

Lecture 2

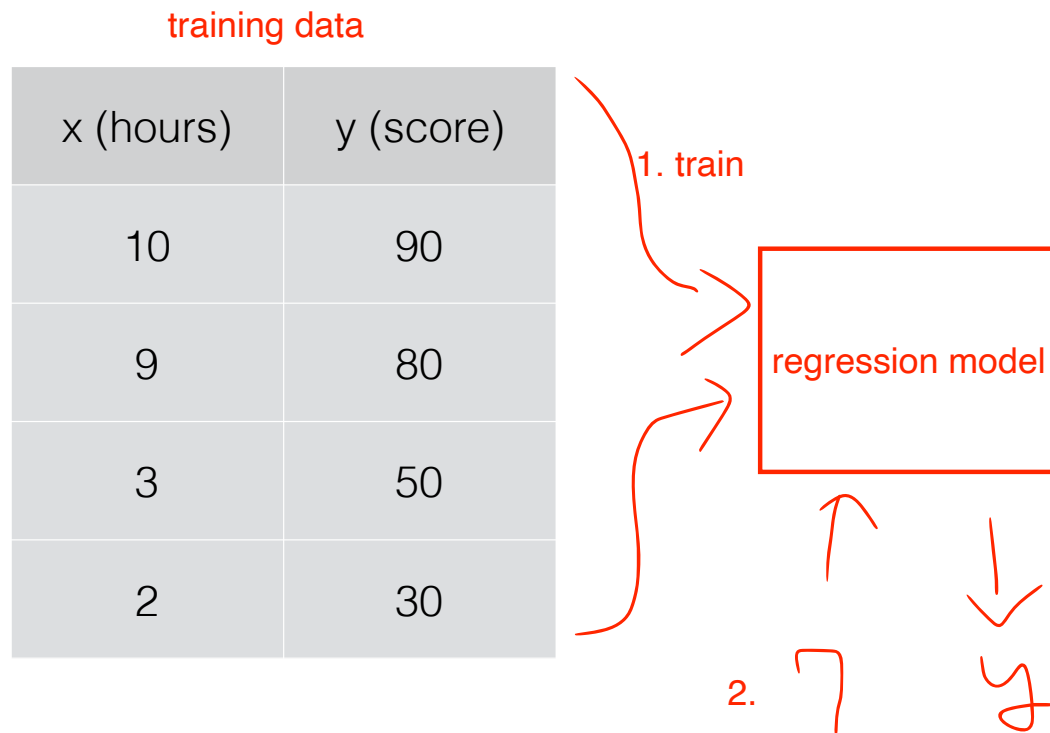
Linear Regression

Sung Kim <hunkim+ml@gmail.com>

Acknowledgement

- Andrew Ng's ML class
 - <https://class.coursera.org/ml-003/lecture>
 - <http://www.holehouse.org/mlclass/> (note)
- Convolutional Neural Networks for Visual Recognition.
 - <http://cs231n.github.io/>
- Tensorflow
 - <https://www.tensorflow.org>
 - <https://github.com/aymericdamien/TensorFlow-Examples>

Predicting exam score: regression



training data를 가지고 학습시켜 regression model을 생성한 후, 어떤 값을 넣으면 학습된 데이터를 바탕으로 y 값을 출력

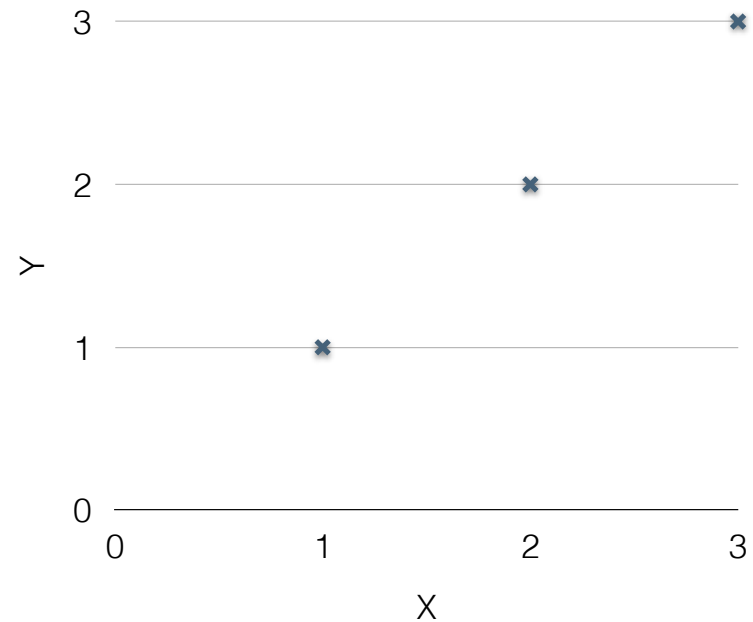
Regression (data)

예측을 해야할 값

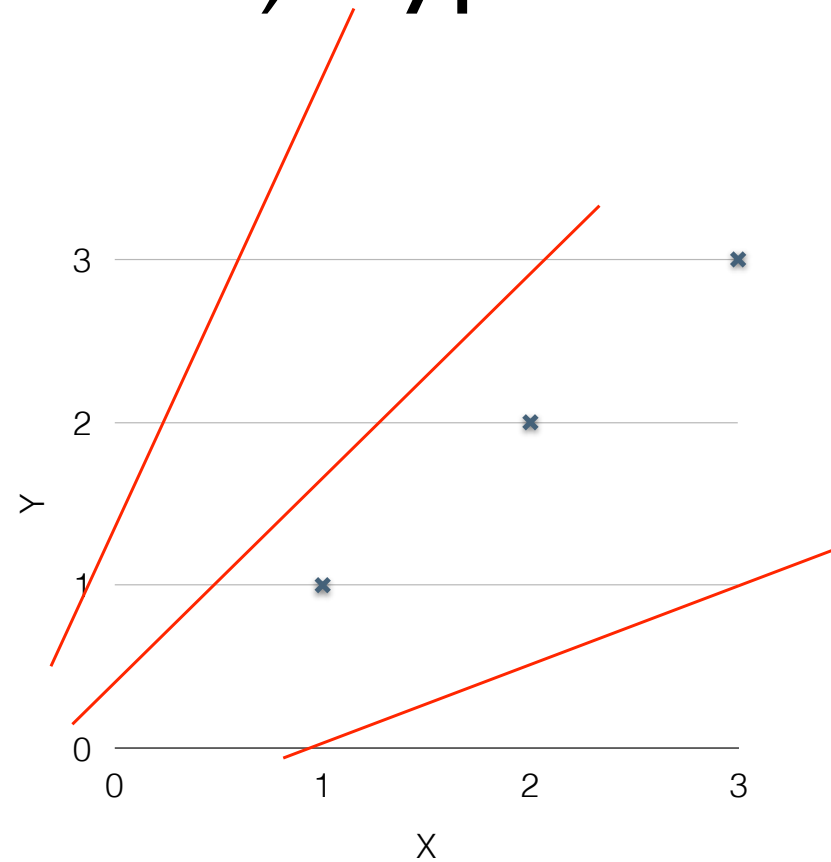
x	y
1	1
2	2
3	3

Regression (presentation)

x	Y
1	1
2	2
3	3

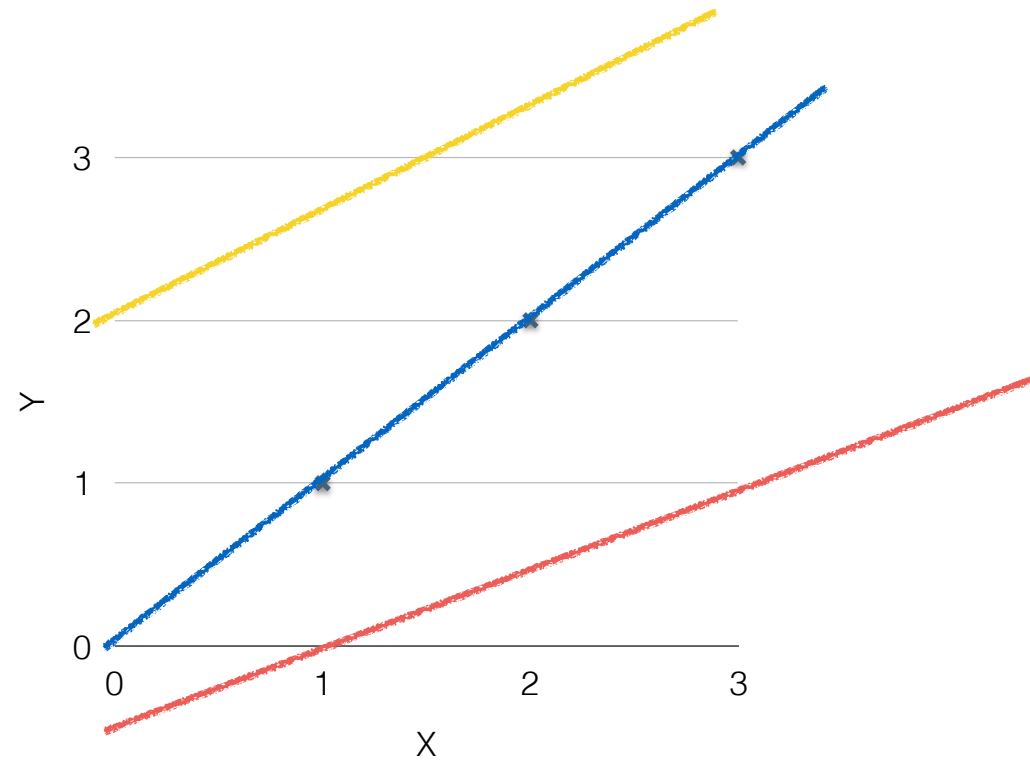


(Linear) Hypothesis



데이터를 설명할 수 있는 선을 찾는 게 학습을 하는 것

(Linear) Hypothesis

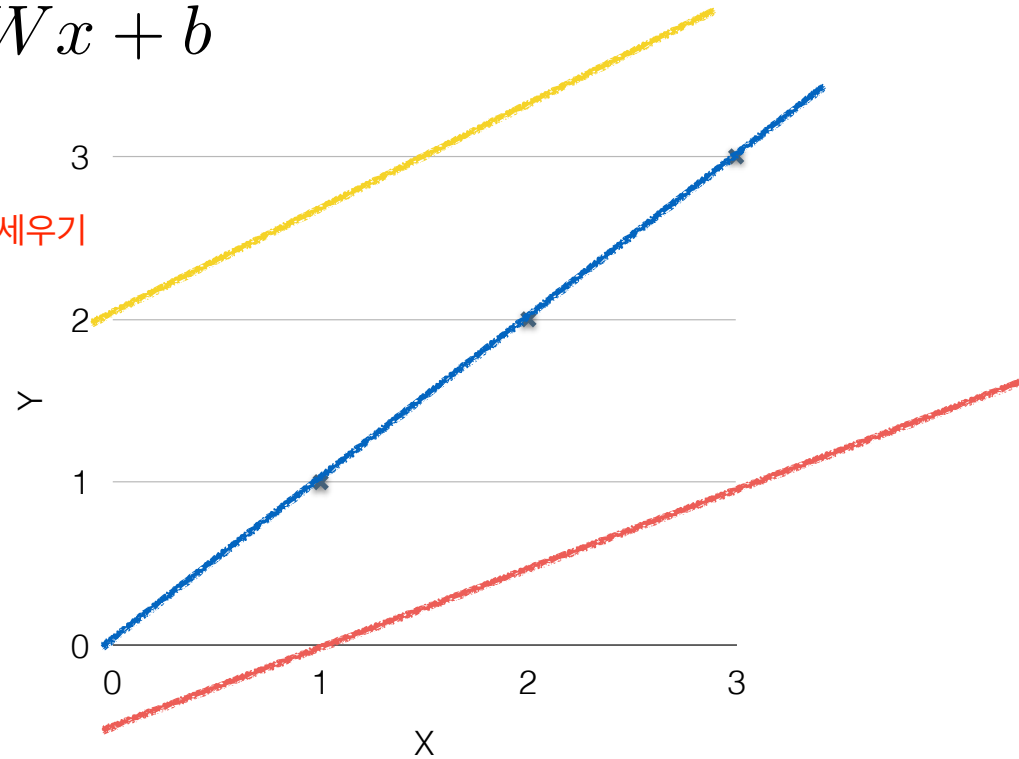


(Linear) Hypothesis

$$H(x) = Wx + b$$

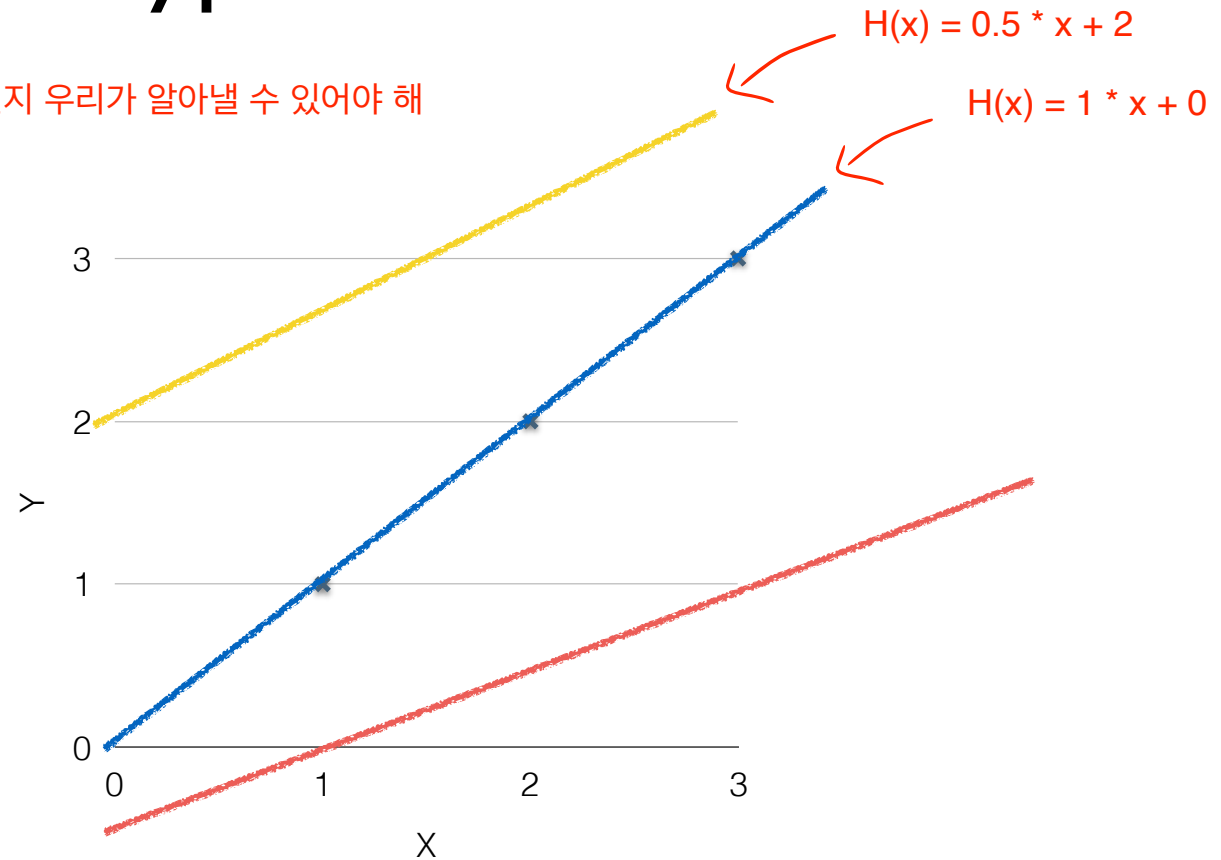
가설

Linear Hypothesis의 첫 번째 단계 - 일차 방정식 세우기



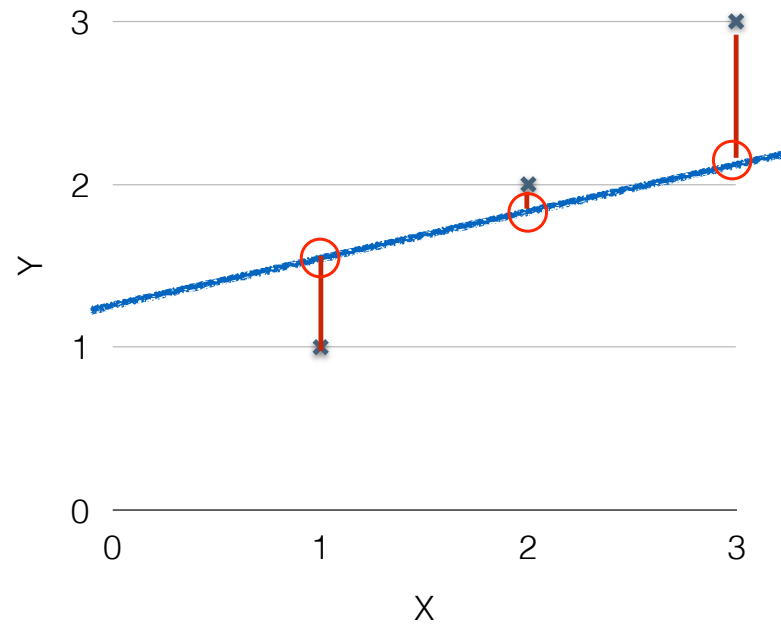
Which hypothesis is better?

어떤 가설이 더 좋은 가설인지 우리가 알아낼 수 있어야 해



Which hypothesis is better?

실제 데이터와 가설이 나타내는 데이터 점들 간의 거리를 계산해서 멀면 안 좋고 가까우면 좋고



^{Loss}Cost function

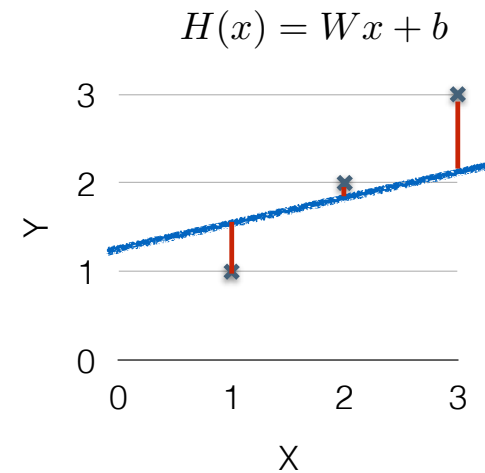
거리를 측정하는 것을 Cost function이라고 부른다

- How fit the line to our (training) data

$$\cancel{H(x) - y}$$

차이가 - 또는 + 일 수 있으므로. 제곱을 한다

$$(H(x) - y)^2$$



Cost function

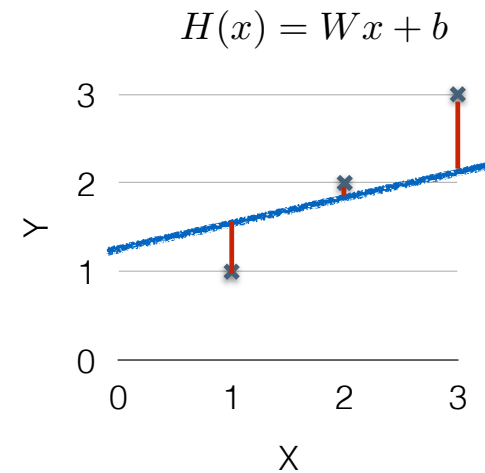
- How fit the line to our (training) data

$$\frac{(H(x^{(1)}) - y^{(1)})^2 + (H(x^{(2)}) - y^{(2)})^2 + (H(x^{(3)}) - y^{(3)})^2}{3}$$

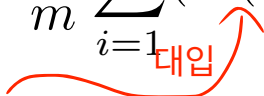
3 값이 세 개 있어서

데이터의 개수

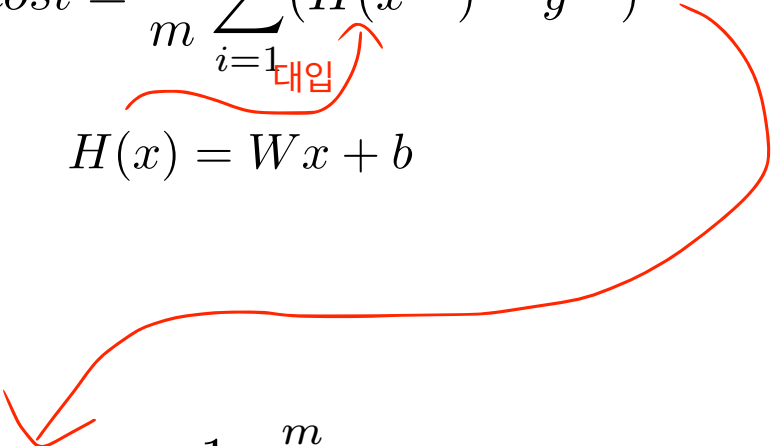
$$cost = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2$$



Cost function

$$cost = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2$$


$$H(x) = Wx + b$$


$$cost(W, b) = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2$$

그러면 W 와 b 의 function이 된다

값(cost)을 가장 작게 하는 W 와 b 를 구하는 것이 Linear Regression의 학습

Goal: Minimize cost

$$\underset{W, b}{\text{minimize}} \text{cost}(W, b)$$

cost를 최소화하는 W 와 b 를 구하는 것이 학습의 목표