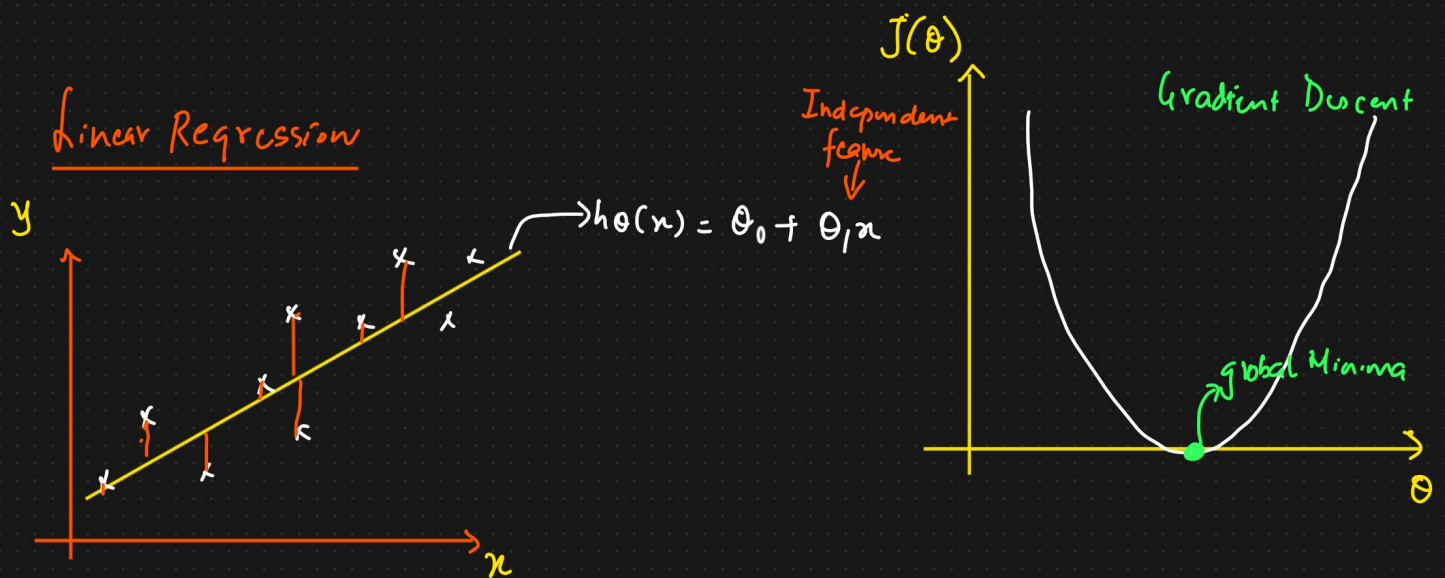
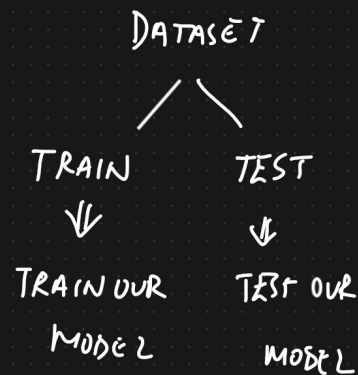
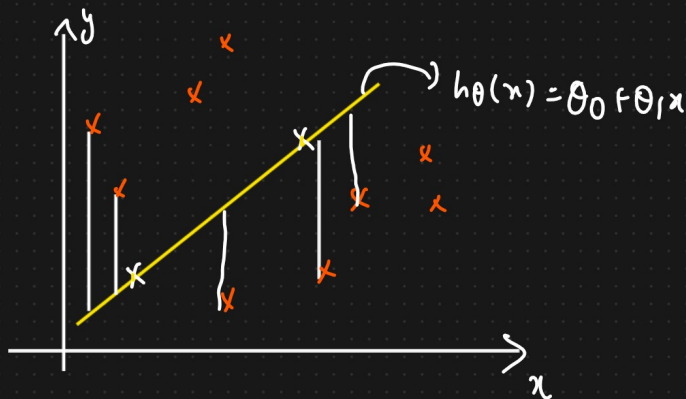


Ridge Regression, Lasso Regression, Elasticnet Regression



$$\text{Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - h_{\theta}(x_i))^2 \quad [\text{Mean Squared Error}]$$

① Ridge Regression (L2 Regularization) \rightarrow Reducing Overfitting.



$$\text{Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - h_{\theta}(x_i))^2$$

\downarrow
0.

Overfitting

Low Bias \downarrow

TRAIN DATASET \rightarrow ACC $\uparrow \uparrow$

TEST DATASET \rightarrow ACC $\downarrow \downarrow$

\downarrow
High Variance

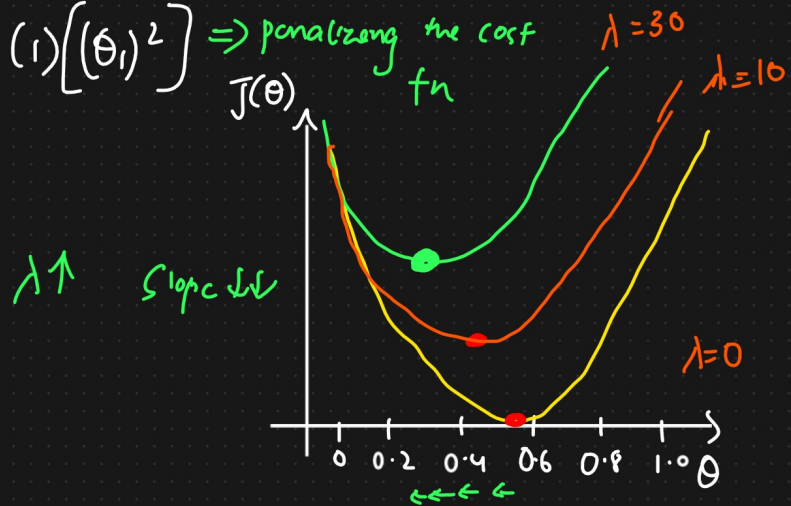
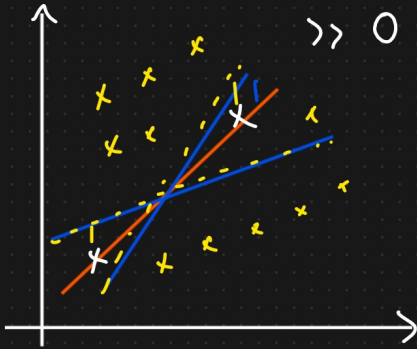
Ridge Regression

hyperparameter λ Let's $\lambda = 1$

$$J(\theta) \text{ Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - h_{\theta}(x)_i)^2 + \lambda \sum_{i=1}^n (\text{slope})^2$$

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

$$= \underbrace{0}_{\gg 0} + (1) [\theta_1^2] \Rightarrow \text{penalizing the cost fn}$$



Simple Linear Regression

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

L2 norm.



$$\text{Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - h_{\theta}(x)_i)^2 + \lambda [\theta_1^2]$$

Multiple Linear Regression

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

$$\text{Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - h_{\theta}(x)_i)^2 + \lambda ((\theta_1)^2 + (\theta_2)^2 + (\theta_3)^2)$$

② Lasso Regression (ℓ_1 Regularization) \rightarrow Feature Selection

$$J(\theta) = \text{Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - h_{\theta}(x_i))^2 + \lambda \sum_{i=1}^n |\text{slope}|$$

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

unit change in x_1
0.65 change in y

$$= 0.52 + 0.65 x_1 + 0.72 x_2 + 0.12 x_3$$

Lasso Regression

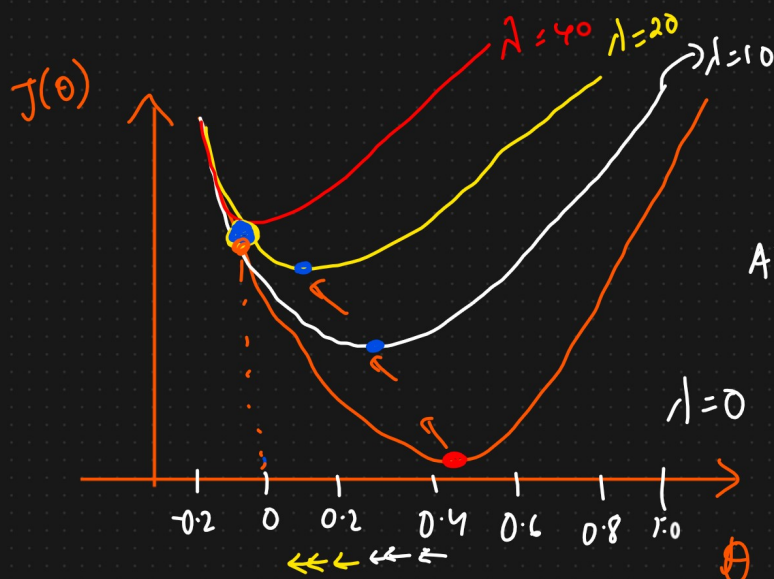
$\lambda = 1$

$$h_{\theta}(x) = 0.52 + 0.65 x_1 + 0.72 x_2 + 0.12 x_3$$

Lasso Regression.

$$\text{Cost fn} = \frac{1}{n} \sum_{i=1}^n (y_i - h_{\theta}(x_i))^2 + \lambda \sum_{i=1}^n |\text{slope}|$$

$$\text{Error} + \lambda [\theta_1 + \theta_2 + \theta_3]$$



$\lambda \uparrow \uparrow$ slope \downarrow

At one point $\boxed{\theta \approx 0}$

$$h_0(x) = 0.52 + 0.65x_1 + 0.72x_2 + \boxed{0.12x_3} \quad \{ \text{Feature Selection} \}$$

$$h_0(x) = 0.52 + 0.65x_1 + 0.72x_2$$

③ Elastic Net Regression

→ ① Reduce Overfitting } Ridge
→ ② Feature Selection } Lasso

$$\text{Cost fn} = \underbrace{\frac{1}{n} \sum_{i=1}^n (y_i - h_0(x)_i)^2}_{\text{MSE}} + \underbrace{\lambda_1 \sum_{i=1}^n (\text{slope})^2}_{\text{Reduce Overfitting}} + \underbrace{\lambda_2 \sum_{i=1}^n |\text{slope}|}_{\text{Feature Selection}}$$

λ_1, λ_2 {Hyperparameter Tunning}