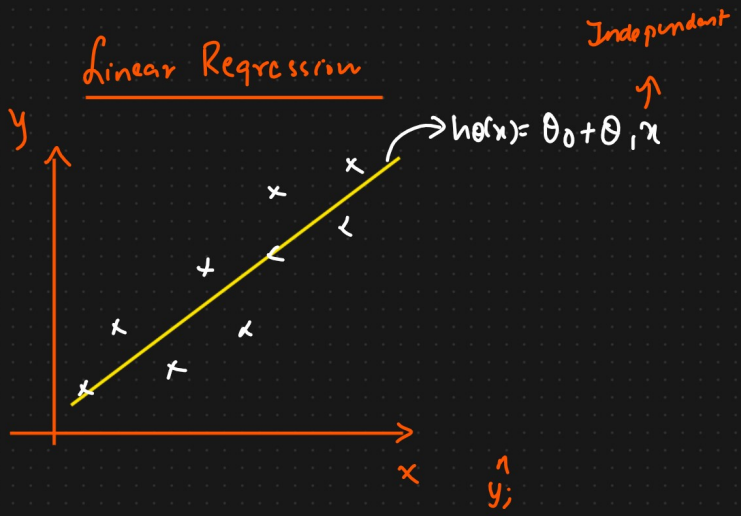
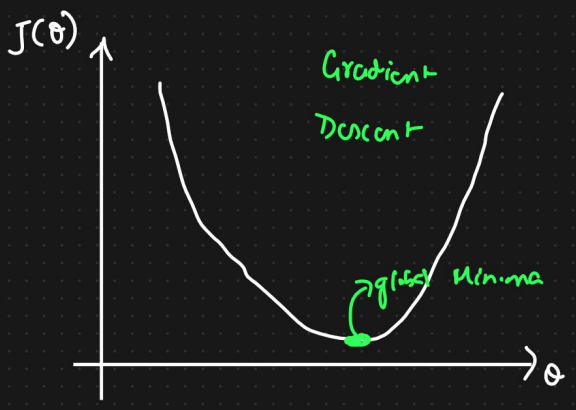
Ridge, Lasso And Elasticnet Regression

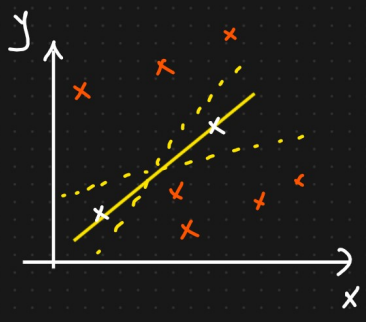
Linear Regression





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

① Ridge Regression (L2 Regularization) ⟶ Reducing Overfitting



Error = 0

Error ↑↑ Test DATA

DATASET

TRAIN TEST

Overfitting

TRAIN ⟶ Acc ↑↑ ⟶ R^2

TEST ⟶ Acc ↓↓ ⟶ R^2

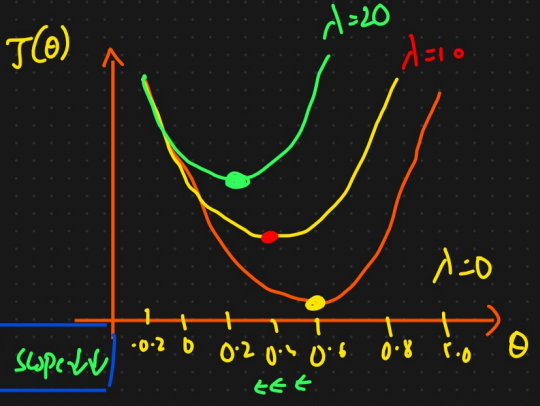
$h_\theta(x) = \theta_0 + \theta_1 x_i$

$\lambda = 1$

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

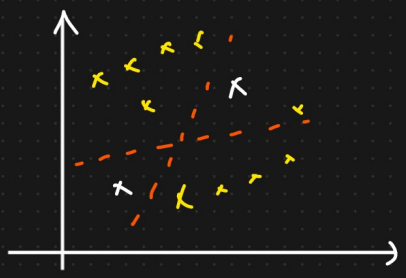
$\downarrow\downarrow$ 0

Hyperparameter



$\lambda \uparrow\uparrow$ slope ↓↓

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



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

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

② Lasso Regression (L1 Regularization) → Feature Selection

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

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

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

feature remove

feature selection

prediction.

$x_2 \rightarrow$ 1 unit

$y \rightarrow 1.5 x_2$

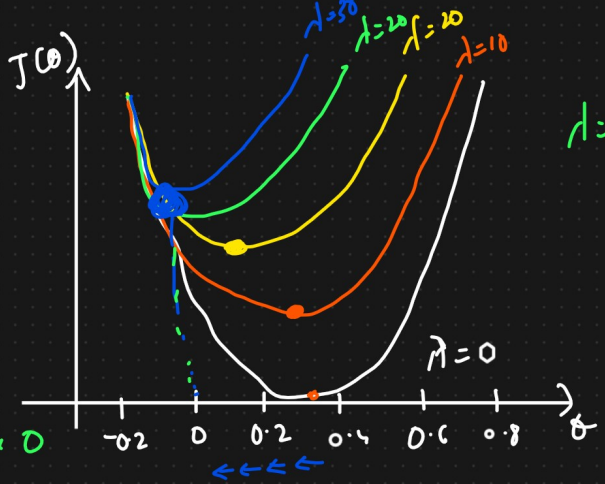
$x_3 \rightarrow$ 1 unit

$y \rightarrow 0.2 x_3$

$\lambda = 1$

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

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



λ = Hyperparameter

θ ↓ λ

λ ↑↑ θ ↓↓

↓

It will became 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}.