Sea r el residuo o distancia entre un punto observado y el correspondiente punto sobre la recta entonces para un iésimo punto ^y_i sobre una recta es válido que

$$r_i = y_i - \hat{y}_i$$

Siendo este ^y_i

$$\hat{y}_i = c_1 x_i + c_o$$

Reemplazando en la primera función se obtiene

$$r_i = y_i - (c_1 x_i + c_o)$$

Para varios puntos de una misma recta entonces el residuo es la sumatoria de todos los residuos, es decir

$$SSR = \sum_{i=1}^{N} r_i^2 = \sum_{i=1}^{N} [y_i - (c_1 x_i + c_o)]^2$$

Esta al cuadrado pues el método que vamos a usar es el de mínimos cuadrados.

Si derivamos la anterior expresión respecto a la constante c0 obtenemos

$$\frac{d\,SSR}{d\,c_o} = \sum_{i=1}^{N} (-2)[y_i - (c_1 x_i + c_o)]$$

Y si lo hacemos respecto a c1

$$\frac{d\,SSR}{d\,c_1} = \sum_{i=1}^{N} (-2x_i)[y_i - (c_1 x_i + c_o)]$$

Como derivamos para optimizar, tenemos que igualar a cero y encontrar los mínimos pues en esto consiste el método de mínimos cuadrados.

$$\sum_{i=1}^{N} (-2)[y_i - (c_1 x_i + c_o)] = 0$$

$$\sum_{i=1}^{N} (-2x_i)[y_i - (c_1 x_i + c_o)] = 0$$

Haciendo a un lado el -2 (pasándolo a dividir al otro lado) y aplicando álgebra obtenemos dos ecuaciones

$$\sum_{i=1}^{N} c_o + \sum_{i=1}^{N} c_1 x_i = \sum_{i=1}^{N} y_i \quad \text{(I)}$$

$$\sum_{i=1}^{N} c_o x_i + \sum_{i=1}^{N} c_1 x_i^2 = \sum_{i=1}^{N} x_i y_i \quad \text{(II)}$$

Aplicando la propiedad que dice que el sumatorio de una constante es igual a la constante por el número de sumandos de las sumatorias en la primera ecuación y la propiedad distributiva de las sumatorias en la ecuación 2 tenemos lo siguiente:

$$c_o N + c_1 \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i \quad \text{(I)}$$

$$c_o \sum_{i=1}^{N} x_i + c_1 \sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} x_i y_i \quad \text{(II)}$$

Dividiendo por N en ambas ecuaciones obtenemos

$$c_o + c_1 \left(\frac{1}{N}\right) \sum_{i=1}^{N} x_i = \left(\frac{1}{N}\right) \sum_{i=1}^{N} y_i \quad \text{(I)}$$

$$c_o \left(\frac{1}{N}\right) \sum_{i=1}^{N} x_i + c_1 \left(\frac{1}{N}\right) \sum_{i=1}^{N} x_i^2 = \left(\frac{1}{N}\right) \sum_{i=1}^{N} x_i y_i \quad \text{(II)}$$

Nótese que para el caso de las sumatorias esto es básicamente un promedio.

$$\bar{x} = \left(\frac{1}{N}\right) \sum_{i=1}^{N} x_i \quad \text{y} \quad \bar{y} = \left(\frac{1}{N}\right) \sum_{i=1}^{N} y_i$$

Reemplazando lo anterior en la ecuación 1 y 2 a la vez que obtenemos el siguiente sistema de ecuaciones lineales de 2x2:

$$\begin{bmatrix} 1 & \bar{x} \\ \bar{x} & (\frac{1}{N})\sum_{i=1}^{N} x_i^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} \bar{y} \\ (\frac{1}{N})\sum_{i=1}^{N} x_i y_i \end{bmatrix}$$

Si resolvemos este sistema (despejamos c0 y c1) y aplicamos la propiedad asociativa de las sumatorias a la vez que reemplazamos c0 por a0 y c1 por a1 demostramos lo que se quería

$$a_0 = \bar{y} - a_1 x \quad \text{y} \quad a_1 = \frac{\sum_{i=1}^{N} \left[(x_i - \bar{x})(y_i - \bar{y}) \right]}{\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

NOTA: para la primera función es a0=promedio de y-a1* promedio de x

Donde a0 representa el corte con el eje “Y” y a1 representa la pendiente.