



## UN NUEVO METODO PARA AJUSTE LINEAL BASADO EN EL EQUILIBRIO DE TORQUES

G. E. Sisniegas Charcape<sup>a</sup>

<sup>a</sup>Departamento de Física Interdisciplinaria. Facultad de Ciencias Físicas,  
Universidad Nacional Mayor de San Marcos. Apartado postal 14 – 0149, Lima 14 - Perú.

---

### RESUMEN

Se presenta un método alternativo para encontrar la ecuación de una línea recta ajustada a una dada población de puntos con cualquier grado de dispersión. Este método da dos soluciones: La ecuación de la línea recta ajustada a tal población de puntos y la ecuación de la perpendicular correspondiente. Este nuevo método contrasta con el método clásico de regresión lineal en su precisión y en su independencia de la dispersión de datos. Mientras que por el método clásico, las soluciones son siempre ambivalentes, aproximadas y afectadas por la dispersión de los datos, el nuevo método da la solución junto a otra identificada como la línea perpendicular ajustada o la ecuación de la transversal a la misma población de puntos.

---

### ABSTRACT

An alternative method is presented here to find the adjusted straight line equation for a population of points with every level of dispersion. This method gives two solutions: The adjusted straight line equation to such population of points and its corresponding perpendicular straight line equation. This new method contrasts with the classical method of linear regression in its accuracy and its independence of dispersion level of data. While by classical method, the solutions are always ambivalent, approached and dependent with dispersion level of data, the new method gives another solution to the adjusted straight line, which is identified as the adjusted perpendicular straight line equation (the transversal equation) to the same population of points.

---

### 1. Introducción

El método presentado aquí se ha desarrollado aplicando el *equilibrio de torques* o *momentos de fuerza* [1]. El método genera dos líneas rectas perpendiculares entre sí cuyo punto de intersección es el propio "centro de masa" de la población de los datos punto, una de las cuales es el eje longitudinal de la población de puntos, y la otra corresponde al eje transversal de los mismos datos. Específicamente, en el caso del resultado referente al eje longitudinal, éste es siempre encontrado en la región más estrecha limitada por las dos ecuaciones del método clásico de regresión lineal [2], donde se supone se encuentra la solución ideal para el ajuste de rectas.

### 2. Planteamiento

Sea  $D_i = (X_i, Y_i)$ , (ver en la Fig.1) donde  $i = 1, 2, \dots, n$ , un punto determinado de un conjunto de datos, constituido de  $n$  puntos cuyos pares de coordenados son

$$(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \dots, (X_n, Y_n). \quad (1)$$

Se busca la línea recta ajustada a los datos cuya ecuación es la siguiente

$$y = a + bx, \quad (2)$$

donde se desea encontrar la pendiente  $b$  y el intercepto  $a$ . También se define la línea recta perpendicular a la línea recta buscada, ec. (2),

que cruza al punto dato  $D_i$  dado y cuya ecuación es la siguiente

$$y = a^i - \frac{1}{b}x \quad (3)$$

donde  $-1/b$  y  $a^i$  son la pendiente y el intercepto respectivamente, el valor del intercepto depende a su vez del punto dato dado  $D_i$  del conjunto de puntos (1).

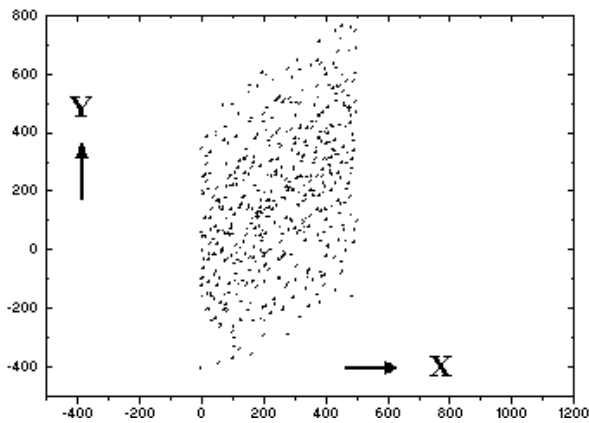


Fig.1. Se muestra una determinada distribución de puntos muy dispersa.

Se define ahora el punto  $P_i = (X^i, Y^i)$  como el punto intersección de las líneas rectas definidas por la ec. (2) y la ec. (3). Éste es el punto más cercano al punto dato dado  $D_i$ , también, se debe considerar a  $M = (X_m, Y_m)$  como el punto que determina el "centro de masa" de todo el conjunto de datos punto definido en (1), ver Fig.2.

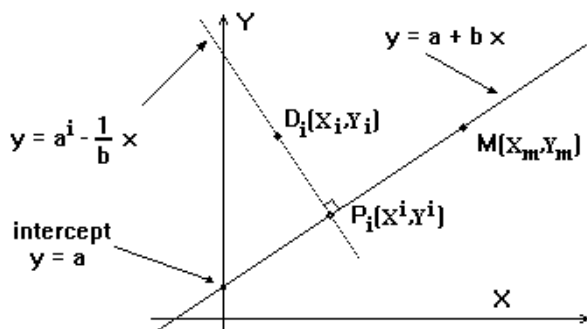


Fig.2. El esquema muestra el centro de masa M del conjunto, un punto dato dado  $D_i$ , la línea recta ajustada buscada  $y = m x + b$ , y el punto  $P_i$  correspondiente al punto de intersección de la línea recta ajustada y la línea recta perpendicular que cruza al punto dato  $D_i$  del conjunto.

Los interceptos  $a$  y  $a^i$  pueden expresarse como funciones de la pendiente  $b$  con ayuda de las coordenadas de los puntos  $M$  y  $D_i$  utilizando respectivamente la ec. (2) y la ec. (3), así

$$Y_m = a + bX_m$$

entonces  $a = Y_m - bX_m \quad (4)$

y

$$Y_i = a^i - \frac{1}{b}X_i$$

entonces  $a^i = \frac{bY_i + X_i}{b} \quad (5)$

Ya que la ordenada del punto intersección  $P_i$  es la misma para la ec.(2) y la ec.(3), entonces

$$bX^i + a = -\frac{1}{b}X^i + a^i,$$

entonces despejando la abscisa  $X^i$  y reemplazando los interceptos con ayuda de la ec. (4) y la ec. (5) se obtiene

$$X^i = \frac{b(Y_i - Y_m) + X_i + b^2 X_m}{b^2 + 1} \quad (6)$$

y para la ordenada correspondiente

$$Y^i = \frac{b^2 Y_i + Y_m + b(X_i - X_m)}{b^2 + 1} \quad (7)$$

entonces las diferencias entre las coordenadas del punto  $D_i$  y del punto  $P_i$ , puntos que se definieron anteriormente, dependerán sólo de  $b$ , y cuyas ecuaciones son las siguientes

$$X_i - X^i = \frac{b^2(X_i - X_m) - b(Y_i - Y_m)}{b^2 + 1}$$

y

$$Y_i - Y^i = \frac{-b(X_i - X_m) + (Y_i - Y_m)}{b^2 + 1} \quad (8)$$

también las ecuaciones para las diferencias entre las coordenadas del punto  $P_i$  y del punto  $M$  dependerán sólo de  $b$ , entonces

$$x^i - x_m = \frac{b(Y_i - Y_m) + (X_i - X_m)}{b^2 + 1}$$

y

$$Y^i - Y_m = \frac{b^2(Y_i - Y_m) + b(X_i - X_m)}{b^2 + 1}$$

De esta forma es posible expresar las distancias correspondientes  $|\overline{MP}_i|$  y  $|\overline{P}_iD_i|$  y como funciones de la pendiente **b**. Estas son:

$$|\overline{MP}_i| = \sqrt{(X^i - X_m)^2 + (Y^i - Y_m)^2} = \frac{(Y_i - Y_m)b + (X_i - X_m)}{\sqrt{b^2 + 1}}$$

y

$$|\overline{P}_iD_i| = \sqrt{(X_i - X^i)^2 + (Y_i - Y^i)^2} = \pm \frac{(X_i - X_m)b - (Y_i - Y_m)}{\sqrt{b^2 + 1}}$$

observe que la ec. (11) usa tanto el signo (+) como el signo (-) pero el resultado de la ec. (11) como se refiere a distancia debe ser positivo.

### 3. Aplicación del concepto de torque

Con la ec. (10) y la ec. (11) es posible definir el torque [1],  $\mathbf{T}_i$ , del punto  $\mathbf{D}_i$  con respecto al punto  $\mathbf{M}$  (Fig.2). Entonces

$$\mathbf{T}_i = |\overline{P}_iD_i| \cdot |\overline{MP}_i|$$

Donde  $|\overline{MP}_i|$  y  $|\overline{P}_iD_i|$  se identifican respectivamente como el brazo del torque y el segmento distancia que fungiría como fuerza del torque. Introduciendo la ec. (10) y la ec. (11) en la ec. (12) se obtiene

$$T_i = \frac{xy b^2 + (x^2 - y^2)b - xy}{b^2 + 1}$$

donde  $x = X_i - X_m$

y  $y = Y_i - Y_m$

De esta manera es posible hacer la sumatoria de los correspondientes torques a cada punto del conjunto de los datos punto. Si consideramos que los puntos del conjunto de datos (1) se

distribuirían alrededor de la línea recta ajustada de modo que habría un equilibrio de torques, entonces de acuerdo al principio de equilibrio de torques la sumatoria debe ser igual a cero. Así

$$F(b) = \sum T_i(b) = 0$$

$$F(b) = \sum \frac{xy b^2 + (x^2 - y^2)b - xy}{b^2 + 1} = 0$$

el denominador de la sumatoria es una constante, entonces es posible extraerlo de la sumatoria y la ecuación se reduce a

$$b^2 \cdot \sum xy + b \sum (x^2 - y^2) - \sum xy = 0$$

asignando un símbolo a cada sumatoria

$$S_a = \sum xy \quad \text{y} \quad S_b = \sum (x^2 - y^2)$$

es posible reescribir la ec. (17) como

$$S_a b^2 + S_b b - S_a = 0$$

Aquí se debe recordar que las variables  $x$  y  $y$  que aparecen en cada elemento de las sumatorias se definen por medio de la ec. (14). Entonces, a partir de la ec. (19) se obtiene

$$b = \frac{-S_b \pm \sqrt{(S_b)^2 + 4(S_a)^2}}{2S_a}$$

la cual es posible reescribir como

$$b = -S \pm \sqrt{S^2 + 1}$$

donde

$$S = \frac{S_b}{2S_a}$$

De este modo se obtienen dos soluciones a partir de la ec. (20): Una es la pendiente de la línea recta ajustada al eje longitudinal y la otra al eje transversal. El valor del intercepto correspondiente estará determinado por la ec. (4) o la ec. (5).

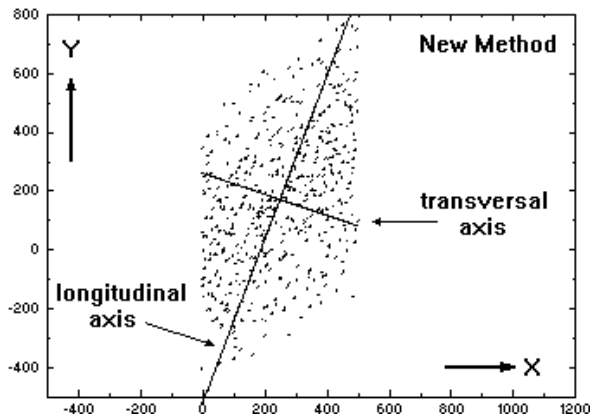


Fig.3. Se muestra la línea recta resultante al aplicar el nuevo método, identificada por HP (siglas del término Hiper Ajuste). Puede apreciarse tanto el eje longitudinal como el eje transversal correspondientes al conjunto de puntos.

#### 4. Comparación con el método de regresión lineal

A fin de efectuar comparaciones se consideró un conjunto de datos punto muy disperso y aleatoriamente distribuido. En la Fig.4 se puede ver el resultado obtenido por el método clásico de regresión lineal [1].

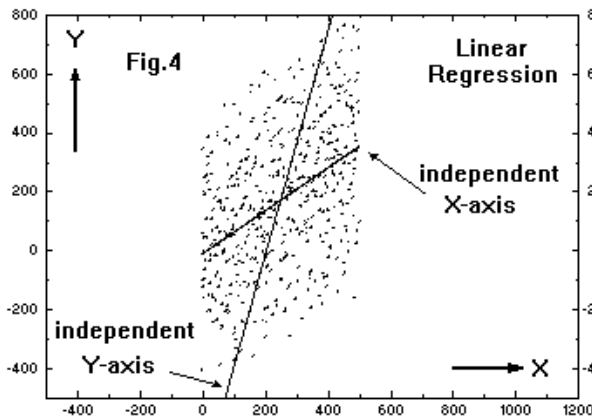


Fig.4. Los resultados aplicando regresión lineal.

La línea de mayor pendiente mostrada en la Fig.4 se genera cuando se elige a la ordenada como variable independiente. De otro modo, cuando se elige la abscisa como variable independiente el resultado es la línea recta con menor pendiente que se muestra en la Fig.4. En la Fig.5 se muestran los resultados de ambos métodos.

#### 5. Conclusiones

El algoritmo presentado en este reporte puede usarse para encontrar los ejes longitudinal y transversal de una distribución cualquiera de datos punto sin ambigüedad. En una futura publicación, se presentará la generalización del método presentado aquí, por el cual es posible obtener la línea recta ajustada de una distribución de datos punto de un sistema n dimensional.

#### Referencias

Bevington Philip R. and Robinson D.Keith, Data Reduction and Error Analysis for the Physical Sciences, 2nd ed.1992, McGraw-Hill, Inc.

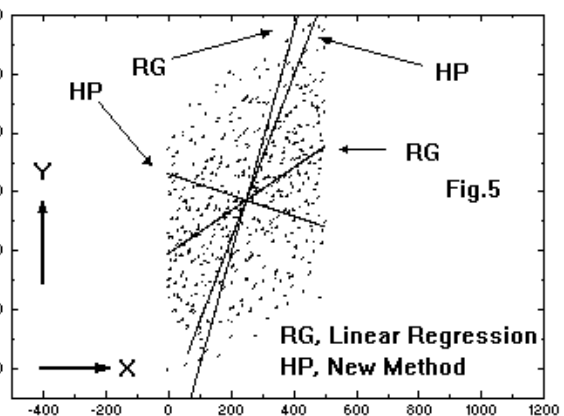


Fig.5. Se confrontan los resultados del método de regresión lineal (RG) y el nuevo método (HP).