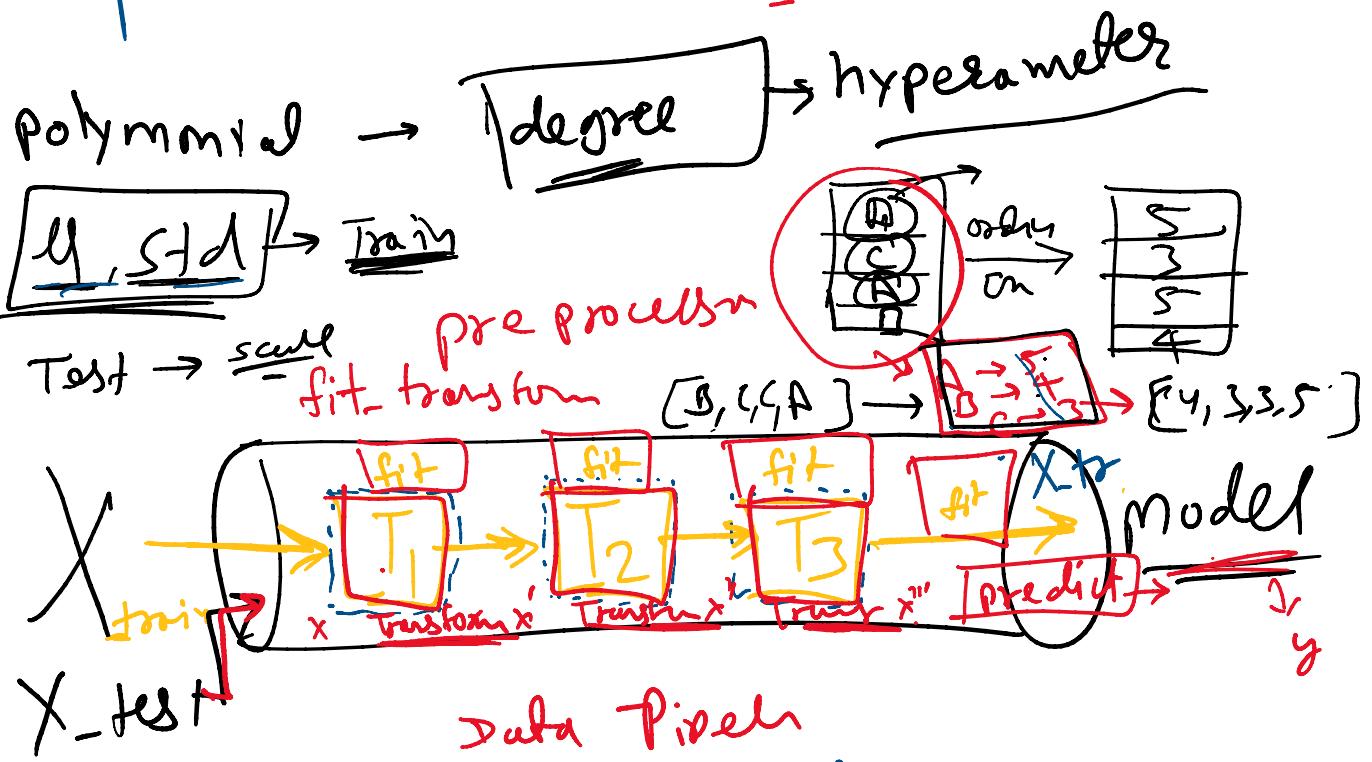
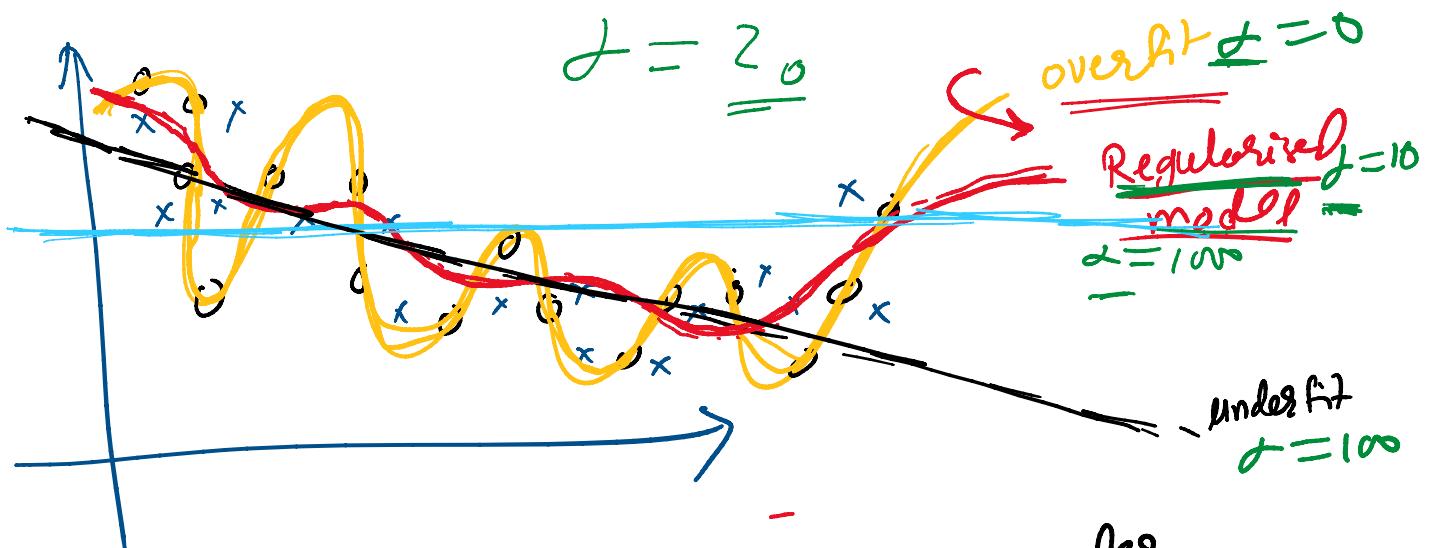


9. Regularization

Monday, August 8, 2022 12:13 PM



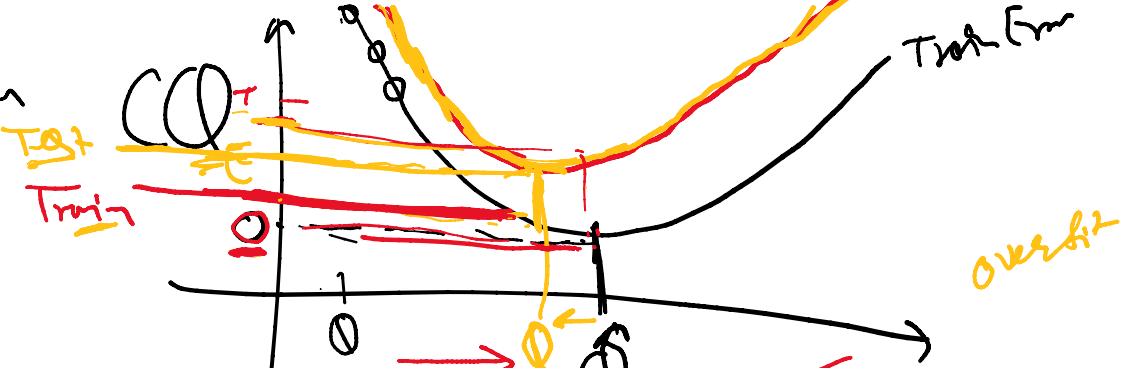
ML
Training model - LR SGD

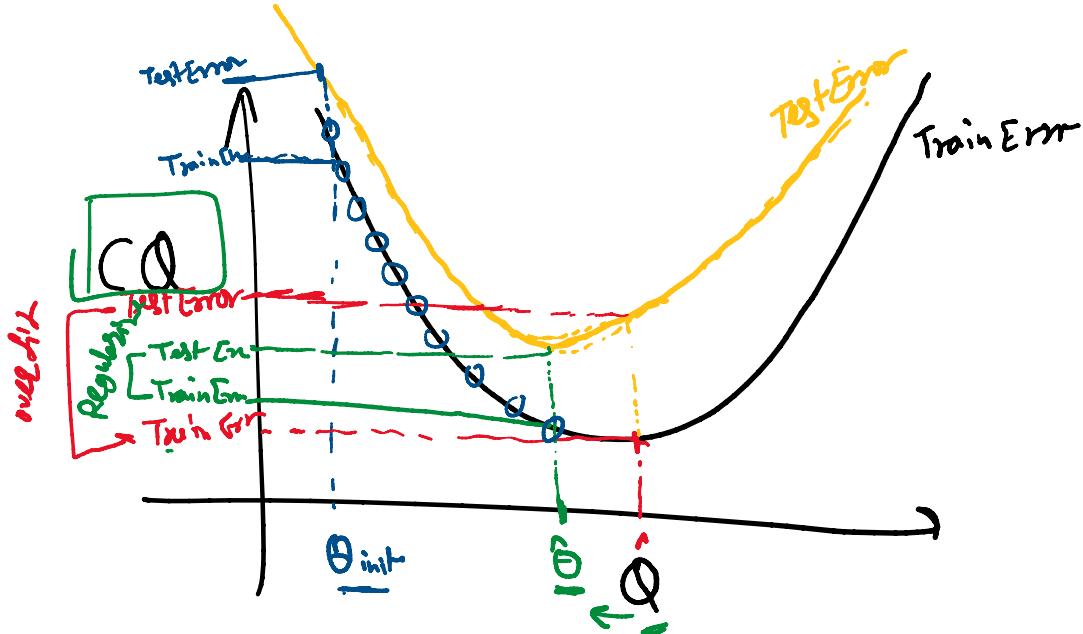
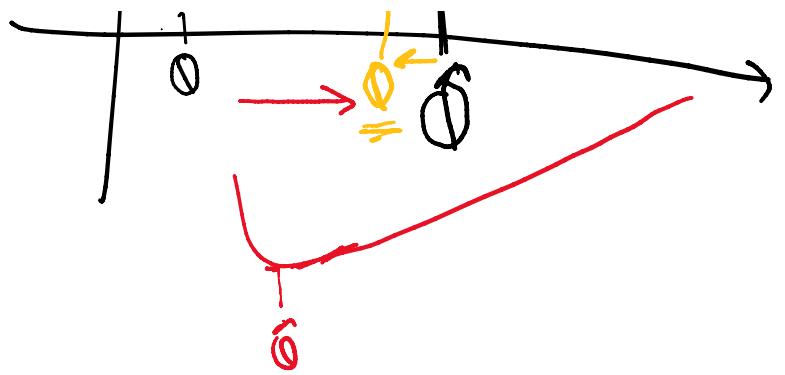
Model

Transform mode - SC, ODE

Train

Error





① Ridge Regression

② LASSO Regression

③ Elastic Net

$$\begin{cases} \hat{w} = \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \dots, \hat{\alpha}_n \\ b = \theta_0 \end{cases}$$

$$\hat{y} = b + w^T X$$

$$y = H(Q) = \theta \cdot X$$

$$\hat{\theta} = (X^T X)^{-1} X^T y$$

Normal Gleich

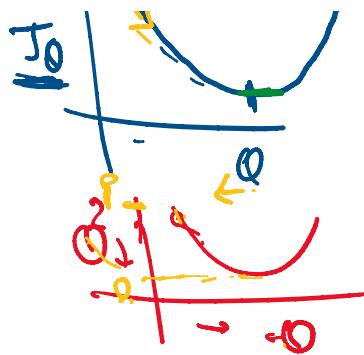
$$J(\theta) = \text{MSE} \theta$$

optimize

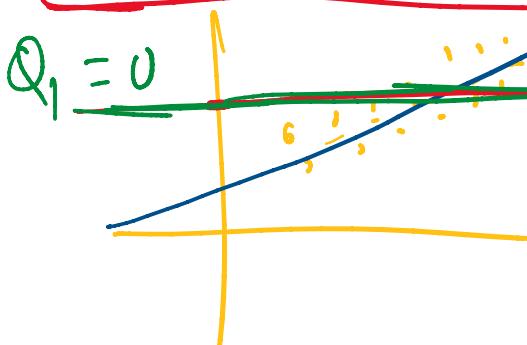


Ridge Regression

Regularization or Hyper Parameter



$$J(\theta) = \text{MSE} \theta + \frac{\lambda}{N} \sum_{i=1}^N \theta_i^2$$



$y = mx + c$
 $y = c$

High Regularization \rightarrow Underfitting

$$\begin{aligned}\hat{\theta}_{\text{next}} &= \hat{\theta}_{\text{prev}} - \nabla J(\theta) \\ &= \hat{\theta}_{\text{prev}} - \nabla \text{MSE}(\theta) - \lambda \sum_{i=1}^N \theta_i\end{aligned}$$

$$\hat{\theta} = (X^T \cdot X)^{-1} \cdot X^T \cdot y$$

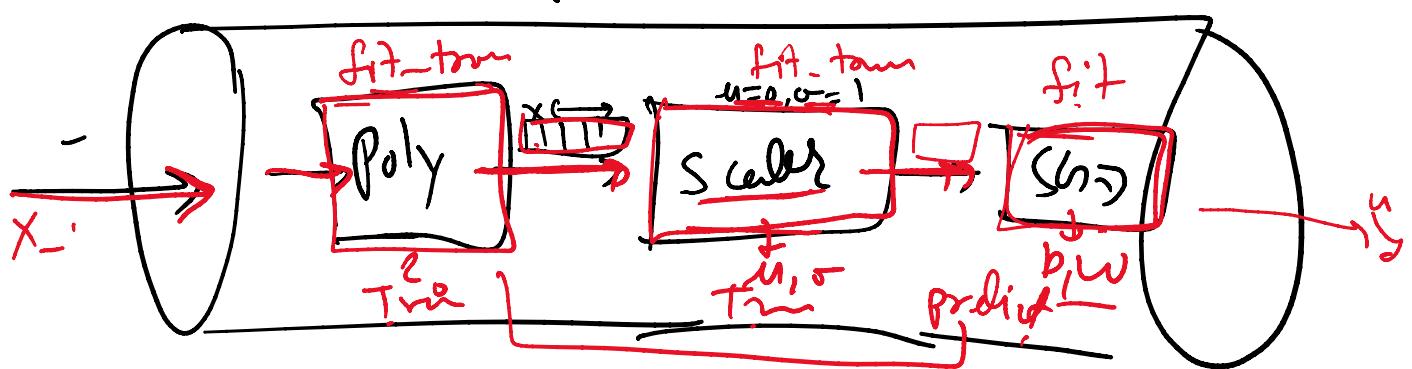
$$\begin{aligned}20 &\rightarrow 10 \\ 20 &\rightarrow 4 \\ 20 &\rightarrow 2\end{aligned}$$

$$\hat{\theta} = (X^T \cdot X + \lambda \cdot I)^{-1} \cdot X^T \cdot y$$

$$I = \lambda \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Pipeline

Pipeline



$X_{\text{train}} \rightarrow \underline{\text{fit}}$ \rightarrow
 $X_{\text{test}} \rightarrow \underline{\text{predict}}$

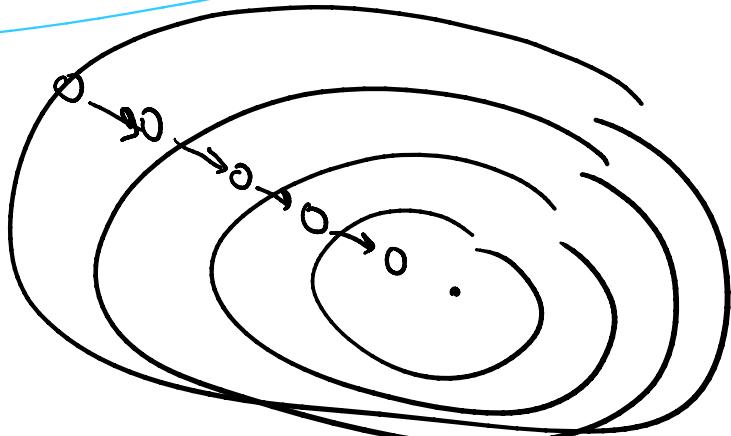
$$\nabla_{\theta} \sum_{i=1}^n \theta_i^2$$

$$\theta_1 = \nabla_{\theta_1} \sum (\theta_1 + \theta_2 + \theta_3)^2$$

$$= \nabla_{\theta_1} 2(\theta_1 + \theta_2 + \theta_3) \cdot \theta_1$$

Ridge

\rightarrow

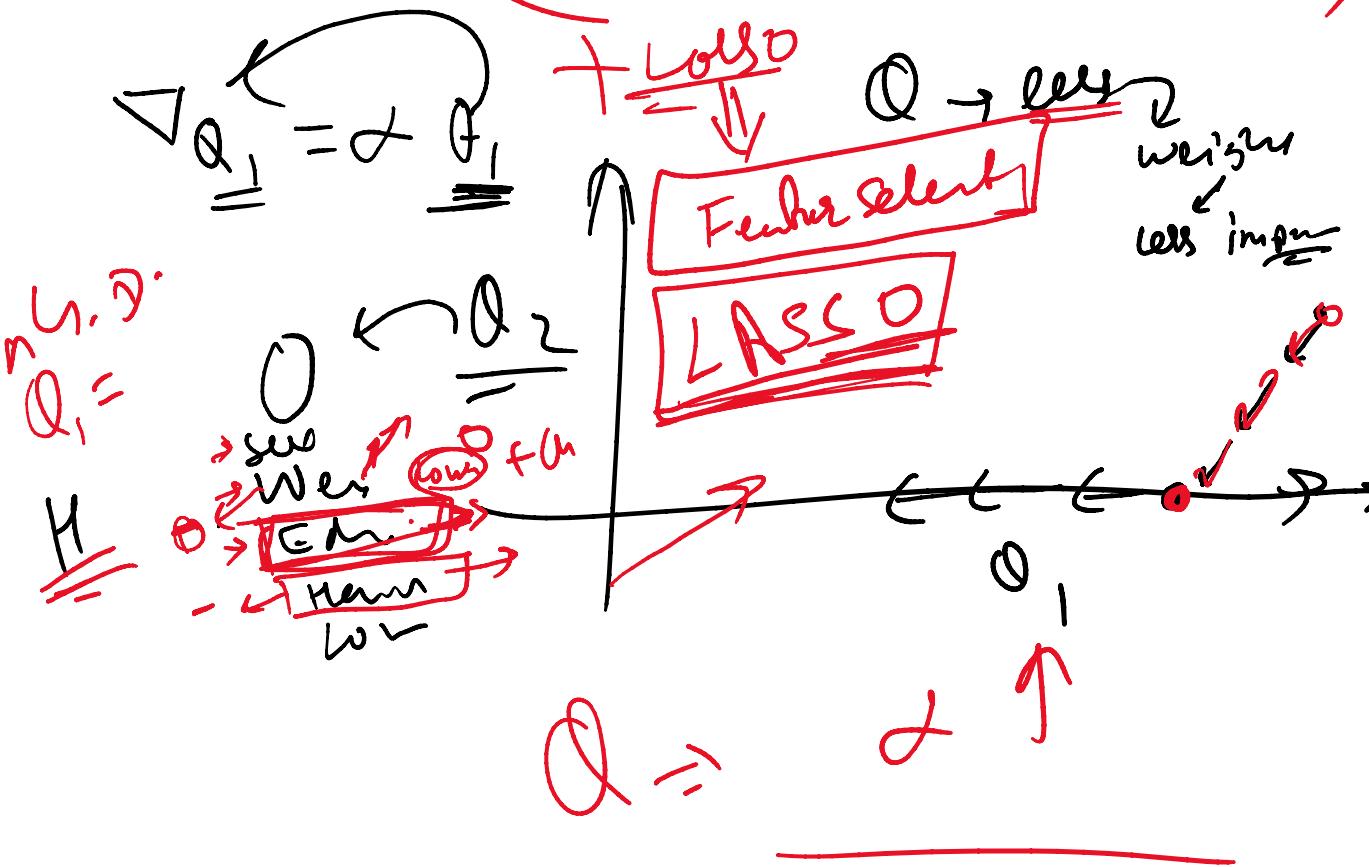


Lasso

Lasso

$$\varepsilon | Q |$$

$$\nabla_{\theta_1} = \frac{1}{Q_1} \left(|\theta_1| + |\theta_2| + |\theta_3| + |\theta_4| \dots \right)$$



Ridge \rightarrow overfit for

Ridge \rightarrow Feature Select

Elastic Net

$\gamma = 1$
 $\gamma = 0$

Elastic Net

$\gamma < 0$

$$\lambda \left[\gamma \cdot |\beta| + (1-\gamma) \beta^2 \right]$$

$\gamma(1|\beta|) \rightarrow \text{LASSO}$ $\boxed{\gamma = 0.5}$

$\gamma(\beta^2) \rightarrow \text{Ridge}$

$\gamma = 0.2, \gamma = 0.7$



ID 4.5

