

디지털미디어랩 머신러닝 여름캠프 3주차

(2) Linear Regression의 Hypothesis와 Cost 설명

지난 시간

- 머신러닝(ML)이란 무엇인가?
- 러닝은 무엇인가?
 - 지도 학습 (Supervised Learning)
 - 비지도 학습 (Unsupervised Learning)
- 회귀 (Regression)
- 분류 (Classification)

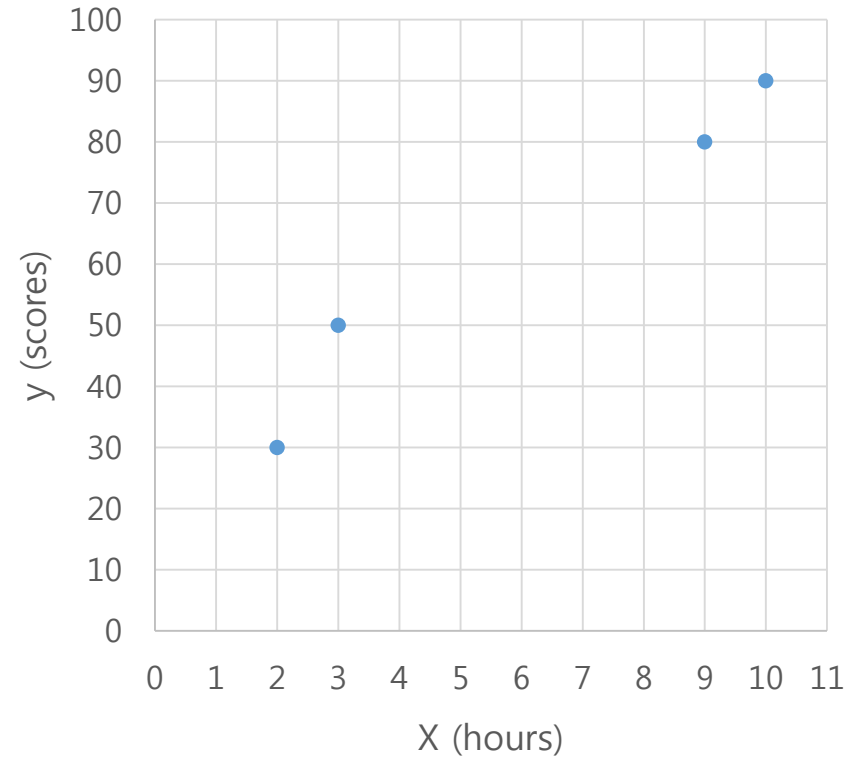
목차

- Linear Regression 모델 생성방법
- Hypothesis
- Cost Function

시험 점수 예측 (Regression)

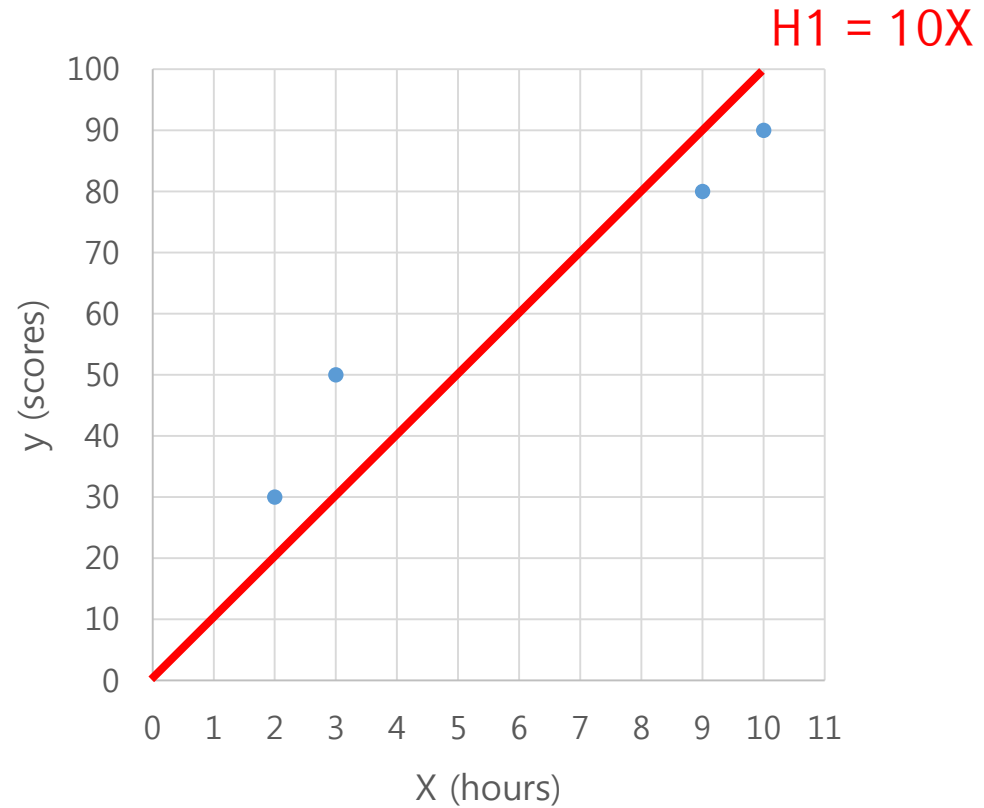
학습 데이터

X (hours)	y (scores)
2	30
3	50
9	80
10	90



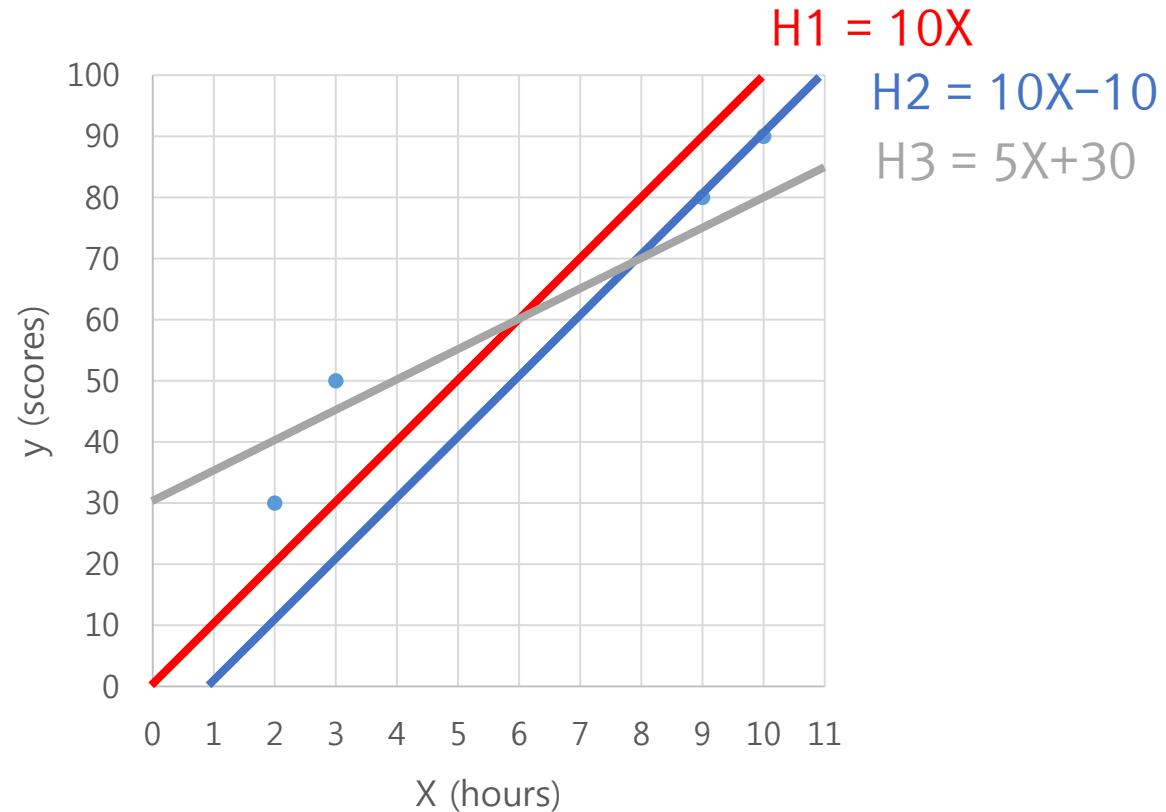
- 학습 데이터를 설명하는 회귀 모델을 만들어 보자.

Hypothesis



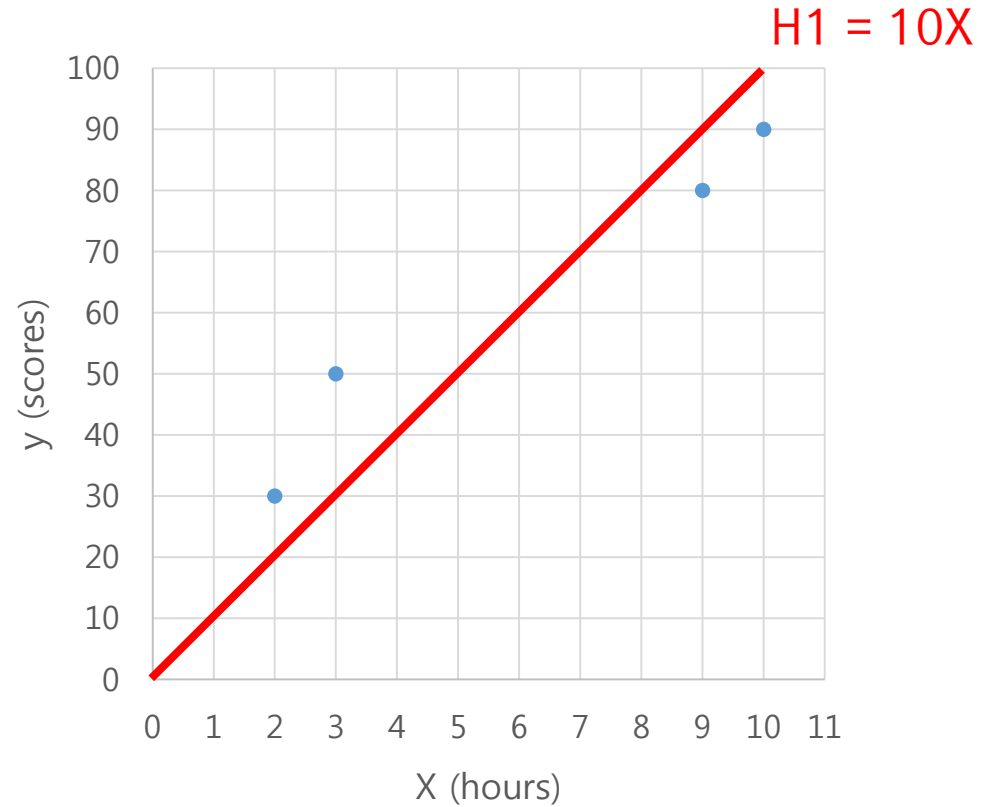
- 임의의 직선을 하나 그린다.
- 이를 가설(Hypothesis)을 하나 세운다고 표현하자.
- 그리고 이 직선을 H1이라 하자.

Hypothesis



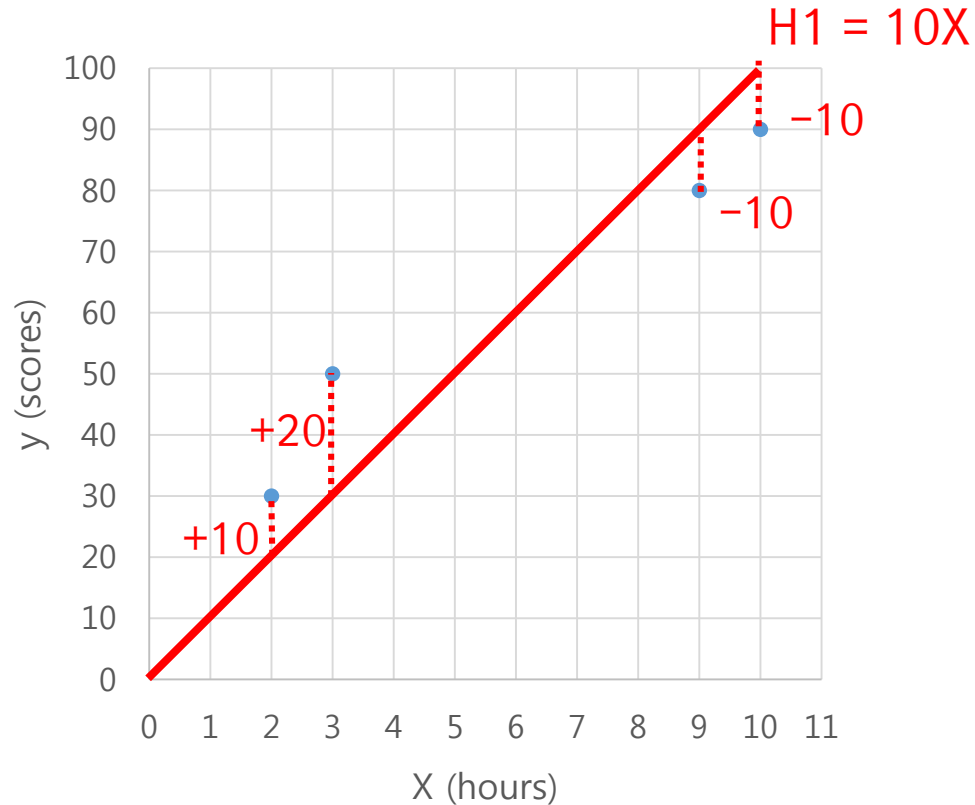
- 임의의 직선을 두 개 더 그리고, 각각 H2, H3이라 하자.
- H1, H2, H3 세 개의 가설 중 기존의 데이터(training data)를 가장 잘 설명하는 가설은 무엇일까?

Cost



- 가장 적합한 가설을 찾기 위해 내가 세운 가설이 실제 데이터와 얼마나 차이 나는지를 계산한다.
- 그리고 그 차이를 비용(Cost)라고 하자.

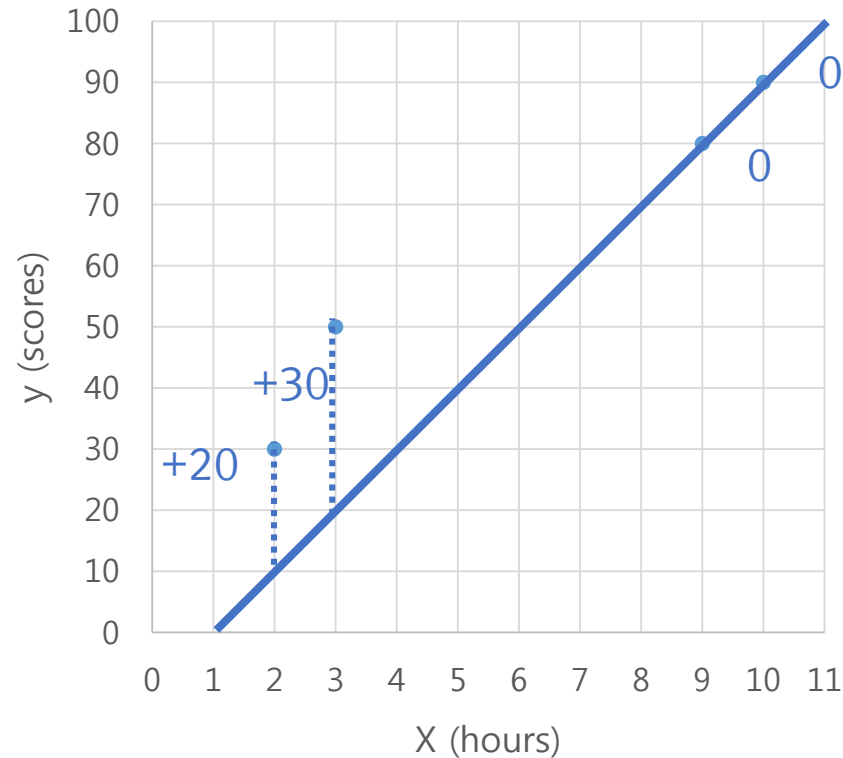
H1의 Cost



$$\begin{aligned} \text{Cost}(H1) &= (10)^2 + (20)^2 + (-10)^2 + (-10)^2 \\ &= 100 + 400 + 100 + 100 = 700 \end{aligned}$$

- 먼저 H1의 비용부터 계산한다.
- 비용을 계산하는 방법은 여러 가지가 있지만, 보통 실제 값의 차이를 제공한 총합을 이용한다.

H2의 Cost

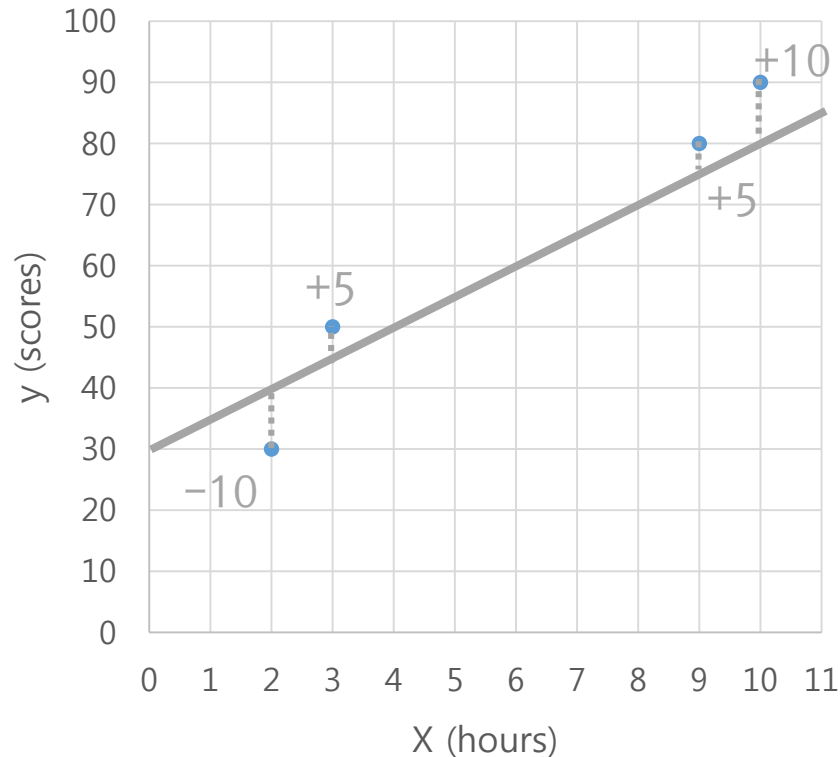


$$H2 = 10X - 10$$

$$\begin{aligned} \text{Cost}(H2) &= (20)^2 + (30)^2 + 0^2 + 0^2 \\ &= 400 + 900 = 1300 \end{aligned}$$

- H1의 비용은 700, H2의 비용은 1300 이므로 H1이 기존의 데이터를 더 잘 설명하는 가설이라고 할 수 있다.

H3의 Cost

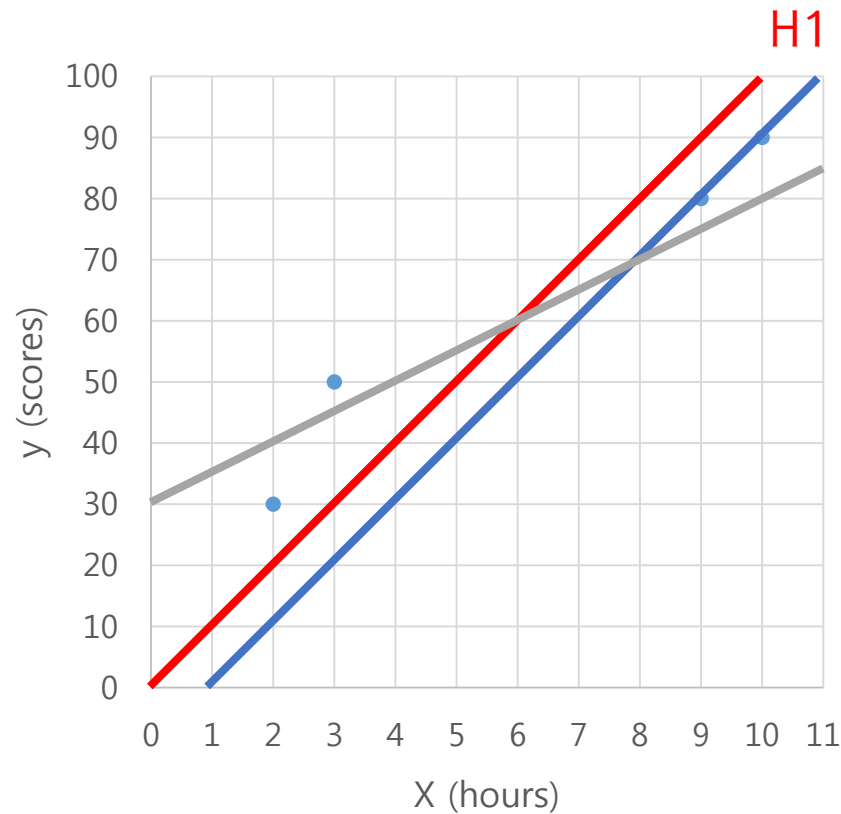


$$H3 = 5X + 30$$

$$\begin{aligned} \text{Cost}(H3) &= (-10)^2 + (5)^2 + (5)^2 + (10)^2 \\ &= 100 + 25 + 25 + 100 = 250 \end{aligned}$$

- H3의 비용은 250으로 가장 낮다. 따라서, H3가 H1, H2보다 더 좋은 가설.
- 차이에 제곱을 해주는 이유는 차이가 클 수록 비용을 크게 한다는 의미.
- 차이가 클 수록 제곱에 비례하여 비용이 커지기 때문에 패널티가 크다.

Hypothesis



$$H2 = 10X - 10$$

$$H3 = 5X + 30$$

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

- 여기서 H는 $H(x) = Wx + b$ 의 형태로 일반화 할 수 있다.
- 따라서,
- H1에서 W는 10, b는 0 이다.
- H2에서 W는 10, b는 -10 이다.
- H3에서 W는 5, b는 30이다.

Cost Function

$$Cost(W, b) = \sum (H(x^{(i)}) - y^{(i)})^2 \quad H(x) = Wx + b$$

- Cost 비용을 계산하는 함수를 Cost Function이라 하며, 위와 같이 일반화 할 수 있다.
- 앞서 계산한 H1, H2, H3의 비용은 다음과 같이 고쳐 쓸 수 있다.
- H1 의 비용 = $Cost(10, 0) = 700$
- H2 의 비용 = $Cost(10, -10) = 1300$
- H3 의 비용 = $Cost(5, 30) = 250$

회귀 모델의 목표

$$\underset{W, b}{\text{Minimize Cost}}(W, b)$$

- 비용이 가장 작은 W 와 b 를 찾는 것.
- 어떻게?
 W , b 를 막 넣어 본다. (비효율적)
 Cost가 작아지는 방향으로 W , b 를 바꾼다. (Gradient Descent 알고리즘)

참고 자료

모두를 위한 머신러닝/딥러닝

<http://hunkim.github.io/ml/>

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