

Question. 4-05

$0 \leq |x| < 1$ 인 Dataset \mathcal{D} 가 다음과 같이 주어졌다.

$$\mathcal{D} = \{(0, 5), (0.1, 5.2), (0.3, 5.6), (-0.1, 4.8), (-0.3, 4.4)\}$$

Dataset을 $y = 2x + 5$ 에서부터 만들었기 때문에, 모델을 $\hat{y} = \theta_1 x + \theta_0$ 로 설정하였다.

initial $\vec{\theta} = (\theta_1, \theta_0) = (1, 1)$ 이고, learning rate $\alpha = 0.1$ 로 주어졌을 때 다음 질문에 답하시오.

1) 각 Data sample에 의한 θ_1, θ_0 의 Update 양을 구하시오.

2) 1)에서의 결과를 통해 $x^{(i)}$ 가 γ 배 되었을 때, θ_1, θ_0 가 Update 되는 양의 변화를 구하시오.

1) $(x, y) = (0, 5)$ 일 때 $\Delta\theta_1 = 2\alpha x(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot 0 \cdot (5 - 0 - 1) = 0$
 $\Delta\theta_0 = 2\alpha(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot (5 - 1) = 0.8$

즉, $x=0$ 일 때 θ_1 은 update 되지 않지만 θ_0 은 update된다.

$(x, y) = (0.1, 5.2)$ 일 때 $\Delta\theta_1 = 2\alpha x(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot 0.1 \cdot (5.2 - 0.1 - 1) = 0.082$
 $\Delta\theta_0 = 2\alpha(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot (5.2 - 0.1 - 1) = 0.82$

즉, $x=0.1$ 일 때 θ_0 의 update 양은 θ_1 의 update 양의 10배이다.

$(x, y) = (0.3, 5.6)$ 일 때 $\Delta\theta_1 = 2\alpha x(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot 0.3 \cdot (5.6 - 0.3 - 1) = 0.258$
 $\Delta\theta_0 = 2\alpha(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot (5.6 - 0.3 - 1) = 0.86$

즉, $x=0.3$ 일 때 θ_0 의 update 양은 θ_1 의 update 양의 $\frac{10}{3}$ 배이다.

$(x, y) = (-0.1, 4.8)$ 일 때 $\Delta\theta_1 = 2\alpha x(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot (-0.1) \cdot (4.8 + 0.1 - 1) = -0.098$
 $\Delta\theta_0 = 2\alpha(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot (4.8 + 0.1 - 1) = 0.78$

즉, $x=-0.1$ 일 때 θ_1 의 update 양은 θ_0 의 update 양의 -10 배이다.

$(x, y) = (-0.3, 4.4)$ 일 때 $\Delta\theta_1 = 2\alpha x(y - \theta_1 x - \theta_0) = 2 \cdot 0.1 \cdot (-0.3) \cdot (4.4 + 0.3 - 1) = -0.222$
 $\Delta\theta_0 = 2\alpha(y - \theta_1 x - \theta_0) = 0.74$

즉, $x=-0.3$ 일 때 θ_1 의 update 양은 θ_0 의 update 양의 $-\frac{10}{3}$ 배이다.

2) 1)의 결과를 통해 $0 \leq |x| < 1$ 인 경우에

θ_1 의 업데이트 양은 $2\alpha x(y - \theta_1 x - \theta_0)$ 이고 θ_0 의 업데이트 양은 $2\alpha(y - \theta_1 x - \theta_0)$

이므로 x 가 γ 배 되었을 때 θ_1 은 $2\alpha \gamma x(y - \theta_1 x - \theta_0)$ 만큼 변화하고

θ_0 은 $2\alpha(y - \theta_1 x - \theta_0)$ 만큼 변화한다.

즉, x 가 γ 배가 되면 θ_1 과 θ_0 의 업데이트 양의 차이도 γ 배가 된다.