

Chapter. 5

Multi-variate Linear Regression

| Theory Assignment

FAST CAMPUS
ONLINE

강사. 신경식

Chapter. 5

Multi-variate Linear Regression

Question. 5-01

Dataset \mathcal{D} 가 다음과 같이 주어졌다.

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

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 $\text{weight}(\theta_2, \theta_1)$ 와 $\text{bias}(\theta_0)$ 의 Update 차이를 설명하시오.

Chapter. 5

Loss Functions

Question. 5-02

Linear regression을 위한 Dataset이 다음과 같이 주어졌다.

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

이때, Dataset $y = ax_2 + bx_1 + c$ 에서부터 만들어졌다. 따라서 Linear regression을 위한 Prediction 모델은 $\hat{y} = \theta_2 x_2 + \theta_1 x_1 + \theta_0$, Loss는 Square Error를 사용할 수 있다.

$\theta_2, \theta_1, \theta_0$ 를 Update하기 위해 n개의 data sample을 이용할 때,

1번의 iteration동안 $\theta_2, \theta_1, \theta_0$ 가 dataset을 잘 표현하는 $\theta_2, \theta_1, \theta_0$ 로 Update되는 과정을

Vector notation을 이용하여 설명하시오.

단, forward/backward propagation을 설명하기 위해 각 연산은 basic building node를 이용하시오.

Question. 5-03

Dataset \mathcal{D} 가 다음과 같이 주어졌다.

$$\mathcal{D} = \left\{ \left(x_2^{(1)}, x_1^{(1)}, y^{(1)} \right), \left(x_2^{(2)}, x_1^{(2)}, y^{(2)} \right), \left(x_2^{(3)}, x_1^{(3)}, y^{(3)} \right) \right\} = \{(1, 1, 4), (3, 2, 5), (6, 4, 5)\}$$

Dataset을 $y = 2x_2 - 3x_1 + 5$ 에서부터 만들었기 때문에, 모델을 $\hat{y} = \theta_2 x_2 + \theta_1 x_1 + \theta_0$ 로 설정하였다.

initial $\vec{\theta}$ 가 $\theta_2 = \theta_1 = \theta_0 = 1$ 이고 learning rate = 0.1로 설정하였을 때, 다음 질문에 답하시오.

- 1) 각 Data sample별로 1 iteration씩 학습을 진행했을 때 각각 $\vec{\theta}$ 의 변화를 구하시오.
- 2) 1)에서의 결과를 참고하여 각 Data sample별로 $\frac{\partial \mathcal{L}(\vec{\theta})}{\partial \theta_2} : \frac{\partial \mathcal{L}(\vec{\theta})}{\partial \theta_1} : \frac{\partial \mathcal{L}(\vec{\theta})}{\partial \theta_0}$ 을 구하시오.
- 3) 위의 결과를 통해 $x_2^{(i)}$ 가 γ 배 되었을 때, $\theta_2, \theta_1, \theta_0$ 가 update되는 양이 몇 배 차이 나는지 구하시오.
- 4) 위의 결과를 통해 $x_1^{(i)}$ 가 γ 배 되었을 때, $\theta_2, \theta_1, \theta_0$ 가 update되는 양이 몇 배 차이 나는지 구하시오.

Chapter. 5

for Several Samples - Theory

Question. 5-04

Question. 5-02로부터

Dataset을 $y = ax_n + \dots + bx_1 + c$ 로, Prediction 모델을 $\hat{y} = \theta_m x_m + \dots + \theta_1 x_1 + \theta_0$ 로 확장하였다.

$\vec{\theta}$ 를 Update하기 위해 n 개의 data sample을 이용할 때, 1번의 iteration동안 $\vec{\theta}$ 가 dataset을 잘 표현하는 $\vec{\theta}$ 로 Update되는 과정을 설명하시오.

단, forward/backward propagation을 설명하기 위해 각 연산은 basic building node를 이용하시오.