

딥러닝 강의 2

Linear regression

Goals

- Linear regression이란
- Cost function
- Minimize cost
- 요약

Linear regression이란?

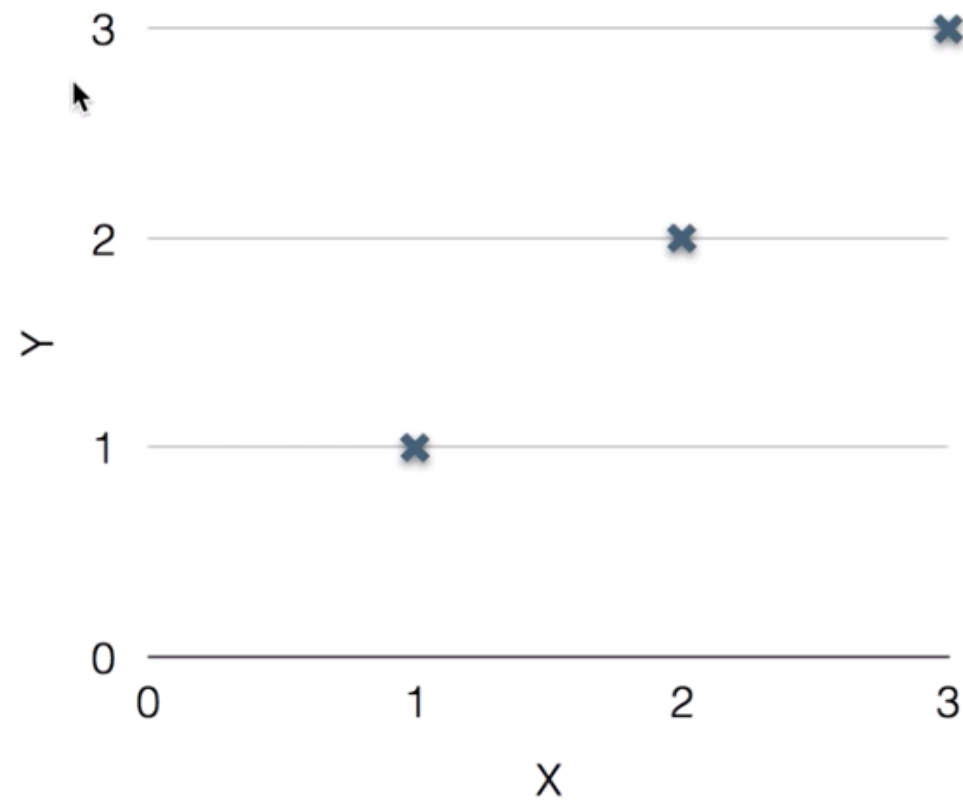
x (hours)	y (score)
10	90
9	80
3	50
2	30

$X = 15, Y = ?$

Linear regression이란?

X	Y
1	1
2	2
3	3

$X = 4, Y = ?$

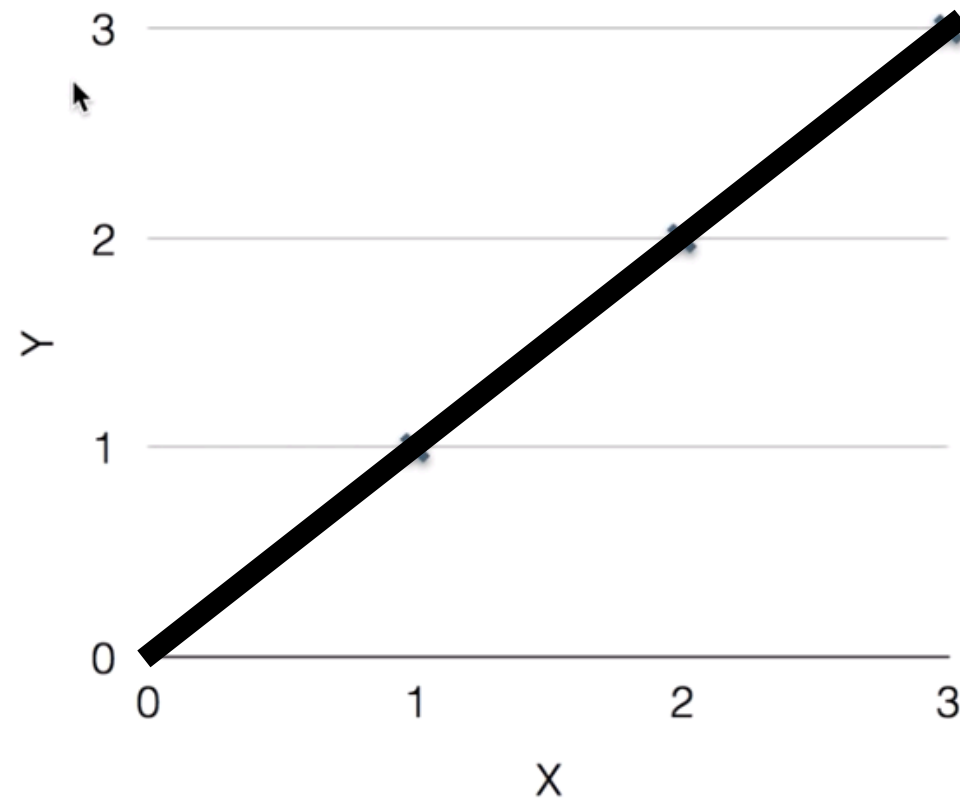


Linear regression이란?

모든 데이터를 가장 공정하게 가로지르는 선을 찾자!

X	Y
1	1
2	2
3	3

$X = 4, Y = 4$

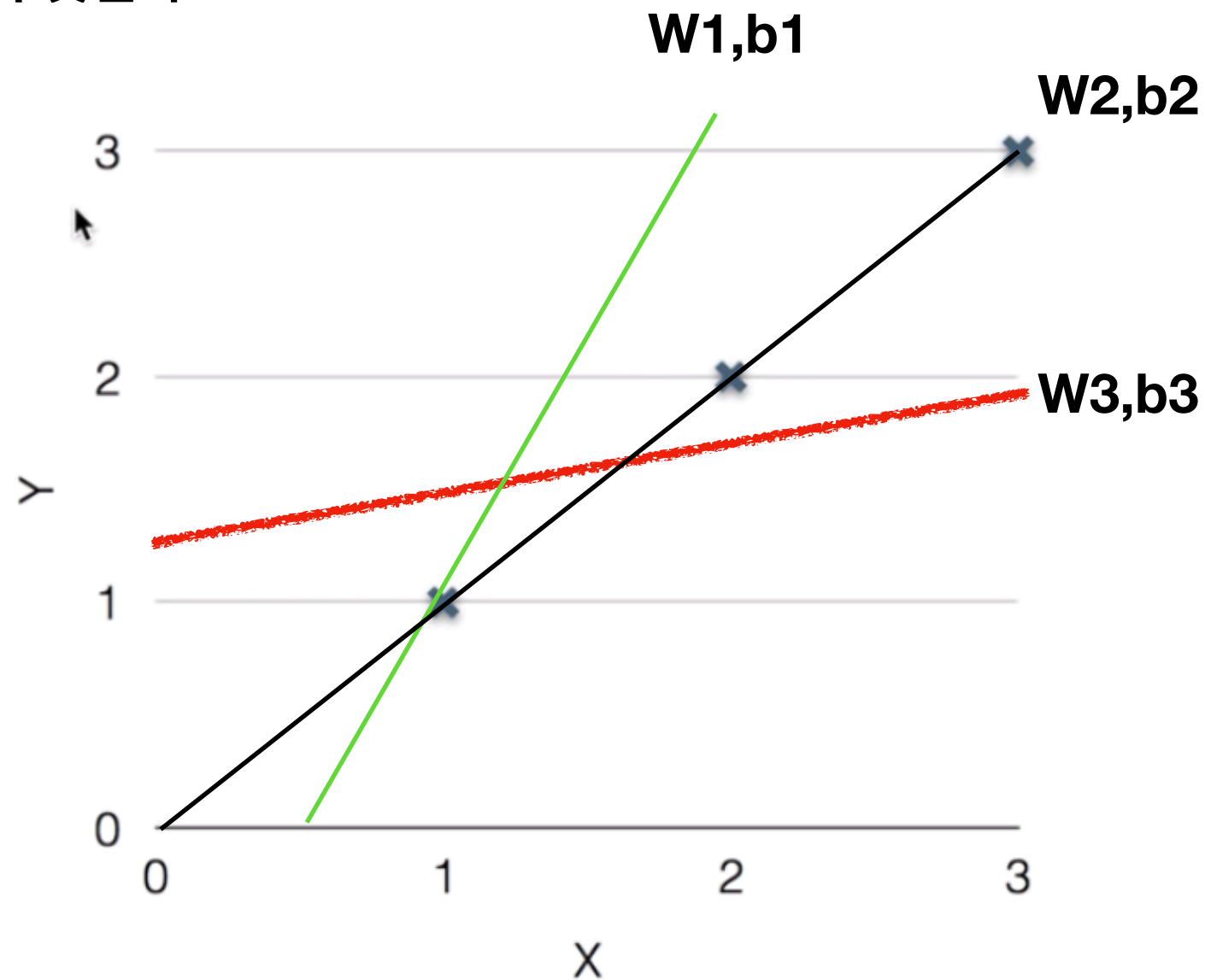


$$Y = AX + B$$

$$Y = X + 0$$

Linear regression이란?

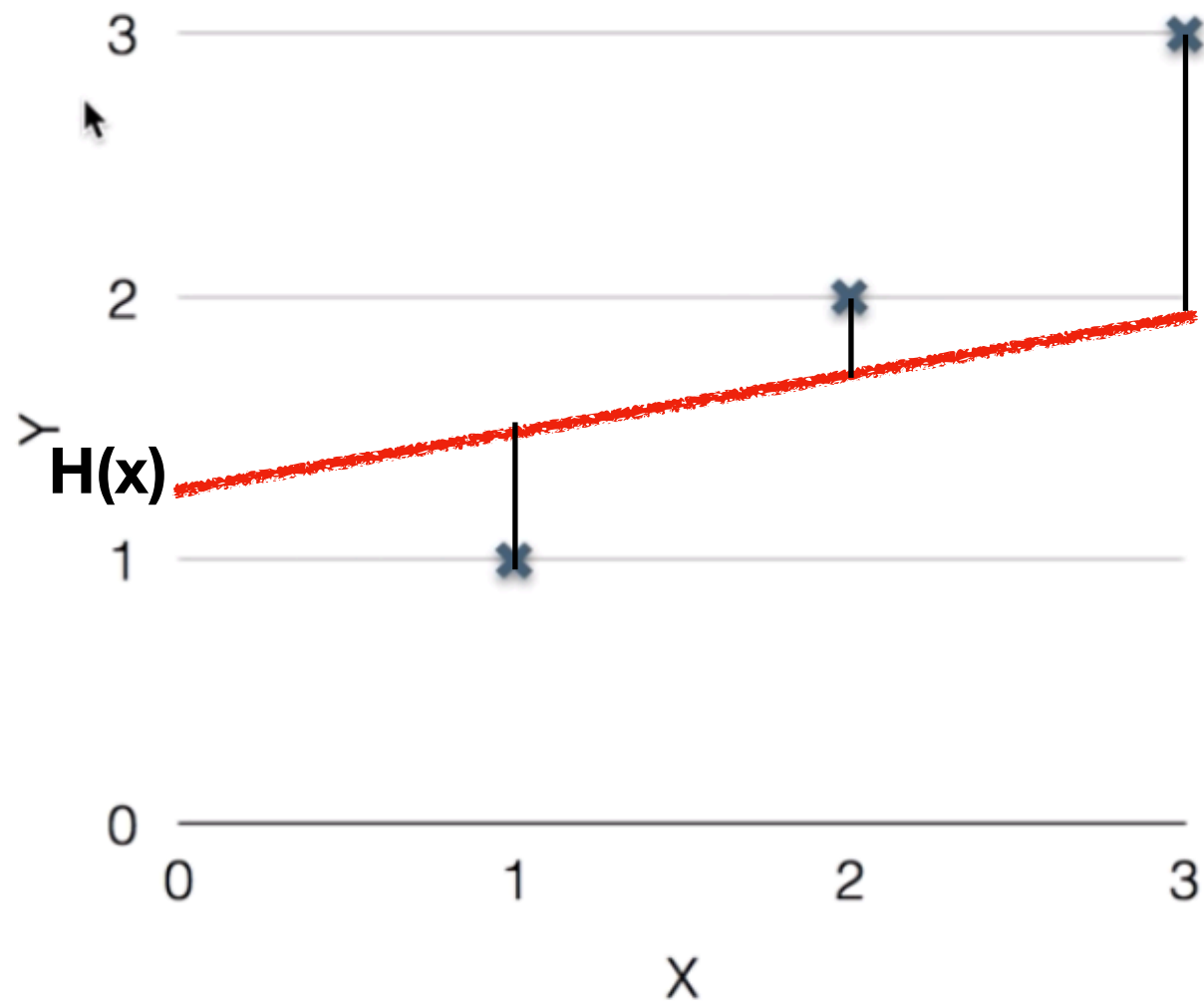
가장 좋은 선은 무엇일까?



Cost function

Cost function

비용함수 평균제곱오차

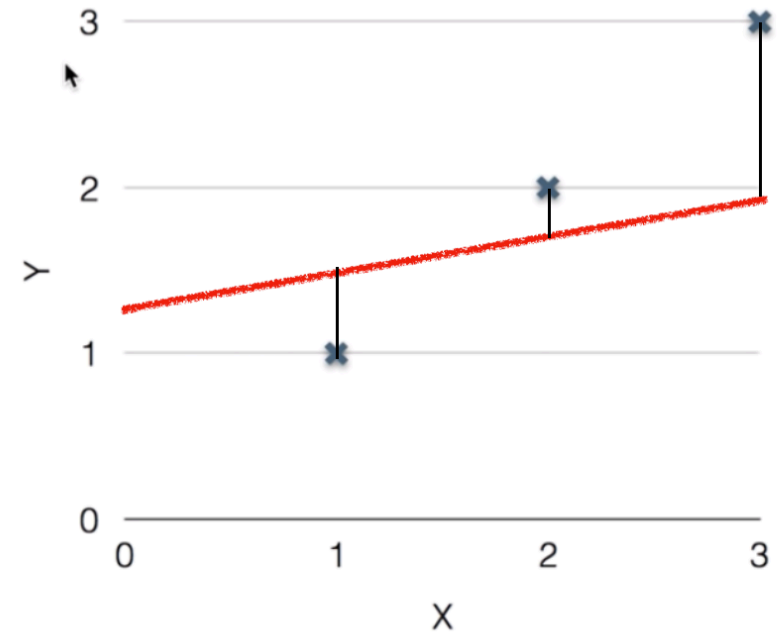


현재의 W, b 로 이루어진 함수가 실제 데이터들과 멀면 멀수록 오차가 크다.

Cost function

비용함수 평균제곱오차

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



예측 선과 실제 데이터의 차이를 제공하면 거리를 쉽게 구할 수 있다.

Q. 제공하면 데이터가 변할 수 있지 않나요?

A. 거리가 멀면 멀수록 더 데이터가 커지고

가까운 데이터는 기존보다 작게 느껴지는 것은 변수의 수정 방향을 확실하게 알 수 있게 해준다.

Cost function

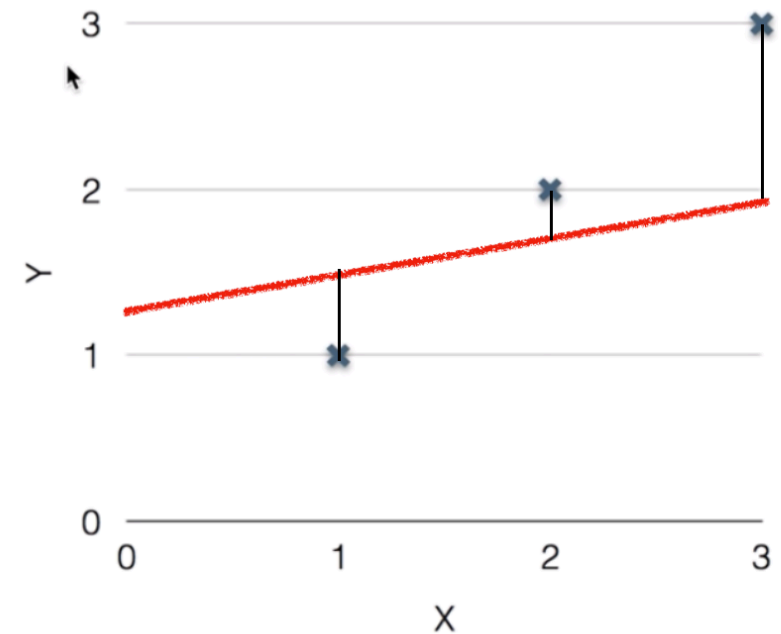
비용함수 평균제곱오차

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

모든 데이터와 선의 거리 평균

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

Formal한 식



Cost function

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

X	Y
1	1
2	2
3	3

$$W = 1$$

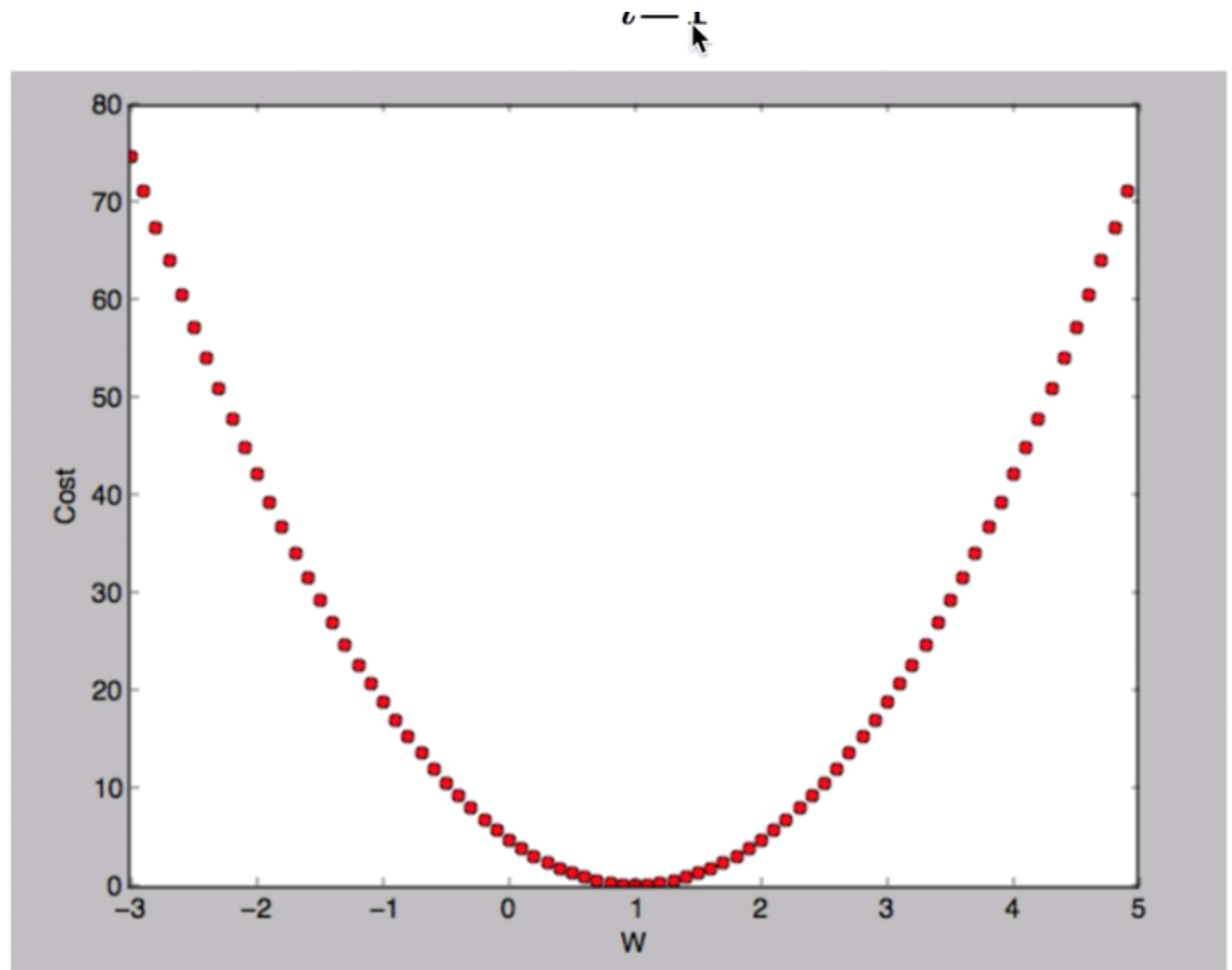
$$\begin{aligned} cost &= ((1*1-1)^2 + (1*2-2)^2 + (1*3-3)^2) / 3 \\ &= (0 + 0 + 0) / 3 \\ &= 0 \quad \text{오차가 없다.} \end{aligned}$$

$$W = 2$$

$$\begin{aligned} cost &= ((2*1-1)^2 + (2*2-2)^2 + (2*3-3)^2) / 3 \\ &= (1 + 4 + 9) / 3 \\ &= 4.67 \quad \text{4.67만큼의 오차} \end{aligned}$$

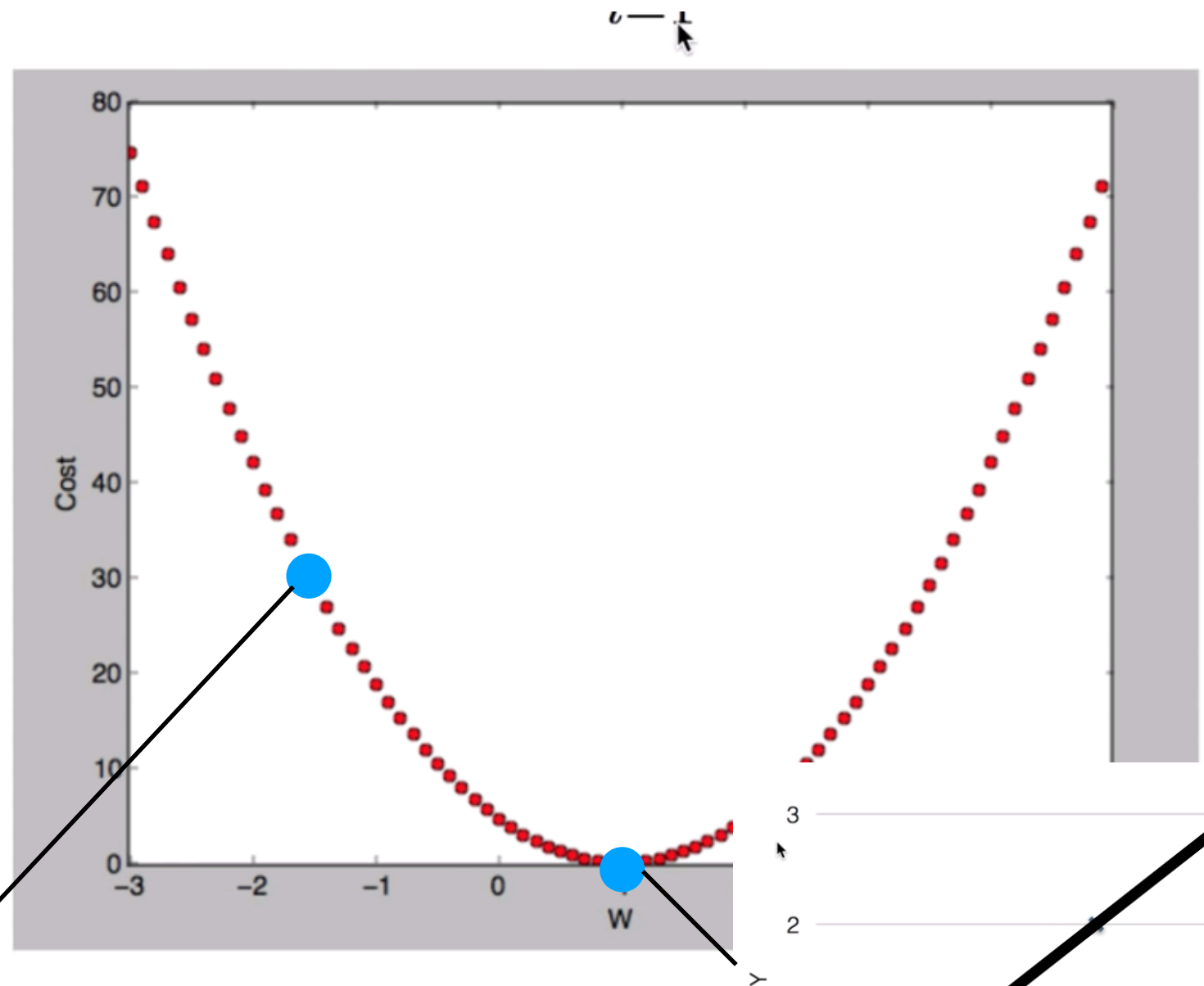
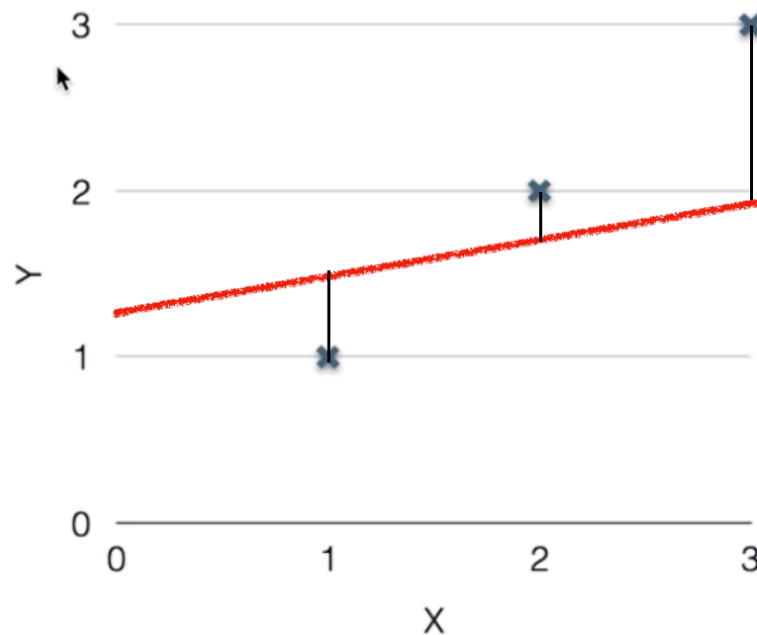
Cost function

- $W = 0$, $\text{cost}(W) = 4.67$
- $W = 1$, $\text{cost}(W) = 0$
- $W = 2$, $\text{cost}(W) = 4.67$
- $W = 3$, $\text{cost}(W) = 16.8?$

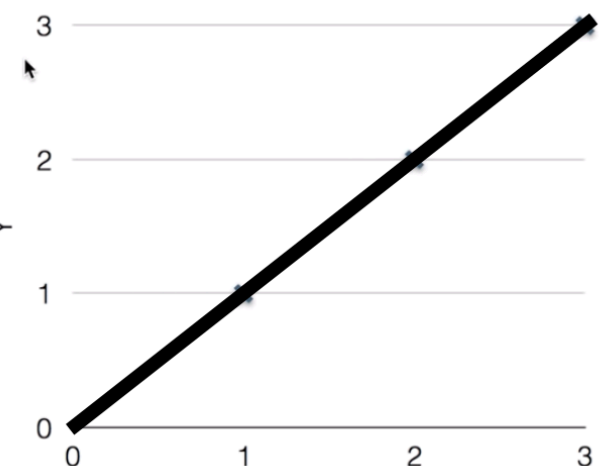


Cost function

- $W = 0$, $\text{cost}(W) = 4.67$
- $W = 1$, $\text{cost}(W) = 0$
- $W = 2$, $\text{cost}(W) = 4.67$
- $W = 3$, $\text{cost}(W) = 16.8$



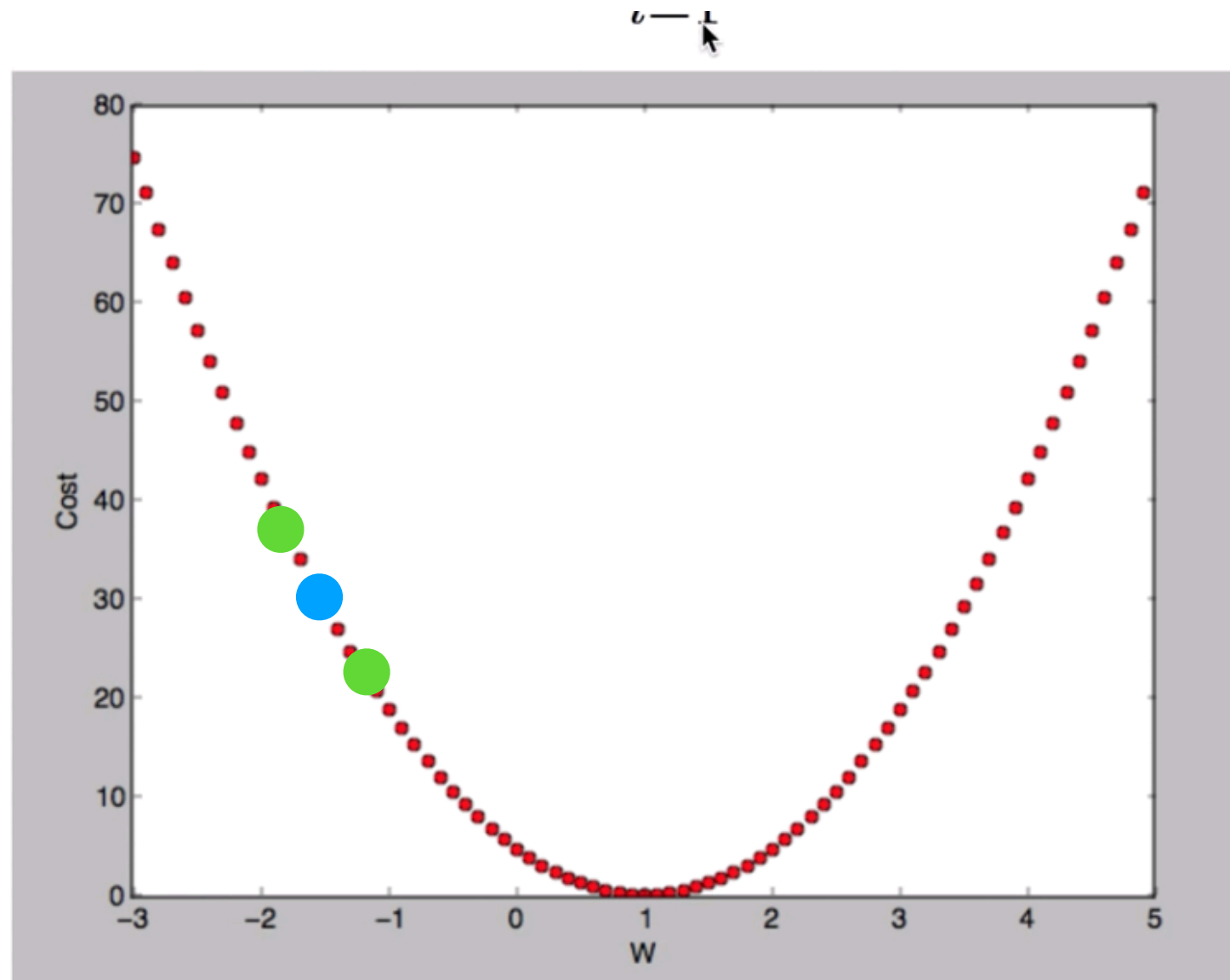
$$\text{cost} = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2$$



Minimize cost

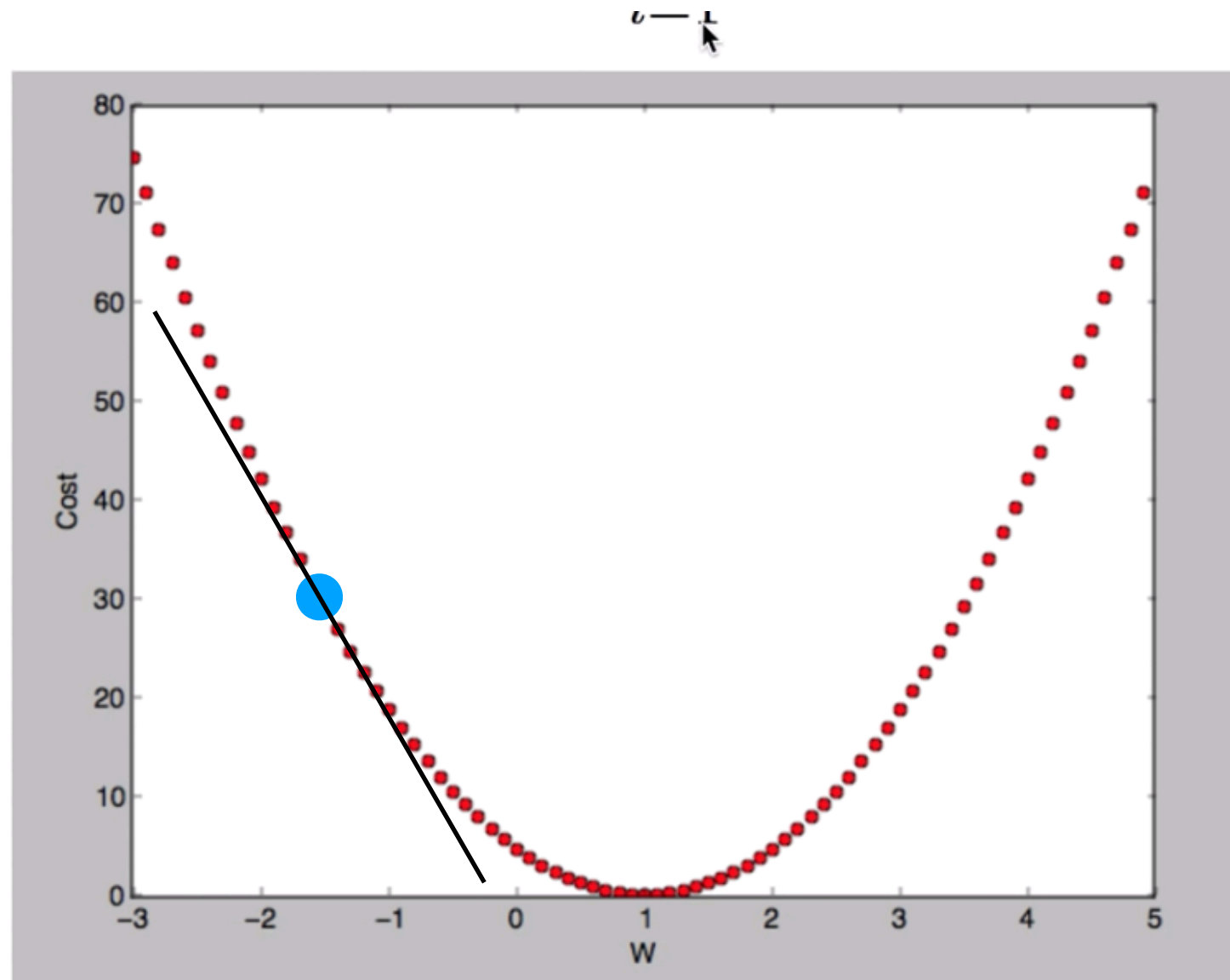
Minimize cost

어디로 가야하는가?



Minimize cost

보다 수학적이고 스마트한 방법 경사 하강법(Gradient descent algorithm)



$$\text{cost} = \frac{1}{m} \sum_{i=1}^m (H(x^{(i)}) - y^{(i)})^2 \quad \xrightarrow{\text{미분 = 기울기}} \quad W := W - \alpha \frac{1}{m} \sum_{i=1}^m (W x^{(i)} - y^{(i)}) x^{(i)}$$

Minimize cost

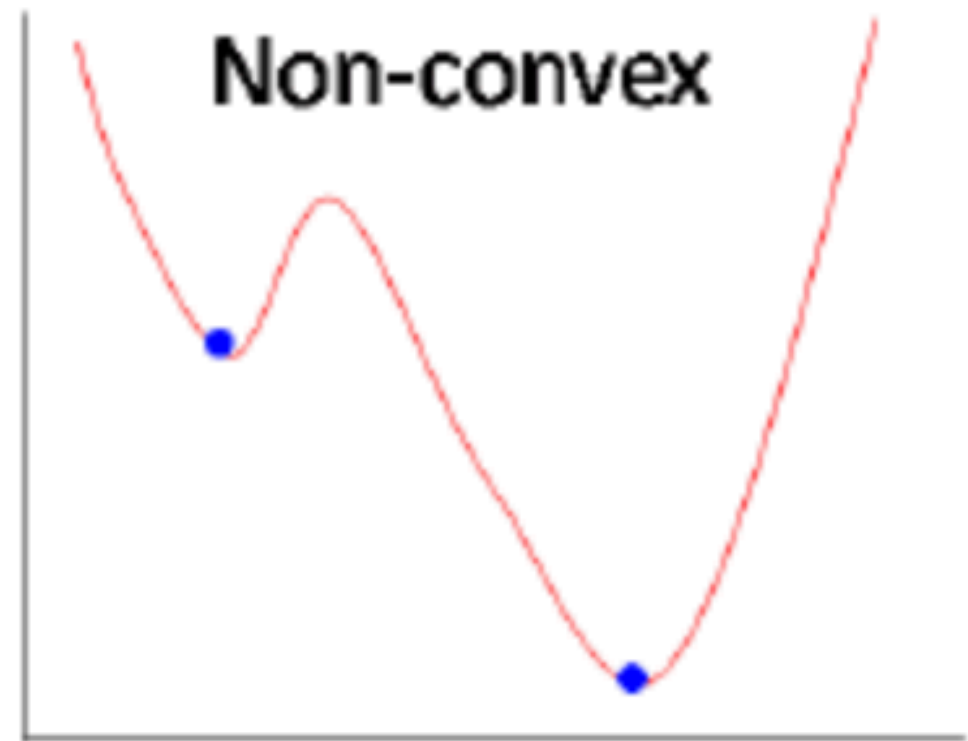
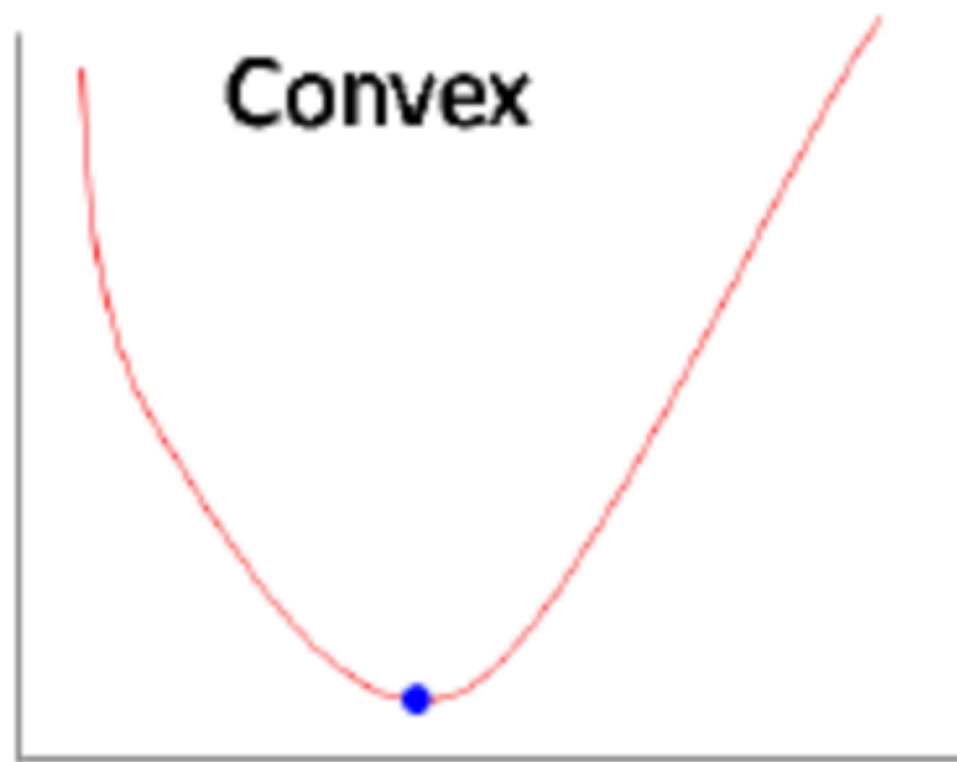
정리 How to work

- 한 지점에서 시작 W, b
- 경사가 완만해지는 방향으로 W, b 값을 변경
- 위의 항목을 경사가 제일 완만한 곳($=0$) 이 될때까지 반복

→ Next 지역 최소화에 대한 문제

Minimize cost

지역 최소화에 대한 문제



Minimize cost

지역 최소화에 대한 문제

다양한 W, b 의 초기값으로
Cost 최소화를 여러번 반복하며 최소 값을 찾는다.

많이 반복할 수록 확실한 최적값을 찾을 수 있겠지만,
시간이 오래걸린다.

= Non-convex 일 경우 Linear regression을 지양한다.

