

Question.3-07

다음과 같이 Dataset이 주어졌다.

$$\mathcal{D} = \{(1,2), (3,6), (4,8)\}$$

이때 다음 질문에 답하시오.

단 Question.3-04, Question.3-05와 마찬가지로 prediction model은 $\hat{y} = \theta x$ 를 사용하고, initial θ 는 1로 설정한다.

1) 각 3개의 data point들의 loss를 이용하여 cost J를 구하고, $\frac{\partial J(\theta)}{\partial \theta}$ 를 구하시오.

2) $\frac{\partial J(\theta)}{\partial \theta}$ 와 gradient descent method를 이용한 θ 의 update equation을 구하고,

$\alpha = 0.1$ 일때 4 iterations에 대한 θ 의 변화를 구하시오.

1) 주어진 data point들에 대한 loss $\mathcal{L}^{(1)}(\theta)$, $\mathcal{L}^{(2)}(\theta)$, $\mathcal{L}^{(3)}(\theta)$ 를 구하면 다음과 같다.

$$\mathcal{L}^{(1)}(\theta) = (y^{(1)} - \theta \cdot x^{(1)})^2 = (2 - \theta)^2$$

$$\mathcal{L}^{(2)}(\theta) = (y^{(2)} - \theta \cdot x^{(2)})^2 = (6 - 3\theta)^2$$

$$\mathcal{L}^{(3)}(\theta) = (y^{(3)} - \theta \cdot x^{(3)})^2 = (8 - 4\theta)^2$$

222고 이를 통해 cost를 구하면

$$\begin{aligned} J(\theta) &= \frac{1}{3} \sum_{i=1}^3 \mathcal{L}^{(i)}(\theta) = \frac{1}{3} [(2-\theta)^2 + (6-3\theta)^2 + (8-4\theta)^2] \\ &= \frac{1}{3} [(2-\theta)^2 + 9(2-\theta)^2 + 16(2-\theta)^2] \\ &= \frac{26}{3} (2-\theta)^2 \end{aligned}$$

이 222 $\frac{\partial J(\theta)}{\partial \theta}$ 를 구해보면

$$\begin{aligned} \frac{\partial J(\theta)}{\partial \theta} &= \frac{\partial}{\partial \theta} \left[\frac{26}{3} (2-\theta)^2 \right] = \frac{52}{3} (2-\theta) \cdot \frac{\partial}{\partial \theta} (2-\theta) \\ &= -\frac{52}{3} (2-\theta) \end{aligned}$$

가 된다.

2) 1)에서 구한 $J(\theta)$ 와 gradient descent method를 이용하여 θ 를 학습시키는 parameter update equation을 구하면

$$\theta := \theta - 0.1 \frac{\partial J(\theta)}{\partial \theta} = \theta + 1.73(2-\theta)$$

가 된다. 이 식을 이용하여 3번 θ 를 update하면 다음과 같다.

$$1^{\text{st}} \text{ iteration: } \theta := 1 + 1.73(2-1) = 2.73$$

$$2^{\text{nd}} \text{ iteration: } \theta := 2.73 + 1.73(2-2.73) = 1.47$$

$$3^{\text{rd}} \text{ iteration: } \theta := 1.47 + 1.73(2-1.47) = 2.39$$

$$4^{\text{th}} \text{ iteration: } \theta := 2.39 + 1.73(2-2.39) = 1.72$$

위의 결과에서 볼수있듯이 θ 가 θ^* 인 2로 수렴하고 있는 것을 알 수 있다.