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

Introducción

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

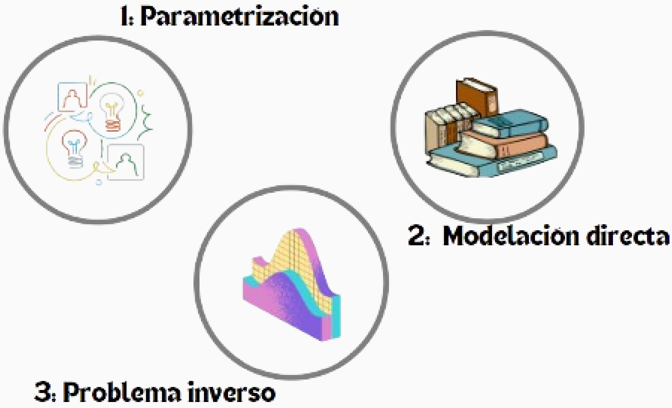
Típicamente, los modelos matemáticos precisan de condiciones iniciales o parámetros para caracterizar la unicidad en su solución. Tales parámetros se conocen como parámetros del modelo y se interpretan como causas del fenómeno modelado. Mientras que la solución del modelo se interpreta como la predicción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

Introducción

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?**

Sistema físico

Estudio de un sistema físico:



Antecedentes

Modelos Descritos por ODEs

Consideramos modelos de la forma

$$G(t, y(t), y'(t), y''(t), ...) = 0$$

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

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

Forward Map

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 Θ a $\mathcal{Y}$ el cual llamaremos **forward map**. Así, el forward map es

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

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

Problema Inverso

Ejemplo

Solución Bayesiana al Problema Inverso

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 $\mathbf{y} = (y_1, ..., y_n)$ son las observaciones de la trayectoria a lo largo del tiempo $t_1, t_2, \cdots, t_n$. 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).

Cuantificación de la Incertidumbre

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

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

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

Distribución Posterior

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

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

simplificando

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

Distribución Posterior

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

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

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.

Método MCMC

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

Algoritmo de Metropolis-Hastings

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 $X_0, X_1, \cdots, X_N$ 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

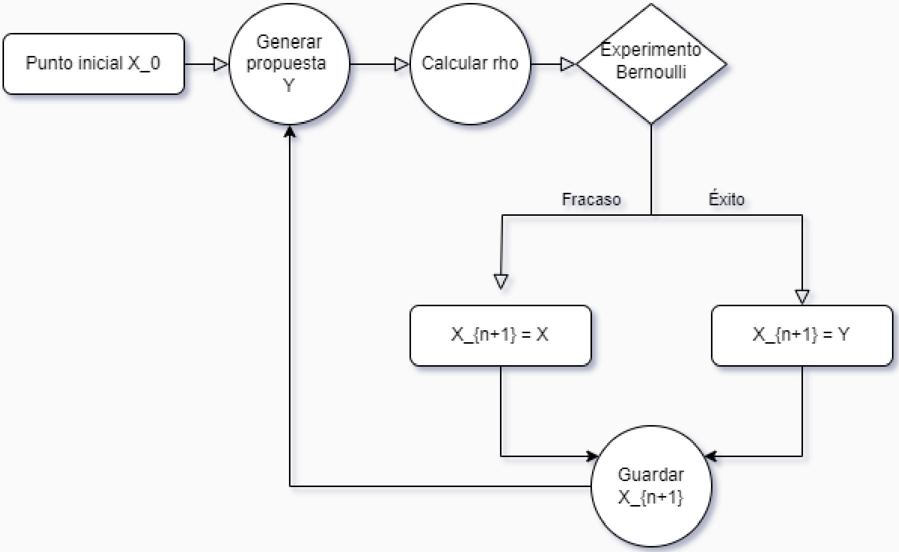
1. Generar una propuesta $Y \sim q(y|X_i)$.
2. Calcular ρ

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

3. Realizar un experimento Bernoulli con probabilidad de éxito ρ.

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

Algoritmo de Metropolis-Hastings



Explicación punto por punto de M-H.

Forward Map Aproximado

- El forward map aproximado F^*_θ 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 Θ.

Construcción del Forward Map Aproximado

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

Consideremos para cada coordenada un intervalo $[\theta_i^{min}, \theta_i^{max}]$ para toda $i \in \{1, \cdots, d\}$, donde θ_i^{min} y θ_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 $\mathcal{M}_i = \{\theta_i^{(1)}, \theta_i^{(2)}, \cdots, \theta_i^{(M-1)}, \theta_i^{(M)}\}$ con $\theta_i^{(1)} = \theta_i^{min}$ y $\theta_i^{(M)} = \theta_i^{max}$.

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

Construcción del Forward Map Aproximado

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 θ, denotando estos k vecinos por $\vartheta^{(1)}, \cdots, \vartheta^{(k)}$ cada uno a una distancia $d_1, \cdots, d_k$ de θ, respectivamente; con $d_1 \leq d_2 \leq \cdots \leq d_k$.

El **forward map aproximado** con k vecinos en una malla con M particiones es

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

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

Consistencia y Utilidad del Forward Map Aproximado

El potencial del uso del forward map aproximado para el estudio del problema inverso reside parcialmente en un menor tiempo de ejecución de la implementación en contraste con su análogo ordinario.

Existen modelos de ecuaciones diferenciales cuya implementación del problema inverso bajo enfoque bayesiano toma días en ejecutarse (Galaviz,2023).

Consistencia y Utilidad del Forward Map Aproximado

Es de interés matemático el estudio de la **consistencia** del método propuesto.

La distribución posterior aproximada para los parámetros del modelo es

$$\tilde{\pi}_M^k(\theta|\mathbf{y}) \propto \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\big(y_i - F_M^k(\theta)\big)^2\right\}\pi(\theta), \tag{4}$$

Consistencia y Utilidad del Forward Map Aproximado

Para formalizar la consistencia del forward map aproximado se hace uso de la *distancia de Kullback-Leibler*.

La distancia de Kullback-Leibler entre f y g se define como

$$D(f,g) = \int f(x) \log\left(\frac{f(x)}{g(x)}\right) dx.$$

Se puede mostrar que $D(f,g) \geq 0$ y $D(f,f) = 0$ (Wasserman,2006).

Consistencia y Utilidad del Forward Map Aproximado

Definamos la distancia entre las distribuciones posteriores como

$$c_M^k = D\left(\pi(\theta|\mathbf{y}), \tilde{\pi}_M^k(\theta|\mathbf{y})\right)$$

Diremos que el forward map aproximado es **consistente** si para una cantidad de vecinos k fijo y conocido, las distancias

$$c_M^k \to 0 \quad \text{a medida que} \quad M \to \infty$$

Forward map

El problema directo:

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

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

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

Ejemplos

Resorte sujeto a fricción
Sea $x(t)$ la posición

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

Forward map

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

Caída sujeto a fricción
Sea $x(t)$ la posición

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

Forward map

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

Crecimiento poblacional
Sea $P(t)$ el tamaño de población

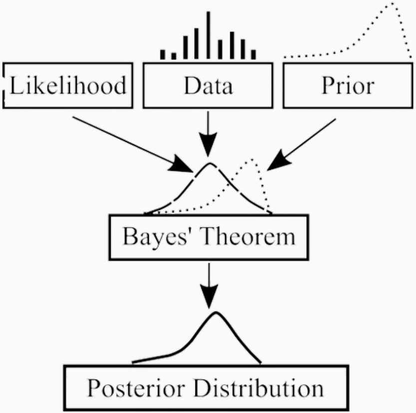
$$\frac{dP}{dt} = rP\left(1 - \frac{P}{K}\right)$$

Forward map

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

La distribución posterior

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



Incertidumbre en los errores

Procedimiento:

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

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

- Un supuesto convencional sobre los errores

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

Inferencia

Distribución posterior

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

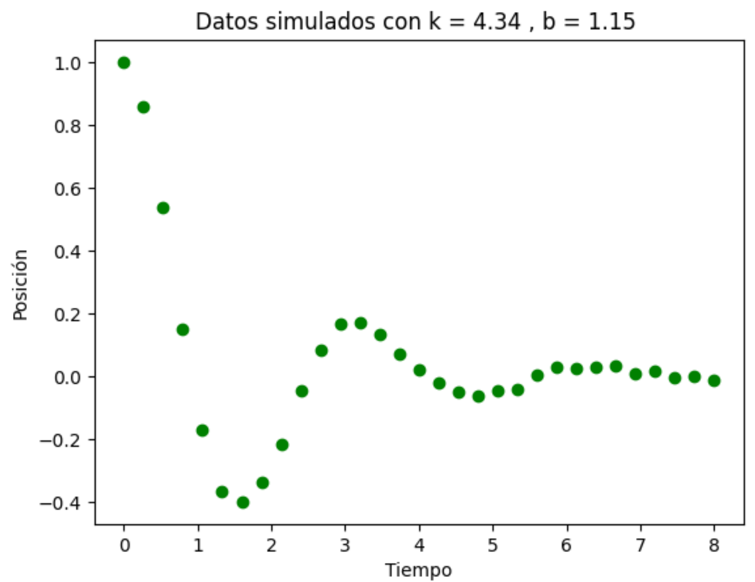
donde $\theta = (\theta_1, ..., \theta_m)$.

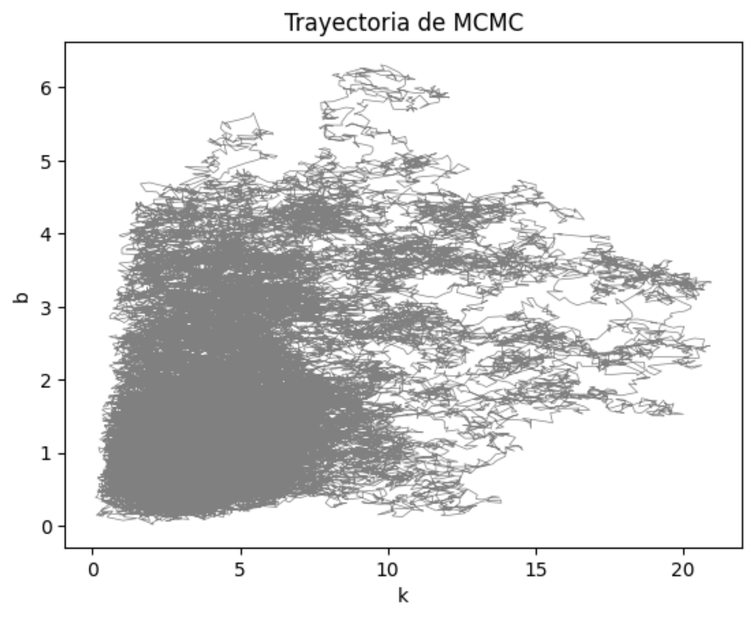
Simular por métodos Monte Carlo

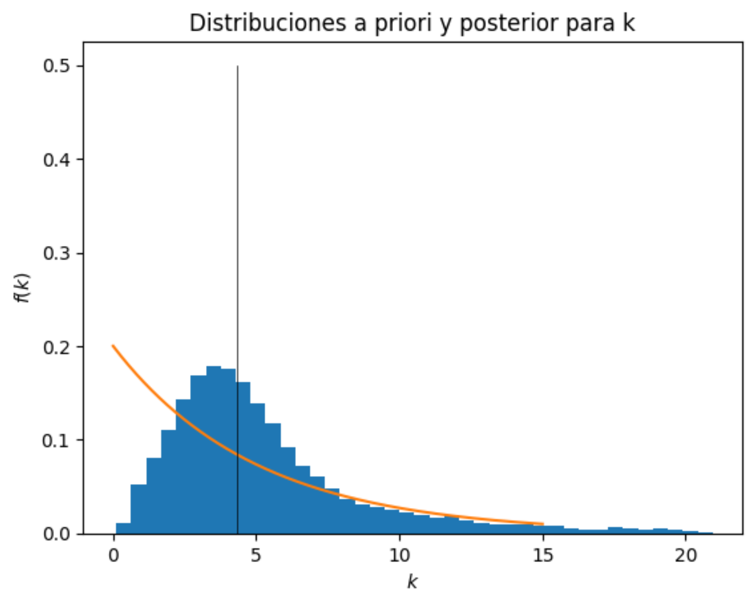
- Se simula por MCMC Metropolis-Hastings [1].

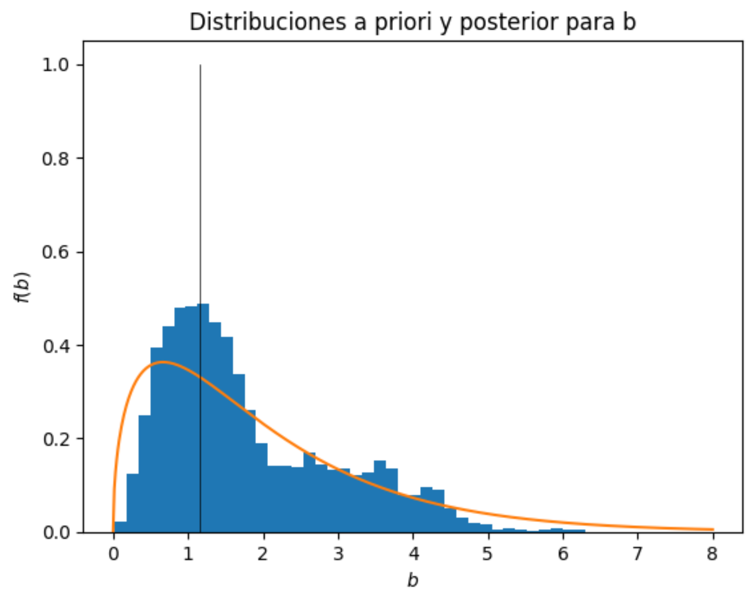
Ejemplo

Resorte sujeto a fricción

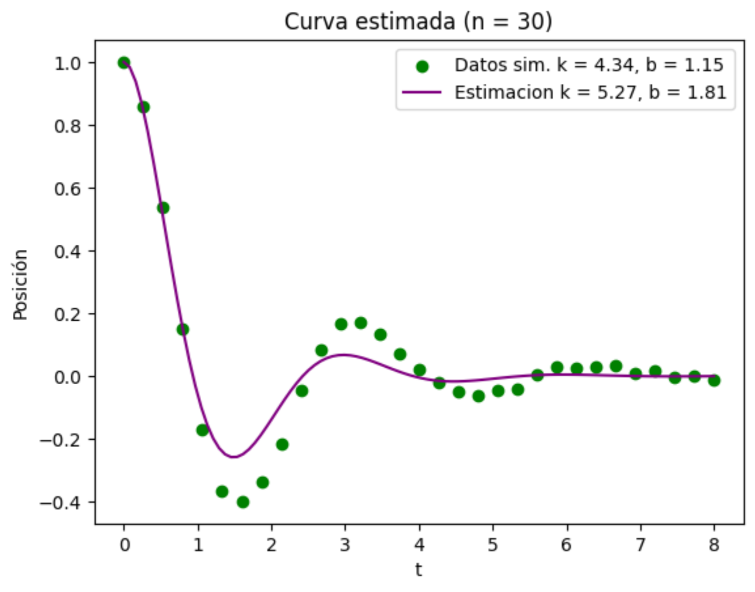



Trayectoria de MCMC
b
0
1
2
3
4
5
6
k
0
5
10
15
20


Distribuciones a priori y posterior para k
0.5
0.4
0.3
0.2
0.1
0.0
f(k)
0
5
10
15
20
k



Distribuciones a priori y posterior para b
1.0
0.8
0.6
0.4
0.2
0.0
f(b)
0
1
2
3
4
5
6
7
8
b

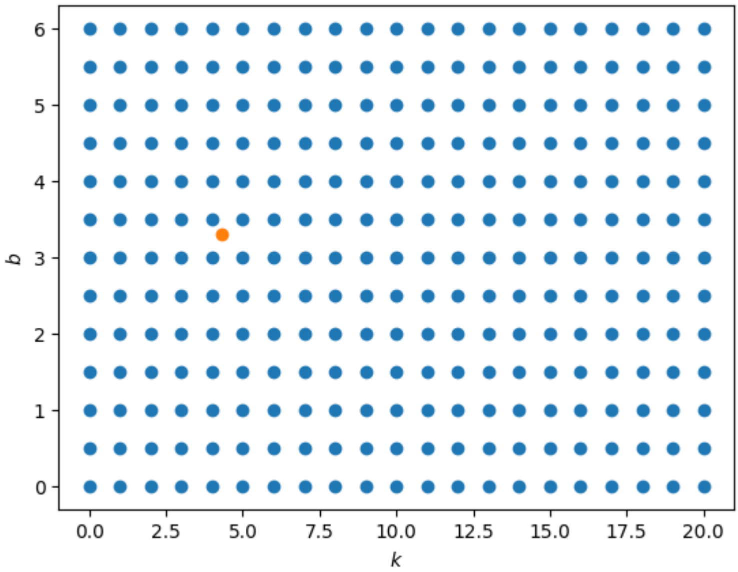

Curva estimada (n = 30)
Datos sim. k = 4.34, b = 1.15
Estimacion k = 5.27, b = 1.81
Posición
t
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
0
1
2
3
4
5
6
7
8

Objetivos

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.

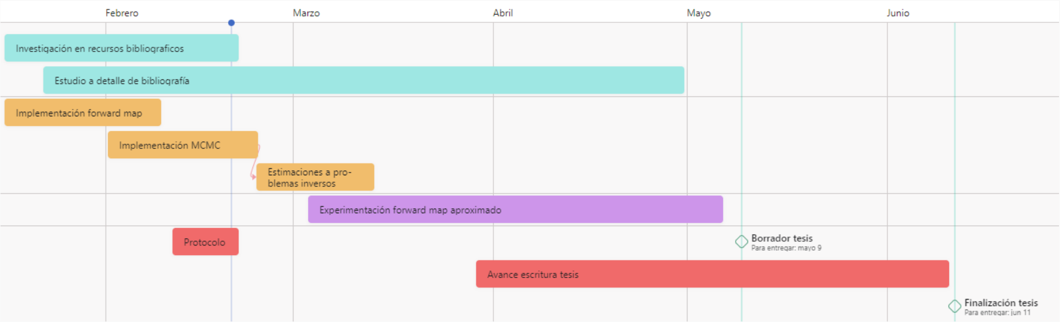
- Se desea encontrar una especie de *interpolación* para solo solucionar una cantidad pequeña de veces el problema forward y para cada punto en el espacio de parámetros se pueda aproximar la solución en función de las soluciones con parámetros cercanos.



6
5
4
3
2
1
0
b
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
k

Plan de trabajo

Plan de trabajo



Referencias

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.