

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

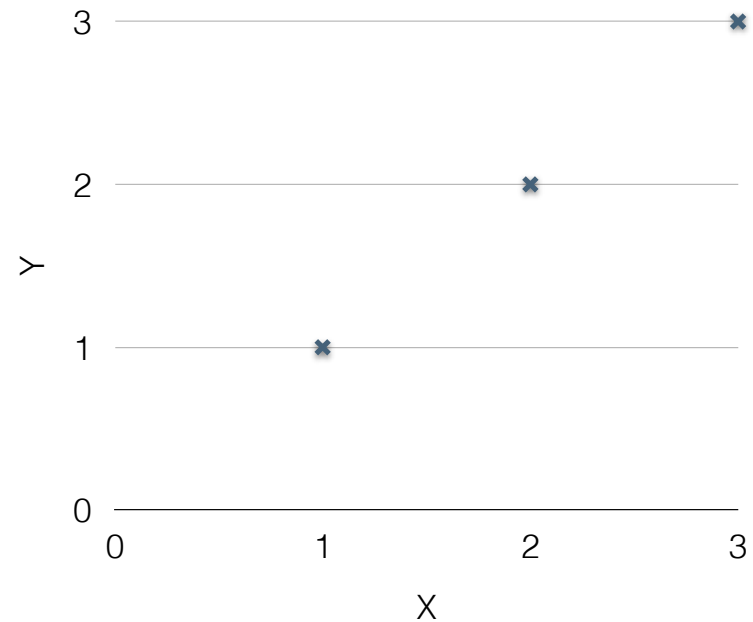
x (hours)	y (score)
10	90
9	80
3	50
2	30

Regression (data)

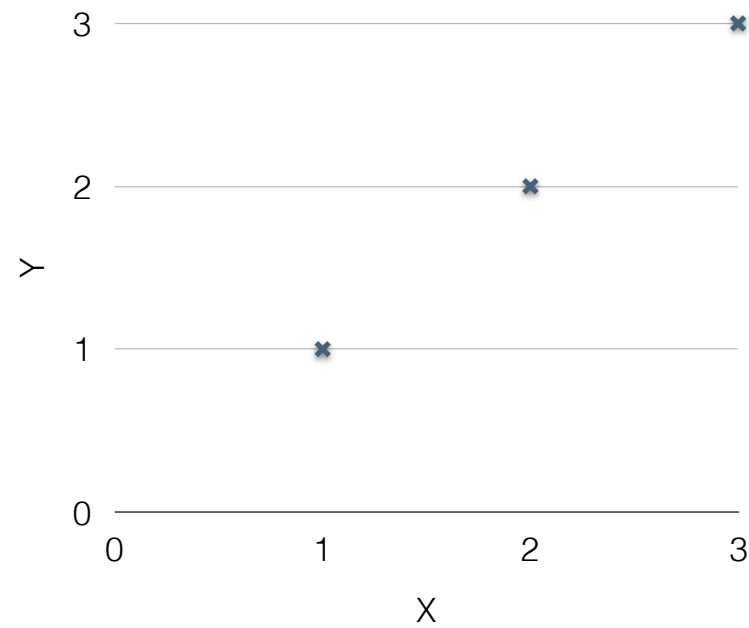
x	y
1	1
2	2
3	3

Regression (presentation)

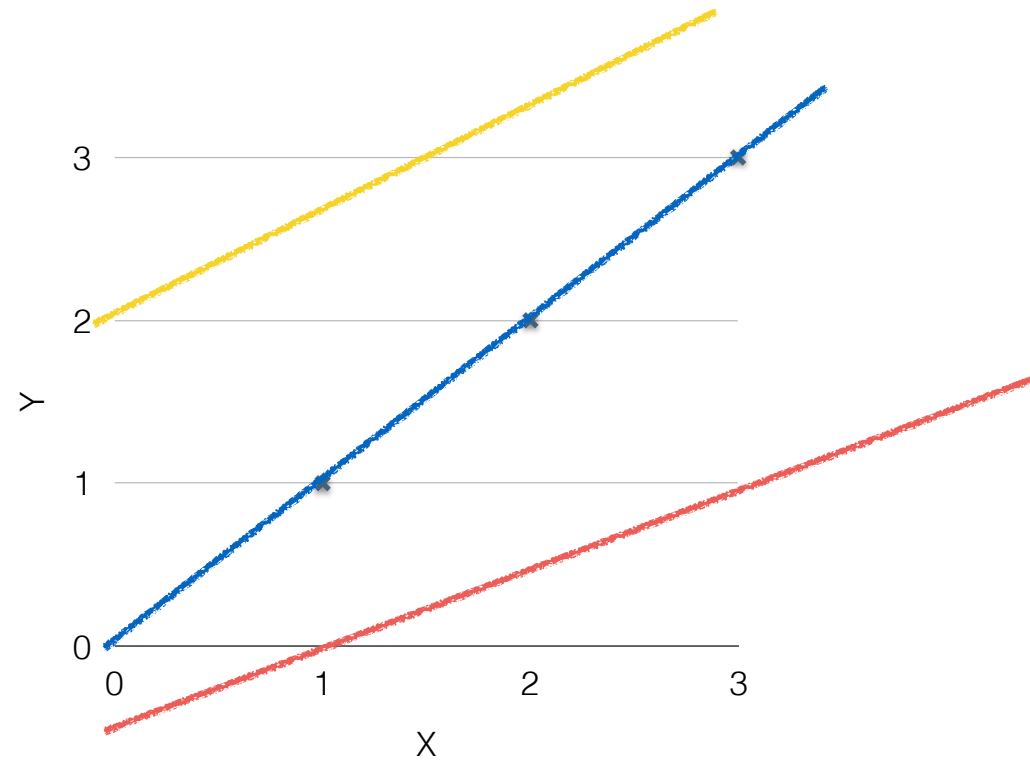
x	Y
1	1
2	2
3	3



(Linear) Hypothesis

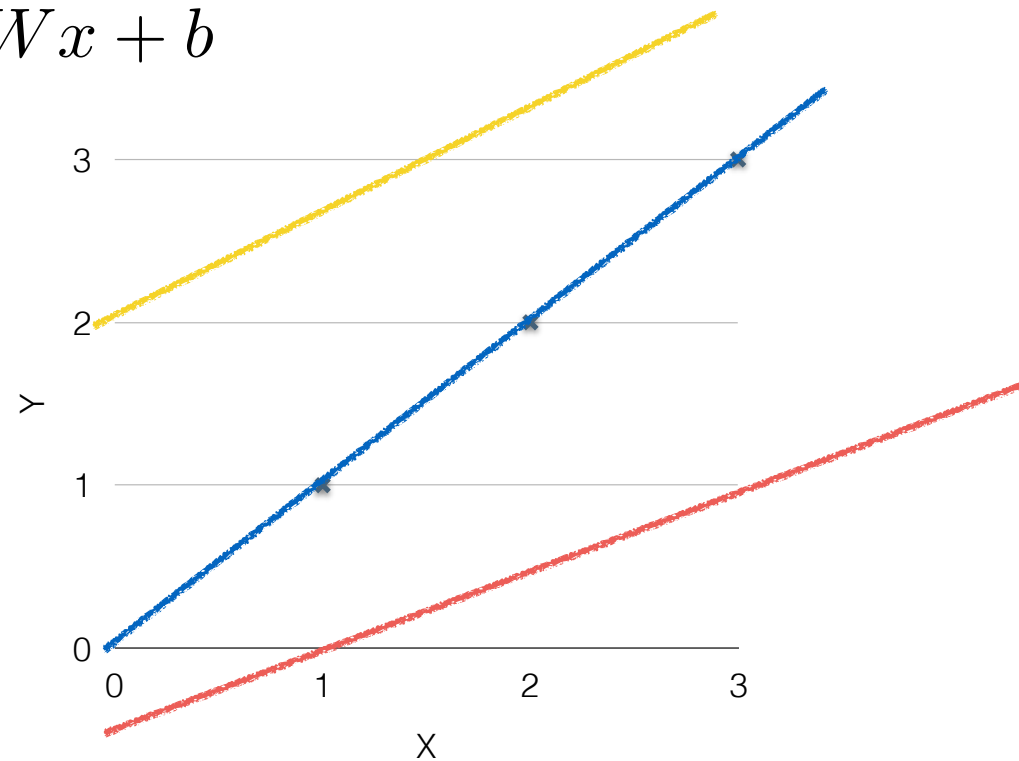


(Linear) Hypothesis

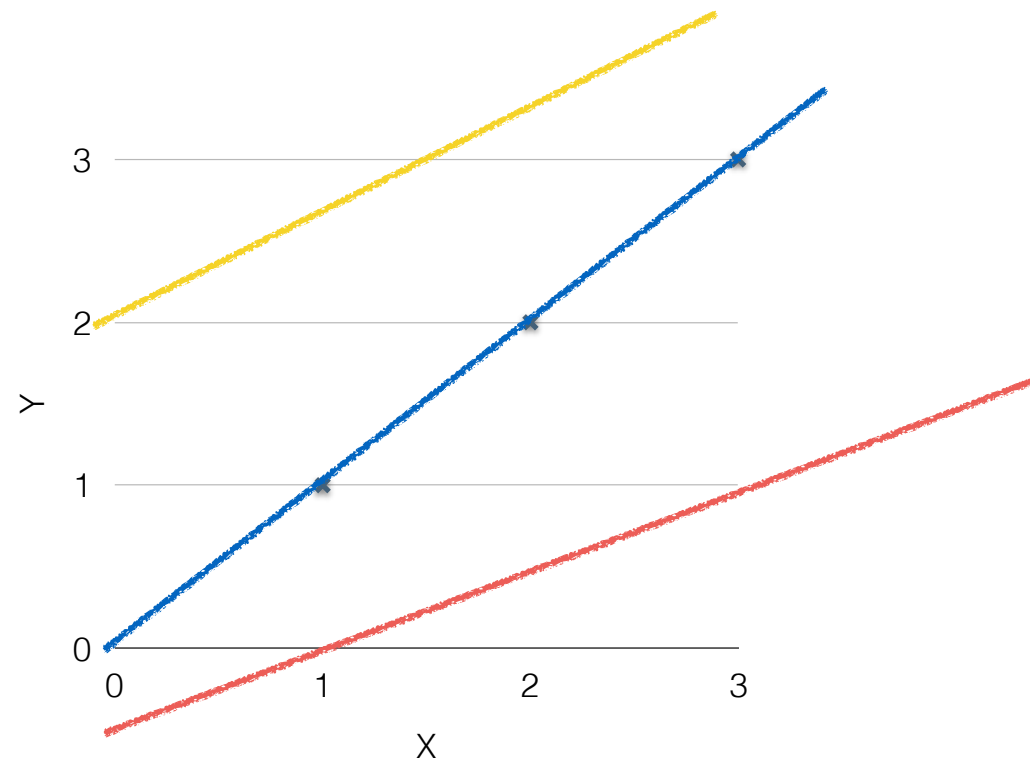


(Linear) Hypothesis

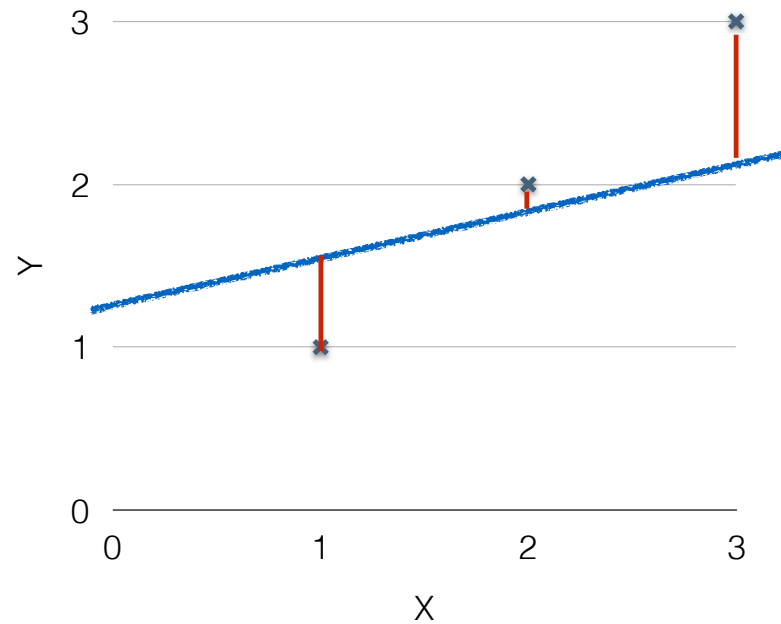
$$H(x) = Wx + b$$



Which hypothesis is better?



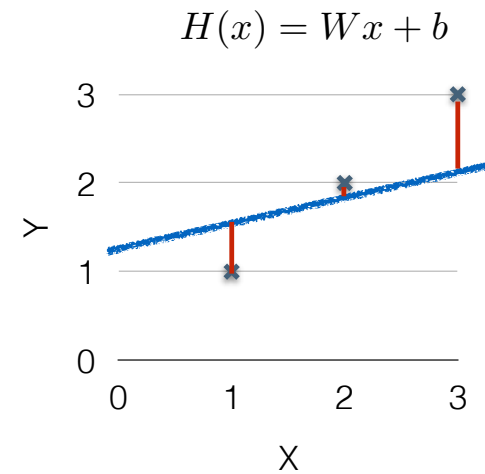
Which hypothesis is better?



Cost function

- How fit the line to our (training) data

$$H(x) - y$$

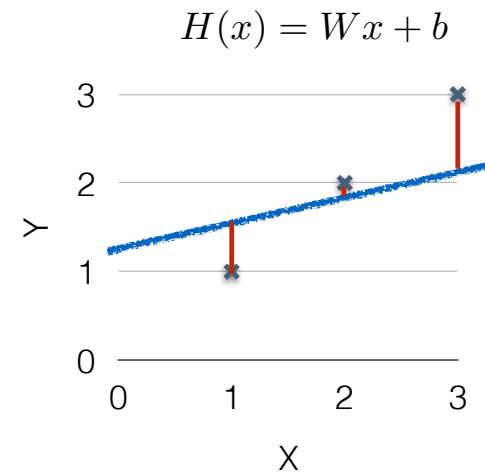


Cost function

- How fit the line to our (training) data

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

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



Cost function

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

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

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

Goal: Minimize cost

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