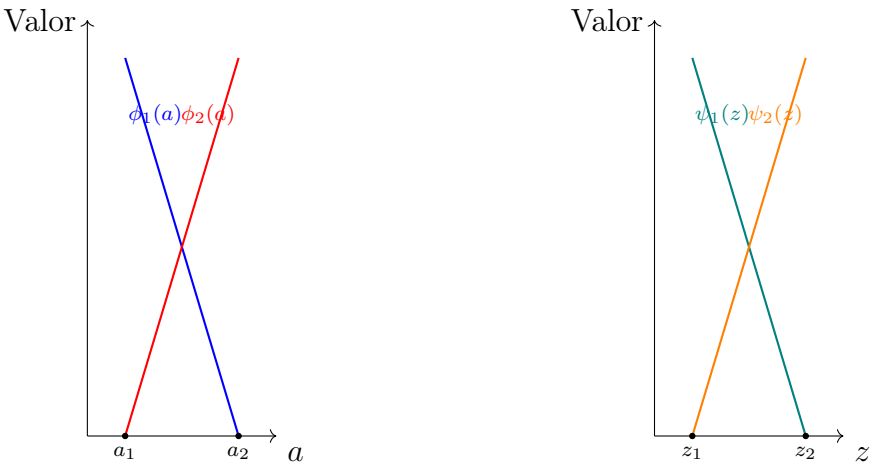


Figura C.2: Funciones base para la interpolación bilineal en las dimensiones a y z.

Implementación General

En términos computacionales, la interpolación en Python se puede utilizar la función:

```
from scipy.interpolate import RegularGridInterpolator
```

Sustituyendo las expresiones (C.2) y (C.3) en la forma de la esperanza de Euler obtenida en el Apéndice B, se tiene:

$$\mathbb{E}_t\left[\frac{a_{t+1}}{c_{t+1}z_{t+1}\pi_{t+1}}\right] \approx \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \frac{e^{(1-\rho_a)\ln a_s + \rho_a \ln a_t + \sqrt{2}x_i}}{\left(\sum_k \theta_k \Psi_k(a^{(i,j)}, z^{(i,j)})\right) e^{\ln z_s + \sqrt{2}x_j} \left(\sum_m \lambda_m \Phi_m(a^{(i,j)}, z^{(i,j)})\right)}. \tag{C.4}$$

Esta expresión muestra cómo las técnicas de interpolación por elementos finitos se integran en la resolución de las esperanzas racionales que aparecen en los modelos DSGE, manteniendo el carácter no lineal del sistema y evitando la necesidad de linearizar.