

$$y = mx + b$$

$m$  = slope

$b$  = intercept

Given matrices:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ and } y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

The equation in matrix form:

$$y = mx + b$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \cdot \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} b + mx_1 \\ b + mx_2 \\ b + mx_3 \end{bmatrix}$$

$$XB = Y$$

$$(X^T X)^{-1} X^T XB = (X^T X)^{-1} X^T Y$$

$$B = (X^T X)^{-1} X^T Y$$

Where:

$X^T$  represents the transpose of matrix  $X$ .

$X^{-1}$  represents the inverse of matrix  $X$ .

$$BIC_k = n \ln(SSE) + k \ln(n)$$

Where,

$k$  = degree

$SSE$  = sum of squared errors

$n$  = number of data points

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Where,

$n$  = number of data points

$Y_i$  = real values

$\hat{Y}_i$  = predicted values

You can rewrite MSE as

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - (m x_i + b))^2$$

$$\text{let, } f(m, b) = (y_i - (m x_i + b))^2$$

Partial derivatives,

$$\text{let, } g(m, b) = y_i - (m x_i + b)$$

$$\text{so, } f(m, b) = g(m, b)^2$$

$$\frac{\partial f}{\partial m} = 2 g(m, b) g'$$

$$\begin{aligned} \text{now, } g' &= \frac{\partial g}{\partial m} = 0 - (1 x_i + 0) \\ &= -x_i \end{aligned}$$

expanding,

$$\begin{aligned} \frac{\partial f}{\partial m} &= 2 (y_i - (m x_i + b))^2 \cdot (-x_i) \\ &= -2 x_i \cdot (y_i - (m x_i + b)) \end{aligned}$$

now,

$$\frac{\partial f}{\partial b} = 2 g(m, b) g'$$

$$\text{now, } g' = \frac{\partial f}{\partial b} = 0 - (0 \cdot 0 + 1) = -1$$

expanding,

$$\begin{aligned} \frac{\partial f}{\partial b} &= 2 \cdot (y_i - (m x_i + b))^2 \cdot -1 \\ &= -2 (y_i - (m x_i + b)) \end{aligned}$$

Final answer,

$$\begin{aligned} \frac{\partial}{\partial m} &= \frac{2}{n} \sum_{i=1}^n -x_i (y_i - (m x_i + b)) \\ \frac{\partial}{\partial b} &= \frac{2}{n} \sum_{i=1}^n -(y_i - (m x_i + b)) \end{aligned}$$