15/10/22

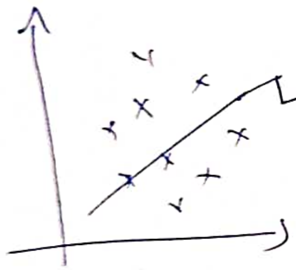
Ridge Regression,

lasso Regression,

Elasticnet Regression

Linear regression :- Find out the Best-fit line to reduce the cost fn.

θ_0 = intercept

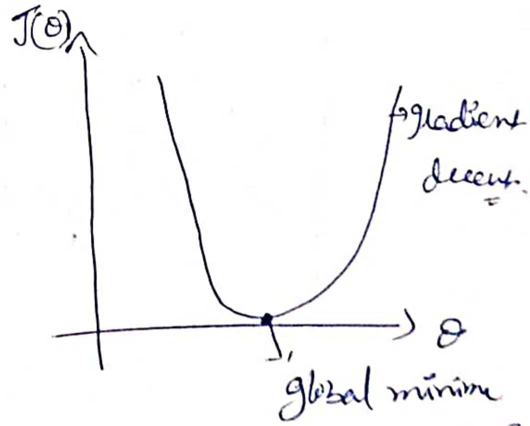


$h_\theta(x) = \theta_0 + \theta(x)$

$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots \theta_n x_n$

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

Mean Squared Error

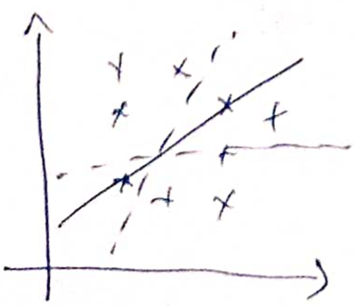


Ridge Regression :- also called as (L2 Regularization) training data → accuracy ↑ >95%

Test data → Accuracy ↓↓ → 60%

to reduce overfitting L2 Norm

overfitting



Cost function = 0

Training data → Low bias

Test data ⇒ Low / High variance

How do we do reduce the overfitting. $\boxed{\text{Hyper Parameter}}$

Cost fn: $\frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^i) - (y^i))^2 + \lambda \sum_{i=1} (\text{slope})^2$

$h_\theta(x) = \theta_0 + \theta_1 x$

$(\text{slope})^2 = \theta_1^2$

here slope $= \theta_1$

if $\lambda = 0$.

if, cost fn $= 0$,

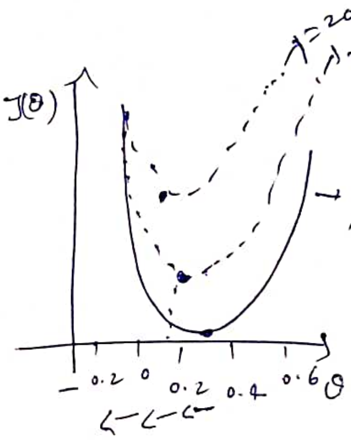
$0 + \lambda(\text{slope})^2$

$$\text{costfn} = \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^i) - y^i\right)^2 + 0 \ldots$$

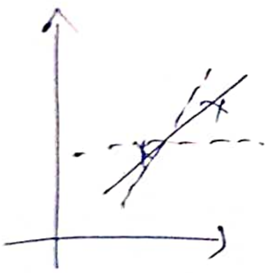
Costfunction of Linear regression = Ridge Regression.

Relation b/n λ and Slope :-

$\lambda \uparrow\uparrow \quad \theta \downarrow$



minimizing cost function



cost fn. $= 0 + [+ve] = [+ve] \downarrow\downarrow\downarrow$

λ is inversely proportional to slope.

→ will θ never become '0';

$h_\theta(x) = \theta_0 + \theta_1(x_1) + \theta_2 x_2 + \theta_3 x_3$

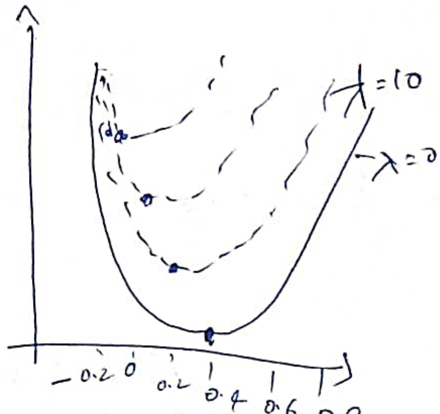
$= \theta_0 + 0.95\,x_1 + 0.82\,x_2 + 1.5\,x_3$

λ won't be -ve.

2. Lasso Regression :- L_1 Norm / L_2 Regularization → Reduce the features ⇓ feature Selection

Cost fn: $\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^i)-(y^i)\right)^2+\lambda\sum_{i=1}^{m}|\text{Slope}|$ magnitude:

relation b/n λ, θ :-



$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$

$= \theta_0 + 0.54x_1 + 0.23x_1 + \boxed{0.10\,x_3}$ $\lambda\uparrow$

$y\uparrow$ 1

$x_1 = 0.54,$

0 → feature removed.

If the feature is having outliers we can use Lasso Regression:

- ElasticNet [L1 and L2 Norm]

Ridge Lasso

Cost fn: $\frac{1}{m}\sum_{i=1}^{m}\left[h_\theta(x^i)+(y^i)\right]^2+\boxed{\lambda_1\sum_{i=1}^{m}(\text{Slope})^2}+\boxed{\lambda_2\sum_{i=1}^{m}|\text{slope}|}$

Practicals