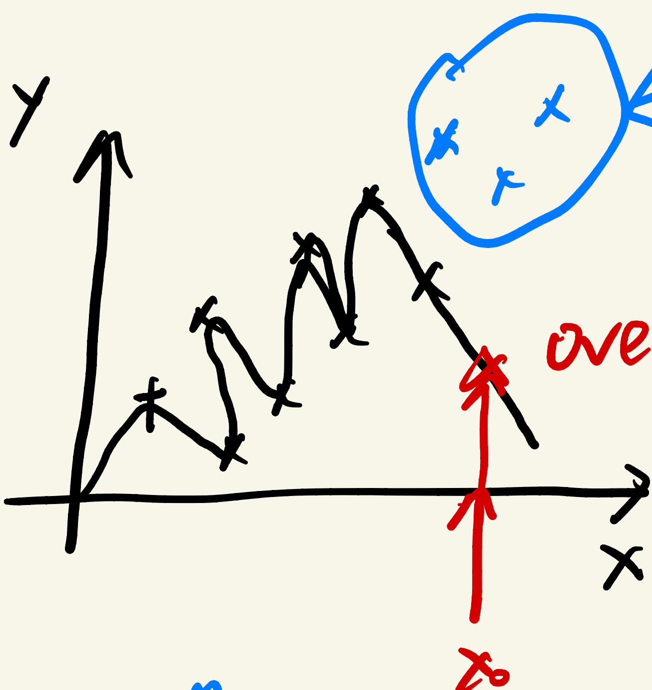
Lecture 2 Linear Regression

metric (valid)



overfitting

overfit on validation

loss (train)

test data

$h(x) = \vec{\theta}^T \vec{X}$ dot product

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_3 \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_3 \end{bmatrix}$$

$$\Rightarrow (X^{(i)}, y^{(i)}) \qquad n$$

linear

$$\underline{\underline{|h_{\theta_1}(x^{(i)}) - y^{(i)}|}} > |h_{\theta_2}(x^{(i)}) - y^{(i)}|$$

$$|h_{\theta_1}(x^{(i)}) - y^{(i)}|^2 > |h_{\theta_2}(x^{(i)}) - y^{(i)}|^2$$

?

$$X\theta \to R^{n\times 1}$$

$$J(\theta) = \frac{1}{2}\underbrace{(h_\theta(x) - y)}_{\text{vector } R^n}^T (h_\theta(x) - y)$$

$X \to R^{n\times(d+1)}$

$n \to$ # samples

$$\nabla_\theta J(\theta) = 0$$

$$\nabla_\theta J(\theta) = \frac{1}{2}\nabla_\theta \left[\theta^T(X^T X)\theta - 2(X^T y)^T\theta + y^T y\right]$$

$$= \frac{1}{2}\left[2X^T X\theta - 2X^T y\right]$$

$$= X^T X\theta - X^T y = 0$$

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

not invertible

$X^T X \to R^{(d+1)\times(d+1)}$

$X \to R^{n\times(d+1)}$

$$\underline{X^T X} \to R^{(d+1)\times(d+1)}$$

$$\Downarrow$$

$$\text{rank}(X^T X) = d+1$$

$\min J = 0$

1. $n < d+1$

$\text{rank}(AB) \leq \min[\text{rank}(A), \text{rank}(B)]$

$\text{rank}(X) \leq \min(n, d+1) < d+1$

$X^T X \leq \text{rank}(X) < d+1$

$X =$ 

2. $\text{rank}(X) < d+1$

$x \quad y$

$x^{(i)} \quad y^{(i)}$

$\varepsilon^{(i)} \sim N(0, \sigma^2)$

$\theta^T x^{(i)} + \varepsilon^{(i)} = y^{(i)}$

$y^{(i)}$ given $x^{(i)}$

maximize $L(\theta)$

$\arg\max L(\theta)$

$= \arg\max \log L(\theta)$

$\log L(\theta) \rightarrow$ log likelihood

$$= n \log \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{\sigma^2} \cdot \frac{1}{2} \sum_{i=1} \left(y^{(i)} - \theta^T x^{(i)}\right)^2$$

$\hat{\theta}$

maximum Likelihood estimator (MLE)

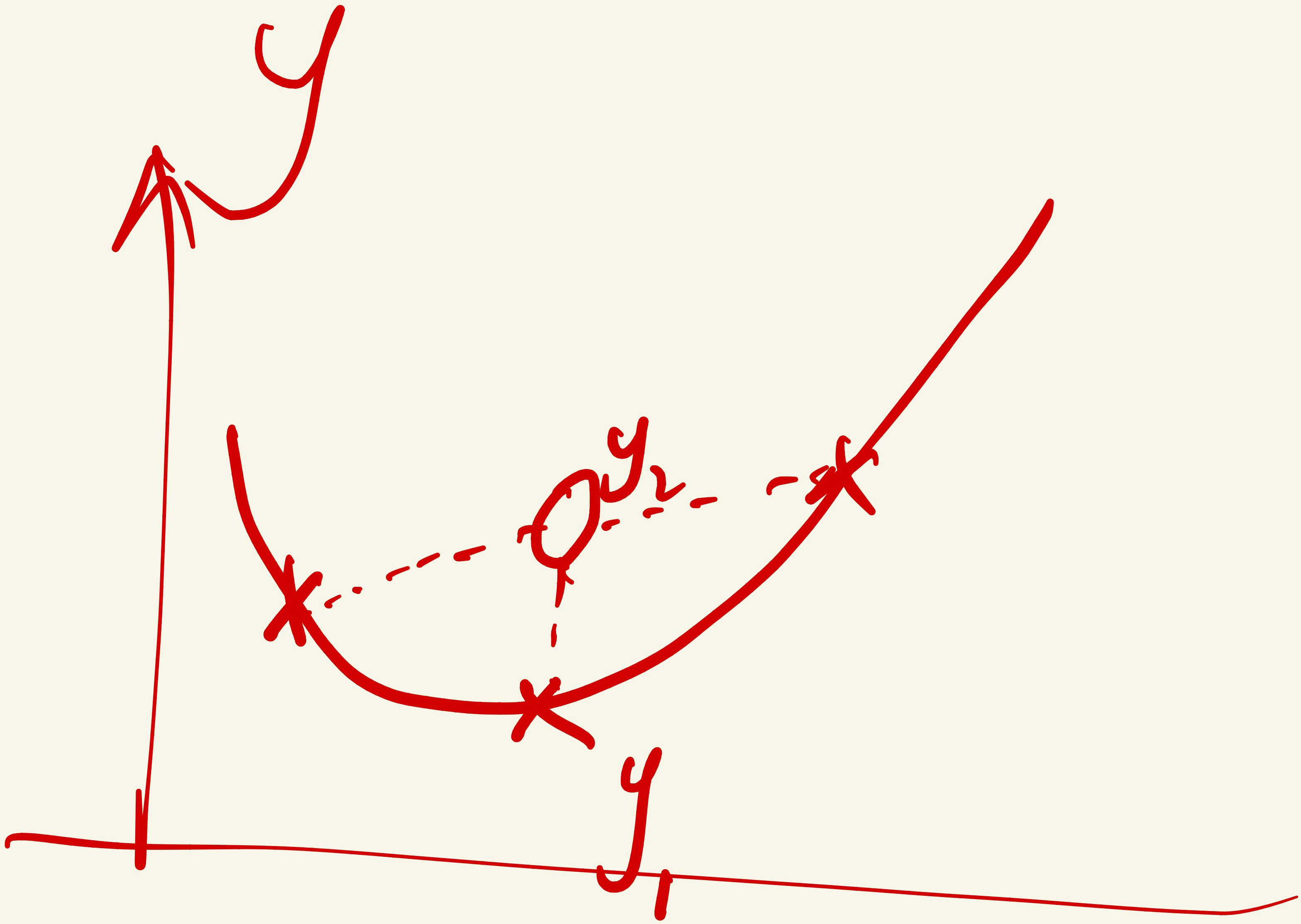
$$\underbrace{\theta^* \qquad \hat{\theta}}_{\text{relation}}$$

$$\boxed{\hat{\theta} \longrightarrow \theta^*}$$

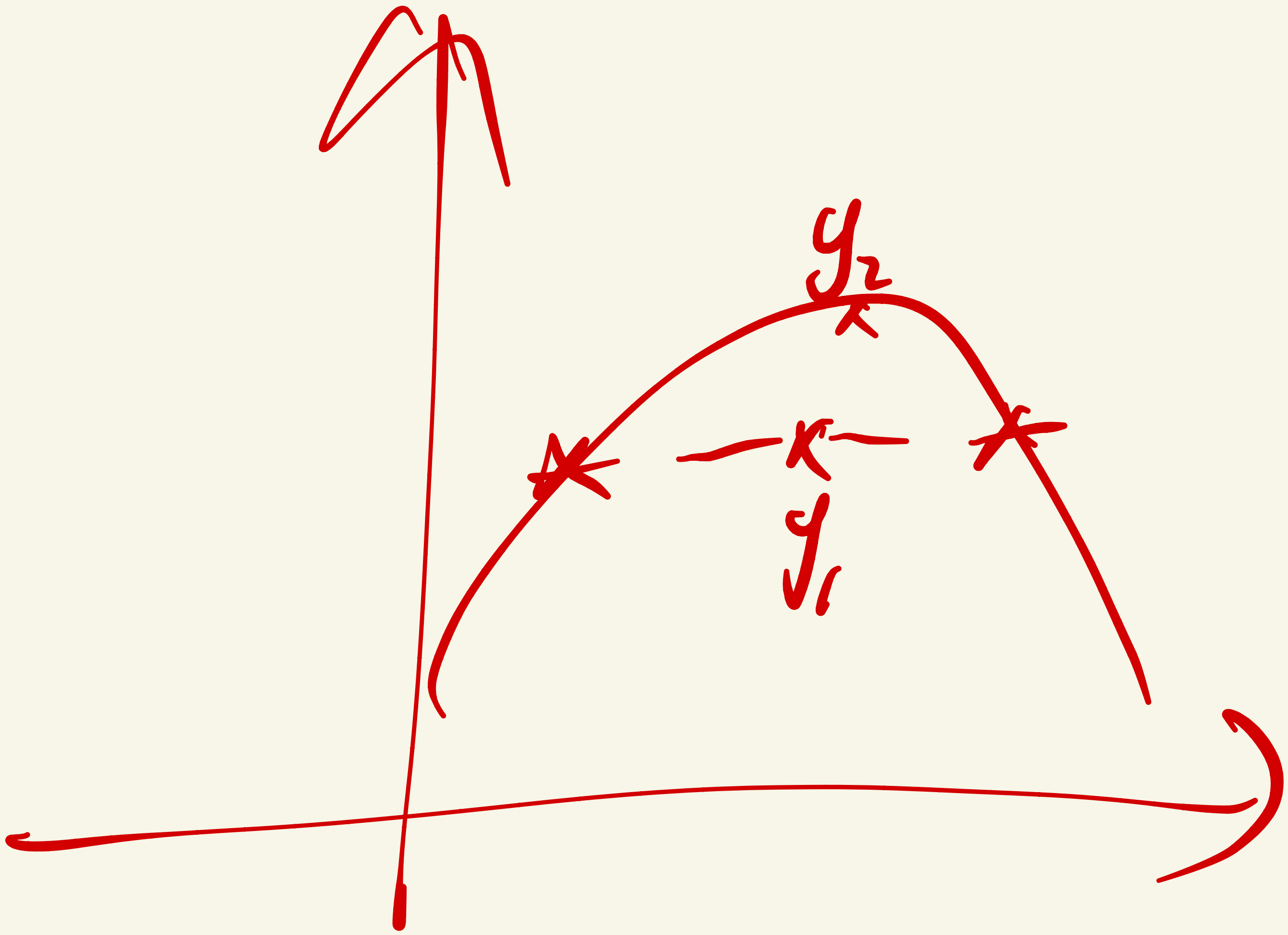
$$\theta_j := \theta_j + \alpha(h_\theta(x^{(i)}) - y^{(i)})\vec{x}_j^{(i)}$$

$$\overline{\theta + 1}$$

$$X_n = \begin{bmatrix} x & x & x \end{bmatrix}$$



$y_2 > y_1$ convex



$y_2 > y_1$ concave