

Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas

FI3104-01 Métodos Numéricos para la Ciencia e Ingeniería

Tarea N°10: Modelos para ensanchamiento de líneas de absorción

Bruno Scheihing, RUT: 18.954.350-6

01 de Diciembre de 2015

Resumen

En la presente tarea se buscó modelar el ensanchamiento de una línea de absorción en un espectro de radiación. Para esto se proponen dos modelos: uno como la superposición de una recta con un perfil gaussiano, y otro como la superposición de una recta con un perfil lorentziano. Al minimizar las diferencias al cuadrado entre estos modelos y los datos, se encontró que de por sí los ajustes no presentan una gran diferencia entre sus valores de χ^2 , aunque el perfil lorentziano es ligeramente mejor en este sentido. Sin embargo, al realizar un test de Kolmogorov-Smirnov sobre las distribuciones resultantes de los datos y de los modelos, se encuentra que ninguno de los modelos describe satisfactoriamente al espectro. Esto es porque para el número de datos disponibles, el test de Kolmogorov-Smirnov requiere una menor distancia D_n entre las distribuciones para poder declarar un modelo como aceptable. El valor relativamente grande de D_n aparece desde el ruido asociado al continuo de los datos.

1. Introducción

En un sistema cuántico (átomos, núcleos atómicos, etc.), las líneas espectrales de absorción son rangos estrechos de frecuencias en los que se observa una carencia de fotones comparado con frecuencias adyacentes. Esto ocurre, por ejemplo, cuando un fotón tiene una energía muy cercana a la necesaria para cambiar el estado del sistema, por ejemplo, excitar un electrón a un nivel de mayor energía.

Teóricamente las líneas de absorción son casi infinitamente delgadas. Los mecanismos de ensanchamiento son múltiples, dentro de los cuales está el ensanchamiento natural debido al principio de incertidumbre de Heisenberg, y el ensanchamiento Doppler térmico.

En la presente tarea se desea ajustar un espectro de cierta fuente que presenta un continuo y una línea de absorción. Se plantea modelar este espectro con la superposición de una recta con: 1) una gaussiana, y 2) un perfil lorentziano, y juzgar así cuál de estos dos modelos representa mejor la situación física de los datos. A partir de los resultados correspondientes se podría hipotetizar acerca del mecanismo físico que subyace a la situación.

Para realizar el ajuste se usará el algoritmo de Levenberg-Marquardt, minimizando una función de mérito χ^2 . Una vez obtenido el mejor *fit* para cada

modelo, se usará el test de Kolmogorov-Smirnov para determinar cuál modelo es mejor, y la aceptabilidad de cada uno de ellos. Para esto se usará el módulo **kstest** de **scipy.stats**.

Se presentarán gráficos con los mejores ajustes, y gráficos de las funciones de distribución representativas del test de Kolmogorov-Smirnov.

2. Procedimiento

Sea $\{x_i, y_i\}_{i=1, \dots, N}$ el conjunto de datos del espectro observado, con x_i las longitudes de onda del dominio y los y_i el flujo por unidad de frecuencia correspondiente. Se desea minimizar la siguiente función:

$$\chi^2 = \sum_{i=1}^N (y_i - f_\alpha(x_i; \vec{a}))^2$$

donde $\alpha \in \{1, 2\}$ y representa el tipo de ajuste a realizar. Denotaremos por $\alpha = 1$ al ajuste con perfil gaussiano y $\alpha = 2$ al ajuste con perfil Lorentziano. En ambos casos, el vector de parámetros \vec{a} estará dado por $\vec{a} = (A, \mu, \sigma, a, b)$. Los modelos se escriben como:

$$f_1(x; \vec{a}) = a \cdot x + b - \frac{A}{\sqrt{2\pi}\sigma^2} \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

$$f_2(x; \vec{a}) = a \cdot x + b - \frac{A}{\pi\sigma} \cdot \frac{1}{1 + \frac{(x - \mu)^2}{\sigma^2}}$$

Estas funciones (el perfil gaussiano y el lorentziano) están implementadas en **scipy.stats** y se usarán directamente desde allí. La recta será agregada en el código. Notar que para el caso gaussiano μ, σ corresponden a la esperanza y a la desviación estándar de la distribución de probabilidad análoga, mientras que en el caso del perfil de Lorentz no es posible definir estas cantidades de forma matemáticamente rigurosa. Sin embargo, los parámetros juegan el mismo rol: μ es el centro de la distribución y σ es una medida de su ancho.

Debido a que los datos originales para el flujo por unidad de frecuencia son del orden de 10^{-16} (en sus unidades respectivas), conviene re-escalarlos por un factor de $\kappa = 10^{16}$, para evitar problemas de precisión a la hora de manejar dichos números. Es evidente que, en términos de los parámetros \vec{a} , es equivalente minimizar

$$\chi_\kappa^2 = \sum_{i=1}^N (\kappa y_i - f_\alpha(x_i; \vec{a}))^2$$

que minimizar

$$\chi^2 = \sum_{i=1}^N (y_i - f_\alpha(x_i; \vec{a})/\kappa)^2$$

Por lo anterior, minimizaremos χ_κ^2 y luego re-escalaremos los parámetros (A, a, b) de acuerdo a κ para obtener el mejor ajuste en términos de las unidades de los datos originales. No tocaremos el parámetro σ pues este indica la dispersión según la longitud de onda, y su influencia sobre la amplitud es meramente para que un cambio en su valor mantenga constante el área bajo la

curva. Se reportarán todos los valores del ajuste en términos de las unidades originales.

Para hacer esta minimización implementaremos el algoritmo de Levenberg-Marquardt. Sean $Y_i = \kappa y_i$, $i = 1, \dots, N$. Dada la función χ_κ^2 a minimizar, se definen (desde Numerical Recipes in C, Cambridge University Press):

$$\beta_j = -\frac{1}{2} \frac{\partial \chi_\kappa^2}{\partial a_j} = \sum_{i=1}^N (Y_i - f_\alpha(x_i; \vec{a})) \frac{\partial f_\alpha(x_i; \vec{a})}{\partial a_j}$$

$$\alpha_{jl} = (1 + \lambda \delta_{jl}) \sum_{i=1}^N \frac{\partial f_\alpha(x_i; \vec{a})}{\partial a_j} \frac{\partial f_\alpha(x_i; \vec{a})}{\partial a_l}$$

donde δ_{jl} es la delta de Kronecker, y λ es el parámetro intrínseco al algoritmo de Levenberg-Marquardt. Los índices j y l van desde 1 a 5, pues corresponden a las componentes del vector de parámetros \vec{a} .

Dado un set de parámetros \vec{a} , el proceso es como sigue: primero se calcula χ_κ^2 y los términos β_j, α_{jl} , dado un valor de λ (inicialmente será 0.001). Luego se resuelve el sistema de ecuaciones:

$$\sum_{l=1}^5 \alpha_{jl} \delta a_l = \beta_j$$

para los δa_l . Luego se evalúa $\chi_\kappa^2(\vec{a} + \delta \vec{a})$. Si el valor de χ_κ^2 disminuye, entonces se actualizan los parámetros \vec{a} por $\vec{a} + \delta \vec{a}$ y χ_κ^2 por su nuevo valor, y se disminuye λ por un factor de 10. Por otra parte, si el valor de χ_κ^2 no disminuye, entonces se aumenta λ por un factor de 10 y se vuelve a resolver el sistema de ecuaciones. Se detendrá el proceso cuando la variación de χ_κ^2 sea menor que 10^{-5} , en el sistema de unidades amplificado por 10^{16} .

Se debe dar una condición inicial para \vec{a} . Los parámetros de la recta (a, b) se estimarán inicialmente haciendo un ajuste lineal a los datos, mientras que los demás serán elegidos a mano para que estén suficientemente cerca del mejor ajuste.

Para hacer el test de Kolmogorov-Smirnov se usa el módulo **kstest** de **scipy.stats**. Este consiste en comparar la distribución empírica de los datos con la distribución generada por el modelo. La distribución de los datos se construye como:

$$F_{sup}^o(y_i) = \frac{i+1}{N} \quad ; \quad F_{inf}^o(y_i) = \frac{i}{N}$$

donde F_{sup}^o y F_{inf}^o son las funciones de distribución “superior” e “inferior” (para asegurarse de contener a la distribución real de los datos). Luego se construye la distribución del modelo, definida para cada valor de y_i como la cantidad de valores generados por el modelo que son menores (o iguales) a ese valor.

El test de Kolmogorov-Smirnov buscará la máxima distancia D_n vertical entre la función de distribución del modelo y las de los datos. En función de este resultado, y dado el número de datos N , se calcula el nivel de confianza asociado al ajuste. Esto permitirá juzgar si los modelos representan bien a los datos o no.

Tabla 1: Resultados ajuste gaussiano

χ^2	$5.2042 \cdot 10^{-35} [erg^2 s^{-2} Hz^{-2} cm^{-4}]$
A	$8.2353 \cdot 10^{-17} [erg s^{-1} Hz^{-1} cm^{-2} \text{Å}]$
μ	6563.224 [Å]
σ	3.2677 [Å]
a	$7.8021 \cdot 10^{-21} [erg s^{-1} Hz^{-1} cm^{-2} \text{Å}^{-1}]$
b	$8.8773 \cdot 10^{-17} [erg s^{-1} Hz^{-1} cm^{-2}]$
D_n	0.1646
Nivel de confianza	0.002342

Tabla 2: Resultados ajuste lorentziano

χ^2	$5.0056 \cdot 10^{-35} [erg^2 s^{-2} Hz^{-2} cm^{-4}]$
A	$1.1132 \cdot 10^{-16} [erg s^{-1} Hz^{-1} cm^{-2} \text{Å}]$
μ	6563.201 [Å]
σ	3.2152 [Å]
a	$7.9229 \cdot 10^{-21} [erg s^{-1} Hz^{-1} cm^{-2} \text{Å}^{-1}]$
b	$8.8113 \cdot 10^{-17} [erg s^{-1} Hz^{-1} cm^{-2}]$
D_n	0.1659
Nivel de confianza	0.002093

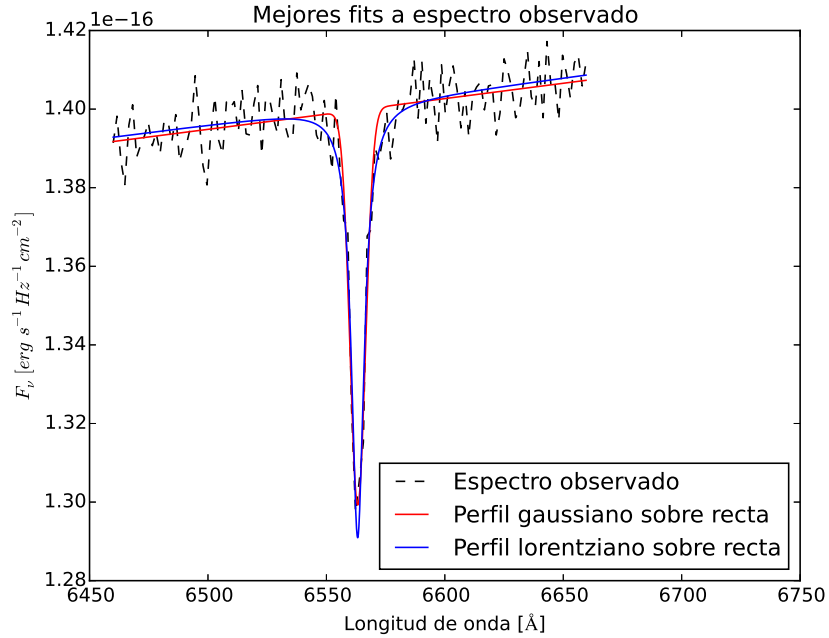


Figura 1: Mejores ajustes a los datos para modelo recta - gaussiana (rojo) y recta - perfil de lorentz (azul)

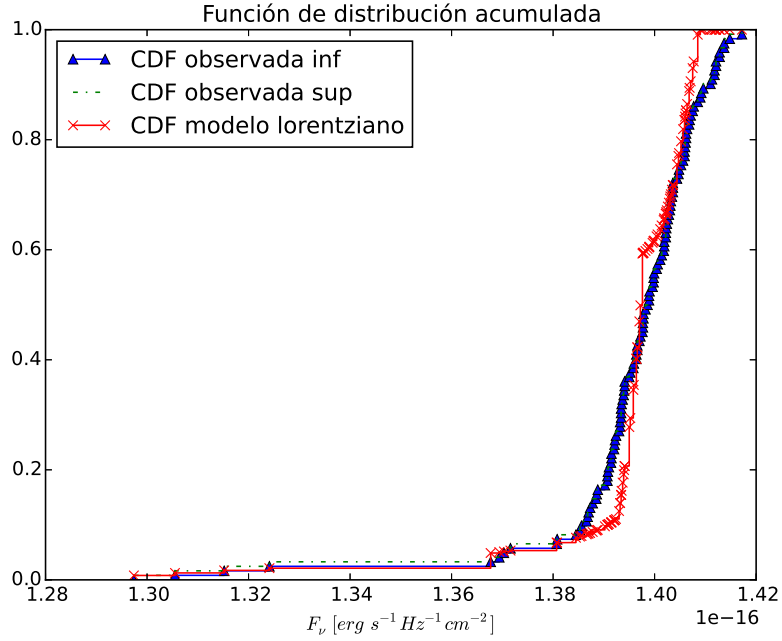
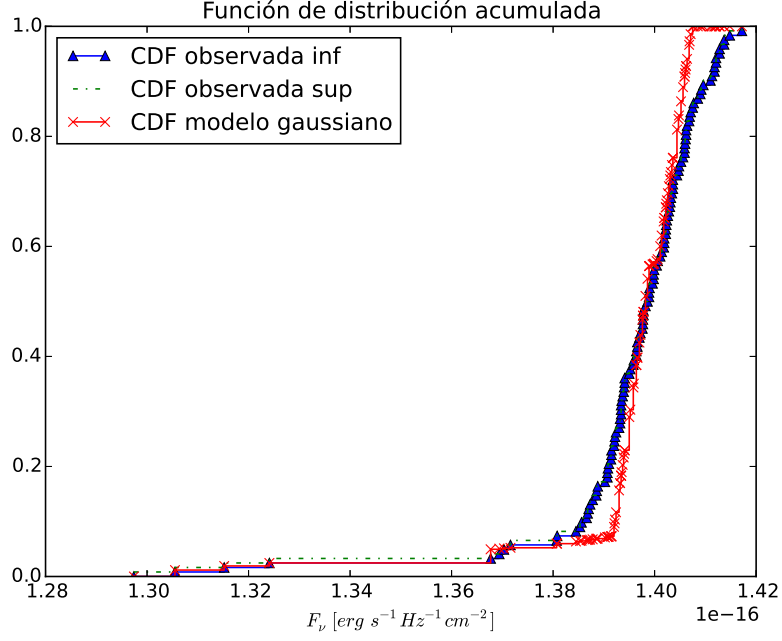


Figura 2: Función de distribución de los valores de F_ν , comparada con la distribución de cada modelo. Arriba: gaussiano, abajo: lorentziano. Las abreviaciones sup. e inf. denotan las funciones F_{sup}^o y F_{inf}^o respectivamente.

3. Resultados

En las Tablas 1 y 2 se presentan los resultados del ajuste con perfil gaussiano y lorentziano respectivamente. En general los parámetros óptimos son similares, y también el valor de χ^2 , aunque en el ajuste con perfil lorentziano este valor es ligeramente más bajo.

La Figura 1 muestra gráficamente los ajustes realizados. A simple vista, no hay muchos argumentos para decidir si un ajuste es mejor que otro o no. Se observa que el perfil de Lorentz suaviza más la transición entre la línea de absorción y el continuo, pero pareciera que en el continuo no sigue la misma línea recta que la gaussiana. Esto es probablemente porque una curva con perfil lorentziano decae mucho más lento que una gaussiana, haciendo que sus efectos sigan siendo perceptibles lejos del centro de su distribución.

Las Tablas 1 y 2 también muestran los valores de D_n para cada modelo y el nivel de confianza asociado. Se observa en ambos casos un valor de D_n similar, ligeramente menor para el caso de la gaussiana, pero ambas curvas tienen un muy bajo nivel de confianza. Estos niveles de confianza son la probabilidad de obtener el resultado dada la hipótesis nula (los datos y el modelo vienen de la misma distribución). Luego se rechaza la hipótesis nula (se puede verificar que el valor crítico de D_n es menor que los obtenidos) en ambos casos, por lo que de acuerdo a este test los modelos no son aceptables.

Para entender el bajo nivel de confianza asociado, observemos la Figura 2. Allí se muestran las funciones de distribución de los modelos (en rojo) y las funciones F_{sup}^o y F_{inf}^o de los datos (en verde y azul, estas son muy similares entre sí dado el número suficientemente grande de datos).

En ambos casos se observa una discrepancia significativa entre las funciones de distribución para valores de F_ν en torno a 1.39 y 1.41 [$10^{-16} \text{ erg s}^{-1} \text{ Hz}^{-1} \text{ cm}^{-2}$]. Si volvemos a observar la Figura 1, vemos que los puntos distribuidos en esos valores corresponden principalmente al ruido del continuo del espectro. Luego se determina que este ruido es la causa del bajo nivel de confianza asociado a los ajustes.

Si hay que elegir uno de los ajustes, como el mejor de acuerdo al test de Kolmogorov-Smirnov, el ajuste gaussiano presenta un nivel de confianza ligeramente superior, por lo que el mejor sería este.

4. Conclusiones

Los resultados de los ajustes fueron en sí satisfactorios, pues se logró obtener curvas que visualmente describen bien los datos, con valores de χ^2 asociados relativamente bajos, y relativamente similares entre sí. El perfil lorentziano mostró minimizar de una forma ligeramente mejor este valor que el perfil gaussiano.

Sin embargo, el test de Kolmogorov-Smirnov arrojó niveles de confianza demasiado bajos comparado con lo deseable para un ajuste. Como se dijo antes, la causa de esto no es la línea de absorción, sino el ruido asociado al continuo de los datos. Este ruido agrega valores por arriba y por debajo del ajuste, que por construcción el *fit* no puede cubrir, haciendo que la distribución de los datos difiera de la de los mejores ajustes.

En ese sentido, tal como indica el test de Kolmogorov-Smirnov, es cierto que el modelo no permite reproducir los valores observados, pues no admite

parámetros que permitan tomar en cuenta el ruido asociado a los datos.

Por otra parte, la línea de absorción pareciera estar bien modelada con ambos ajustes. Sería interesante re-hacer el test de Kolmogorov-Smirnov para un intervalo más pequeño de longitudes de onda en torno a la línea de absorción, y ver cuál es el nivel de confianza asociado, pues allí el ruido tiene menor relevancia en el comportamiento de los datos.