



TAREA 12

Fecha de entrega: 18/12/2016 23:59 hrs

Problema

En esta tarea, Ud. modelará una línea de absorción de una observación espectroscópica similar a tareas anteriores pero esta vez utilizando técnicas Bayesianas.

Ésta vez también haremos un par de simplificaciones con respecto a la tarea anterior:

1. El nivel del continuo es una constante $= 10^{-16}$.
2. La longitud de onda del centro de la línea es conocido: 6563 Å.

El espectro que debe modelar se encuentra en el archivo 'espectro.dat', en unidades de flujo por unidad de frecuencia $f_\nu [\text{ergs}^{-1}\text{Hz}^{-1}\text{cm}^{-2}]$ vs. longitud de onda en [Å].

La línea que debe modelar, al igual que en la tarea anterior, no es una gaussiana sino que algo más complicado. Ud. nuevamente seguirá dos procedimientos para modelar el espectro y luego comparará los resultados de los dos modelos.

1. **Línea gaussiana simple.** Note que como la longitud de onda central es conocida, el modelo tiene sólo dos parámetros libres.
2. **Línea gaussiana doble.** Cuando realice el modelo anterior se dará cuenta de que el modelo falla levemente al reproducir las *alas* de la línea. Para mejorar el modelo, se puede intentar modelar la línea como la suma de dos gaussianas. Un modelo como ese está representado en la Figura 1.

Para cada uno de los dos modelos estime, usando métodos Bayesianos, los parámetros (reporte, por ejemplo, la esperanza $E[\theta]$, de los parámetros), y sus intervalos de 68 % de confianza, que en estadística Bayesiana se llaman intervalos de credibilidad.

NOTA 1.

Ud. deberá definir probabilidades a priori para cada uno de los parámetros de los modelos. Sea explícito en su informe sobre cuáles fueron sus elecciones y su justificación.

NOTA 2.

Para calcular la verosimilitud necesitará los errores asociados a cada punto (pixel). Esta vez asumiremos que los errores son gaussianos y constantes a lo largo del espectro. Dado que el continuo es conocido, Ud. puede estimar el *ruido* mirando la variación de los puntos con respecto al valor del continuo conocido. Explícite en su informe el valor que determinó.

Segunda Parte.

Como se mencionó antes, el modelo de dos gaussianas debe producir un mejor fit de los datos, pero a cambio de tener el doble de parámetros libres que el modelo de una gaussiana simple. Utilice métodos de selección Bayesiana de modelos para decidir cuál de los dos modelos es una mejor representación de los datos. Justifique su decisión.

NOTA 1.

Ud. deberá realizar una integral en 4 dimensiones en algún momento. Describa claramente el procedimiento que eligió para hacer dicha integral.

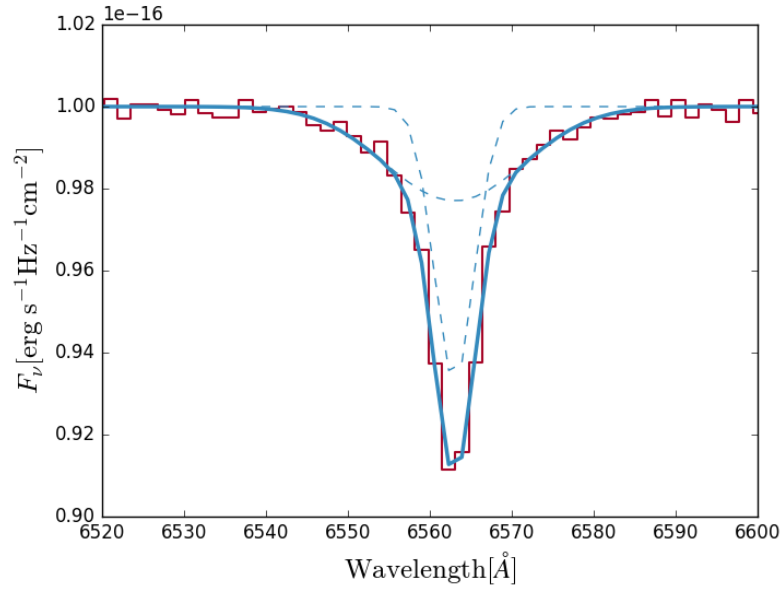


Figura 1: Las líneas punteadas son las dos gaussianas simples cuya suma corresponde a la línea azul continua. La curva roja representa el espectro. Como se aprecia, la suma de las dos gaussianas simples representa bastante bien el espectro (el cual no es originalmente una suma de dos gaussianas). El mejor fit se obtiene a cambio de un modelo más complejo: esta vez son dos gaussianas de centro conocido por lo que el modelo tiene 4 parámetros.

Instrucciones Importantes.

- **NO USE JUPYTER NOTEBOOKS.** Estamos revisando en serio el diseño del código por lo que es imprescindible que entregue su código en un archivo de texto `.py`.
- Evaluaremos su uso correcto de python. Si define una función relativamete larga o con muchos parámetros, recuerde escribir el *docstring* que describa los parámetros que recibe la función, el output, y el detalle de qué es lo que hace la función. Recuerde que generalmente es mejor usar varias funciones cortas (que hagan una sola cosa bien) que una muy larga (que lo haga todo). Utilice nombres explicativos tanto para las funciones como para las variables de su código. El mejor nombre es aquel que permite entender qué hace la función sin tener que leer su implementación ni su *docstring*.
- Su código debe aprobar la guía sintáctica de estilo ([PEP8](#)). Lleva puntaje.
- Utilice `git` durante el desarrollo de la tarea para mantener un historial de los cambios realizados. La siguiente [cheat sheet](#) le puede ser útil. **Revisaremos el uso apropiado de la herramienta y asignaremos una fracción del puntaje a este ítem.** Realice cambios pequeños y guarde su progreso (a través de *commits*) regularmente. No guarde código que no corre o compila (si lo hace por algún motivo deje un mensaje claro que lo indique). Escriba mensajes claros que permitan hacerse una idea de lo que se agregó y/o cambió de un *commit* al siguiente.

- Para hacer un informe completo Ud. debe decidir qué es interesante y agregar las figuras correspondientes. No olvide anotar los ejes e incluir una *caption* o título que describa el contenido de cada figura. Tampoco olvide las unidades asociadas a las cantidades mostradas en los diferentes plots (si es que existen).
- La tarea se entrega subiendo su trabajo a github. Clone este repositorio (el que está en su propia cuenta privada), trabaje en el código y en el informe y cuando haya terminado asegúrese de hacer un último `commit` y luego un `push` para subir todo su trabajo a github.
- El informe debe ser entregado en formato `pdf`, este debe ser claro sin información de más ni de menos. **Esto es muy importante, no escriba de más, esto no mejorará su nota sino que al contrario.** La presente tarea probablemente no requiere informes de más de 3 o 4 páginas en total (dependiendo de cuántas figuras incluya; esto no es una regla estricta, sólo una referencia útil). Asegúrese de utilizar figuras efectivas y tablas para resumir sus resultados. Revise su ortografía.