Practica I - Parte 3

Método de Ajuste por Mínimos Cuadrados

Introducción a las Ciencias de la Tierra y el Espacio I - 2011

Índice

1. Introducción

Hasta ahora hemos aprendido a medir magnitudes físicas y a expresarlas correctamente. Ahora vamos a aprender algo más interesante. Vamos a tratar de hacer algo que es un objetivo primario para los investigadores en serio, que es intentar establecer interdependencias causales entre dos variables.

1.1. Fuerza de un Resorte (*Ley de Hooke*)

Por ejemplo, es conocida la relación entre la longitud de un resorte l y la fuerza que aplicada F

$$F = -kl$$

donde k es una constante llamada constante elástica del resorte. Alguien debió deducir esta relación que no es caprichosa, es decir, a partir de l y F conocer su relación. La relación es de la forma lineal. Esto es, que se parece a la ecuación de una recta

$$y = bx + a \tag{1}$$

en ella podemos reconocer las siguientes correspondencias

$$\begin{cases} y & = F \\ b & = -k \\ x & = l \\ a & = 0 \end{cases}$$

Supongamos que postulamos una relación de este tipo para dos variables de las cuales tenemos una serie de medidas hechas en el laboratorio. Nuestra meta es determinar a y b. Debido a errores experimentales, los puntos (x_i, y_i) no estarán perfectamente alineados. Estos errores son de origen estadístico, instrumental, etc., pero no es nuestra intención estudiarlos aquí. En pocas palabras, para determinar los coeficientes a y b debemos encontrar la recta de mejor ajuste a todos los pares de puntos (x_i, y_i).

> Nota: Acá estamos suponiendo que la relación entre las variables es lineal. Podría no serlo y entonces el ajuste que hagamos va a diferir bastante de nuestros datos experimentales.

1.2. Mínimos Cuadrados

Queda claro entonces que cualquier ajuste que hagamos, por mejor que éste resulte, no demuestra ninguna relación entre dos variables *a priori*. Un ejemplo de esto es simplemente imaginar un ajuste que lineal, en una función que puede tener una leve curvatura, y no darnos cuenta simplemente por el estrecho dominio de nuestros datos experimentales. El ajuste lineal siempre será "válido" en la región donde se han tomado los datos.

Llamemos $\overline{y}$ a la imagen del valor experimental x_i por la recta de mejor ajuste (cuyos coeficientes son a y b). Luego definimos ε_i a la diferencia para cada x_i, del valor experimental y_i y el correspondiente por la recta de mejor ajuste $\overline{y}_i$.

$$\varepsilon_i = y_i - \underbrace{(bx_i + a)}_{\overline{y}_i} \tag{2}$$

$$\varepsilon_i = y_i - \overline{y}_i$$

Vemos que para cada par (x_i, y_i) esa diferencia en general no va a ser nula (ver figura). Vamos a imponer que la suma de los cuadrados de esas diferencias ε_i sea la menor posible. Para esto es necesario derivar e igualar a cero para luego encontrar el mínimo. Hay que recordar que estamos derivando una función de dos variables a y b, las incógnitas que queremos determinar, así que hay que hacer derivadas parciales. Nuestra función es la siguiente:

$$\sum_{i=1}^{N} \varepsilon_i^2 = \sum_{i=1}^{N} (y_i - bx_i - a)^2 \tag{3}$$

y desarrollando obtenemos

$$\sum_{i=1}^{N} \varepsilon_i^2 = b^2 \sum_{i=1}^{N} x_i^2 + a^2 N - 2b \sum_{i=1}^{N} x_i y_i - 2a \sum_{i=1}^{N} y_i + 2ab \sum_{i=1}^{N} x_i + \sum_{i=1}^{N} y_i^2$$

Para hallar el mínimo derivamos con respecto a nuestras variables e igualamos a cero

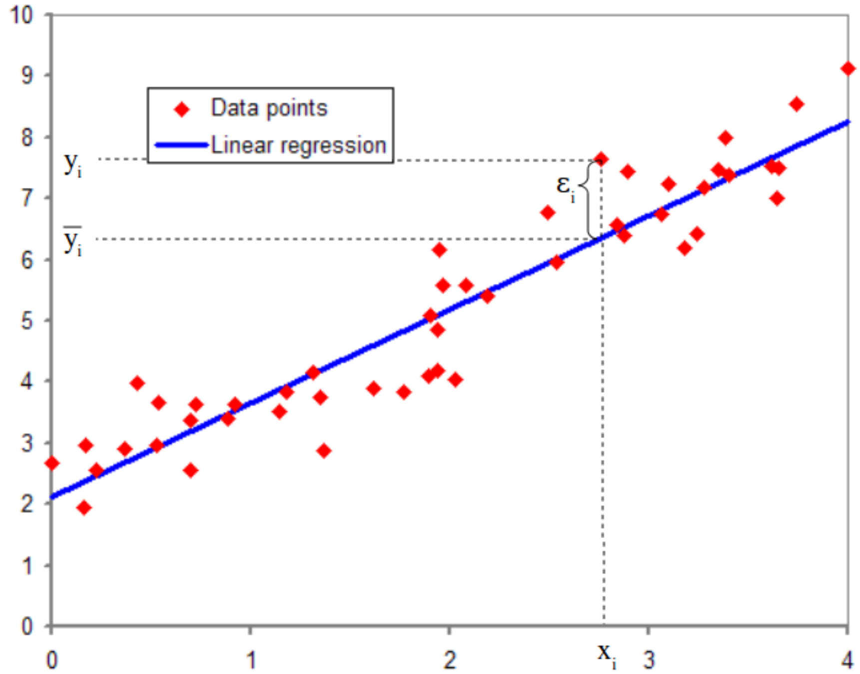
$$\frac{\partial}{\partial a} \left(\sum_{i=1}^{N} \varepsilon_i^2 \right) = 0$$

$$\frac{\partial}{\partial b} \left(\sum_{i=1}^{N} \varepsilon_i^2 \right) = 0$$

Esto se transforma en un sistema de ecuaciones con a y b como incógnitas a determinar

$$2aN - 2 \sum_{i=1}^{N} y_i + 2b \sum_{i=1}^{N} x_i = 0$$

Figura 1: Ejemplo de un ajuste lineal por mínimos cuadrados. Observe que la diferencia de los puntos experimentales ha sido tomada en la dirección vertical, dado que hemos supuesto que no hay error en los valores del eje de las x.



$$2b\sum_{i=1}^{N} x_i - 2\sum_{i=1}^{N} x_i y_i + 2a\sum_{i=1}^{N} x_i = 0$$

Este sistema tiene solución analítica y la misma es

$$b = \frac{N\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2} \tag{4}$$

$$a = \frac{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i x_i}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2} \tag{5}$$

Con este método podemos relacionar magnitudes físicas a través de su determinación experimental.

2. Errores Asociados y Bondad del Ajuste

2.1. Coeficiente de Correlación de *Pearson*

Existe un parámetro que nos dice que tan acertada fue la elección de la recta como curva de mejor ajuste. El mismo se denomina *Coeficiente de Correlación de Pearson* r y cumple que $r \in [-1, 1]$ siendo los valores de mejor correlación, aquellos cercanos al 1 ó -1 y los de correlación muy pobre (o no hay correlación alguna) cuando r es muy cercano a 0. El coeficiente de correlación r puede calcularse mediante la siguiente expresión

$$r = \frac{N\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sqrt{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}\sqrt{N\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i\right)^2}} \tag{6}$$

2.2. Error en el Ajuste (Δa y Δb)

Cabe también preguntarnos cual es el error asociado a los valores de a y b calculados. Los errores Δa y Δb pueden ser fácilmente calculados mediante estas expresiones

$$\Delta b = \sqrt{\left(\frac{N}{N-2}\right)\frac{\sum_{i=1}^{N}(y_i - bx_i - a)^2}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}} \tag{7}$$

$$\Delta a = \sqrt{\left(\frac{1}{N-2}\right)\frac{\sum_{i=1}^{N}(y_i - bx_i - a)^2 \sum_{i=1}^{N} x_i^2}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}} \tag{8}$$

3. Linealización

Pero en muchos casos podemos sospechar que la relación no es lineal; por ejemplo, podría ser del tipo exponencial

$$y = a\mathrm{e}^{bx} \tag{9}$$

Para confirmar esto por mínimos cuadrados, y en tal caso hallar a y b, debemos llevar esta ecuación exponencial a la de una recta o ecuación lineal. A este proceso se le llama linealización.

Para linealizar esta ecuación, aplicamos el logaritmo a ambos lados de la ecuación. De esta manera obtenemos la siguiente expresión

$$\ln(y) = \ln(a) + bx \tag{10}$$

Nuestras nuevas variables a graficar serán “$\ln(y)$” y “x”, y esta gráfica sí nos dará una recta con pendiente b y ordenada en el origen $\ln(a)$.

Ejercicio 1 ¿Cuáles serán las constantes a determinar si linealizamos la siguiente relación?

$$y = \frac{a}{b+x}$$

Respuesta: Invirtiendo la expresión es posible llegar a la ecuación de una recta donde las nuevas constantes a determinar serán

$$\begin{cases} \frac{1}{a} = & \text{pendiente} \\ \frac{b}{a} = & \text{ordenada en el origen} \end{cases}$$

4. Mínimos Cuadrados en Matlab

Una vez cargados en memoria los vectores de datos con los que se desea realizar un ajuste por mínimos cuadrados, es posible realizar el mismo de una manera muy simple. Supongamos que los vectores x e y definidos son los siguientes. Arbitrariamente x podría ser el tiempo e y una magnitud física cualquiera, por ejemplo, una temperatura o una velocidad.

$$x = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$$

$$y = (3{,}2, 4{,}9, 7{,}1, 8{,}8, 11{,}3, 13{,}1, 14{,}5, 16{,}8, 19{,}7, 22{,}1)$$

Tales vectores deberán ser ingresados en Matlab de la siguiente manera

```
x=[1 2 3 4 5 6 7 8 9 10]
y=[3.2 4.9 7.1 8.8 11.3 13.1 14.5 16.8 19.7 22.1]
```

En la ventana de comandos, ejecutamos la tarea *polyfit*

```
p=polyfit(x,y,1)
```

Esto creara una variable nueva P que es un vector que contiene 2 elementos. Los mismos son el coeficiente angular y la ordenada en el origen de la recta que mejor ajusta los datos. La sintaxis de la tarea *polyfit* involucra los vectores x e y y el grado del polinomio al cual se quieren ajustar los datos. En un caso mas general la tarea puede aplicarse para encontrar la mejor parábola que ajusta los datos, o polinomios de mayor grado, simplemente el vector P tendrá $n + 1$ elementos donde n es el grado del polinomio. En nuestro ejemplo $n = 1$. Luego para conocer el valor de las constantes ajustadas simplemente nos fijamos en los valores de los elementos del vector P.

4.1. Bondad del Ajuste (*corrcoef*)

Finalmente podemos computar el Coeficiente de Correlación de Pearson r de la siguiente manera

```
corrcoef(x,y)
```

Este comando devolverá una matriz cuyos elementos de la diagonal serán 1. Los elementos en los extremos de la anti-diagonal contienen el valor de r. Para este ejemplo es fácil corroborar que $r = 0{,}9978$. Se trata de un buen ajuste?