

MODELADO MATEMÁTICO 2

Análisis Galaxia - Parte 2

Carlos Andrés Bautista Torres

Aspectos Generales

En esta segunda parte del análisis de la galaxia se tienen en cuenta cuatro aspectos básicos. El primero es el **análisis de ruido**, es decir, la filtración de los datos para diferenciar los datos provenientes de la galaxia y los provenientes de fuentes externas (ruido de fondo). Posteriormente se realiza un **ajuste del modelo**, teniendo en cuenta las diferentes funciones que se socializaron en clase, haciendo mapas de contorno y perfiles unidimensionales además de un análisis de el valor de intensidad de ruido de fondo que genera menor error χ^2 . Después se realiza una **validación de la metodología de ajuste**, donde se generan datos artificiales de posibles galaxias y se determina la efectividad de la optimización que ajusta la función con los datos. Finalmente se realiza una **exploración de la incertidumbre en los datos**, en donde se calculan los parámetros óptimos a mano.

Estructura del Análisis

Esta carpeta consta de un archivo .ipynb, en donde se encuentra el análisis previamente explicado.

RUIDO:

Primero se cargan los archivos que contienen los datos con los cuales se cotejará el modelo. Posteriormente se recortan muestras del ruido de fondo de las cuatro esquinas de la imagen, cada recorte con una dimensión de 25x25. Una vez concatenados los recortes, es posible realizar el histograma, obteniendo los valores de amplitud, media y varianza.

Luego se optimiza una función gaussiana de tal forma que represente apropiadamente la distribución del ruido. Este proceso dio como resultado un $\chi^2 = 0.0099$, lo que implica que la distribución gaussiana es una buena aproximación.

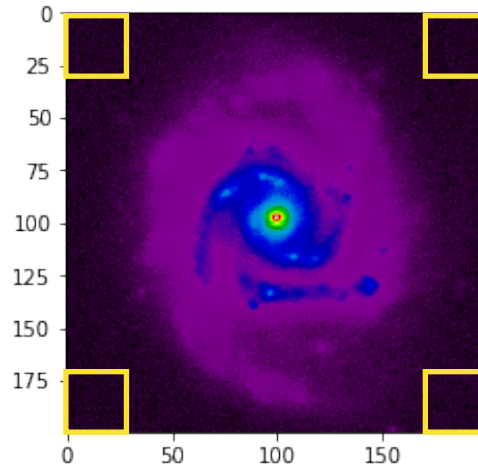


Figura 1: Imagen que muestra las cuatro regiones que se utilizaron para realizar el análisis del ruido de fondo.

Para determinar qué tan efectivo es modelar la dispersión de los datos a partir de una función gaussiana, se generaron datos artificiales y se modelaron con dicha distribución, obteniendo los valores “experimentales” para la media y la varianza:

Nº	μ_{exp}	$Error_r$	σ_{exp}	$Error_r$
1	0.521	0.008	2.406	0.028
2	0.474	0.081	2.273	0.028
3	0.559	0.082	2.372	0.014
4	0.503	0.027	2.284	0.023
5	0.519	0.006	2.472	0.057
6	0.429	0.169	2.244	0.041
7	0.525	0.015	2.178	0.069
8	0.391	0.243	2.383	0.019
9	0.574	0.110	2.247	0.039
10	0.461	0.108	2.411	0.031

En promedio, el parámetro μ obtenido mediante la distribución gaussiana tiene un error relativo de **0.0849** con respecto al parámetro μ de los datos (considerado como μ_{teo}). Por otra parte, el promedio del error relativo del parámetro σ es de **0.039**.

A pesar de la efectividad de la distribución gaussiana para modelar el ruido de fondo, existe otra función de distribución la cual es bueno estudiar para determinar su eficiencia en este tipo de análisis, y es la distribución de Poisson.

$$f(k, \lambda) = \frac{e^{-\lambda} \lambda^k}{k!}$$

Sin embargo, como se aprecia en la ecuación, la distribución de Poisson solo depende de dos parámetros y uno de ellos (k), [está restringido a valores enteros y positivos](#), lo que realmente limita la optimización de la función.

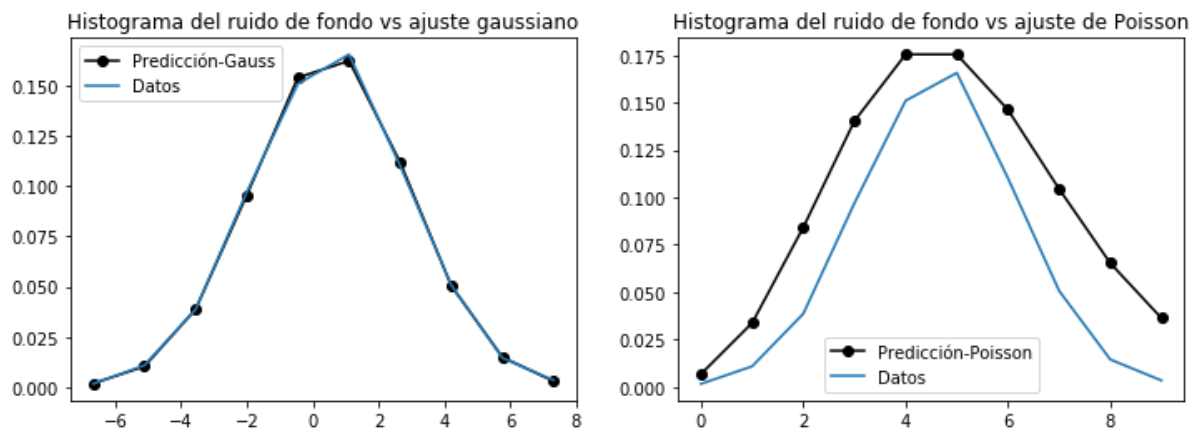


Figura 2: Comparación entre la efectividad de la distribución de Poisson (a la derecha) y la distribución gaussiana (a la izquierda).

AJUSTE DEL MODELO:

En esta sección se realizó un proceso similar al descrito en la primera entrega, pero esta vez se tuvo en cuenta una técnica diferente de visualización que es el gráfico de contorno. También se halló el valor óptimo del ruido de fondo para

obtener el menor χ^2 , determinando en el proceso, que la convolución es una herramienta bastante útil para suavizar los datos.

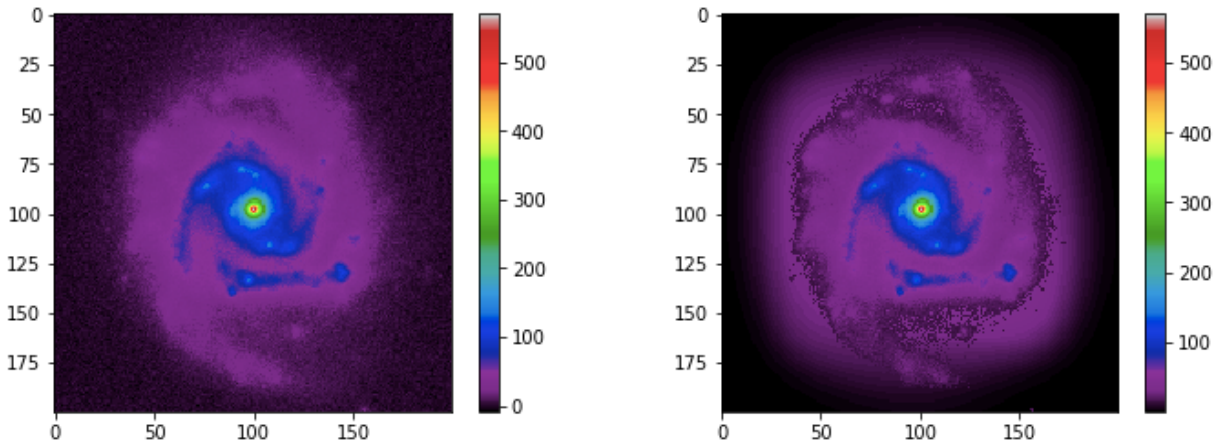
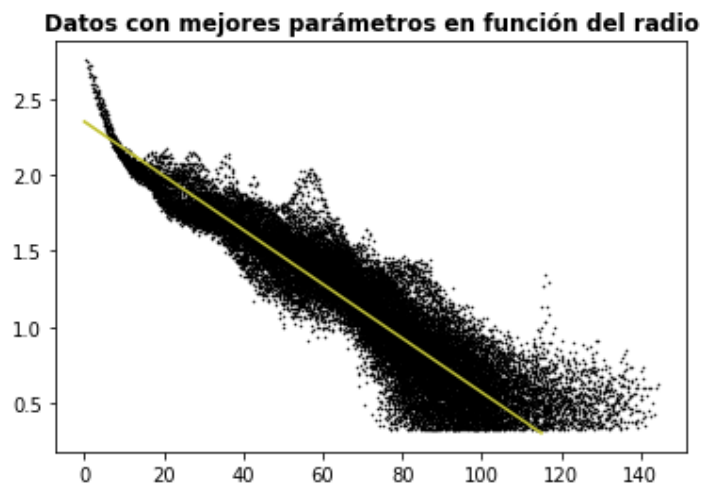


Figura 3: Foto de la galaxia antes y después de aplicarle una convolución.

Al finalizar el análisis se determinó que el valor más apropiado a considerar como ruido de fondo es $\approx \mu + \frac{2}{3}\sigma$. Finalmente se grafica el perfil unidimensional para los datos filtrados (sin ruido), dando como resultado la siguiente gráfica:



Al final se intentó generar una función de dos funciones exponenciales para corregir la diferencia de intensidad en el pico, pero los resultados fueron decepcionantes :(

VALIDACIÓN DE LA METODOLOGÍA DE AJUSTE:

Aquí se generaron datos ficticios de galaxias con ruido consistentes con los datos originales y a partir de los datos artificiales, se optimizaron los parámetros y se calculó el error relativo de cada uno de ellos, teniendo en cuenta que los valores teóricos eran aquellos que se introducían a mano en la creación de los datos.

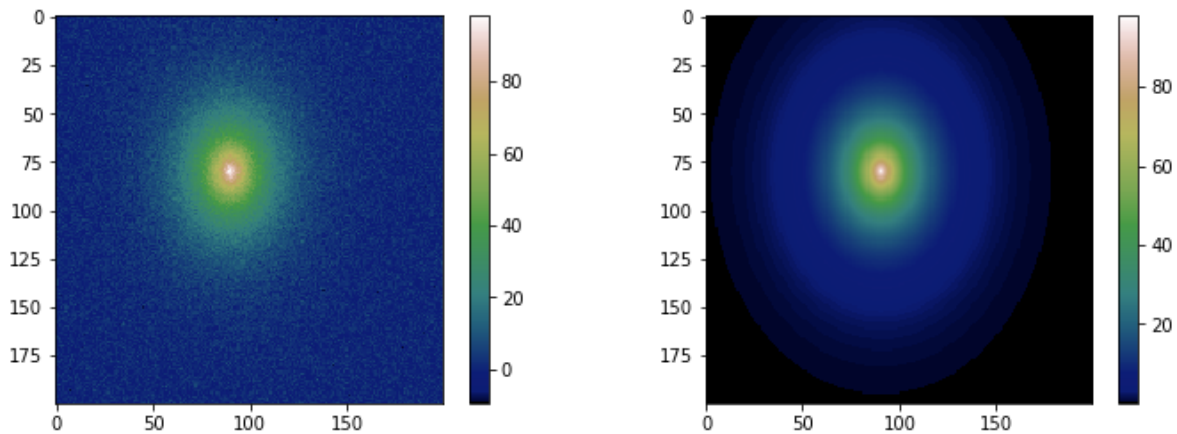


Figura 4: A la izquierda está la galaxia artificial con ruido incorporado, y a la derecha el modelo exponencial que mejor se acopla a los datos.

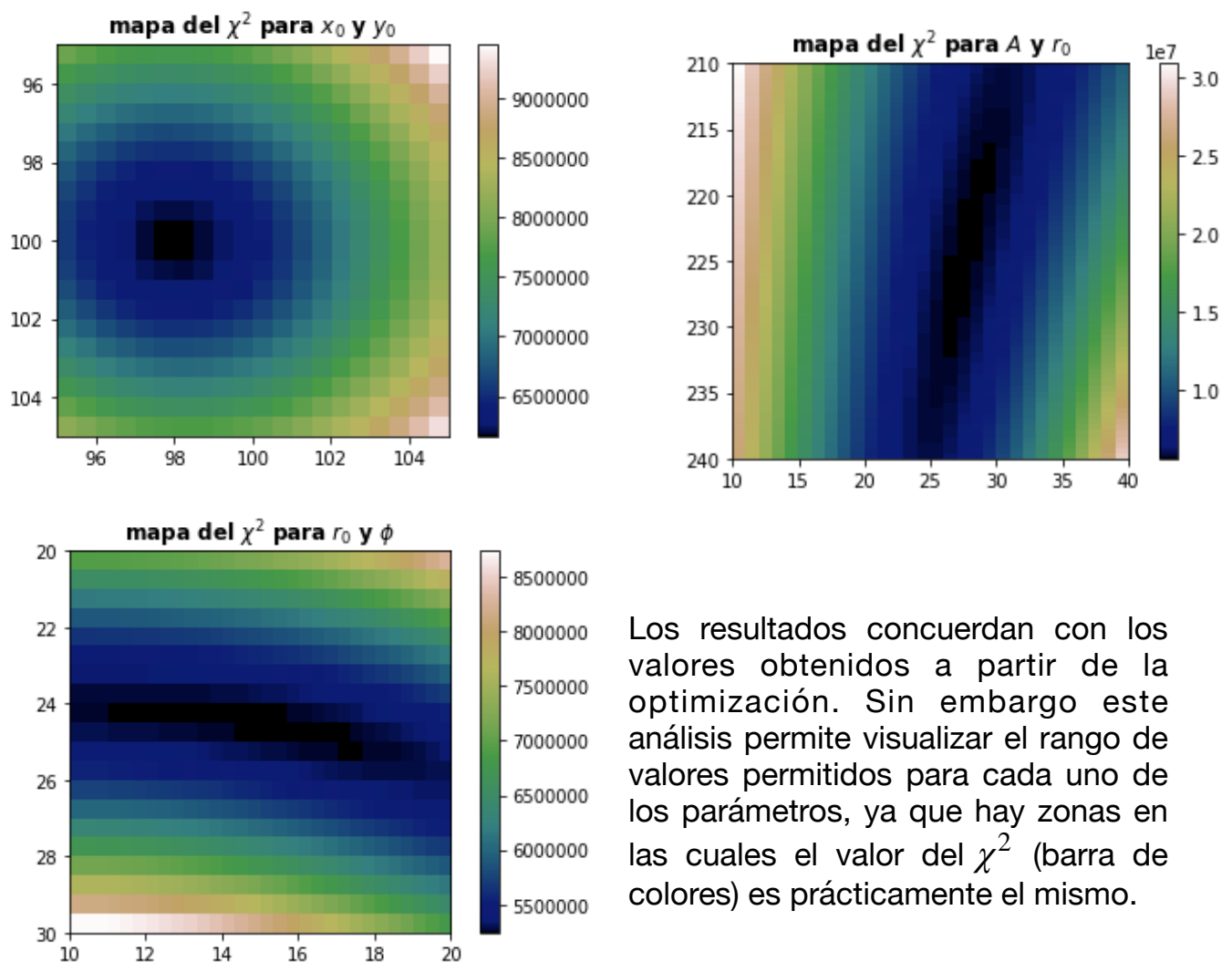
Los resultados después de repetir el procedimiento con valores aleatorios para los diferentes parámetros fueron los siguientes:

Parámetro optimizado	Error relativo promedio
x_0	9.801×10^{-4}
y_0	4.855×10^{-4}
A	1.053×10^{-2}
r_0	1.966×10^{-2}
ϕ	1.488×10^{-2}

Si bien el modelo es bastante eficaz para todos los parámetros, presenta una efectividad mucho mayor en x_0 y y_0 . Esto puede ser debido a que tales parámetros especifican el punto de mayor luminosidad, y dicho punto al ser el más brillante, no se ve tan afectado por el ruido de fondo.

EXPLORACIÓN DE LA INCERTIDUMBRE EN LOS RESULTADOS:

Se realizó lo que la función de scipy hace al optimizar los parámetros, pero ahora manualmente y con rangos limitados que reduzcan el tiempo computacional debido a la cantidad de bucles. Son 5 parámetros, entonces la exploración fue de a dos parámetros.



Desarrollo del Código

Para el desarrollo de este trabajo se requirieron cuatro librerías:

- Numpy
- Astropy.io
- Matplotlib.pyplot
- Scipy.optimize

La primer librería se usó para la declaración de los arreglos matriciales y otras funciones útiles para el desarrollo del código. La segunda se utilizó para cargar los datos de los espectros de la galaxia, es una librería especial que se utiliza en astrofísica, también posee la función *astropy.convolution* con la cual se realizó la convolución. Para graficar los datos y los modelos se utilizó la tercera librería y por último, Scipy.optimize, la librería fundamental para realizar la optimización de parámetros, mediante a la función *leastsq*, la cual permite minimizar el error cuadrático del modelo con respecto a los datos.