

Question. 4-08

target function $y = 3x + 2$ 의 이상적인 Dataset이 다음과 같다.

$$\mathcal{D} = \{(1, 5), (2, 8), (5, 17)\}$$

그러나 noise에 의해 Dataset이 다음과 같이 왜곡된 경우를 가정하자.

$$\mathcal{D} = \{(1, 6), (2, 5), (5, 15)\}$$

θ_1, θ_0 이 (3,2)로 학습이 끝난 상태에서 각 data sample에 의해 θ_1, θ_0 가 update되는 값들을 구하고, bias가 있는 경우에서 noise가 학습에 미치는 영향을 분석하시오.

1) θ_1 과 θ_0 가 target에 도달한 상태로 이상적인 데이터에선 $J_{loss} = 0$ 이 되어 더 이상의 학습이 진행되지 않는다.

그러나 noise에 의해 θ 가 target에 도달하였음에도 J_{loss} 가 발생하게 된다.

이를 확인하기 위해 각 data sample을 이용해 θ_1, θ_0 를 어떻게 학습시키는지 확인한다.

아래의 연산에서 왜곡된 data point의 영향을 보기 위해 실제 θ 는 update하지 않는다.

$$\text{data point } (1, 6) \quad J = (y - \theta_1 x - \theta_0)^2 = (6 - 3 - 2)^2 = 1 \quad \Delta \theta_1 = 2x \times (y - \theta_1 x - \theta_0) = 2x$$

$$\Delta \theta_0 = 2 \times (y - \theta_1 x - \theta_0) = 2x$$

$$\text{data point } (2, 5) \quad J = (y - \theta_1 x - \theta_0)^2 = (5 - 8)^2 = 9 \quad \Delta \theta_1 = 2x \times (y - \theta_1 x - \theta_0) = -12x$$

$$\Delta \theta_0 = 2 \times (y - \theta_1 x - \theta_0) = -6x$$

$$\text{data point } (5, 17) \quad J = (y - \theta_1 x - \theta_0)^2 = (15 - 17)^2 = 4 \quad \Delta \theta_1 = 2x \times (y - \theta_1 x - \theta_0) = -20x$$

$$\Delta \theta_0 = 2 \times (y - \theta_1 x - \theta_0) = -4x$$

위에서 알 수 있듯 noise가 섞인 data point는 J_{loss} 를 발생시켜 target θ 에서 멀어지게 한다.

이는 θ_1 과 θ_0 모두에게 영향을 끼치며 $|x| > 1$ 의 경우 θ_1 이 noise에 더 민감하게 반응하고

$0 < |x| < 1$ 의 경우 θ_0 가 noise에 더 민감하게 반응한다.