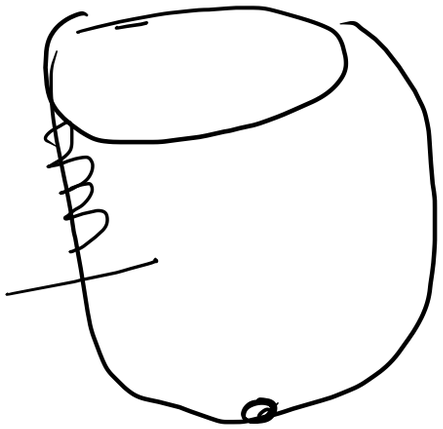
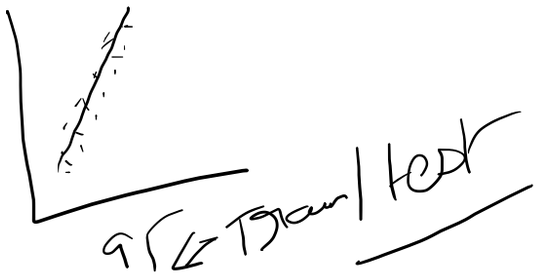
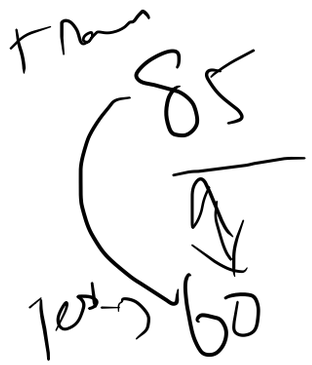






1. Cross validation
2. dimension reduction
3. Regularization
   - L1 & L2
   - L3

Regularization

Overfitting

$$\hat{y} = 349.97 + 2.0117\,x_1 + 30\,x_2 + 60\,x_3 + 70\,x_4$$

After adding pen

(1)

$$\hat{y} = 349.97 + 2.0117 + 20\,x_2 + 50\,x_3 + 60\,x_4$$

$$\hat{y} = 349.97 + 2.0097\,x_1 + 20\,x_2 + 0\,x_3 + 0\,x_4$$

L1

Euclidean distance

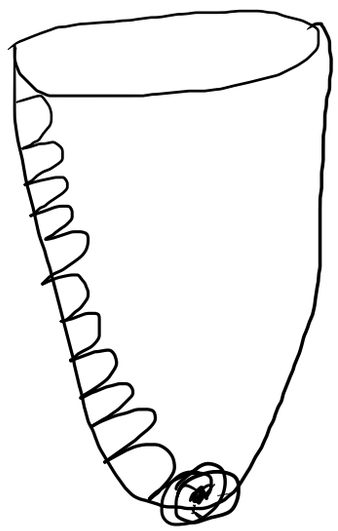
(x_1, y_1) (x_2, y_2)

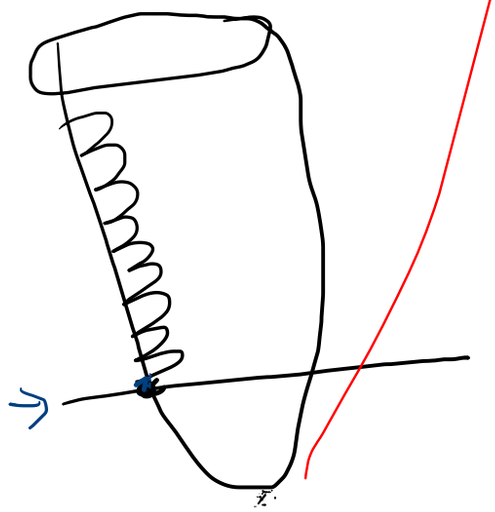
$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

manhattan is 2nd ans

$$d = |x_2 - x_1| + |y_2 - y_1|$$

overfit





L_1 ∈ Manhattan distance → Lasso → Least absolute shrinkage and Selection Operator

L_2 ∈ Euclidean distance → Ridge → Square magnitude of coefficient

θ_0 → Coefficient

[0.01, 0.02, 0.3]

Ridge

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x) - y)^2 + \lambda \sum \theta_j^2$$

$\frac{\lambda}{2m}$

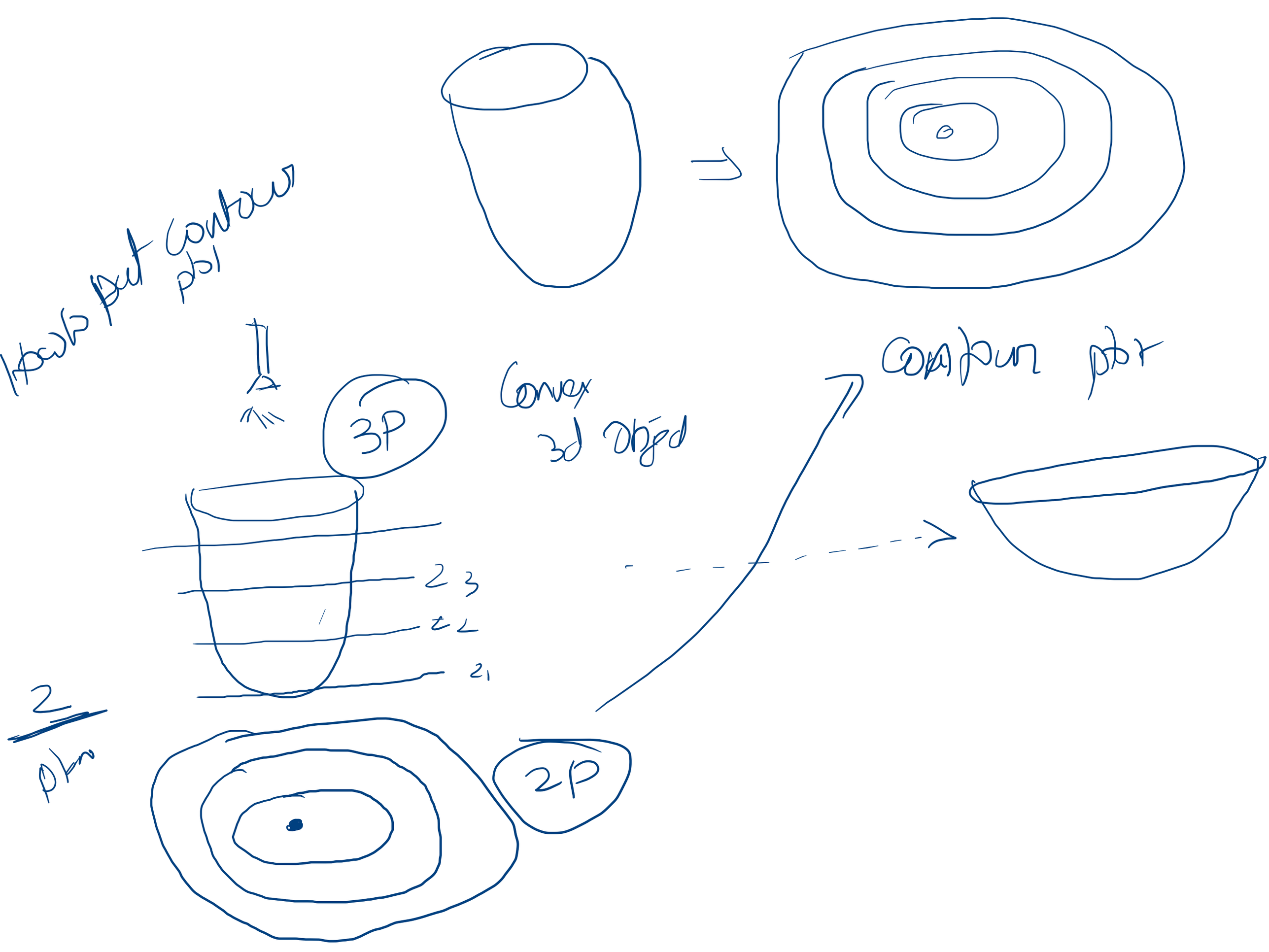
L2

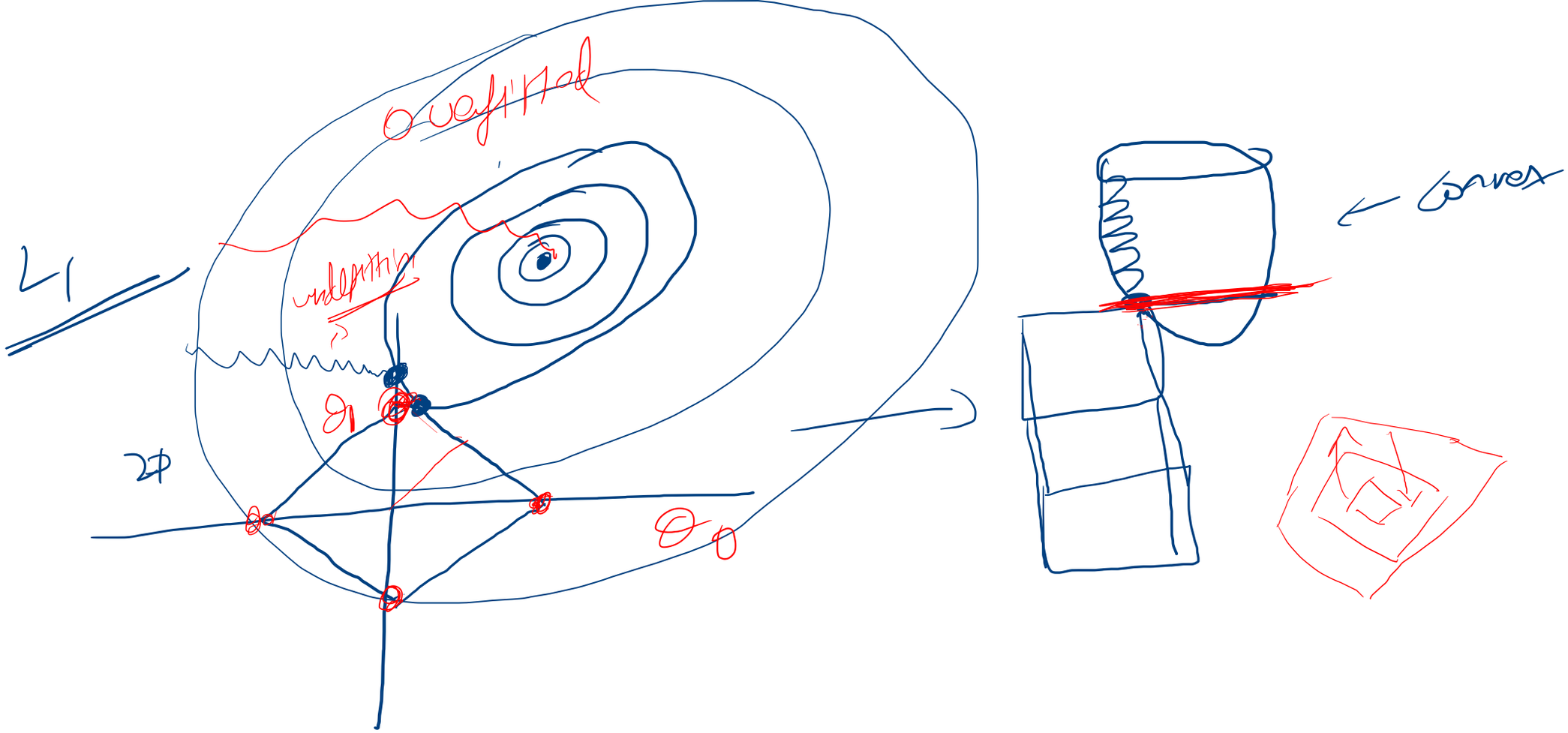
Lasso regression

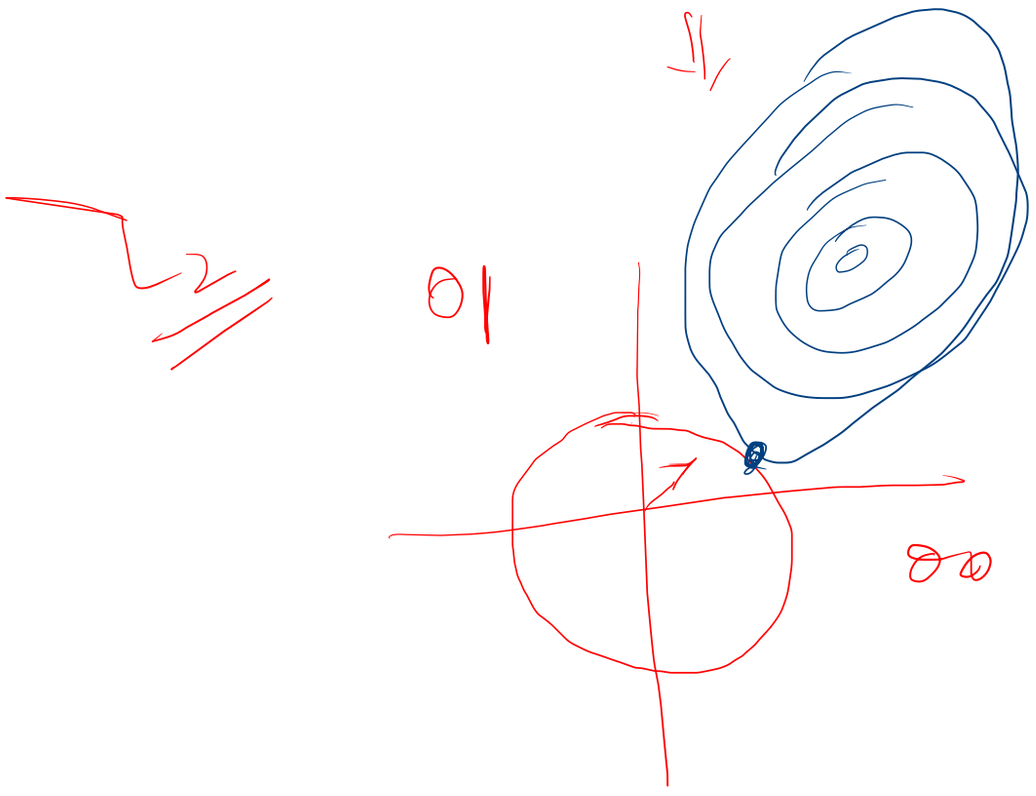
$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x) - y)^2 + \frac{\lambda}{2m} \sum |\theta|$$

(Sparse coefficient)

polynomial regression

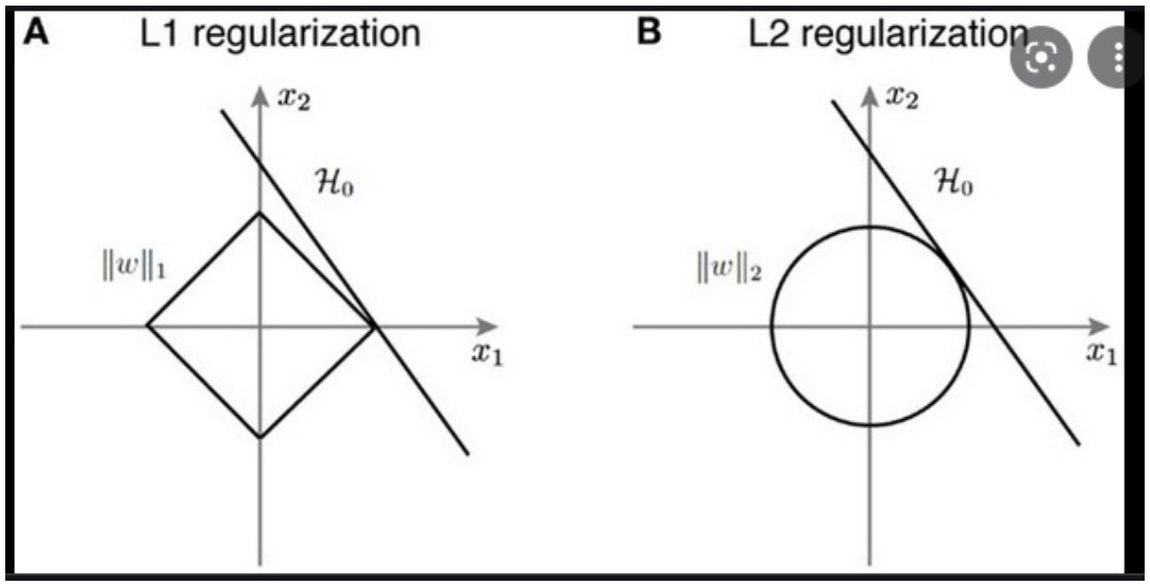

Convex
3d Object
3P
2P
z_3
z_2
z_1

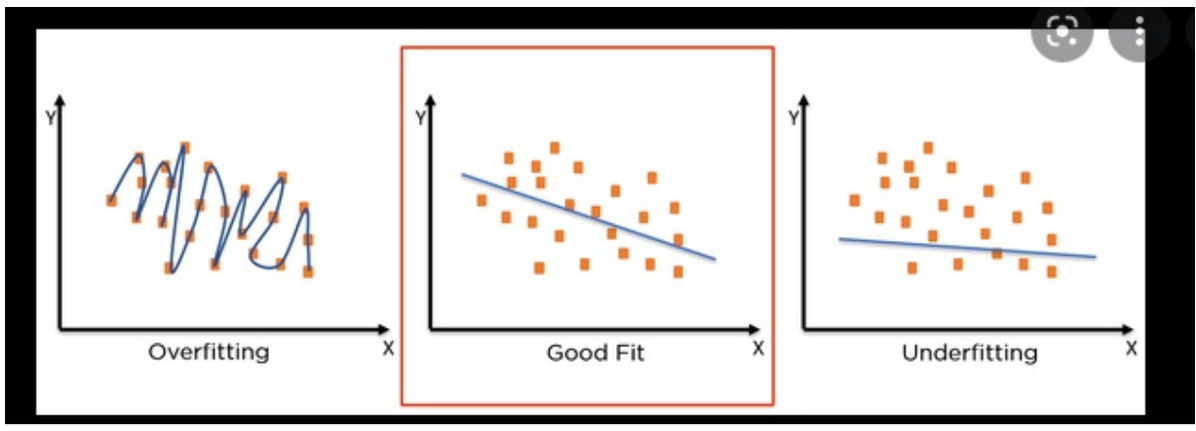




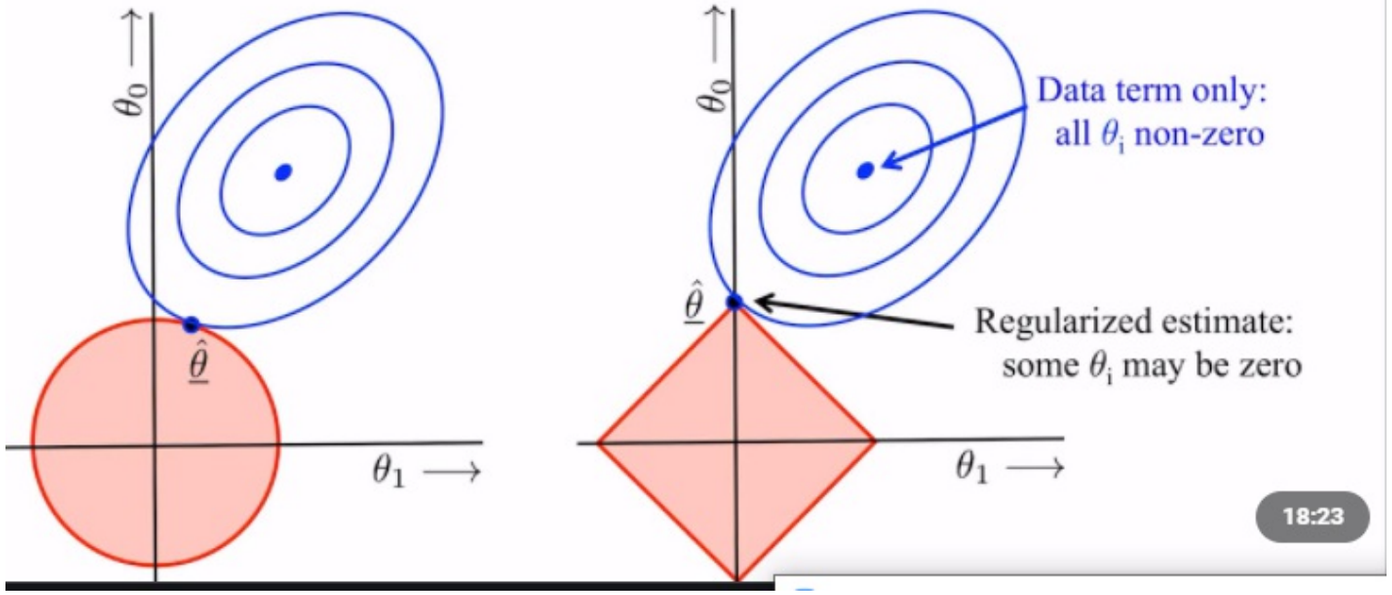
x^2+y^2

$\theta_1^2+\theta_2^2$





L1 tends to generate sparser solutions than a quadratic regularizer



18:23