



Física Experimental 1



Método de Mínimos cuadrados

Supongamos que tenemos un modelo de un sistema físico descrito por una función $y = f(x; a, b, c, \dots)$, donde a, b, c , etc. son parámetros, y un conjunto de medidas experimentales del sistema físico (x_i, y_i) . El problema entonces, es determinar los parámetros a, b, c, \dots , que hagan que la función se *ajuste lo mejor posible* a los datos adquiridos. Tenemos que definir o convenir qué significa “ajustar lo mejor posible”. Para ello, buscaremos minimizar la suma de los cuadrados de las “distancias” (ver fig.) entre los puntos experimentales y la curva teórica. Esto es:

$$\sum_i E_i^2 = \sum_i (y_i - f(x_i, a, b, \dots))^2 \quad (1)$$

Para minimizar esta función con respecto a los parámetros, buscaremos los valores de los parámetros a_{min}, b_{min}, \dots , que cumplan:

$$\begin{aligned} \frac{\partial \sum_i (y_i - f(x_i; a, b, \dots))^2}{\partial a} &= 0 \\ \frac{\partial \sum_i (y_i - f(x_i; a, b, \dots))^2}{\partial b} &= 0 \end{aligned} \quad (2)$$

y así para los otros parámetros (esta condición es análoga a la que debe satisfacer una función de una variable, $f(x)$, para presentar un extremo en $x = x_0$; la derivada en este punto debe ser nula: $f'(x_0) = 0$). Los valores que obtengamos del sistema de ecuaciones 2 serán los valores buscados.

Vamos a ilustrar ahora lo anterior con un ejemplo. Supongamos que modelamos la relación entre dos magnitudes x e y por:

$$y = f(x; a, b) = ax + b \quad (3)$$

En general no existirán valores de a y b que logren que la recta por ellos definida pase por *todos* los puntos medidos, debido a los diferentes tipos de errores cometidos al medir.

El problema es entonces, dado el conjunto de medidas (x_i, y_i) , con $i = 1, 2, \dots, n$; evaluar del mejor modo posible los parámetros a y b , de manera de obtener la recta que *mejor* se aproxime a todos los puntos.

Sea entonces una recta de coeficientes a y b como se observa en la figura 1. Llamemos y_{iteo} a la ordenada en el punto de la recta que tiene abscisa x_i y cumple la relación teórica: $y_{iteo} = ax_i + b$. La distancia que hay de ese punto al valor medido será $E_i = |y_i - y_{iteo}|$. Tomaremos entonces como una buena estimación de la desviación de las observaciones con respecto a esta recta, el valor:

$$\sum E_i^2 = \sum (y_i - (ax_i + b))^2 \quad (4)$$

donde 4 es un caso particular de 1. Observe que es función solamente de a y b , porque los x_i e y_i son conocidos.

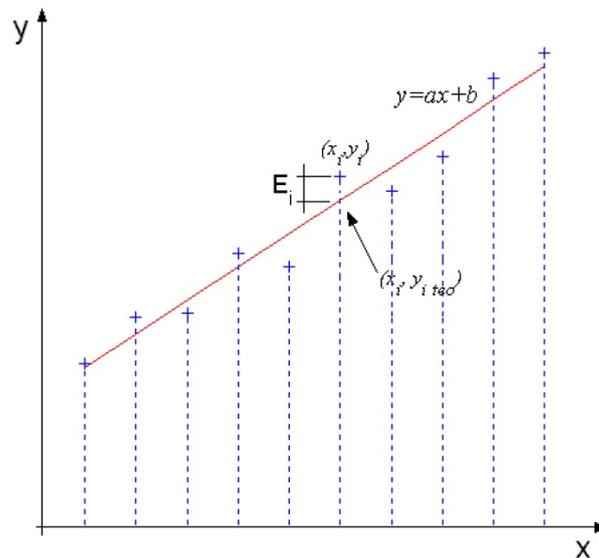


Figura 1: Distancias entre los puntos experimentales y la curva teórica.

Los valores de a y b para los cuales esta expresión es mínima son aquellos que cumplen:

$$\begin{aligned} \frac{\partial \sum E_i^2}{\partial a} &= 0 \\ \frac{\partial \sum E_i^2}{\partial b} &= 0 \end{aligned} \quad (5)$$

de donde, sustituyendo por 4 y resolviendo estas ecuaciones obtenemos:

$$a_{min} = \frac{(n \sum x_i y_i - \sum x_i \sum y_i)}{(n \sum x_i^2 - (\sum x_i)^2)} \quad (6)$$

$$b_{min} = \frac{(\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i)}{(n \sum x_i^2 - (\sum x_i)^2)} \quad (7)$$

Se sugiere al estudiante verificar estos resultados.

Veamos ahora cómo saber en forma cuantitativa si el ajuste por el modelo lineal es *bueno*. Para ello se define el *coeficiente de correlación*:

$$r = \frac{cov(x, y)}{\sigma_x \sigma_y} \quad (8)$$

donde $cov(x, y)$ representa la covarianza, definida por:

$$cov(x, y) = \frac{\sum x_i y_i - (1/n) (\sum x_i \sum y_i)}{(n - 1)} \quad (9)$$

y σ_x y σ_y son las desviaciones cuadráticas medias de x e y respectivamente:

$$\sigma_x^2 = \frac{\sum_{i=1}^n \left(x_i - \frac{\sum_{i=1}^n x_i}{n} \right)^2}{n - 1}$$

y una relación análoga para σ_y .

Se puede ver que por su definición $0 \leq |r| \leq 1$, y dentro de este intervalo podemos distinguir tres casos:

- $|r| \approx 1$ implica correlación lineal fuerte ($r = -1$ implica una relación lineal con pendiente negativa).
- $|r| \approx 0$ implica correlación nula, esto es, que las magnitudes x e y no están relacionadas.
- $0 < |r| < 1$ implica correlación estadística.

Las definiciones anteriores nos permiten obtener expresiones sencillas para las incertidumbres de a y b . Usando los resultados anteriores se puede encontrar una expresión para σ_a y σ_b de la forma:

$$\sigma_a = |a| [(1/r^2) - 1]^{1/2} / \sqrt{(n-2)} \quad (10)$$

y para b :

$$\sigma_b = \sigma_a \left[\left(\sum x_i^2 \right) / n \right]^{1/2} \quad (11)$$

En el ejemplo que se trabajó anteriormente, el modelo teórico es lineal. En los casos en que el modelo no sea lineal, es posible de todas formas aplicar el método de mínimos cuadrados, minimizando las distancias de los puntos experimentales a la función que corresponda. En muchos casos es posible mediante un cambio de variable, reescribir la ecuación teórica en forma lineal. De esa forma podemos aplicar el método de mínimos cuadrados tal como fue explicado en el ejemplo.

Ejercicio: Verifique que las siguientes relaciones se pueden llevar a una forma $u = Av + B$. Determine A y B en función de a y b :

- | | |
|--|--------------------------|
| 1. $Y = bX^a \Rightarrow \log Y = \log b + a \log X$ | $u = \log Y; v = \log X$ |
| 2. $Y = be^{aX} \Rightarrow \ln Y = \ln b + aX$ | $u = \ln Y; v = X$ |
| 3. $Y = \frac{a}{b+X} \Rightarrow \frac{1}{Y} = \frac{b}{a} + \frac{X}{a}$ | $u = \frac{1}{Y}; v = X$ |