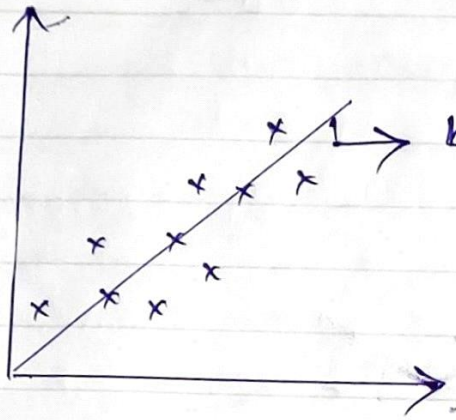
Notes

ML Class 4

Ridge Regression, Lasso Regression, Elasticnet Regression

Linear Regression



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

or

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

↓

Multiple Independent feature

Cost function : (MSE)

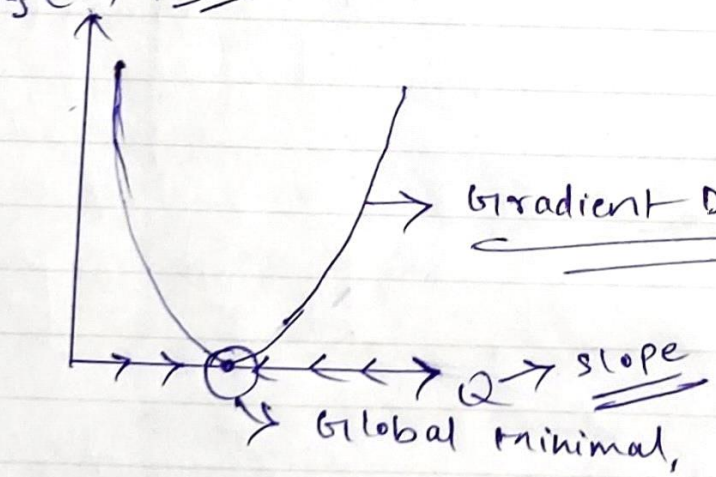
$h_\theta(x) = \theta_0 + \theta_1 x$

$$= \frac{1}{n} \sum_{i=1}^{n} \left(h_\theta(x)^{(i)} - y^{(i)} - y^{(i)} \right)^2$$

or

$$\frac{1}{n} \sum_{i=1}^{n} \left(y - \hat{y} \right)^2$$

Actual ↑ ↓ predicted



Expt. No..... Date.....

Page. No.....

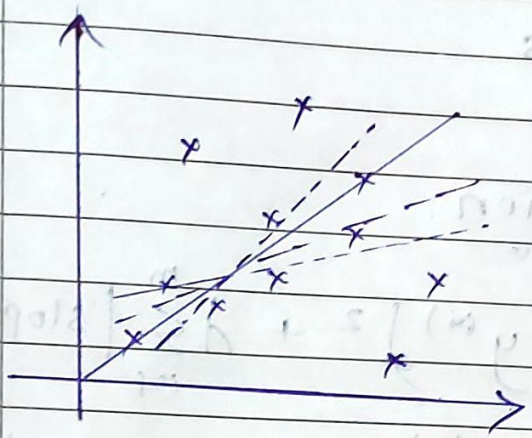
(i) Ridge Regression (L_2 Regularization) → Reduce Overfitting

It is used to Reduce Overfitting

Overfitting

Training Data → Accuracy ↑↑ 95% (Low Bias)

Test Data → Accuracy ↓↓ 60% (High variance)



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

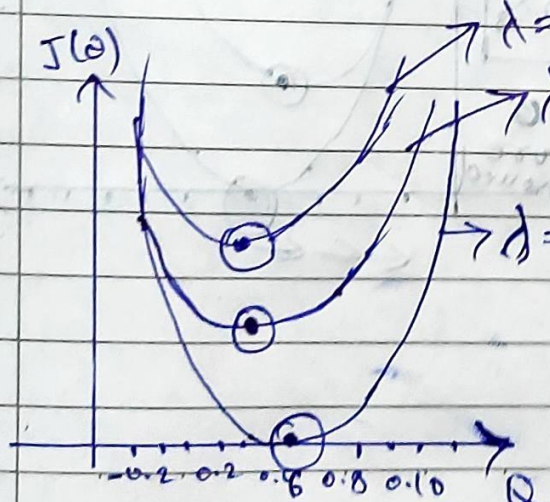
⇓

Ridge Regression

$$\text{cost fn} = \frac{1}{n}\sum_{i=1}^{n}\left(h\theta(x)^{(i)} - y^{(i)}\right)^2 + \lambda\sum_{i=1}^{n}(\text{slope})^2$$

λ → Hyperparameter (~~Hyperparameter~~)

θ → θ_1

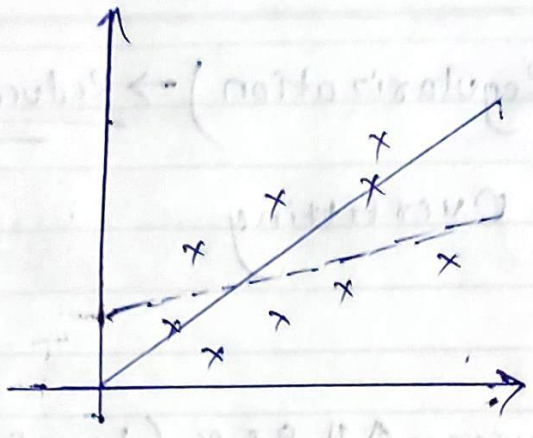


$$h\theta(x) = \theta_0 + \theta_1 x + \theta_2 x_2 + \theta_3 x_3$$
$$= \theta_0 + 0.95 x_1 + 0.82 x_2 + 1.5 x_3$$

* if $\theta_1 = 0$ then it will deleted

λ ↑	θ ↓

Teacher's Signature.....



Cost fn = 0 + λ (slope)²

= +ve ↓↓↓

(2) Lasso Regression [L1 Norm
L1 Regularization]

It Reduces the features

⇓

feature selection.

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

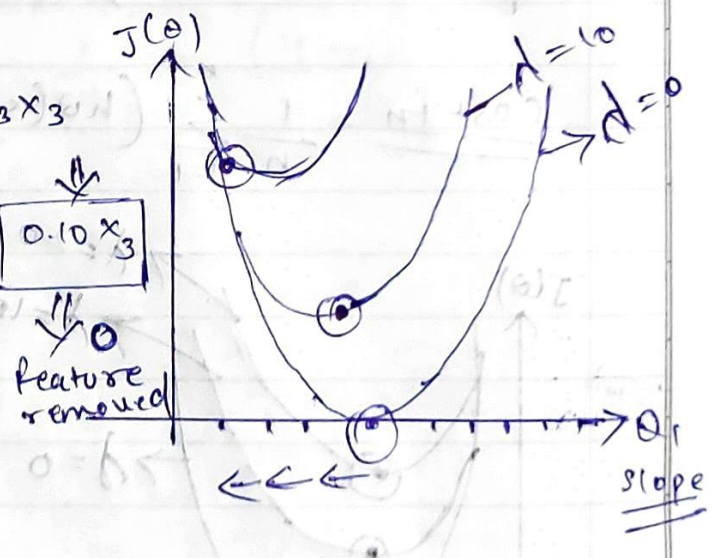
$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$

$$= \theta_0 + 0.54 x_1 + 0.23 x_2 + \boxed{0.10 x_3}$$

⇓

0

Feature removed



* ElasticNet Regression [L_1 and L_2 Norm]

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

$\lambda_1 \sum_{i=1}^{m}(\text{slope})^2 +$

$\lambda_2 \sum_{i=1}^{m}|\text{slope}|$

RMSE

MAE

Ridge

Lasso

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