Práctico 4

Inferencia Bayesiana con métodos MonteCarlo: Cadenas de Markov

Introducción

En este trabajo práctico implementaremos distintos conceptos y técnicas estudiados, relacionados con la inferencia estadística, ajuste de funciones, selección de modelos, cuadrados mínimos, interpolación y minimización. Las actividades a realizar son:

- Inferencia Bayesiana, para realizar el ajuste de un modelo paramétrico a un conjunto de datos.
- Exploración del espacio de parámetros para estimar la función de *Likelihood* usando Cadenas de Markov Monte Carlo (MCMC).
- Implementación del algoritmo de Metrópolis-Hastings para llevar a cabo realizaciones de MCMC.
- Interpolación de datos para construir una función contínua y derivable que pase por un conjunto de puntos.
- Minimización de funciones, mediante la técnica del gradiente descendente.
- Implementación de funciones en Python

La inferencia estadística se puede llevar a cabo como una aplicación del teorema de Bayes. Si tenemos un conjunto de datos d que se puede describir por un modelo m con parámetros ϕ, queremos calcular el mejor modelo que puede dar lugar a esos datos, es decir, maximizar la probabilidad posterior de los parámetros dados los datos para un modelo m, $p(\phi|d, m)$. Esta probabilidad es proporcional al *Likelihood* $p(d|\phi, m)$ por la función distribución de la probabilidad anterior (*prior*, $p(\phi, m)$).

$$p(\phi|d, m) = \frac{p(d|\phi, m)\, p(\phi|m)}{p(d|m)} \tag{1}$$

y está normalizada por la *evidencia*, es decir, la probabilidad marginal del *Likelihood* para el modelo m:

$$p(d|m) = \int_\Omega p(d|\phi, m)p(\phi|m)d\phi, \tag{2}$$

donde Ω denota el espacio de parámetros.

Cuando se ajusta un modelo a un conjunto de datos, se quiere conocer la función de *Likelihood*, $p(d|\phi, m)$, que depende de los parámetros ϕ. Existen varios métodos para llevar esto a cabo, entre ellos las Cadenas de Markov Monte Carlo (MCMC). En particular, el algoritmo de Metrópolis-Hastings es un método de MCMC que se utiliza para simular distribuciones multivariadas.

Una version simple de este algoritmo se puede escribir de la siguiente manera (pseudocódigo):

```
mientras eslabon es menor que el largo de la cadena
   hacer un salto
   evaluar el likelihood del nuevo punto
   sortear un numero aleatorio
   si el numero aleatorio es menor que la probabilidad
      aceptar el nuevo valor
   si no
      realizar un nuevo sorteo
   fin si
```

```
fin mientras
```

Ejercicios

▶ **1.** Leer y graficar los datos pertenecientes a la función de luminosidad de galaxias obtenida por Blanton et al. (2001). La figura de interés se reproduce a continuación. Los datos correspondientes a la función de luminosidad se pueden descargar del aula virtual.

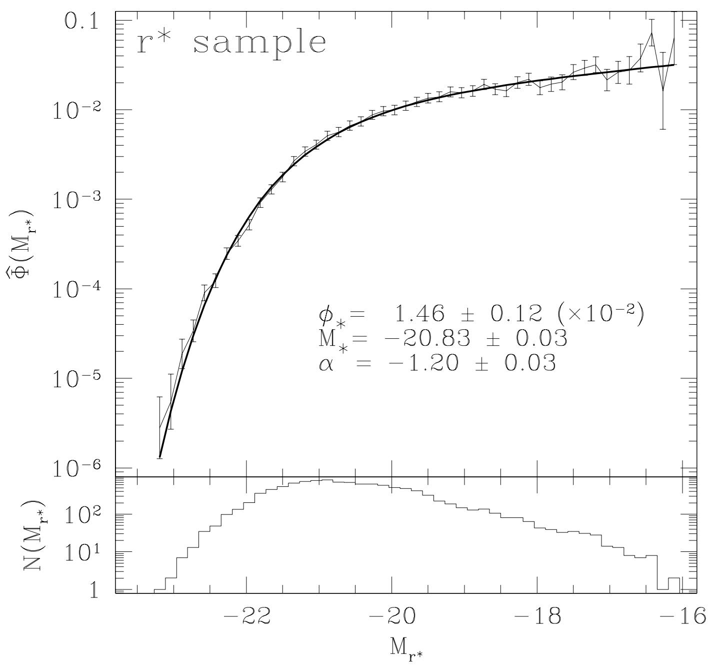

Figure 1: Función de luminosidad de las galaxias del Sloan Digital Sky Survey en la banda r. Imagen tomada del trabajo de Blanton et al. (2001).

▶ **2.** Se desea ajustar el modelo m al conjunto de datos d, mediante un análisis Bayesiano. Para el caso de los datos del Ej. 1, se utilizará como modelo la función de Schechter:

$$\Phi(M)dM = 0.4\,ln(10)\Phi_*\,10^{-0.4(M-M_*)(\alpha+1)} \times exp(10^{-0.4(M-M_*)})\,dM$$

.

Los parámetros que se intentan determinar son ϕ_*, M_* y α. Para ello, escribir en PYTHON las siguientes funciones:

- la función de *likelihood*
- la función de *priors*
- la función de *probabilidad posterior*

Estas funciones se necesitarán cuando se explore el espacio de parámetros. Determine claramente los parámetros de entrada y de salida de estas funciones e impleméntelas en un programa. Explicite las hipótesis que intervienen en este planteo (especialmente respecto de los errores), la motivación de dichas hipótesis y las simplificaciones que se consiguen. Utilizando estas funciones, implementar un algoritmo de Metrópolis-Hastings para explorar el espacio de parámetros. Los parámetros del algoritmo son el largo de la cadena (un número entero) y las funciones *prior*. Tener en cuenta que los *priors* contienen el conocimiento previo sobre los datos. Los *priors* más comunes son funciones planas y funciones gaussianas. Explicar qué parámetros intervienen, cómo y porqué se eligen sus valores. Experimentar con distintas funciones de *prior* y evaluar su impacto en la realización de las cadenas.

► **3.** Visualizar las propiedades de convergencia de las cadenas. Encontrar condiciones de la simulación para las cuales se produce un mal y un buen mezclado (*mixing*) de las cadenas. Para ello realize varias cadenas y compárelas.

Sugerencia: Se pueden realizar los siguietes gráficos:

- *Likelihood* vs. valor de un parámetro
- valor de un parámetro vs. eslabón de la cadena
- Distribución marginal de los valores de un parámetro (en este caso tener en cuenta que la cadena sea "estable")
- Curvas de nivel de la función de *Likelihood*.
- Máxima separación entre cadenas como función del número de pasos. Para esto es necesario contar con varias cadenas independientes, inicializadas con posiciones aleatorias de acuerdo a los *priors*

► **4.** Para una elección de parámetros que produzca un buen mezclado, correr varias cadenas y compararlas.

► **5.** Implemente el método de gradiente descendente para encontrar el mínimo de *Likelihood* marginalizado a cada uno de los parámetros. Discuta posibles cuestiones de precisión, conveniencias e inconveniencias del método.

▷ **6.** Compare los resultados del ejercicio anterior con los obtenidos de implementar el método del gradiente descendente en el espacio de parámetros. Establezca un conjunto de parámetros iniciales y visualize los distintos modelos obtenidos en los sucesivos pasos iterativos.

Deberá entregar un informe sobre los ejercicios marcados con "►" mediante un repositorio de git. En dicho informe deberá incluir la resolución de los ejercicios, junto con los códigos empleados y los gráficos correspondientes, acompañados de una breve introducción y conclusiones.

Fecha límite de entrega de informes: Miércoles 18 de Noviembre, por medio del Aula Virtual.

Material de consulta:

- Material de clases (Aula virtual)
- Numerical Recipes, Press et al.
- Mathematics of scientific computing, Kincaid & Ward
- Trotta, R., "Bayes in the sky: Bayesian inference and model selection in cosmology", arXiv:0803.4089
- Blanton et al. (2001), AJ, 121, 2358