textos\ física 1 | lin-reg.tex

AJUSTE DE UNA LINEA RECTA POR EL METODO DE LOS MINIMOS CUADRADOS

RAFAEL GARCIA MOLINA, *Departamento de Física, Universidad de Murcia*

Dado un conjunto de n parejas (x_i, y_i) de datos experimentales, hemos de encontrar la ecuación de la recta $y = ax + b$ que pasa lo más cerca posible de los puntos experimentales (de forma que éstos estén aproximadamente equirrepartidos alrededor de la recta).

La técnica del ajuste por mínimos cuadrados (o regresión lineal) permite obtener la pendiente a de la recta y la ordenada b en el origen, correspondientes a la recta teórica $y = ax + b$ que mejor se ajusta a los n datos experimentales (x_i, y_i). Las expresiones que permiten calcular la pendiente, a, y la ordenada en el origen, b, son las siguientes:[1]

$$a = \frac{1}{D}\sum_{i=1}^{n}(x_i - \langle x\rangle)y_i \qquad\qquad b = \langle y\rangle - a\langle x\rangle$$

donde $\langle x\rangle$ e $\langle y\rangle$ representan los valores medios de los datos x_i e y_i, respectivamente:

$$\langle x\rangle = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad\qquad \langle y\rangle = \frac{1}{n}\sum_{i=1}^{n} y_i$$

El valor de D necesario para los cálculos anteriores se define como $D = \sum_{i=1}^{n}(x_i - \langle x\rangle)^2$.

Los errores absolutos correspondientes a la pendiente, ε_a, y a la ordenada en el origen, ε_b, son:

$$\varepsilon_a \approx \sqrt{\frac{1}{D}\frac{\sum_{i=1}^{n} d_i^2}{n-2}} \qquad\qquad \varepsilon_b \approx \sqrt{\left(\frac{1}{n} + \frac{\langle x\rangle^2}{D}\right)\frac{\sum_{i=1}^{n} d_i^2}{n-2}}$$

donde $d_i = y_i - ax_i - b$.

Puesto que la pendiente y la ordenada en el origen de una recta son parámetros que suelen estar asociados a algunas (combinaciones de) magnitudes de interés físico, después de realizar las operaciones anteriores obtendremos los valores de a y b junto con sus respectivos errores: $a \pm \varepsilon_a$ y $b \pm \varepsilon_b$.

En la representación gráfica aparecerán los puntos experimentales (x_i, y_i), junto con sus errores absolutos. La recta $y = ax + b$ ajustada por mínimos cuadrados (donde los valores de a y b se habrán calculado como se indicó anteriormente) aparecerá superpuesta a los datos experimentales.

En el procedimiento expuesto, para representar la recta no se suelen tener en cuenta los errores de a y b, éstos sólo intervienen para caracterizar a y b a efectos de posteriores cálculos numéricos.

Un parámetro que se emplea para caracterizar la calidad de un ajuste es el coeficiente de regresión lineal r, definido como sigue:

$$r = \sqrt{\frac{\sum_{i=1}^{n}(ax_i + b - \langle y\rangle)^2}{\sum_{i=1}^{n}(y_i - \langle y\rangle)^2}}$$

El valor de r nos indica lo bien (o mal) que la recta $y = ax + b$ obtenida mediante ajuste por mínimos cuadrados se ajusta a los n puntos experimentales (x_i, y_i). El coeficiente de regresión lineal está comprendido entre 0 y 1. Un buen ajuste está indicado por un valor de r muy próximo a la unidad: $r \approx 1$.

Veamos un caso práctico. Supongamos que medimos la diferencia de potencial V entre los extremos de un cable conductor cuando variamos la intensidad de corriente I que circula por él. Los datos experimentales aparecen en la tabla 1. De acuerdo con la ley de Ohm, cabe esperar un comportamiento lineal de la forma $V = RI$; es decir, la pendiente de la recta nos daría la resistencia R del conductor y la ordenada en el origen debería de ser cero.

Procesando estos datos de acuerdo con lo expuesto anteriormente,[2] obtenemos[3] que $a = 2.46 \pm 0.06$, $b = 0.09 \pm 0.08$ y el coeficiente de regresión lineal vale $r = 0.998$. Interpretando en términos físicos estos resultados podemos decir que la resistencia del conductor es $R = 2.46 \pm 0.06\ \Omega$ y, aunque la ordenada en el origen no es nula, vale aproximadamente cero dentro de los límites de su error; el valor del coeficiente de regresión lineal, $r = 0.998$, nos indica que el ajuste es relativamente bueno.

TABLA 1

$I \pm 0.02$ (A)	$V \pm 0.1$ (V)
0.40	1.1
0.60	1.5
0.82	2.2
0.98	2.5
1.20	2.9
1.38	3.5
1.60	4.1
1.80	4.5

En la Fig. 1a hemos representado los datos experimentales (x_i, y_i), que en este caso corresponden a (I_i, V_i). Nótese que no han de unirse mediante una línea quebrada que pase por todos los puntos, como en la Fig. 1b.

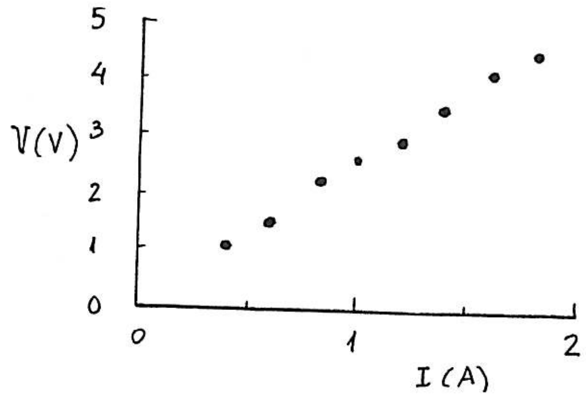


FIG. 1A

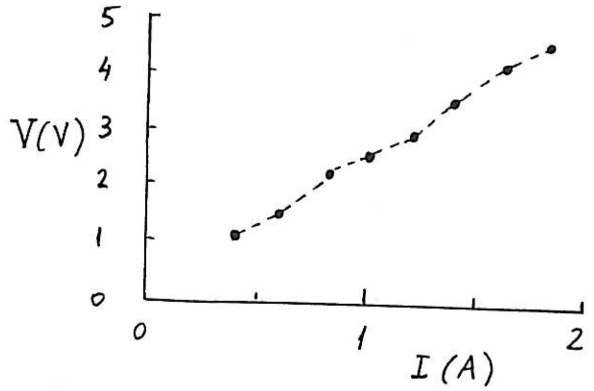


FIG. 1B

En la Fig. 2 se ha representado la recta $y = ax + b$ obtenida mediante la técnica del ajuste por mínimos cuadrados. Por último, la Fig. 3 corresponde a la representación final, en la que aparecen conjuntamente los puntos experimentales y la recta del ajuste.

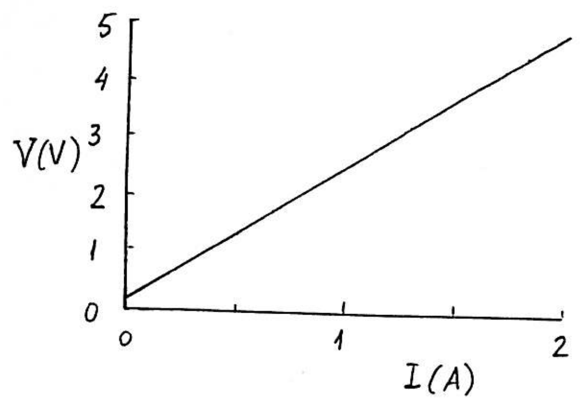
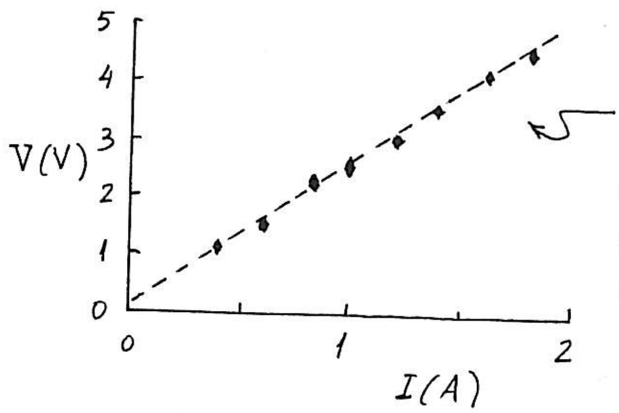


FIG. 2



El tamaño de los puntos apenas deja ver las barras de error

FIG. 3

Referencias

[1] G. L. Squires, *Practical Physics* (McGraw-Hill, London, 1968).

[2] Los resultados presentados se han obtenido con el programa SQUIRES.MCD, del cual se adjunta una copia para el caso discutido en este ejemplo.

[3] Hemos aplicado los convenios de la teoría de errores sobre la cantidad de cifras significativas que ha de presentar un resultado.