SUBJECT: Ridge Regression (L2 Regularization) DATE:

* Ridge regression is used when the model is overfitting
  - will help us hyperparameter tune the linear regression

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

(hyper parameter (default = 1))

- Linear regression cost function
- Ridge Regression cost function

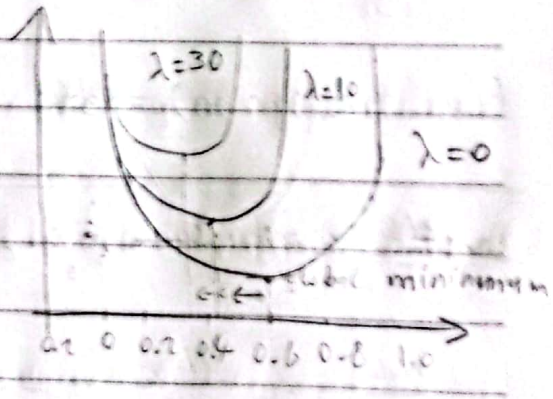
↓

it make the error never equal to zero ($\because > 0$)

* Relationship between (λ & θ)

$\lambda \uparrow$, $\theta \downarrow$ but it will never be zero

$\lambda \uparrow$ slope $\downarrow$



If the θ's movement is high

→ this Features are highly correlated with the output feature

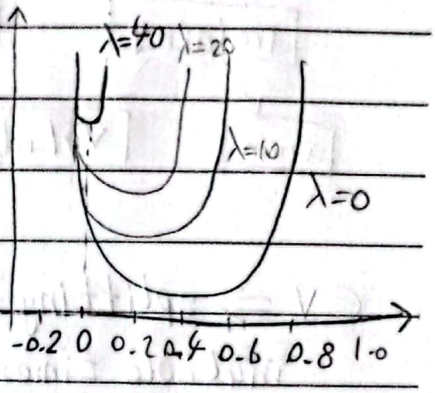
* Small Feature values doesn't affect the movement much

SUBJECT: Lasso Regression (L1 Regularization) DATE:

→ used as feature selection

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

* At a certain λ → the coeffient = 0 → we remove that specific Feature



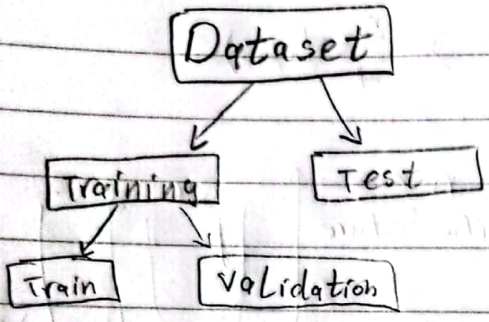
Elasticnet regression → the combination of ridge and lasso regression

$$\text{cost function} = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x)^{(i)} - y^{(i)}) + \lambda_1 \sum_{i=0}^{m} (\text{slope})^2 + \lambda_2 |\text{slope})$$

↓ Reduce overfitting

↓ Feature selection

SUBJECT: Types of Cross Validation DATE:



CV ⇒ splitting the dataset into training and validation sets multiple times and evaluating everytime then getting the average

⇒ Prevents overfitting, ensures better generalization

① Leave One Out CV (LOOCV) ⇒ Uses 1 one sample as val. set and rest for training

② Leave P out CV ⇒ Similar to LOOCV but removes P samples at a time for validation

③ K-Fold CV ⇒ Splits the dataset into K equal-sized Fold

④ Stratified K-Fold CV ⇒ Similar to K-Fold but ensures that each Fold has the same propotion of each class (useful For imbalanced data)

⑤ Time series CV ⇒ Used for time-dependent data