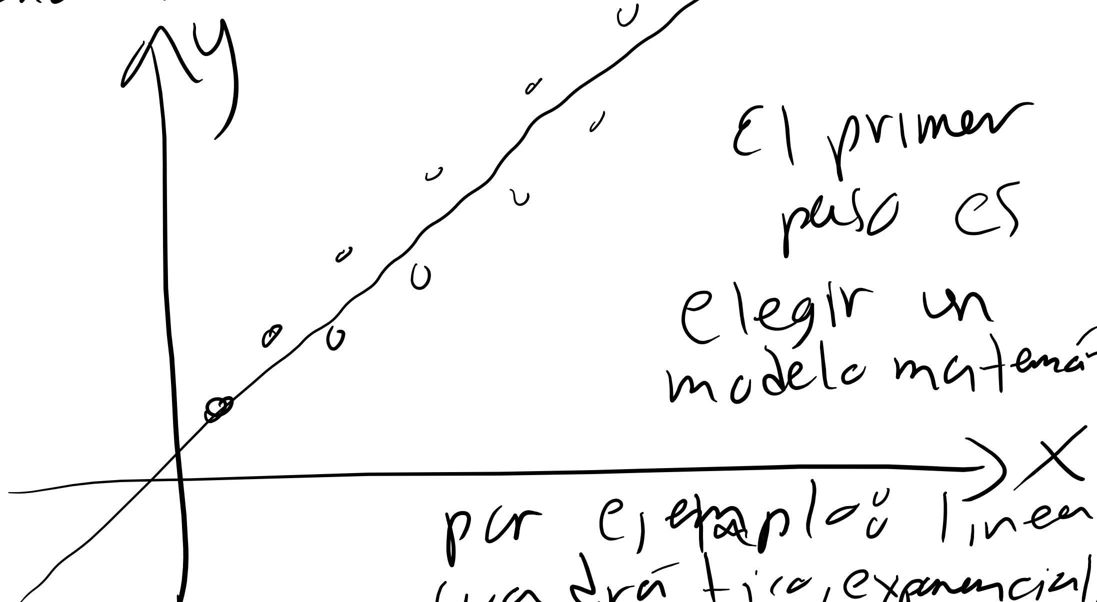
Recapitulando...

Estabamos discutiendo el problema de asociar una ecuación a una tabla de datos (ajuste de curva), existen varias metologías: ajuste por mínimos cuadrados, interpolación de Newton o de Lagrange.

Método de mínimos cuadrados

Es una de las técnicas más populares para construir la ecuación de una tabla de valores relacionados.

x	y
x_1	y_1
x_2	y_2
x_3	y_3
$\vdots$	$\vdots$
x_n	y_n

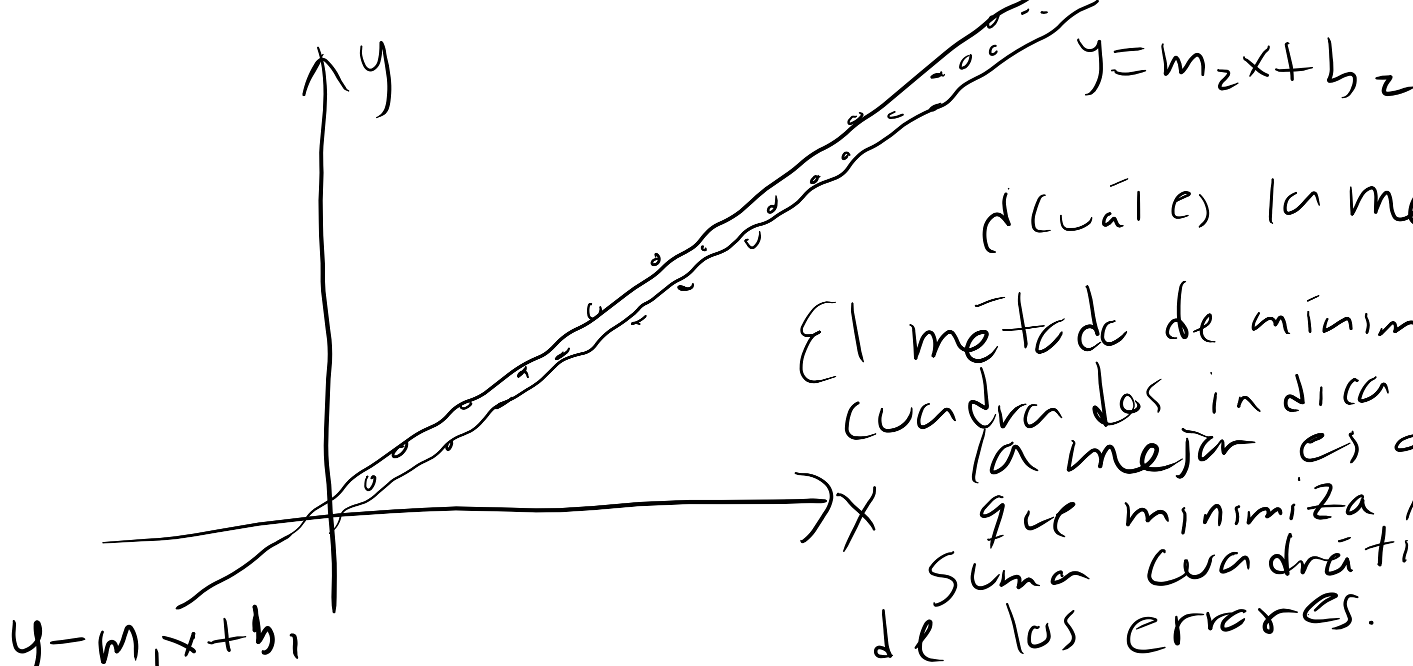


El primer paso es elegir un modelo matemático

por ejemplo: lineal, cuadrático, exponencial, trigonométrico, etc.

Supongamos que es lineal la tendencia

De todas las posibles rectas ¿cuál es la mejor?



¿Cuál es la mejor?

El método de mínimos cuadrados indica que la mejor es aquella que minimiza la suma cuadrática de los errores. que se define como

$$\boxed{S = \sum_{i=1}^{n} \left(y_{real} - y_{aprox}\right)^2}$$

donde $y_{aprox} = mx + b$

Ejemplo:

x	y
1	3.2
2	4.9
3	7.4
4	8.8

Dados los siguientes valores y los modelos matemáticos

$$y = 2x + 1$$

$$y = 1.9x + 1.1$$

¿Cuál es el mejor según el criterio de mínimos cuadrados?

x	y_{real}	$y_{aprox_1} = 2x+1$	$y_{real} - y_{aprox}$	$(y_{real} - y_{aprox})^2$
1	3.2	3	0.2	0.04
2	4.9	5	-0.1	0.01
3	7.4	7	0.4	0.16
4	8.8	9	-0.2	0.04
				$\boxed{S = 0.25}$

x	y_{real}	$y_{aprox_2} = 1.9x+1.1$	$y_{real} - y_{aprox}$	$(y_{real} - y_{aprox})^2$
1	3.2	3	0.2	0.04
2	4.9	4.9	0	0
3	7.4	6.8	0.6	0.36
4	8.8	8.7	0.1	0.01
				$\boxed{S = 0.41}$

En base al método de mínimos cuadrados el mejor sería el primero pues tiene más chica la S.

Pero.

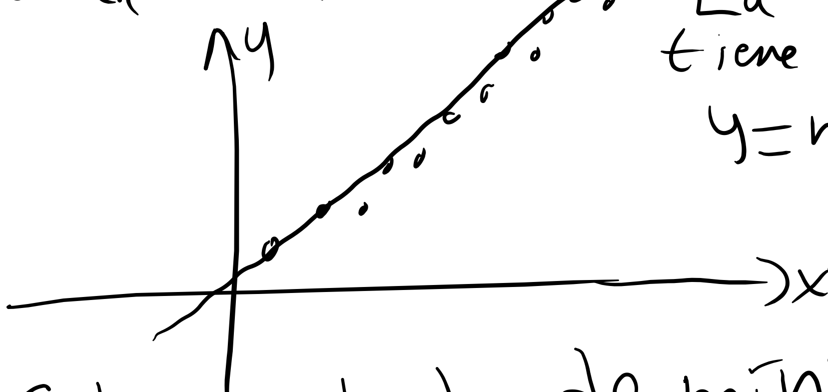
¿Existirá alguno mejor?, es decir, que tenga más chica la S.

Según Excel la mejor recta es $\boxed{y = \underbrace{1.93}_{m}x + \underbrace{1.25}_{b}}$

x	y_{real}	$y_{aprox} = 1.93x + 1.25$	$y_{real} - y_{aprox}$	$(y_{real} - y_{apr})^2$
1	3.2	3.18	0.02	0.0004
2	4.9	5.11	-0.21	0.0441
3	7.4	7.04	0.36	0.1296
4	8.8	8.97	0.17	0.0289
				$S = 0.203$

Estamos interesados en determinar la mejor recta por mínimos cuadrados para un conjunto de datos.

x	y
x_1	y_1
x_2	y_2
$\vdots$	$\vdots$
x_n	y_n



La ecuación tiene la forma $y = mx + b$

El método de mínimos cuadrados te da la mejor "m" y la mejor "b". Vamos a deducir las fórmulas de m, b.

Queremos minimizar la suma cuadrática de errores

$$S = \sum_{i=1}^{n} (y - y_{aprox})^2, \quad y_{aprox} = mx + b$$

$$S = \sum (y - mx - b)^2$$

para minimizar S derivo con respecto a m y luego con respecto a b.

$$\frac{\partial S}{\partial m} = 2\sum(y - mx - b)(-x) = -2\sum(xy - mx^2 + bx)$$

$$\frac{\partial S}{\partial b} = 2\sum(y - mx - b)(-1) = -2\sum(y - mx - b)$$

Igualo a cero

$$\begin{cases} -2\sum(xy - mx^2 - bx) = 0 \\ -2\sum(y - mx - b) = 0 \end{cases}$$

divido cada ecuación entre -2 y abro las sumas

$$\begin{cases} \sum xy - m\sum x^2 - b\sum x = 0 \\ \sum y - m\sum x - b\underbrace{\sum 1}_{n} = 0 \end{cases}$$

rescribo

$$\begin{cases} \sum xy = m\sum x^2 + b\sum x \\ \sum y = m\sum x + nb \end{cases}$$

Sistema lineal para m y b resuelvo por Crammer

$$m = \frac{\begin{vmatrix} \sum xy & \sum x \\ \sum y & n \end{vmatrix}}{\begin{vmatrix} \sum x^2 & \sum x \\ \sum x & n \end{vmatrix}} = \frac{n\sum xy - \sum x\sum y}{n\sum x^2 - (\sum x)^2}$$

$$\therefore \boxed{m = \frac{n\sum xy - \sum x\sum y}{n\sum x^2 - (\sum x)^2}}$$

$$b = \frac{\begin{vmatrix} \Sigma x^2 & \Sigma xy \\ \Sigma x & \Sigma y \end{vmatrix}}{\begin{vmatrix} \Sigma x^2 & \Sigma x \\ \Sigma x & n \end{vmatrix}} = \frac{\Sigma x^2 \Sigma y - \Sigma x \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2}$$

$$\therefore \boxed{b = \frac{\Sigma x^2 \Sigma y - \Sigma x \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2}}$$

B	C	D	E	F
	x	**y**	xy	x^2
	1	**3.2**	3.2	1
	2	**4.9**	9.8	4
	3	**7.4**	22.2	9
	4	**8.8**	35.2	16
sumas	10	24.3	70.4	30
	Σx	Σy	Σxy	Σx^2

$n = 4$

¡cuatro parejas!

J	K
m	b
1.93	1.25