

Question. 5-01

Dataset \mathcal{D} 가 다음과 같이 주어졌다.

$$\mathcal{D} = \{(x_2^{(1)}, x_1^{(1)}, y^{(1)}), (x_2^{(2)}, x_1^{(2)}, y^{(2)}), \dots, (x_2^{(n)}, x_1^{(n)}, y^{(n)})\}$$

Dataset을 $y = ax_2 + bx_1 + c$ 에서부터 만들었기 때문에, Linear regression을 위한

Prediction 모델을 $\hat{y} = \theta_2 x_2 + \theta_1 x_1 + \theta_0$ 로, Loss를 Square Error로 설정하였다.

이 때, 다음 질문에 답하시오.

1) $\hat{y} = \theta_2 x_2 + \theta_1 x_1 + \theta_0$ 에 대한 Loss를 구하고 $\theta_2, \theta_1, \theta_0$ 의 Update Equation을 구하시오.

2) 1)을 토대로 weight(θ_2, θ_1)와 bias(θ_0)의 Update 차이를 설명하시오.

1) Square Error를 이용하여 Loss를 구한다.

$$\hat{y} = \theta_2 x_2 + \theta_1 x_1 + \theta_0 \text{ 이므로 } \mathcal{L} = (y - \hat{y})^2 = (y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))^2$$

이때, Loss의 $\theta_2, \theta_1, \theta_0$ 에 대한 gradient를 구하면 다음과 같다.

$$\frac{\partial \mathcal{L}}{\partial \theta_2} = \frac{\partial}{\partial \theta_2} [(y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))^2] = 2(y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0)) \frac{\partial}{\partial \theta_2} [y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0)] = -2x_2 (y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))$$

$$\frac{\partial \mathcal{L}}{\partial \theta_1} = \frac{\partial}{\partial \theta_1} [(y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))^2] = 2(y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0)) \frac{\partial}{\partial \theta_1} [y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0)] = -2x_1 (y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))$$

$$\frac{\partial \mathcal{L}}{\partial \theta_0} = \frac{\partial}{\partial \theta_0} [(y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))^2] = 2(y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0)) \frac{\partial}{\partial \theta_0} [y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0)] = -2(y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))$$

gradient descent method를 이용하여 $\theta_2, \theta_1, \theta_0$ 의 update equation을 구하면 다음과 같다.

$$\theta_2 := \theta_2 - \alpha \frac{\partial \mathcal{L}}{\partial \theta_2} = \theta_2 + 2\alpha x_2 (y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))$$

$$\theta_1 := \theta_1 - \alpha \frac{\partial \mathcal{L}}{\partial \theta_1} = \theta_1 + 2\alpha x_1 (y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))$$

$$\theta_0 := \theta_0 - \alpha \frac{\partial \mathcal{L}}{\partial \theta_0} = \theta_0 + 2\alpha (y - (\theta_2 x_2 + \theta_1 x_1 + \theta_0))$$

2) Weight와 bias update equation을 보면 $2\alpha(y - \hat{y})$ 를 공통하게 가져 learning rate와 error rate에 비례하여 값이 이루어짐을 알 수 있다.

그러나 weight의 경우 θ_2, θ_1 의 update equation에 각각 x_2, x_1 이 곱하여 input x_2, x_1 에 선형하게 반응하고

bias의 경우 weight에 비해 상대적으로 input에 선형하게 반응하여 update를 할 수 있다.