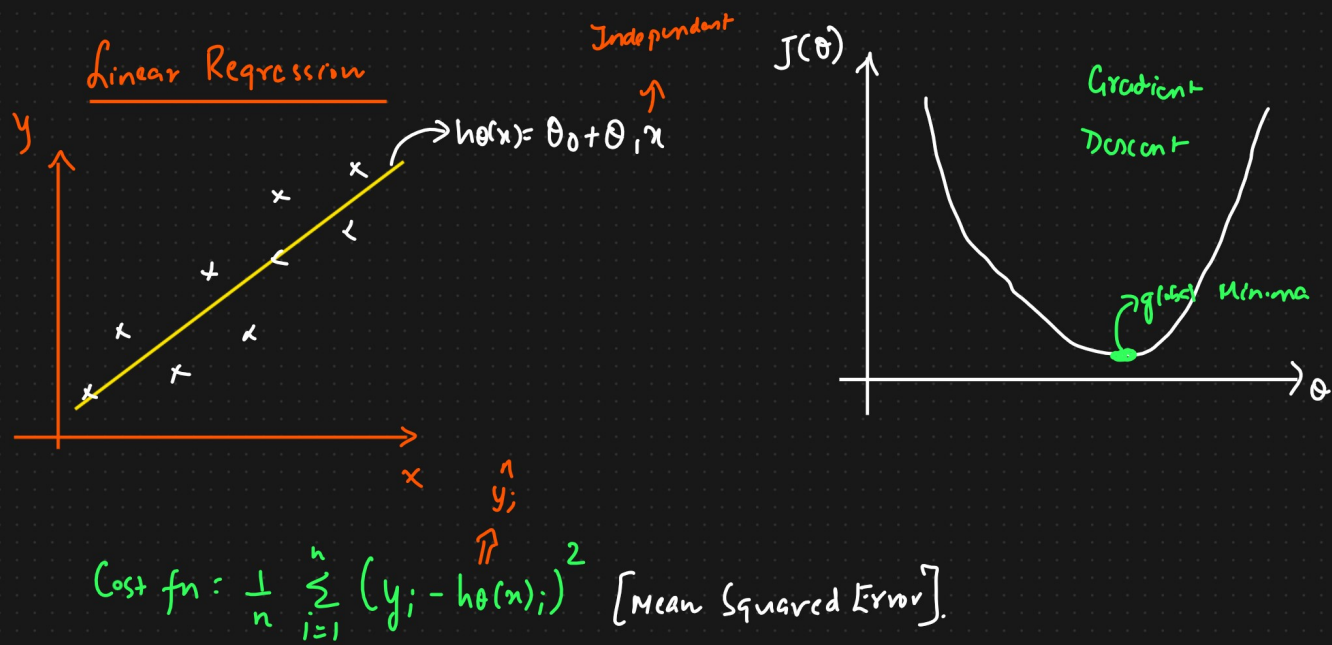
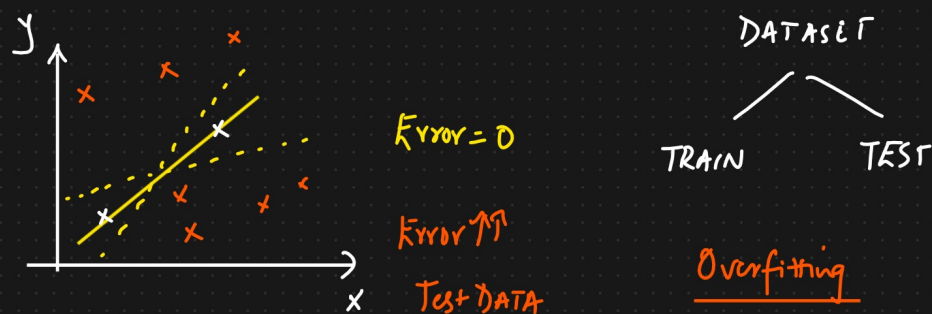


Ridge, Lasso And Elasticnet Regression

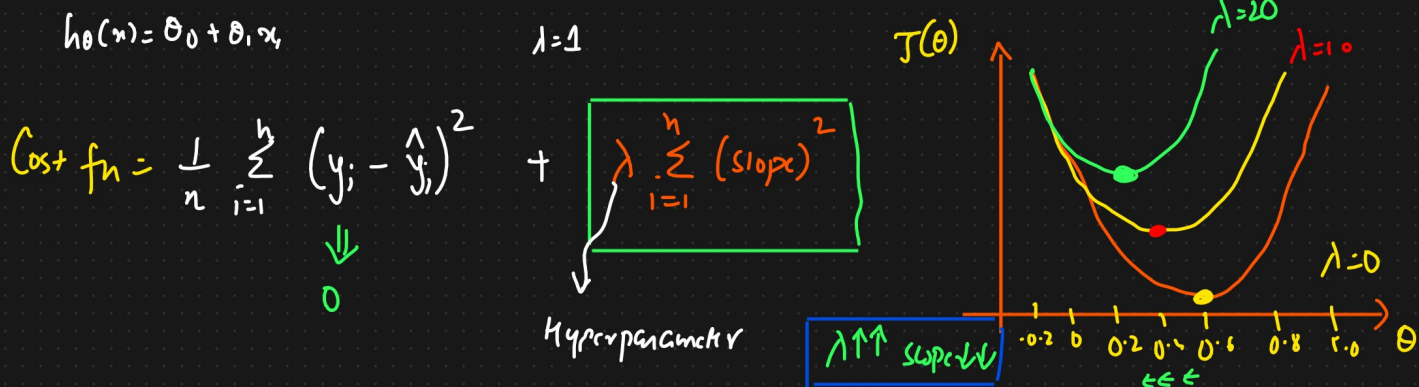


① Ridge Regression (L_2 Regularization) \rightarrow Reducing Overfitting

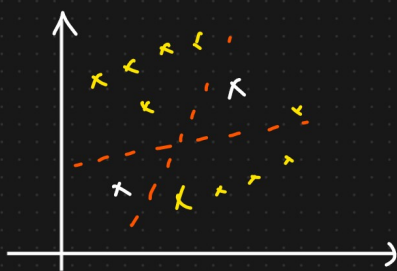


TRAIN \rightarrow Acc $\uparrow\uparrow \rightarrow R^2$

TEST \rightarrow Acc $\downarrow\downarrow \rightarrow R^2$



$= 0 + 1 [(\theta_1)^2] \Leftarrow$ penalize the cost function



$$h_0(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$

$$\text{Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda [(\theta_1)^2 + (\theta_2)^2 + (\theta_3)^2]$$

② Lasso Regression (ℓ_1 Regularization) \rightarrow Feature Selection

$$\text{Cost fn} : \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^n |\text{slope}|$$

$$h_0(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$

$$= 0.52 + 0.65 x_1 + 1.5 x_2 + 0.2 x_3$$

prediction

feature remove

feature selection

$x_2 \rightarrow 1 \text{ unit}$

$y \rightarrow 1.5 x_2$

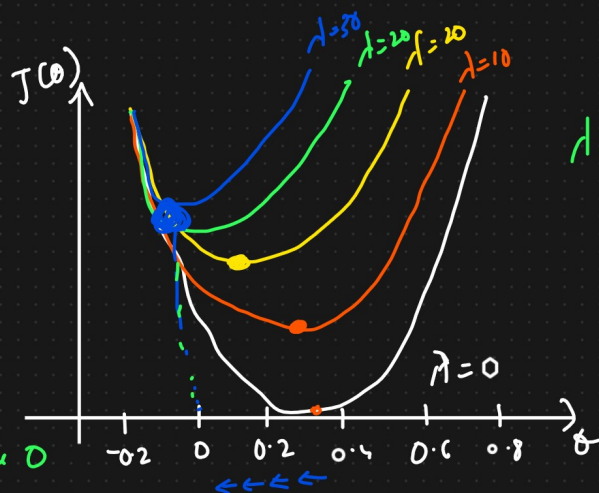
$x_3 \rightarrow 4 \text{ unit}$

$y \rightarrow 0.2 x_3$

$\lambda = 1$

$$\text{Cost fn} : \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^n |\text{slope}|$$

$$= \text{Error} + 1 [|\theta_1| + |\theta_2| + |\theta_3|]$$



$\lambda = \text{hyperparameter}$

θ & λ

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

θ will become 0

③ ElasticNet Regression

→ Reducing Overfitting → Ridge
→ Feature Selection → Lasso

$$\text{Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda_1 \sum_{i=1}^n (\text{slope})^2 + \lambda_2 \sum_{i=1}^n |\text{slope}|.$$

↓ ↓ ↓
MSE + Reducing Overfitting + Feature Selection

λ_1, λ_2 { Hyperparameter Tuning }.