midterm = Thursday
May 9th
in Owen 217

Wednesday = "review"
day

Linear Regression

high. dim.

$p >> n$ $\quad (X^T X)^{-1}$ not invertable

$p \approx n$ $\quad (X^T X)^{-1}$ may not work

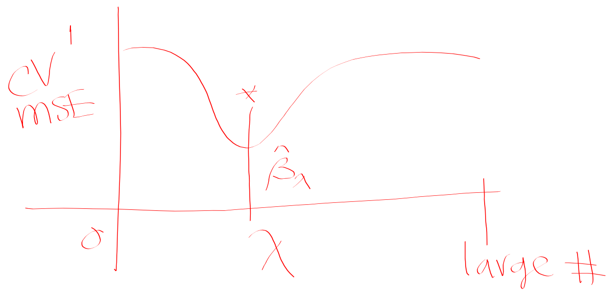
$p << n$ $\quad \hat{\beta}_{LR} = \widetilde{(X^T X)^{-1}} \vec{X}^T \vec{y}$

low dim.

$$\frac{1}{0+1}$$

$$\hat{\beta}_{ridge} = (X^T X + \lambda I_p)^{-1} \vec{X}^T \vec{y}$$

grid search of λ in $(0, \infty)$
with cross-valided MSE



Ridge Regression vs Linear Regression

$\boxed{n >> p}$

$MSE(\hat{f}) = bias(\hat{f})^2 + Var(\hat{f})$

f is ~ linear

$\hat{\beta}_{ridge}$ and $\lambda = 0$ then $\hat{\beta}_{ridge} = \hat{\beta}_{LR}$

optimal λ

$\hat{\beta}_{ridge} = 0$, $\lambda = \text{huge \#}$

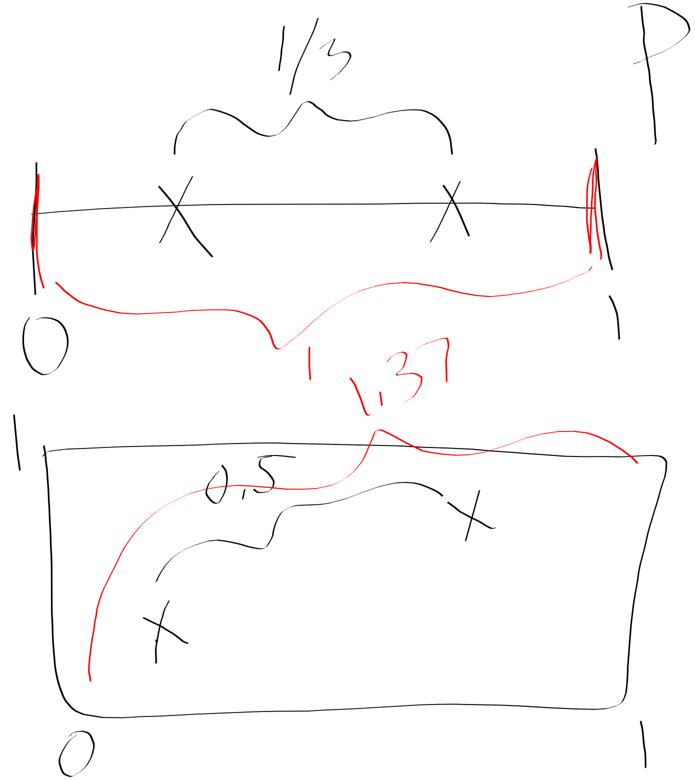
Curse of dimensionality

kNN

p = predictors

$p \gg n$ = # example

$p = 1$



$p = 2$

$$\hat{\beta}_{LR} = (\vec{X}^T\vec{X})^{-1}\vec{X}^T\vec{y}$$

was minimizing value

$$MSE = \frac{1}{n}(\vec{y} - \hat{y})^T(\vec{y} - \hat{y})$$

$$= \frac{1}{n}(\vec{y} - \vec{X}\hat{\beta})^T(\vec{y} - \vec{X}\hat{\beta})$$

$\vec{X}$ is $\underset{\#\text{example}}{n} \times \underbrace{(p+1)}_{\#\text{ of predictors}}$

$$\frac{\partial MSE}{\partial \beta} = -2\underset{p+1 \times n}{\vec{X}^T}\underset{n \times 1}{(\vec{y}} - \vec{X}\hat{\beta}) = 0$$

$$(p+1) \times 1 = 0$$

Ridge Regression

< should work well if

$p >> n$ or $p \approx n$

and $Y = f(X)$ is roughly linear.

$\hat{\beta}$ to minimize

$$(\vec{y}-\vec{X}\hat{\beta})^T(\vec{y}-\vec{X}\hat{\beta})$$

subject to $\hat{\beta}^T\hat{\beta} = \beta_1^2+\beta_2^2+\cdots+\beta_p^2 \le c$

Lagrange Multilipers

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

$$\frac{\partial C(\hat{\beta})}{\partial\hat{\beta}} = 0 \Rightarrow \text{solve for } \hat{\beta}$$