LASSO

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p$$

Find $\hat{\beta}$ to minimize

$$h(\hat{\beta}) = \underbrace{(\vec{y} - \vec{X}\hat{\beta})^T(\vec{y} - \vec{X}\hat{\beta})}_{\text{MSE on training set}} + \underbrace{\lambda \sum_{i=1}^{p} |\beta_i|}_{\text{penalty on size of } \hat{\beta}}$$

Solve for β if $\beta < 0$

$$-2\vec{x}^T(\vec{y} - \vec{x}\beta) - \lambda = 0$$

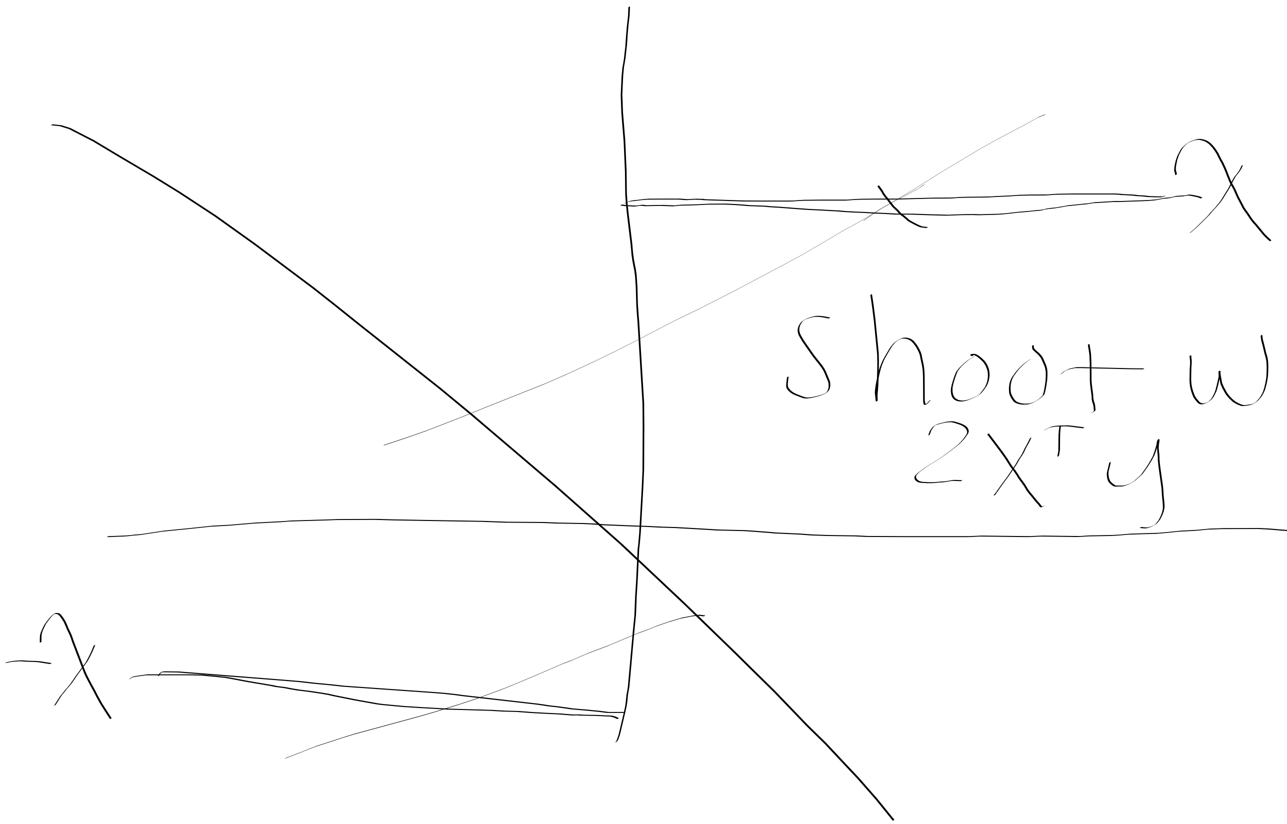
$$\cancel{-2\vec{x}^T\vec{y}} + 2\vec{x}^T x\beta = \lambda + 2\vec{x}^T\vec{y}$$

$$\beta = \frac{\lambda + 2\vec{x}^T\vec{y}}{2x^Tx}$$

$$\beta < 0 \quad \frac{\lambda + 2\vec{x}^T\vec{y}}{2x^Tx} < 0 \Rightarrow \lambda + 2\vec{x}^T\vec{y} < 0$$

$$2\vec{x}^T\vec{y} < -\lambda$$

$$[x_1\ x_2 \cdots x_n]\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1^2 + x_2^2 + \cdots + x_n^2$$



Shoot with $2X^Ty$

<u>$p = p$</u>

Start $\hat{\beta}^{(0)} = \hat{\beta}_{LR}$

Iterate through p coordinates of $\hat{\beta}$

<u>Step m</u> $\beta_j^{(m)}$ "shoot" to find update formula
$j = 1, \ldots, p$

Stop when $\hat{\beta}^{(m)}$ converges

$$\hat{\beta}^{(m-1)} \approx \hat{\beta}^{(m)}$$

Taylor Polynomial

$f(x) = \sin x$ approximate near a

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

$\boxed{\text{math } 253}$

Ridge Regresson

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p$$

training data $S = \{\vec{X}, \vec{y}\}$

Find $\hat{\beta}$ minimizing value

$$\underbrace{\cancel{\frac{1}{n}}(\vec{y} - \vec{X}\hat{\beta})^T(\vec{y} - \vec{X}\hat{\beta})}_{\text{MSE on training set}} + \underbrace{\lambda \sum \beta_i^2}_{\text{penalty on size of } \hat{\beta}}$$

$$\hat{\beta} = (\qquad + \lambda I)^{-1} \vec{X}^T \vec{y}$$

$\underline{p=1}$

$$\frac{\partial}{\partial\beta}\left[\lambda|\beta|\right]$$

$$\frac{\partial h(\hat{\beta})}{\beta} = -2X^T(\vec{y}-\vec{X}\hat{\beta}) + \underline{\qquad\qquad}$$

$$= \begin{cases} -2X^T(\vec{y}-\vec{X}\hat{\beta}) - \lambda & \text{if } \beta < 0 \\ ? & \text{if } \beta = 0 \\ -2X^T(\vec{y}-\vec{X}\hat{\beta}) + \lambda & \text{if } \beta > 0 \end{cases}$$

Solution

$$\beta \begin{cases} \dfrac{\lambda + 2\vec{x}^T\vec{y}}{2X^TX} & \text{if } \cancel{\beta < 0} \quad 2\vec{x}^T\vec{y} < -\lambda \\ 0 & \text{if } \cancel{\beta = 0} \quad -\lambda \le 2x^Ty \le \lambda \\ \dfrac{-\lambda + 2\vec{x}^T\vec{y}}{2X^TX} & \text{if } \cancel{\beta > 0} \quad 2\vec{x}^T\vec{y} > \lambda \end{cases}$$