

Markov Chain Monte-Carlo con enfoque bayesiano en análisis de relación astrofísica MA4402 - Simulación Estocástica: Teoría y Laboratorio

Departamento de Ingeniería Matemática
Universidad de Chile
Profesor: Joaquín Fontbona
Auxiliares: Pablo Araya - Bruno Hernández
Autor: Kurt Walsen

I. ABSTRACT

A continuación se presenta *Markov Chain Monte-Carlo con enfoque bayesiano en análisis de relación astrofísica*, proyecto realizado para el curso MA4402 - *Simulación Estocástica: Teoría y Laboratorio* y cuyo objetivo consiste en emplear técnicas y metodologías aprendidas durante el curso para estudiar la relación *Masa - Dispersión de velocidades* (referenciada en [1]), la cual corresponde a una relación astrofísica que cumplen galaxias, donde existe una relación lineal entre la masa del agujero negro central M una galaxia y su dispersión de velocidades σ .

Se presenta un set de datos con 46 galaxias entre espirales(S), elípticas(E) y lenticulares(S0). Los datos contienen las columnas $\log M$ (logaritmo de la masa del agujero negro central M), $\log \sigma$ (logaritmo de la dispersión de velocidades σ), $\text{err_log} M$, $\text{err_log} \sigma$ que corresponden a las incertezas de $\log M$ y $\log \sigma$, respectivamente.

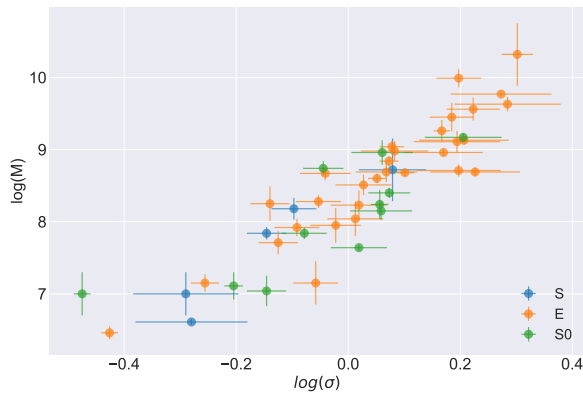


Fig. 1. Gráfico $\log(M)$ vs $\log(\sigma)$ para los datos disponibles, con barras de error bidimensionales.

Se plantea el siguiente modelo lineal para describir la relación:

$$\log\left(\frac{M}{M_{\odot}}\right) = m \cdot \log\left(\frac{\sigma}{\sigma_0}\right) + b + \mathcal{N}(0, \omega^2) \quad (1)$$

donde (m, b, ω^2) corresponden a los parámetros del modelo, M esta en unidades de masa solar $M_{\odot} = 2 \cdot 10^{30} [kg]$, σ esta en unidades de $\sigma_0 = 200 [km/s]$, y ω^2 corresponde a la varianza de la dispersión interna del modelo, dada por una variable aleatoria gaussiana centrada.

Para deducir la relación lineal entre $\log(M)$ y $\log(\sigma)$ se estudiará la distribución de los parámetros que define (1) a través de un modelo generativo basado en un análisis bayesiano con muestreos a través de *Markov Chain Monte-Carlo (MCMC)*.

El modelo bayesiano implementado considera el hecho de que los datos presentan outliers, incertezas bi-dimensionales y dispersión interna, efectos que se traducen en la cantidad de parámetros utilizados y la función a posteriori considerada (ver [2]).

El muestreo a través de *MCMC* se realiza a mano con *Metropolis Hastings*, donde debido a problemas de estabilidad de la Cadena de Markov en una de las variables motiva a estudiar con la librería *emcee* que es más robusta y utiliza la técnica de *stretch* (ver [3]) para muestrear de la distribución de los parámetros.

REFERENCES

- [1] Harris (2013) A Catalog of Globular Cluster Systems: What Determines the Size of a Galaxy's Globular Cluster Population?
- [2] Hogg (2010) Data analysis recipes: Fitting a model to data
- [3] Foreman-Mackey et al. (2013) emcee: The MCMC Hammer