

# Protocolo de Tesis

Cuantificación de incertidumbre bayesiana aproximada en problemas inversos de ODE

---

César Isaí García Cornejo

Asesor: Dr. José Andrés Christen Gracia

CIMAT

# Introducción

---

Los modelos que pretendan describir los fenómenos naturales deben ser causales, respetando orden entre causa y efecto.

Típicamente, los modelos matemáticos precisan de condiciones iniciales o parámetros para caracterizar la unicidad en su solución. Tales parámetros se conocen como parámetros del modelo y se interpretan como causas del fenómeno modelado. Mientras que la solución del modelo se interpreta como la predicción del fenómeno, que se asocia a los efectos.

El proceso anteriormente descrito se llama problema directo o *forward problem* ya que sigue la dirección de causalidad. Sin embargo, es interesante también el problema inverso, dada ciertas observaciones de las cualidades de un fenómeno (efectos), **¿es posible calcular las causas del modelo que rige el fenómeno?**

Estudio de un sistema físico:

## 1: Parametrización



## 2: Modelación directa



## 3: Problema inverso



# Antecedentes

---

Consideramos modelos de la forma

$$G(t, y(t), y'(t), y''(t), \dots) = 0$$

con  $y(0) = y_0, y'(0) = y'_0, \dots$  sus condiciones iniciales.

Más aún, se puede generalizar a sistemas de ecuaciones diferenciales (Apostol, 2019).

El **problema directo** es entonces aquel que dado un  $\theta \in \Theta$  obtiene la trayectoria o solución de la ecuación diferencial.

Denotemos por  $\mathcal{Y}$  al espacio de las soluciones posibles de la ecuación diferencial. De esta forma existe un mapeo del espacio  $\Theta$  a  $\mathcal{Y}$  el cual llamaremos **forward map**. Así, el forward map es

$$\theta \mapsto F[\theta]$$

donde  $\theta \in \Theta$  y  $F[\theta] \in \mathcal{Y}$ .



Ejemplo

Principalmente existen dos razones para que las observaciones de un modelo no concuerden exactamente con las predicciones.

1. Errores de medición por la incertidumbre de los aparatos de medición.
2. Defectos propios del modelo.

El paradigma bayesiano para problemas inversos se centra en cuantificar la incertidumbre en los parámetros del modelo. Es decir, su objetivo es establecer una medida de probabilidad posterior a las observaciones de la trayectoria del modelo. Equivalentemente, se busca la distribución de probabilidad  $\pi(\theta|\mathbf{y})$  donde  $\mathbf{y} = (y_1, \dots, y_n)$  son las observaciones de la trayectoria a lo largo del tiempo  $t_1, t_2, \dots, t_n$ . Partiendo de la información acerca de los parámetros previo a cualquier observación, se propone una distribución de probabilidad  $\pi(\theta)$  llamada distribución a priori, y tras aplicar el Teorema de Bayes obtener la distribución posterior (Wasserman, 2013).

De esta forma, se establece que las discordancias entre observaciones y predicciones siguen una distribución normal de la forma

$$y_i = F_{\theta}(t_i) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2),$$

donde las  $y_i$ 's son las observaciones del fenómeno.

Se ha establecido una distribución para las observaciones  $y_i$  las cuales tienen asociada una verosimilitud sobre  $\theta$  y  $\sigma$ . Como se consideran errores condicionalmente independientes, pues se tratan de errores de medición, la verosimilitud se sigue de

$$\mathcal{L}(\theta, \sigma) = f(\mathbf{y}|\theta, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (y_i - F_{\theta}(t_i))^2 \right\},$$

simplificando

$$f(\mathbf{y}|\theta, \sigma) = \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - F_{\theta}(t_i))^2 \right\}, \quad (1)$$

Del teorema de Bayes, se obtiene la distribución posterior para los parámetros

$$\pi(\theta, \sigma | \mathbf{y}) = \frac{f(\mathbf{y} | \theta) \pi(\theta)}{\int f(\mathbf{y} | \theta) \pi(\theta) d\theta}. \quad (2)$$

donde la constante de integración  $h(\mathbf{y}) = \int f(\mathbf{y} | \theta) \pi(\theta) d\theta$  es la constante de normalización para la distribución posterior.

Simulación de la posterior.  $1/\sqrt{n}$

El algoritmo de Metropolis-Hastings es un método de MCMC para generar muestras de una distribución objetivo  $f(x)$  partiendo de muestras de una distribución propuesta  $q(y|x)$ .

el algoritmo de Metropolis-Hastings genera una cadena  $X_0, X_1, \dots, X_N$  construyendo recursivamente la cadena dependiendo solamente del estado previo, es decir preservando la propiedad de Markov.



# Algoritmo de Metropolis-Hastings

El estado inicial  $X_0$  se toma aleatoriamente. Luego, teniendo hasta el estado  $X_i$ , el estado  $X_{i+1}$  se obtiene siguiendo

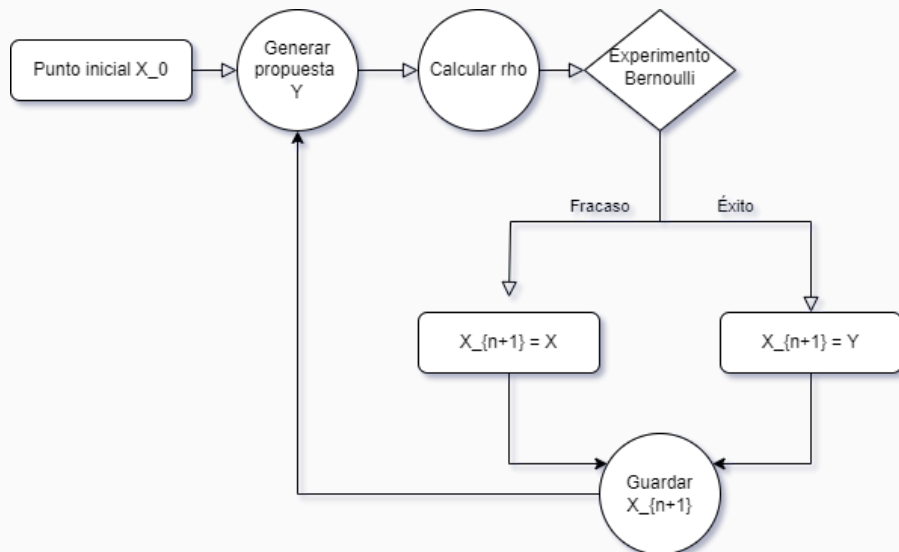
1. Generar una propuesta  $Y \sim q(y|X_i)$ .
2. Calcular  $\rho$

$$\rho = \min \left\{ \frac{f(y)q(x|y)}{f(x)q(y|x)}, 1 \right\}$$

3. Realizar un experimento Bernoulli con probabilidad de éxito  $\rho$ .

$$X_{i+1} = \begin{cases} Y & , \text{ con probabilidad } \rho \\ X_i & , \text{ con probabilidad } 1 - \rho \end{cases}$$

# Algoritmo de Metropolis-Hastings



Explicación punto por punto de M-H.

- El forward map aproximado  $F_{\theta}^*$  es una construcción multifacética basada en aproximaciones de estadística espacial.
- Tal aproximación pretende crear una distribución posterior aproximada  $\tilde{\pi}(\theta|\mathbf{y})$  con la sustitución del forward map a su versión aproximada.
- Existen diferentes maneras de hacer las aproximaciones del forward map. La forma propuesta es considerar una discretización del espacio de parámetros  $\Theta$ .

## Construcción del Forward Map Aproximado

Para cada  $\theta \in \Theta \subset \mathbb{R}^d$  se puede escribir como  $\theta = (\theta_1, \theta_2, \dots, \theta_d)$ .

Consideremos para cada coordenada un intervalo  $[\theta_i^{min}, \theta_i^{max}]$  para toda  $i \in \{1, \dots, d\}$ , donde  $\theta_i^{min}$  y  $\theta_i^{max}$  son cotas para el espacio de parámetros que contenga la masa de probabilidad.

Tomemos una partición equidistante en  $M$  puntos para cada coordenada. La partición de la coordenada  $i$ -ésima es el conjunto  $\mathcal{M}_i = \{\theta_i^{(1)}, \theta_i^{(2)}, \dots, \theta_i^{(M-1)}, \theta_i^{(M)}\}$  con  $\theta_i^{(1)} = \theta_i^{min}$  y  $\theta_i^{(M)} = \theta_i^{max}$ .

De esta forma, la partición crea una malla  $\mathcal{M} = \mathcal{M}_1 \times \dots \times \mathcal{M}_d$  del espacio parametral  $\Theta$ .

Sea  $\vartheta \in \mathcal{M}$  elemento de la malla  $\mathcal{M}$ . El forward map para cada elemento de  $\mathcal{M}$  es,

$$y_j = F[\vartheta_j],$$

con  $y_j = y_j(t)$  funciones continuas

## Construcción del Forward Map Aproximado

Para parámetros  $\theta \notin \mathcal{M}$  se propone buscar a los  $k$  vectores de parámetros  $\vartheta \in \mathcal{M}$  más cercanos en distancia euclidiana a  $\theta$ , denotando estos  $k$  vecinos por  $\vartheta^{(1)}, \dots, \vartheta^{(k)}$  cada uno a una distancia  $d_1, \dots, d_k$  de  $\theta$ , respectivamente; con  $d_1 \leq d_2 \leq \dots \leq d_k$ . La aproximación al forward map con  $k$  vecinos en una malla con  $M$  particiones es

$$\tilde{F}_M^k(\theta) = \sum_{j=1}^k \omega_j(\theta) F(\vartheta^{(j)}) \quad (3)$$

donde  $\omega_j(\theta) = d_j^{-1} / \sum_{i=1}^k d_i^{-1}$











## El problema directo:

Predecir los valores de las observables físicos  $\mathbf{d}$  que corresponde a un modelo  $\theta$

$$\theta \mapsto \mathbf{d} = \mathbf{F}(\theta)$$

Incertidumbre de las mediciones e imperfecciones del modelo [2].

# Ejemplos

## Resorte sujeto a fricción

Sea  $x(t)$  la posición

$$m\ddot{x} = -kx + b\dot{x}$$

## Caída sujeto a fricción

Sea  $x(t)$  la posición

$$m\ddot{x} = g - b\dot{x}$$

## Crecimiento poblacional

Sea  $P(t)$  el tamaño de población

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K}\right)$$

## Forward map

$$\theta = (k, b) \mapsto F(\theta) = x(t)$$

## Forward map

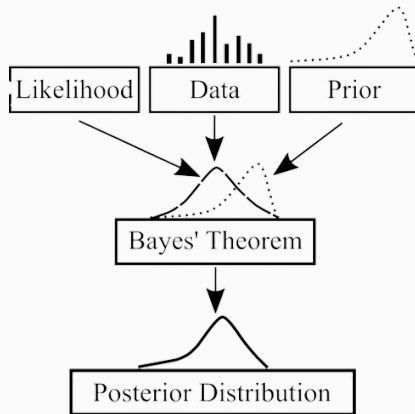
$$\theta = (g, b) \mapsto F(\theta) = x(t)$$

## Forward map

$$\theta = (r, K) \mapsto F(\theta) = P(t)$$

## La distribución posterior

$$\pi(\theta|x^n) \propto \mathcal{L}(\theta|x^n)\pi(\theta)$$



Procedimiento:

- **Observaciones:**  $(x_1, t_1), (x_2, t_2), \dots, (x_n, t_n)$
- Cuantificar el error mediante la relación

$$x_i = F_{\theta}(t_i) + \varepsilon_i, \quad i = 1, \dots, n$$

- Un supuesto convencional sobre los errores

$$\varepsilon_i \sim N(0, \sigma^2)$$

## Distribución posterior

$$\begin{aligned}\pi(\theta|x^n) &\propto \mathcal{L}(\theta|x^n)\pi_{\Theta}(\theta) \\ &\propto \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (x_i - F_{\theta}(t_i))^2\right) \pi_{\Theta}(\theta) \\ &\propto \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - F_{\theta}(t_i))^2\right) \pi_{\Theta}(\theta)\end{aligned}$$

donde  $\theta = (\theta_1, \dots, \theta_m)$ .

## Simular por métodos Monte Carlo

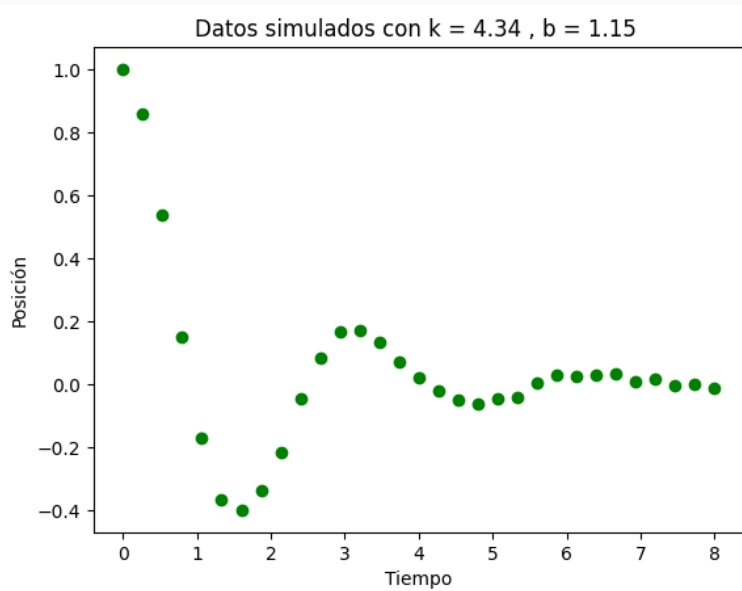
- Se simula por MCMC Metropolis-Hastings [1].



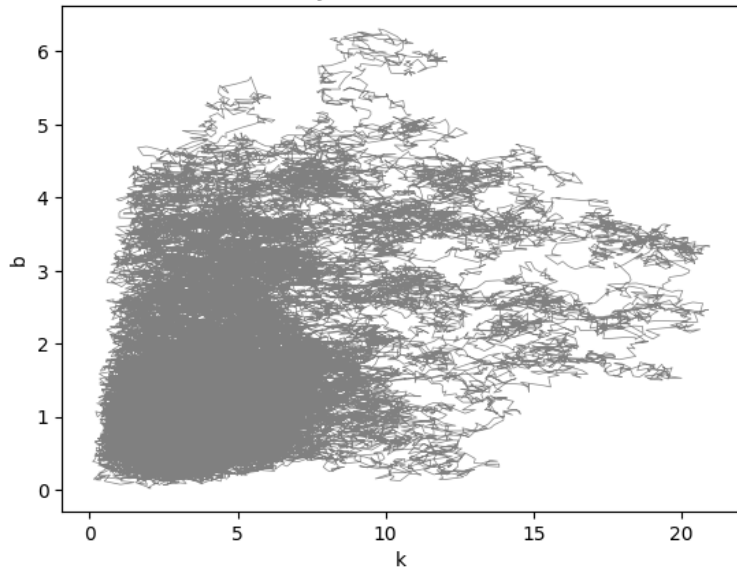
## Ejemplo

---

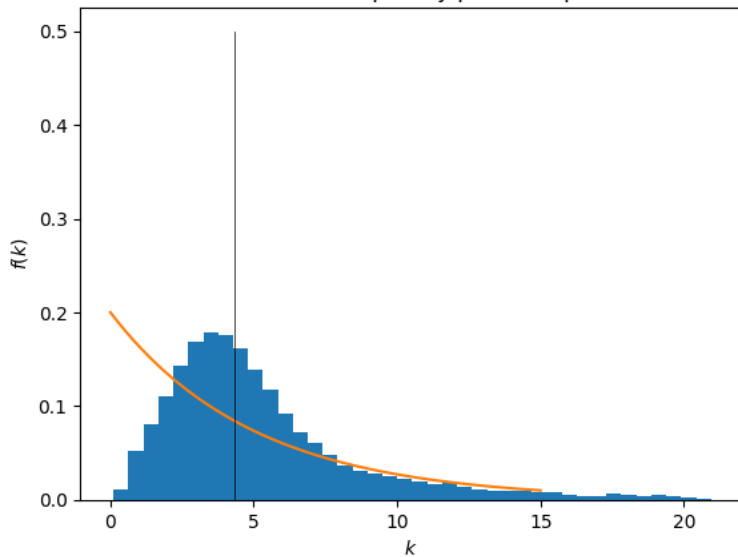
## Resorte sujeto a fricción



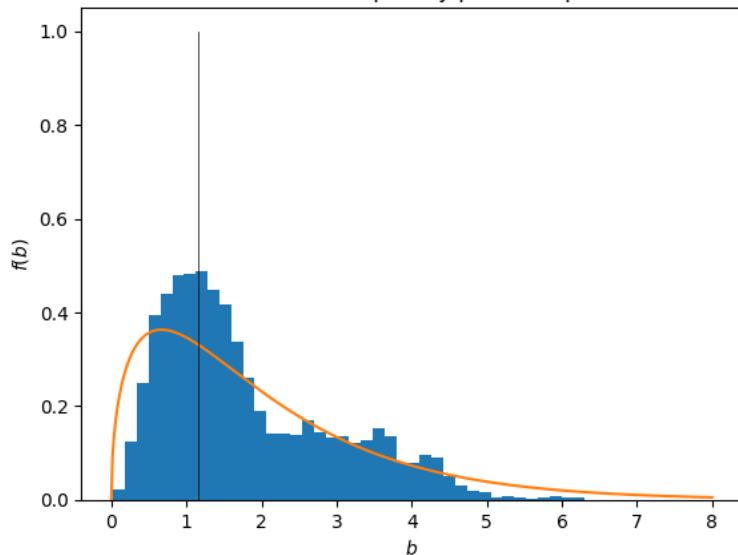
## Trayectoria de MCMC



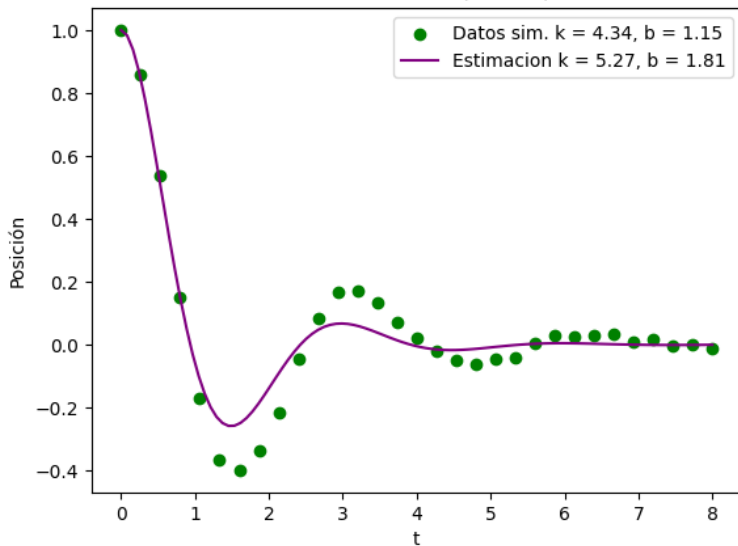
Distribuciones a priori y posterior para  $k$



Distribuciones a priori y posterior para  $b$



Curva estimada ( $n = 30$ )



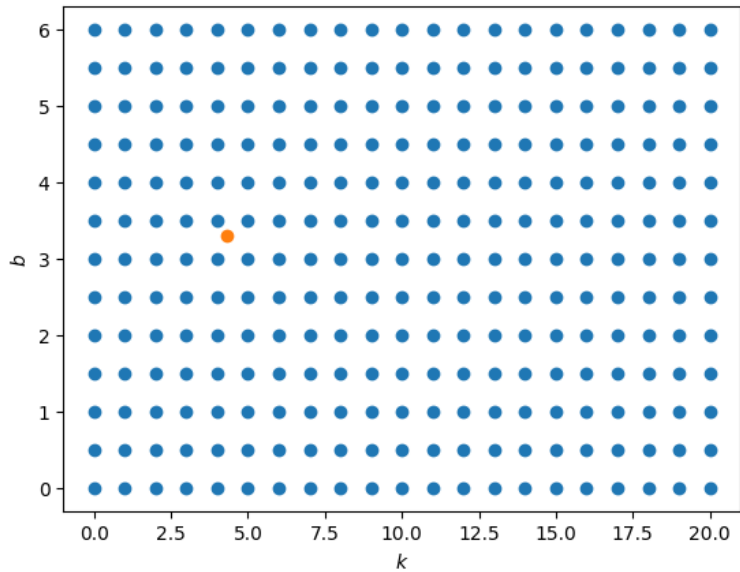
# Objetivos

---

A pesar de que el análisis previo nos permite obtener estimaciones bastante precisas para el problema inverso. Sin embargo, son **computacionalmente pesadas**, esto debido a que a cada paso en la cadena requiere que se solucione el problema forward, es decir, se resuelven ecuaciones diferenciales tantas veces como se deje correr la cadena.

- Se desea encontrar una especie de *interpolación* para solo solucionar una cantidad pequeña de veces el problema forward y para cada punto en el espacio de parámetros se pueda aproximar la solución en función de las soluciones con parámetros cercanos.

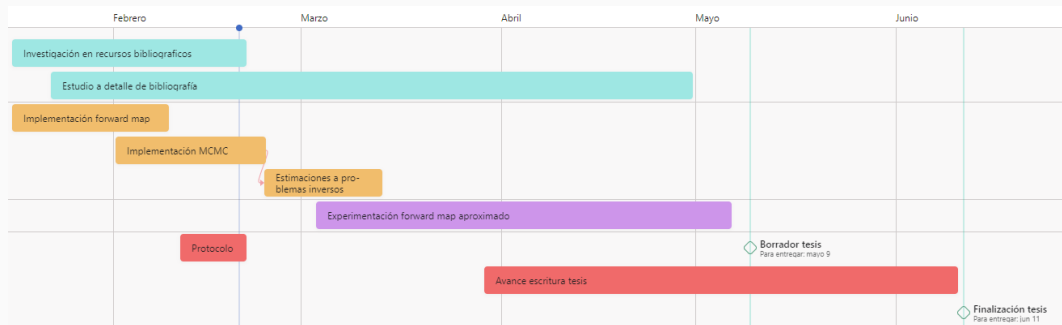




# Plan de trabajo

---

# Plan de trabajo





C. P. Robert, G. Casella, and G. Casella.  
***Monte Carlo statistical methods, volume 2.***  
Springer, 1999.



A. Tarantola.  
***Inverse problem theory and methods for model parameter estimation.***  
SIAM, 2005.