Inferencia Bayesiana en modelos de ecuaciones diferenciales: Un enfoque computacional

Erick Iván Guerrero Flores

Universidad Autónoma Metropolitana, Unidad Iztapalapa; Maestría en Ciencias (Matemáticas Aplicadas e Industriales)

Asesores: Dr. Gabriel Núñez Antonio & Dr. José Héctor Morales Bárcenas

Objetivos

Inferencia guiada por datos en el contexto de soluciones numéricas de sistemas de ecuaciones diferenciales y modelos de datos circulares, para así lograr una mejor descripción (análisis de la variabilidad) del fenómeno real que genera los datos

Estadística bayesiana

La estadística bayesiana es una metodología que emplea los elementos de la teoría de la decisión y el enfoque subjetivo de la probabilidad para construir una metodología unificada y coherente. Esta permite:

- Incorporar la información que el investigador tenga del problema, al análisis formal de inferencia estadística.
- Distribuciones de probabilidad para describir el grado de creencia sobre los parámetros de un modelo.
- El análisis natural de la distribución de la variable de respuesta en un modelo (la distribución predictiva).

La **distribución final** emplea el teorema de Bayes y es el punto de partida para realizar inferencias bajo un enfoque bayesiano.

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

Una equivalencia a la ecuación de la **distribución final** seria considerar factor $p(y)$ como un valor constante:

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

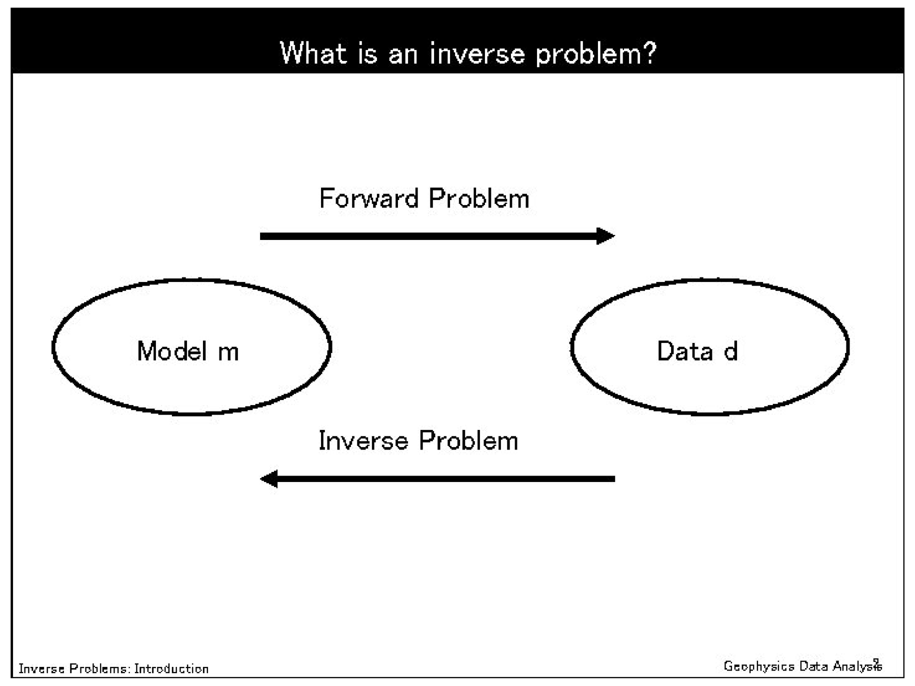

Figure: Problema directo y problema inverso

Modelos de EDO

Las ecuaciones diferenciales ordinarias tienen aplicaciones importantes y son una herramienta poderosa en el estudio de muchos problemas en las ciencias naturales y en la tecnología, esto nos invita a tener un interés especial en los sistemas de ecuaciones diferenciales para la **modelación matemática**.

$$\begin{cases} y' = f(t, y(t)) \\ y(t_0) = y_0 \end{cases}$$

Métodos MCMC

Los métodos MCMC utilizados para simular muestras de nuestra distribución final son:

1. The Robust Adaptive Metropolis (RAM)
2. The Hamiltonian Monte-Carlo Algorithm (HMC)
3. Aproximate Bayesian Computation (ABC)

Planteamiento del problema

Dado el conjunto de datos y_{ij} y las predicciones (soluciones) de las ecuaciones diferenciales, se considera el siguiente modelo estadístico:

$$y_{ij} = \hat{y}_j(t_i|\theta) + \epsilon_{ij}$$

Donde:

- $\hat{y}_j(t_i|\theta)$ es la solución numérica del sistema de ecuaciones del estado j en el tiempo t_i con valor de los parámetros θ .
- ϵ_{ij} es el error de medición de la observación i del estado j, $\epsilon_{ij} \sim N(0, \sigma_j)$

El parámetro σ_j nos explica la variabilidad de los datos en cada uno de los estados.

La distribución final resultante es:

$$p(\theta|y) \propto p(y|\theta) \times p(\theta)$$

Programación en Paralelo

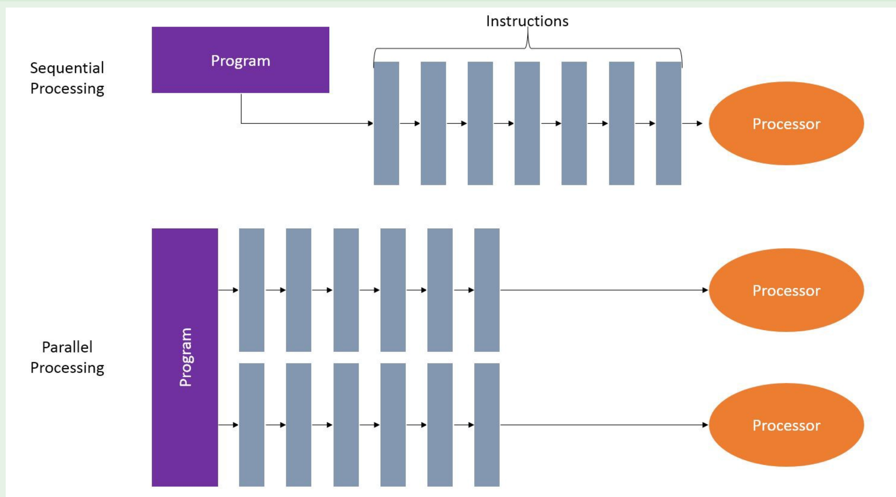

Figure: Programación en paralelo

1o modelo: FH-N

El modelo FitzHugh–Nagumo es una versión simplificada del modelo de Hodgkin y Huxley, que describe la dinámica de la neurona más en detalle. Su sistema de ecuaciones diferenciales es:

$$\frac{dV}{dt} = \gamma\left(V - \frac{V^3}{3} + R\right)$$
$$\frac{dR}{dt} = -\frac{1}{\gamma}(V - \alpha + \beta R)$$

Con $\alpha, \beta, \gamma \in \mathbb{R}^+$, donde se implementaron los tres métodos MCMC

1o modelo: FH-N

Parámetro	Verdadero valor	RAM	HMC	ABC
α	0.3	(0.14 ; 0.40)	(0.13 ; 0.36)	(0.15 ; 0.50)
β	0.3	($2.54e^{-05}$; 0.43)	(0.21 ; 0.58)	(0.15 ; 0.45)
γ	2	(1.80 ; 2.27)	(1.73 ; 2.27)	(1.55 ; 2.19)
V_0	-1	(-1.21 ; -0.44)	(-1.39 ; -0.65)	(-1.04 ; 0.42)
R_0	1	(0.41 ; 1.268)	(0.80 ; 1.60)	(-0.17 ; 2.09)
σ_1^2	0.25	(0.19 ; 0.48)	(0.15 ; 0.36)	(0.10 ; 1.40))
σ_2^2	0.25	(0.19 ; 0.48)	(0.13 ; 0.33)	(0.21 ; 1.41)

Table: Intervalos de Probabilidad

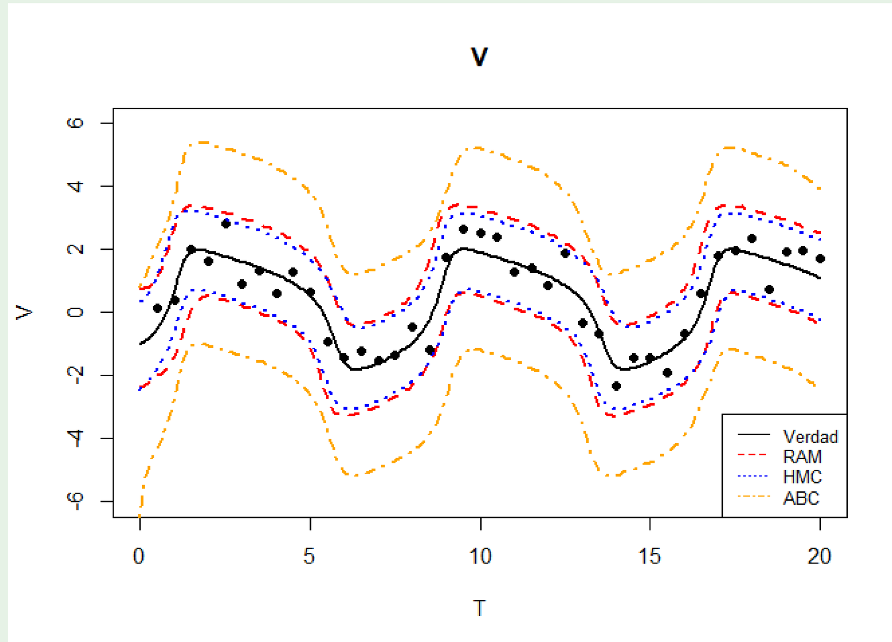

Figure: Verdadera solución para es estado V, las bandas de probabilidad respectivas a cada método y los 40 datos simulados

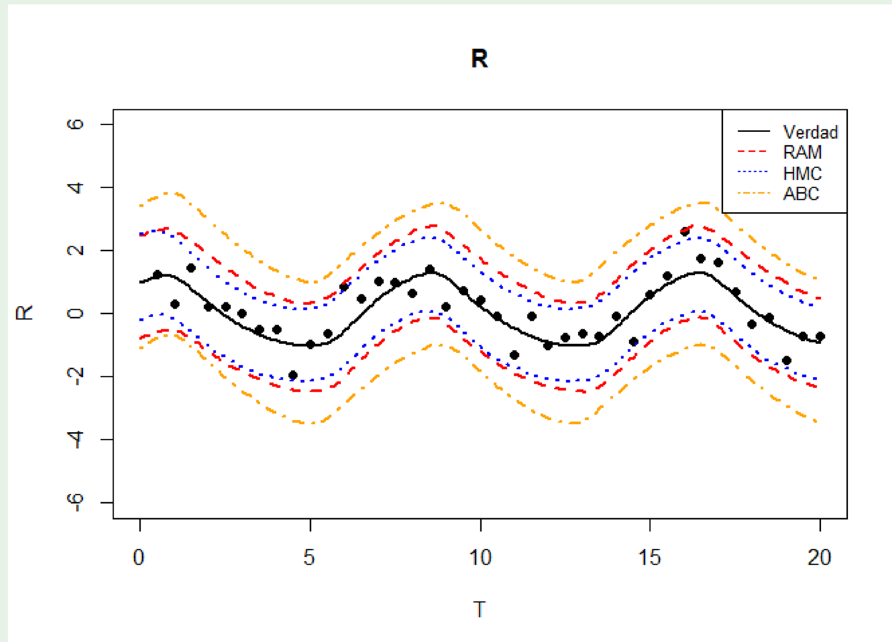

Figure: Verdadera solución para es estado R, las bandas de probabilidad respectivas a cada método y los 40 datos simulados

2o modelo: Datos circulares

La función de densidad del angulo aleatorio θ, esta dada por la normal proyectada:

$$PN(\theta|\mu, I) = C_6[1 + \frac{b}{\phi(b)}\Phi(b)]I_{(0,2\pi)}(\theta)$$

donde $b = v'\mu$ con $v = (Cos\ \theta, Sen\ \theta)$ y:

$$C_6 = \frac{\exp(-\frac{1}{2}\mu'\mu)}{2\pi v' v}$$

Donde se implementaron los tres métodos MCMC y el algortimo Metropolis-Hastings, con los siguientes resultados:

Parámetro	Verdadero valor	MH	RAM	HMC	ABC
μ_1	1	(0.96 ; 1.47)	(0.91 ; 1.42)	(0.94 ; 1.42)	(0.70 ; 1.77)
μ_2	1	(0.76 ; 1.25)	(0.74 ; 1.22)	(0.73 ; 1.19)	(0.69 ; 1.34)

Table: Intervalos de probabilidad del escenario μ=(1,1)

2o modelo: Datos circulares

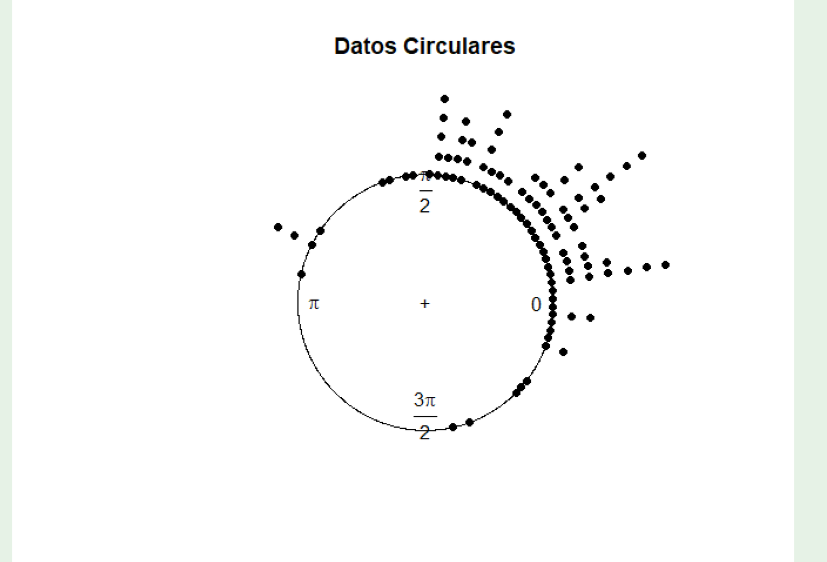

Figure: 100 datos circulares del angulo $\mu = (1, 1)$

Conclusión

Métodos MCMC	Ventajas	Desventajas
RAM	Fácil de implementar. Método adaptativo. Precio justo entre tiempo de cómputo y calidad de estimaciones.	Requiere de varias iteraciones y éstas aumentan según la dificultad del modelo a analizar.
HMC	El gradiente nos ayuda a mover el vector de parámetros de forma más eficiente. No requiere tantas iteraciones. Rápido si el gradiente se puede calcular analíticamente.	Requiere el cálculo del gradiente y si éste es numérico puede ser muy costoso computacionalmente.
ABC	No requiere verosimilitud. Bastante rápido.	Depende de las distribuciones iniciales. Depende del valor de ϵ. Necesita que la medida de comparación de los datos sea suficiente. Complicado estimar la varianza.
MH	Fácil de implementar. Método clásico para la inferencia bayesiana.	Depende fuertemente de la elección de la distribución propuesta. Puede tener tasas de aceptación muy baja.

Figure: Ventajas y desventajas

Algo que debe de quedar claro con el análisis presentado es que los métodos MCMC **no** son cajas negras que van a arrojar resultados positivos en todos los escenarios o modelos de interés, y no nos debemos quedar con un solo método, sino que uno debe de elegir adecuadamente el método MCMC que más se adapte al modelo matemático que se planea estudiar, esto para poder hacer más eficiente el análisis e inferencia que se propone hacer.

Extensión de las aplicaciones

- Inferencia sobre un modelo de conteo de macrofagos (células) y *Mycobacterium tuberculosis*.
- Inferencia con datos asociados al covid-19.

Referencias

- Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of Markov Chain Monte Carlo.* CRC press, 2011.
- K.V. Mardia and P.E. Jupp. *Directional Statistics.* Wiley Series in Probability and Statistics. Wiley, 2009.
- M. Vihola. Robust adaptive metropolis algorithm with coerced acceptance rate. *Stat Comput 22*, page 997–1008, 2012.

Contacto

- Web: http://mat.izt.uam.mx/mcmai/
- Email: erickivan0723g@gmail.com