Punto 2

Se tiene un conjunto de puntos sobre un plano 2D y se quiere encontrar la recta que mejor modele esos números. Para hacerlo, se debe encontrar la recta que tenga la mínima distancia entre los puntos y ella misma. El residuo r del iesimo punto se puede calcular de la siguiente manera

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

donde y_i el punto iesimo del conjunto de puntos y $\hat{y}_i$ es el iesimo punto de la recta que modela el conjunto de puntos. Reemplazando la ecuación de una recta de la forma $\hat{y}(x) = a_1 x + a_0$ se obtiene

$$r_i = y_i - (a_1 x + a_0)$$

como hasta ahora tenemos la ecuación para el iésimo punto, debemos encontrar la ecuación que contenga todos los residuos acumulados. Para hacer esto toca sumar todos los r_i

$$\sum_1^N r_i = \sum_1^N y_i - (a_1 x + a_0)$$

Sin embargo, si se suman los residuos de esta manera se tendrán inconsistencias, pues un punto que se aleje x unidades y otro que se aleje $-x$ unidades, darán un residuo de 0. Para evitar esto y para minimizar los puntos que se acerquen a la recta y maximizar los puntos que se alejen de la misma, la distancia se eleva al cuadrado

$$SSR = \sum_1^N r_i^2 = \sum_1^N [y_i - (a_1 x + a_0)]^2$$

Esta ecuación se debe minimizar para asegurar la recta que mejor modele los puntos. Para hacer esto, se debe derivar, igualar a 0 y despejar para las dos constantes a_0 y a_1

$$\frac{SSR}{d a_0} = \sum_1^N (-2)[y_i - (a_1 x + a_0)] = 0$$

$$\frac{SSR}{d a_1} = \sum_1^N (-2 x_i)[y_i - (a_1 x + a_0)] = 0$$

despejando se optienen las siguientes dos ecuaciones

$$a_0 + a_1 \bar{x} = \bar{y} \quad \text{y}$$

$$a_0 \bar{x} + a_1 \left(\frac{1}{N}\right) \sum_{i=1}^N x_i^2 = \left(\frac{1}{N}\right) \sum_{i=1}^N x_i y_i$$

para resolver el sistema de ecuaciones se plantea una matriz de la siguiente forma

$$\begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \left(\frac{1}{N}\right)\sum_{i=1}^{N} x_i^2 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}=\begin{pmatrix} \bar{y} \\ \left(\frac{1}{N}\right)\sum_{i=1}^{N} x_i y_i \end{pmatrix}$$

aplicando gauss queda

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}=\begin{pmatrix} \bar{y} \\ \dfrac{\sum_{i=1}^{N} [(x_i-\bar{x})-(y_i-\bar{y})]^2}{\sum_{i=1}^{N} (x_i-\bar{x})^2} \end{pmatrix}$$

despejando tanto para a_0 como para a_1 , queda

$$a_0=\bar{y}-a_1 x$$

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

entonces queda demostrado que para un conjunto de puntos (x,y), las anteriores ecuaciones modelan la linea optima $\hat{y}(x)=a_1 x+a_0$