Date......................

Maths behind Lasso Regression

As, we know, → In Ridge as well as in Lasso also in normal Linear Regression the Formula to Calulate the "b" is same...

$$\boxed{b = \bar{Y} - m\bar{X}}$$

Calulating "m" for Lasso Regression...

$$\boxed{\text{Loss Function} \Rightarrow \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 + \lambda|m|} \Rightarrow \sum_{i=1}^{n}(Y_i - mX_i - b)^2 + \lambda|m|$$

Derivating in terms of "m"...

$$\Rightarrow \sum_{i=1}^{n}(Y_i - mX_i - \bar{Y} - m\bar{X})^2 + \lambda m \Rightarrow \frac{dL}{dm}$$

Assuming that the $m > 0$

$$\frac{dL}{dm} \Rightarrow \sum_{i=1}^{n}(Y_i - mX_i - \bar{Y} - m\bar{X})(-X_i - \bar{X}) + 2\lambda m = 0$$

Multiplying '2' with 'λ'.

$$\frac{dL}{dm} \Rightarrow 2\sum_{i=1}^{n}\left[(Y_i - \bar{Y}) - m(X_i - \bar{X})\right](X_i - \bar{X}) + \lambda = 0$$

$$\Rightarrow -2\sum_{i=1}^{n}\left[(Y_i - \bar{Y})\cdot(X_i - \bar{X}) - m(X_i - \bar{X})^2\right] + \lambda = 0$$

Dividing with '2' the complete equation...

$$\Rightarrow -\sum_{i=1}^{n}\left(Y_i - \bar{Y})\cdot(X_i - \bar{X}) + m(X_i - \bar{X})^2 + \lambda = 0\right.$$

$$\Rightarrow m(X_i - \bar{X})^2 + \lambda = \sum_{i=1}^{n}(Y_i - \bar{Y})\cdot(X_i - \bar{X})$$

$$\Rightarrow \boxed{m = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})\cdot(X_i - \bar{X}) - \lambda}{(X_i - \bar{X})^2}}$$

Date..........

As per the Lasso Regression is concerned, "m" can fall into 3 cases:-

When $\boxed{m > 0}$ Case-1

$$m = \boxed{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X}) - \lambda}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}$$

When $m = 0$ Case-2

$$m = \boxed{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}$$

Case-3

When $\boxed{m < 0}$

$$m = \boxed{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X}) + \lambda}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}$$

Q Why there is Sparsity in the coefficient values in Lasso Regression??

Ans Meaning of Sparsity ⇒ Values of the coef_ is equal to Zero '0'.

⊙ Let's proof this with the help of example:-
As, per the Formula where $\boxed{m > 0}$

Let's say $\boxed{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X}) = a}$, $\boxed{\sum_{i=1}^{n}(X_i - \bar{X})^2 = b}$

$a = 100,\ b = 50$

$m = \dfrac{100 - \lambda}{50}$

Case 1:- 'λ' is equal to '0'.
Case 2:- 'λ' is equal to '50'.
Case 3:- 'λ' is equal to '100'.
Case 4:- 'λ' is equal to '150'.

$\Rightarrow m = \dfrac{100}{50} \Rightarrow \boxed{2}$ $\boxed{\text{Case-1}}$

$\Rightarrow m = \dfrac{100 - 50}{50} \Rightarrow \boxed{1}$ $\boxed{\text{Case-2}}$ [As 'λ' increases, 'm' decreases...]

$\Rightarrow m = \dfrac{100 - 100}{50} \Rightarrow \boxed{0}$ $\boxed{\text{Case-3}}$ [This also leads to Feature Selection]

If the Value of 'λ' is increased to a level that, it matches the value of "a". Then it leads to Sparsity in the values of coef_.

$\Rightarrow m = \dfrac{100 - 150}{50} \Rightarrow \boxed{-1}$

∴ This is not acceptable because, here only positive values of "m" is included.

Spiral

Date.....................

• So, if the value of 'λ' increases too high and exceeds the value of 'a' then automatically Case-3 is called..

$$m = \boxed{\frac{\sum_{i=1}^{n} (Y_i - \bar{Y})(X_i - \bar{X}) + \lambda}{\sum_{i=1}^{n} (X_i - \bar{X})^2}}$$

Case 4:- 'λ' is equal to '150'.

$$m \Rightarrow \frac{100 + 150}{50} \Rightarrow \boxed{5}$$

"Therefore, we can conclude that, as the value of "λ" increases, the value of "m" decreases..."

"and the "m" value decreases to Zero '0', it can never be negative".