then previous one.

Ridge & Lasso Regression

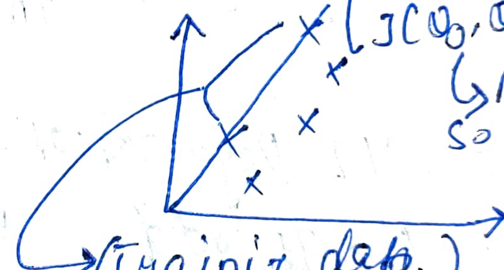
Cost function $= \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2$

↓ $J(\theta_0, \theta_1)$

↓ give gradient descent

$\left\{ \begin{array}{l} \theta_0 = 0. \\ J(\theta_0, \theta_1) = 0 \end{array} \right\}$

↳ No diff so 0.



(training data)

Overfitting: (My model performs well with training data) but [fails to perform well with test data]

↳ Condition - Low Bias

↳ High variance.

Bias - Talk about training data.
Variance - Test data.

Underfitting

① Model accuracy is bad with training data.
② Model accuracy is also bad with test data.
{High Bias & High Variance}

Model-1	Model-2	Model-3
Training Acc = 90%	Training Acc = 92%	Training Acc = 70%
Test Acc = 80%	Test Acc = 91%	Test Acc = 65%
⇓	↓	
Overfitting	Generalised Model	Underfitting
{Low Bias, High variance}	{Low Bias, Low Variance}	{High Bias, High Variance}

$$J(\theta_0,\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2$$
$$= (\hat{y}^{i} - y^{(i)})^2 = 0$$

Ridge (L2 Regularisation)
↳ It add $+(\text{slope})^2$
So, $0 + 1(2) = 2$ ↓↓
⇒ $0 + 1(2)^2 = 4$ ↓↓

we use it.

$\theta_j = 0$
$h_\theta(x) = \theta_0 + \theta_1 x$
$h_\theta(x) = \theta_1 x$
↳ slope.

To prevent overfitting.

$(\hat{y}^{(i)} - y^{(i)})^2 + \lambda(\text{Slope})^2$
⇓
Small value $+ \lambda(\text{Slope})$

Lasso (L1 Regularization) :-

$= (\hat{y}-y)^2 + \lambda \boxed{|slope|}$ → Feature selection.

$h_0(x) = \hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \dots + \theta_n x_n$

So, ~~$h_0(x) = \hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \dots \theta_n x_n$~~

$|\theta_0 + \theta_1 + \theta_2 + \theta_3 + \dots + \theta_n|$.

→ We are neglecting which are not important using modulus.

L1 → { ① Preventing overfitting.
L2 → ② Feature selection }

{ λ → Crossvalidation }

Performance metric good we used that.

Ridge Regression (L2 Norm) :-

Cost function $= (h_0(x)^{(i)} - y^{(i)})^2 + \lambda (slope)^2$

Purpose : Preventing overfitting.

Lasso Regression (L1 Reg) :-

Cost function $= (h_0(x^{(i)})^2 - y^{(i)})^2 + \lambda |slope|^2$

Purpose : ① Prevent overfitting.

② Feature selection.

Assuption of Linear Regression

① Normal / Gaussian Distribution → Model will get trained well.

② Standardisation. {Scaling your data} → Z-score
($\mu = 0$ & $\sigma = 1$)

③ Linearity

④ Multi Collanarity.

→ 95% correlated

X_1 X_2 Y

X_3

But highly
correlated with
But we drop
one features.

Variation Inflation factor :- Find how?

→ It is use to solve multi collinarity...