RIDGE REGRESSION

For ridge regression

We know → $L = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda m^2$

Therefore to find the value of slope & intercept i.e m and b.

$$L = \sum_{i=1}^{n} (y_i - mx_i - b)^2 + \lambda m^2$$

Differentiating w.r.t b

$$\frac{dL}{db} = \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - mx_i - b)^2 = 0$$

$$\Rightarrow \sum \frac{\partial}{\partial b} (y_i - mx_i - b)^2 = 0$$

$$\Rightarrow \sum (-1) 2 (y_i - mx_i - b) = 0$$

$$\Rightarrow \sum (-2)(y_i - mx_i - b) = 0$$

Dividing by n b/s

$$\Rightarrow \frac{\sum y_i}{n} - \frac{\sum mx_i}{n} - \frac{\sum b}{n} = 0$$

We know, $\frac{\sum y_i}{n} = \bar{y}$

$$\Rightarrow \bar{y} - m\bar{x} - \frac{nb}{n} = 0$$

$$\Rightarrow y - m\bar{x} - b = 0$$

$$\Rightarrow \boxed{b = y - m\bar{x}}$$

$$L = \sum_{i=1}^{n} (y_i - mx_i - \bar{y} - m\bar{x})^2 + \lambda m^2$$

Putting the value of b in loss function

Now differentiating w.r.t m

$$\frac{\partial L}{\partial m} = 2\sum_{i=1}^{n} (y_i - mx_i - \bar{y} + m\bar{x})(-x_i + \bar{x}) + 2\lambda m = 0$$

$$\Rightarrow -2\sum_{i=1}^{n} (y_i - \bar{y} - mx_i + m\bar{x})(x_i - \bar{x}) + 2\lambda m = 0$$

$$\Rightarrow \lambda m - \sum_{i=1}^{n} (y_i - \bar{y} - m(x_i - \bar{x}))(x_i - \bar{x}) = 0$$

$$\Rightarrow \lambda m - \sum_{i=1}^{n} ((y_i - \bar{y})(x_i - \bar{x}) - m(x_i - \bar{x})^2) = 0$$

$$\Rightarrow \lambda m - \sum_{i=1}^{n} ((y_i - \bar{y})(x_i - \bar{x})) + m\sum_{i=1}^{n} (x_i - x)^2 = 0$$

$$\Rightarrow \lambda m + m\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})$$

$$\boxed{\Rightarrow m = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2 + \lambda}}$$

* If the value of λ is set to 0 then ridge regression performs similar to linear regression.
* Greater the value of lamda lesser will be the value of our slope.