

Question.3-01

Dataset이 다음과 같이 단 하나의 data sample로만 이루어졌다고 하자.

$$\mathcal{D} = \{(x^{(1)}, y^{(1)})\} = \{(1, 2)\}$$

위의 dataset은 $y = 2x$ 에서부터 만들어졌기 때문에, model을 $\hat{y} = \theta x$ 로 설정할 때 다음 문제들에 답하시오.

1) θ 가 1, 1.5, 2일 때의 square error를 이용한 loss를 각각 구하고, θ 가 2에 가까워질 때 loss의 변화를 비교하시오.

(square error는 y 와 \hat{y} 를 이용하여 $(y - \hat{y})^2$ 로 구한다.)

2) 임의의 θ 에 대한 loss를 algebraic equation으로 표현하고 그래프를 그리시오.

1) 먼저 θ 가 1, 1.5, 2일 때 predictor는 다음과 같다.

$$\begin{array}{ccc} \hat{y} = x & \hat{y} = 1.5x & \hat{y} = 2x \\ \text{그리고 } x^{(1)} = 1 \text{에 대해 각각 predictor의 prediction을 다음과 같다.} & & \\ \hat{y}|_{x=1} = 1 & \hat{y}|_{x=1} = 1.5 & \hat{y}|_{x=1} = 2 \end{array}$$

따라서 각각의 loss는 L_1, L_2, L_3 가 하며

$$L_1 = (2 - 1)^2 = 1 \quad L_2 = (2 - 1.5)^2 = 0.25 \quad L_3 = (2 - 2)^2 = 0$$

이 된다.

즉, predictor가 실제 dataset의 만들어진 $y = 2x$ 에 가까워질수록 loss는 감소한다.

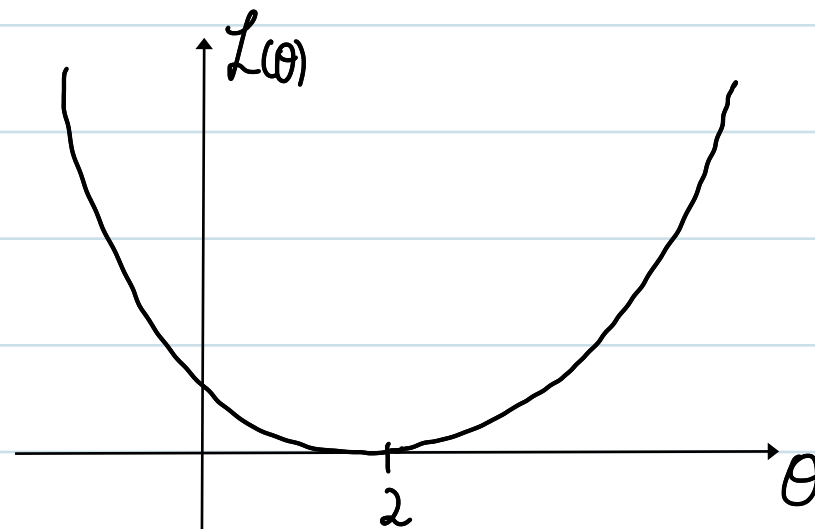
2) 임의의 θ 에 대해 predictor는 $\hat{y} = \theta x$ 가 되고 loss L 은

$$\begin{aligned} L &= (y - \hat{y})^2 \\ &= (y - \theta x)^2 \end{aligned}$$

이 되고, data sample (1, 2)에 대해

$$L(\theta) = (2 - \theta)^2$$

이 되고 $L(\theta)$ 의 graph는



위와 같이 $\theta = 2$ 에서 최솟값을 가지는 2차함수가 된다.