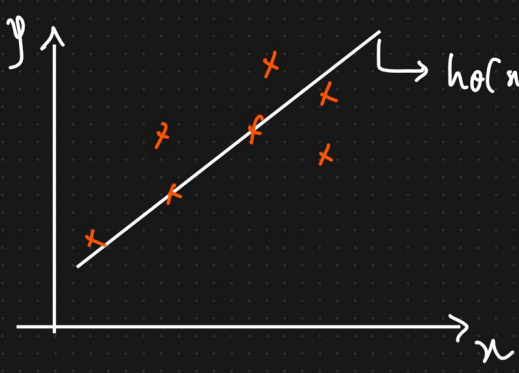
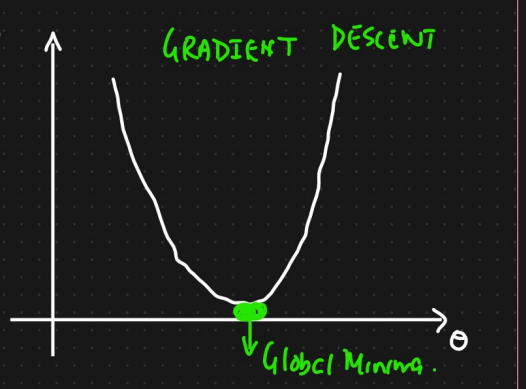
Ridge Regression, Lasso Regression, Elasticnet Regression

Linear Regression





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

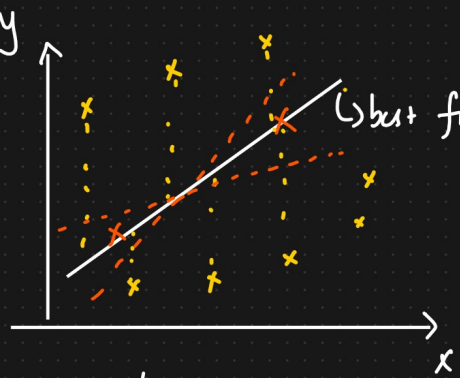
⇓

Mean Squared Error

- **Slope :-** It is movement in y if unit movement in x.

① Ridge Regression (L2 Regularization) → Reduce Overfitting

Overfitting :→



- Traindata → Acc ↑ → low Bias
- Test data → Acc ↓ → High Variance

$h_\theta(x) = \theta_0 + \theta_1 x_1$

$\lambda = 0$

$\lambda = 1$

$\lambda = 10$

Hyperparameter

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

⇓

$$= 0 + (1)\left[(\theta_1)^2\right]$$

$\lambda \uparrow$ Slope ↓

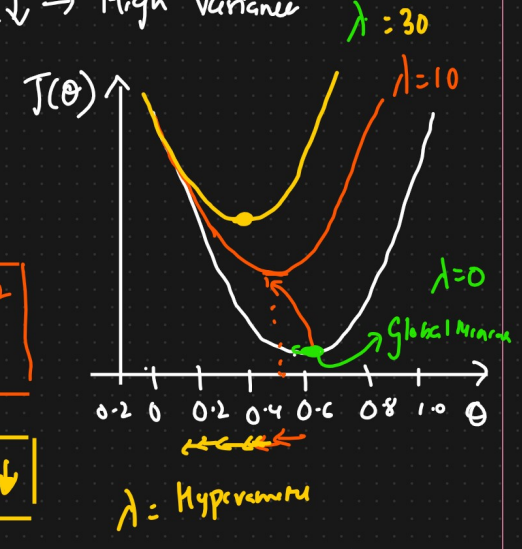
$\lambda = 1$

> 0

$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$ → 0.24

$= 0.34 + 0.52x_1 + 0.48x_2 + 0.24x_3$

(0.52 ⇓, 0.48 ⇓, 0.24 ⇓⇓)



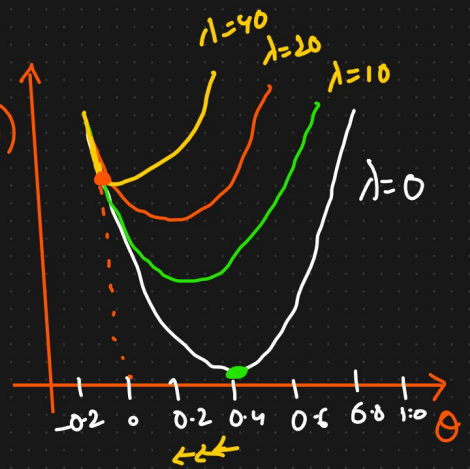
As increasing λ, θ is decreasing. But it'll never be 0.

$$= 0.34 + 0.40x_1 + 0.38x_2 + \boxed{0.14x_3}$$

The features that are not that important will automatically get deleted, and features that are very, very important it will be considered

② Lasso Regression (L_1 Regularization) → Feature Selection

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



$$h_\theta(x) = \theta_0 + \theta_1x_1 + \theta_2x_2 + \theta_3x_3 + \theta_4x_4$$

$$h_\theta(x) = 0.52 + 0.65x_1 + 0.72x_2 + 0.34x_3 + \boxed{0.12x_4}$$

Lasso Regression → 0

$$= 0.52 + 0.51x_1 + 0.60x_2 + 0.14x_3 + \boxed{0x_4}$$

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

But after one point θ will become 0.

In short we are actually trying to remove that specific feature.

③ ElasticNet Regression

Combination of both ridge and Lasso.

→ ① Reduce Overfitting

→ ② Feature Selection

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

λ_1 → Reduce Overfitting

λ_2 → Feature Selection

Hyperparameter Tuning the Linear Regression

Hyperparameter tuning in linear regression refers to the process of selecting the best values for parameters that are not learned from the data but control the learning process. In simple linear regression, there are no hyperparameters to tune since it only estimates coefficients using the least squares method. However, in regularized linear regression models (like Ridge and Lasso regression), hyperparameter tuning is crucial.

Hyperparameter:- It is a parameter that is set before the learning process begins and controls how a machine learning model is trained.

Learning rate (in gradient descent):- Controls how fast the model updates its weights
Regularization strength (A) (in Ridge/Lasso regression)

Tuning :- Tuning is the process of finding the best values for hyperparameters to improve model performance. The goal is to optimize a model's accuracy, minimize error, or balance bias-variance trade-off