

Práctica 4. Ajuste de datos, ajuste de mínimos cuadrados.

Rosa Luz Zamora Peinado

16 de febrero de 2016

Introducción

El presente reporte corresponde a la práctica #4 del curso de Física Computacional. La práctica consistió en hacer ajustes con análisis de regresión, en este caso, lineal para datos de la temperatura en Nueva York en invierno desde 1900 hasta 1999; y exponencial para la presión atmosférica en función de la altitud. Ambos conjuntos de datos obtenidos de *Cuantitative Environment Learning Project* [3][4]. Los métodos de ajuste fueron hechos en IPython Notebook.

En estadística, el análisis de la regresión es un proceso estadístico para la estimación de relaciones entre variables. Incluye muchas técnicas para el modelado y análisis de diversas variables, cuando la atención se centra en la relación entre una variable dependiente y una o más variables independientes. El análisis de regresión es ampliamente utilizado para la predicción y previsión, donde su uso tiene superposición sustancial en el campo de aprendizaje automático. El análisis de regresión se utiliza también para comprender que cuales de las variables independientes están relacionadas con la variable dependiente, y explorar las formas de estas relaciones. [2] En este reporte se habla de dos tipos de regresiones:

Regresión Lineal

La regresión lineal o ajuste lineal es un modelo matemático usado para aproximar la relación de dependencia entre una variable dependiente Y , las variables independientes X_i . Este modelo puede ser expresado de la forma:

$$Y = mx + b$$

Ajuste Exponencial Cuando la curva de regresión de y sobre x es exponencial, es decir para cualquier x considerada, la media de la distribución está dada por la siguiente ecuación predictora:

$$y = \alpha \beta^x$$
$$\log y = \log \alpha + x \log \beta$$

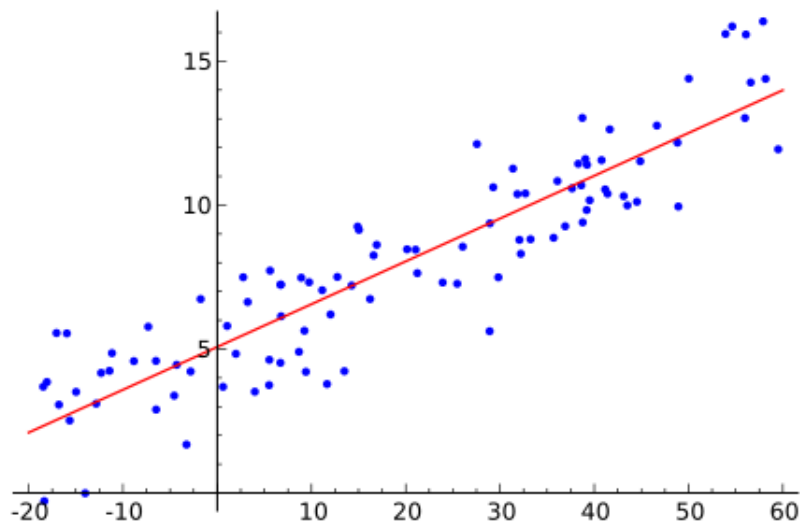


Figura 1: Ejemplo de una regresión lineal con una variable dependiente y una independiente. [5]

Actividades realizadas

- **Ajuste Lineal para la Temperatura de invierno en Nueva York (1900-1999)**

Una vez con los datos obtenidos de [3], a continuación se muestra el código, así como los resultados (Figura.2) del mismo para el ajuste lineal de los datos.

```
from scipy import optimize
import matplotlib.pyplot as plt
import numpy as np
datos = np.loadtxt('temperaturedata.txt')
x=datos[:,0]
y=datos[:,1]
print x
print y
plt.plot(x,y,'yo')

fitfunc = lambda p, x: p[0]*x + p[1]

errfunc = lambda p, x, y: fitfunc(p, x) - y

p0 = [10, 5]

p1, success = optimize.leastsq(errfunc, p0[:], args=(x, y))

time = np.linspace(x.min(), x.max(), 100)
plt.plot(x, y, "co", label="Datos")
plt.plot(time, fitfunc(p1, time), "m-", label="Curva ajustada")
```

```
plt.title("Temperatura de invierno en Nueva York (1900-1999)")
plt.grid()
plt.legend()
plt.xlabel("Año")
plt.ylabel("Temperatura (F)")
```

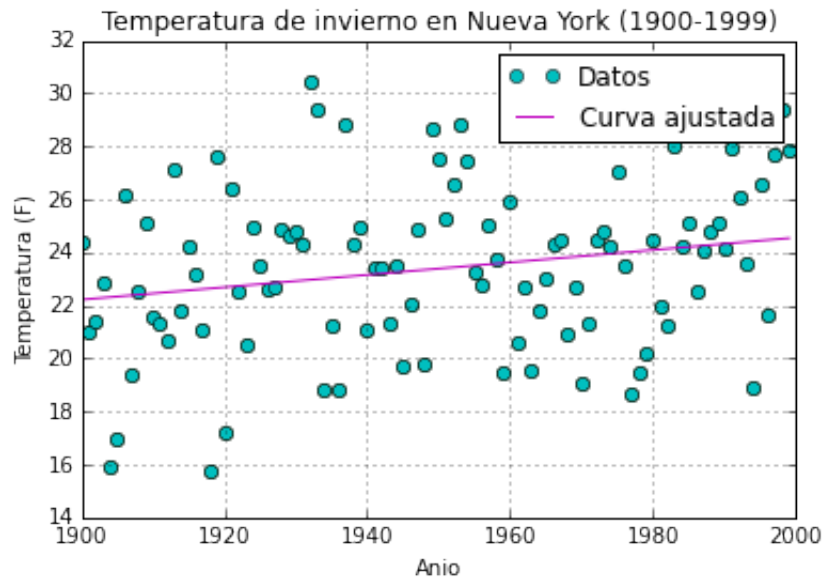


Figura 2: Gráfica de Temperatura en invierno en Nueva York (1900-1990)

■ Ajuste Exponencial para la Presión en función de la Altitud

Una vez con los datos obtenidos de [4], a continuación se muestra el código, así como los resultados(Figura.3) del mismo para el ajuste exponencial de los datos.

```
file = open('pressuredata.txt', 'r+')

with open('pressuredata.txt', 'r') as f:
    data = f.readlines()

    for line in data:
        columns = line.split(",")
        print columns

from scipy.optimize import curve_fit
import matplotlib.pyplot as plt
import numpy as np

datos = np.loadtxt('pressuredata.txt')
```

```

x=datos[:,0]
y=datos[:,1]
print x
print y

def func(x, A, B, C):
    return A * np.exp(-B * x) + C

Y = y + 0.2*np.random.normal(size=len(x))

popt, pcov = curve_fit(func, x, Y)

plt.figure()
plt.plot(x, Y, 'yo', label="Datos")
plt.plot(x, func(x, *popt), 'g-', label="Curva ajustada")

plt.grid()
plt.legend()
plt.title("Presion atmosferica vs. Altitud")
plt.xlabel("Altitud (Km)")
plt.ylabel("Presion (Mb)")

plt.show()

```

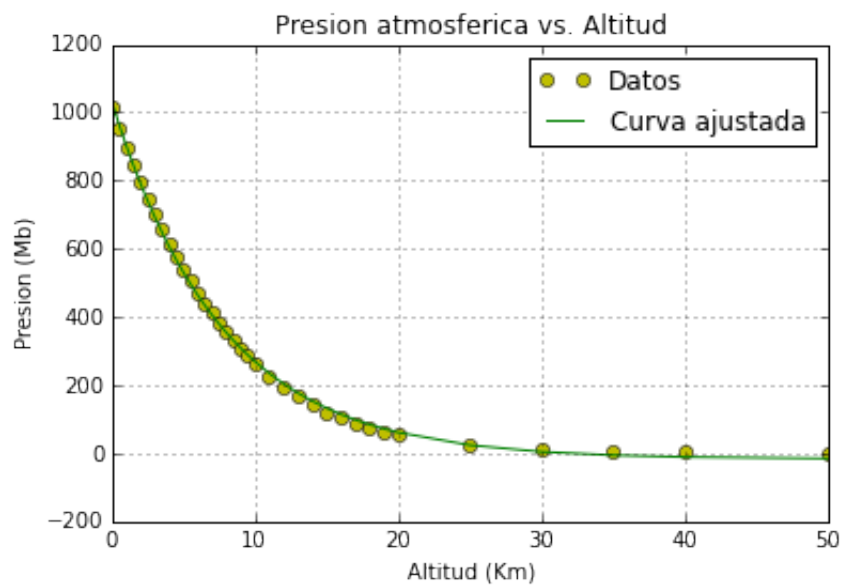


Figura 3: Gráfica de Presión Atmosférica

Bibliografía

- [1] Física Computacional(2016-2) *Actividad 4*. Recuperado el 15 de febrero de 2016 de [http://computacional1.pbworks.com/w/page/105016164/Actividad%204%20\(2016-1\)](http://computacional1.pbworks.com/w/page/105016164/Actividad%204%20(2016-1))
- [2] Wikipedia *Análisis de la Regresión*. Recuperado el 15 de febrero de 2016 de https://es.wikipedia.org/wiki/An%C3%A1lisis_de_la_regresi%C3%B3n
- [3] Quantitative Environmental Learning Project *New York's Winter Mean Temperature* Recuperado el 15 de febrero de 2014 de <http://www.seattlecentral.edu/qelp/sets/048/048.html>
- [4] Quantitative Environmental Learning Project *Earth's Standard Atmosphere* Recuperado el 15 de febrero de 2016 de <http://www.seattlecentral.edu/qelp/sets/024/024.html>
- [5] Wikipedia *Regresión Lineal* Recuperado el 15 de febrero de 2016 de https://es.wikipedia.org/wiki/Regresi%C3%B3n_lineal