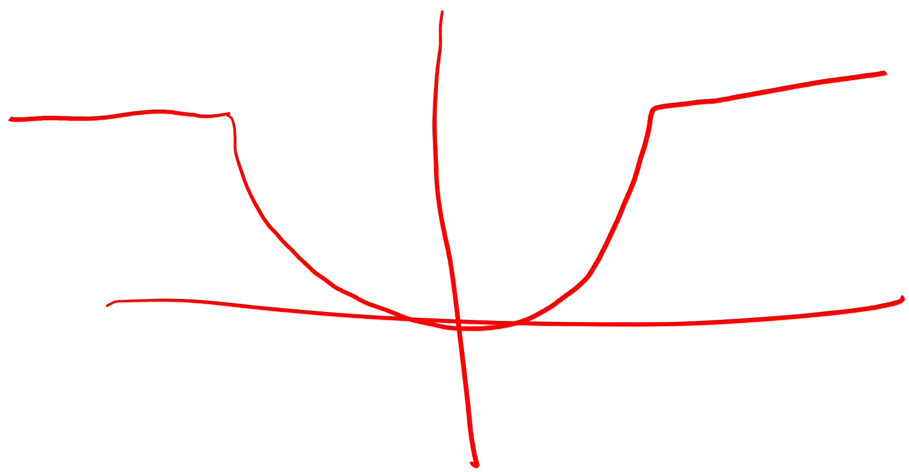
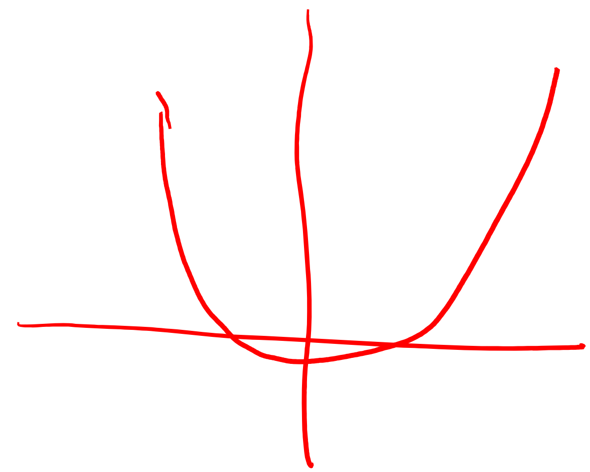




$$f(x_i) = w^T \varphi(x_i)$$

$$\min_w \frac{\lambda}{2} \|w\|_p + \sum_{i=1}^{N} L\left(y_i - \left(w^T \varphi(x_i)\right)\right)$$

Ridge Regression

$$\min_w \frac{\lambda}{2} \|w\|^2 + \sum_{i=1}^{N} \left(y_i - w^T \varphi(x_i)\right)^2$$

LASSO

$$\min_w \frac{\lambda}{2} \|w\|_1 + \sum_{i=1}^{N} \left(y_i - w^T \varphi(x_i)\right)^2$$

$$\min_{(\beta_1, \beta_0)} \sum_{i=1}^{N} \left(y_i - (\beta_1 x + \beta_0)\right)^2 + \lambda K$$

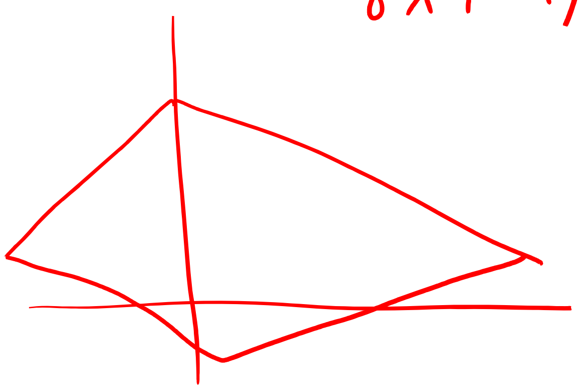
Subject to, $|\beta_1| + |\beta_0| \le K$

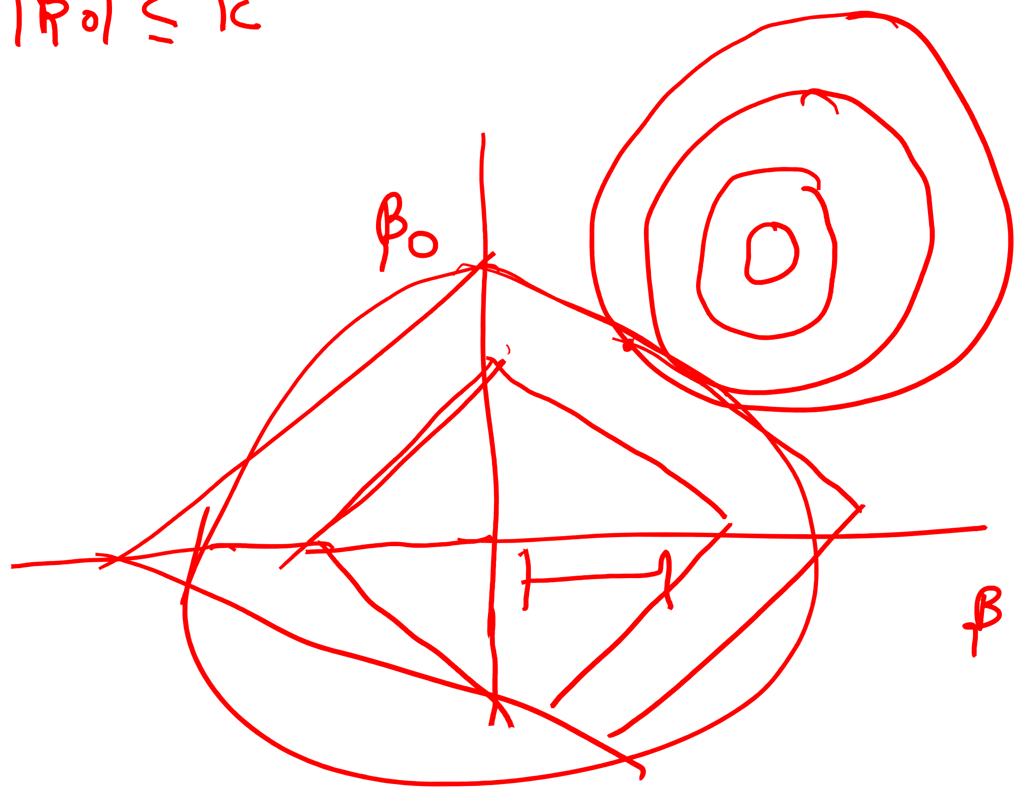
$$\min \quad ax + by$$

Subject to

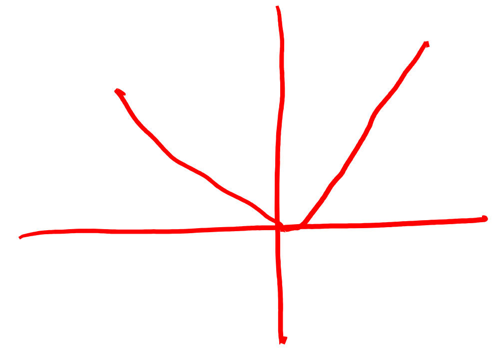
$$cx + dy \le e$$

$$\gamma x + my \le c$$





$L(y) = |y|$



L₁-Norm Regression model

$$\underset{w}{\min} \underbrace{\frac{\lambda}{2}\|w\|^2 + \sum_{i=1}^{N} |y_i - w^T\varphi(x_i)|}_{J(w)}$$

$$\delta_w J(w) = \lambda w + \sum_{i=1}^{N} g(i)$$

$$g(i) = \delta_w\left(|y_i - w^T\varphi(x_i)|\right) = -\mathrm{sign}(y_i - w^T\varphi(x_i))\,\varphi(x_i) \quad \text{if } |y_i - w^T\varphi(x_i)| > 0$$

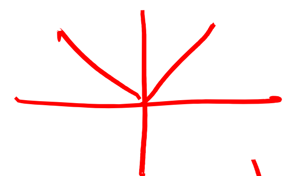
$$= \gamma\,\varphi(x_i), \text{ otherwise}$$

$$\gamma \in [-1, 1]$$

$$\min_{w} \frac{\lambda}{2} \|w\|_1 + \sum_{i=1}^{N} \left| y_i - w^T \varphi(x_i) \right|$$

Elastic Net

$$\min_{w} \frac{\lambda\alpha}{2} \|w\|_1 + \frac{\lambda}{2}(1-\alpha) \|w\|^2 + \sum_{i=1}^{N} \left(y_i - w^T \varphi(x_i) \right)^2$$

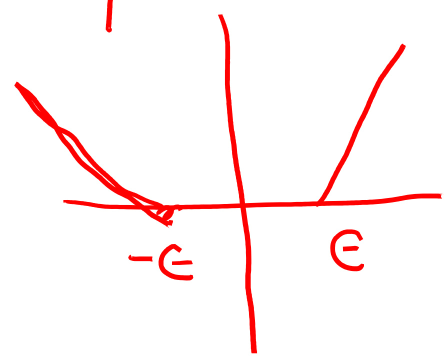


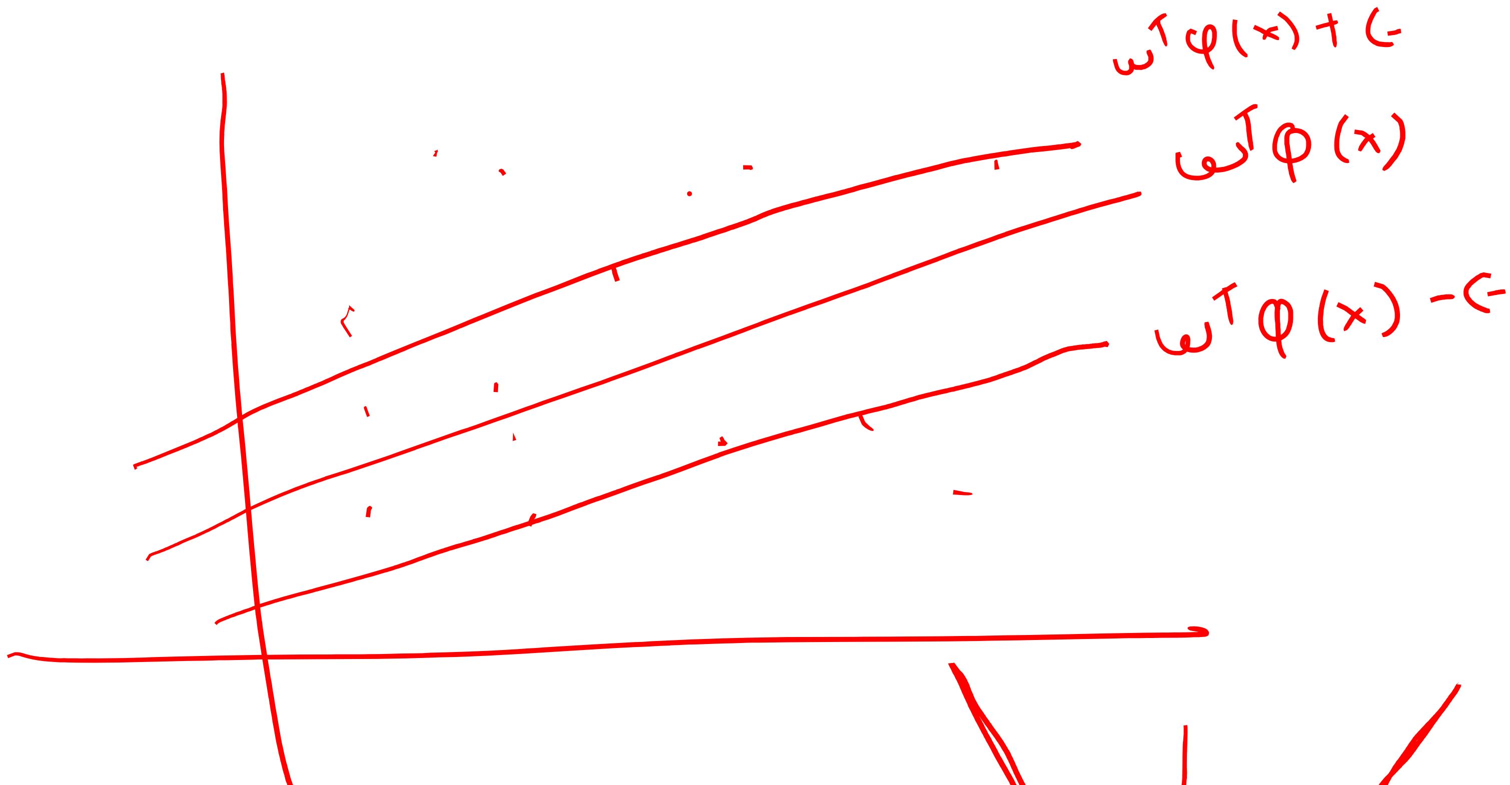
ϵ- insensitive Regression

$$\min_{w} \frac{\lambda}{2} \|w\|^2 + \sum_{i=1}^{N} \left| y_i - w^T \varphi(x_i) \right|_{\epsilon}$$

where

$$|u|_{\epsilon} = \begin{cases} 0 & |u| < \epsilon \\ |u| - \epsilon & \text{otherwise} \end{cases}$$




$\omega^T \varphi(x) + \epsilon$
$\omega^T \varphi(x)$
$\omega^T \varphi(x) - \epsilon$

Huber