Question.3-03

문제에서의 상황이 Question.3-02와 동일할 때, 다음의 물음에 답하시오.

1) Question.3-02의 1) 상황에서 1)의 결과를 이용하여 MSE cost를 구하시오.

2) MSE cost를 algebraic equation으로 표현하고, MSE cost를 최소로 만드는 θ를 구하시오.

3) 2)에서 구한 algebraic equation이 θ에 대한 몇차인지 구하고, convexity를 말하시오.

1) 먼저 Question.3-02 1)의 결과는 다음과 같았다.

$$\mathcal{L}^{(1)} = (y^{(1)} - \hat{y}^{(1)})^2 = (2-1)^2 = 1$$
$$\mathcal{L}^{(2)} = (y^{(2)} - \hat{y}^{(2)})^2 = (4-2)^2 = 4$$
$$\mathcal{L}^{(3)} = (y^{(3)} - \hat{y}^{(3)})^2 = (6-3)^2 = 9$$

그리고 cost J는 loss들의 평균을 의미하므로 다음과 같다.

$$J = \frac{1}{3}\sum_{i=1}^{3} \mathcal{L}^{(i)} = \frac{1}{3}(1+4+9) = 4.67$$

2) Cost J에 대한 식을 변수 θ와 data sample $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), (x^{(3)}, y^{(3)})$에 대해 표현하면

$$J = \frac{1}{3}\sum_{i=1}^{3} \mathcal{L}^{(i)} = \frac{1}{3}\sum_{i=1}^{3} (y^{(i)} - \hat{y}^{(i)})^2$$
$$= \frac{1}{3}\sum_{i=1}^{3} (y^{(i)} - \theta \cdot x^{(i)})^2$$

이 된다. 이때 Question.3-02 2)의 결과에서 알 수 있듯이 각각의 $(y^{(i)} - \theta \cdot x^{(i)})^2$을 최소로 만드는 θ는

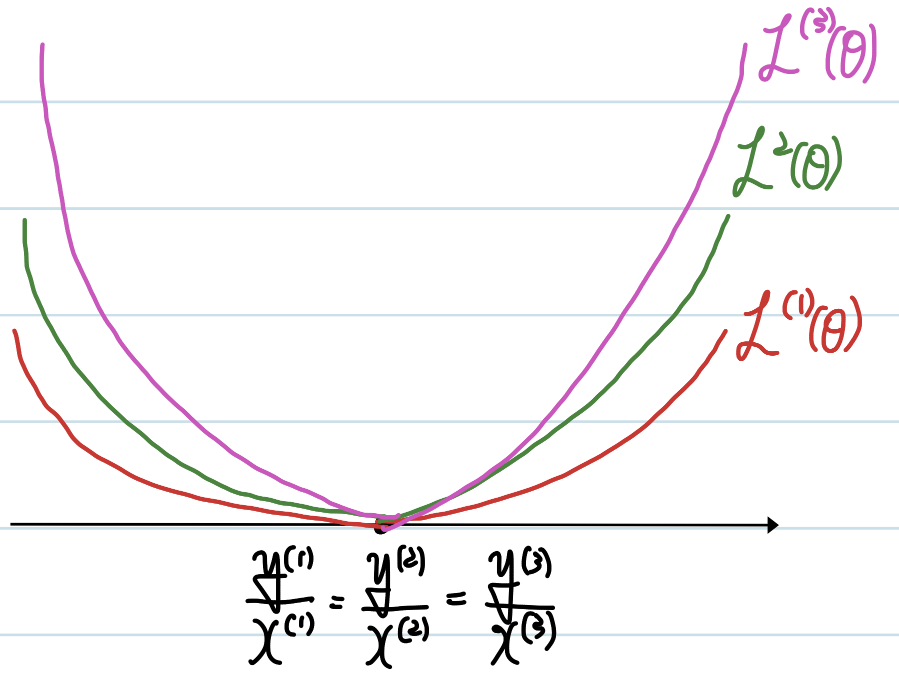
$$\frac{y^{(1)}}{x^{(1)}}, \quad \frac{y^{(2)}}{x^{(2)}}, \quad \frac{y^{(3)}}{x^{(3)}}$$

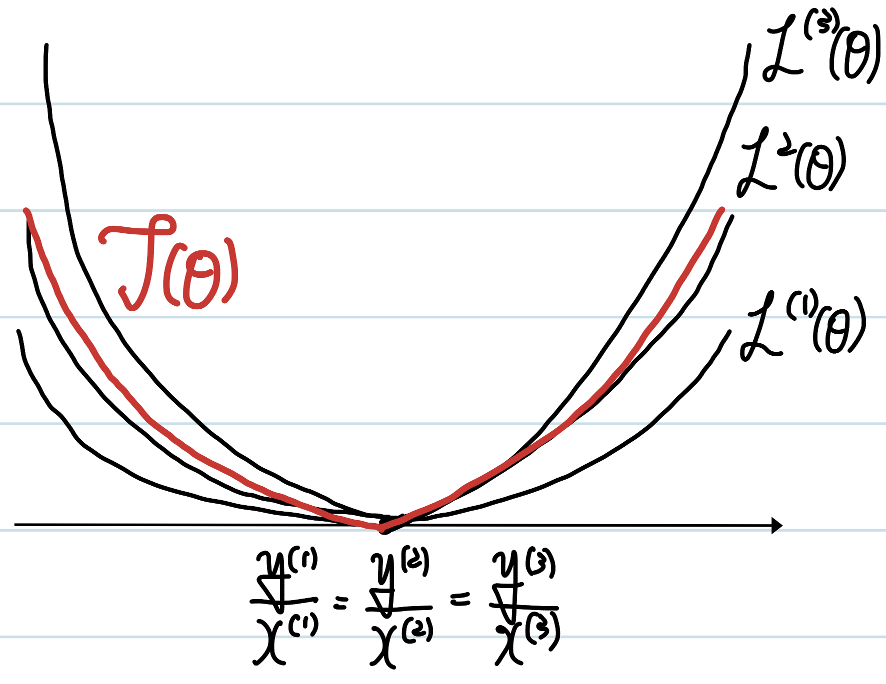
이고 이는 모두 dataset이 만들어진 $y = ax$의 a와 같을 때이다.

3) 1)에서 구한대로 cost는 loss의 산술평균값이다. 이때 각각의 $\mathcal{L}^{(i)}$ 함수를 변수 θ에 대해 표현하면

$$\mathcal{L}^{(1)}(\theta) = (2-\theta)^2 \qquad \mathcal{L}^{(2)}(\theta) = (4-2\theta)^2 = 4(2-\theta)^2 \qquad \mathcal{L}^{(3)}(\theta) = (6-3\theta)^2 = 9(2-\theta)^2$$

이 된다. 이를 그래프로 표현하면 아래, 왼쪽의 그래프가 된다. 그리고 각 그래프들의 평균을 그린 $J(\theta)$는 아래, 오른쪽의 그래프와 같다.





즉, $J(\theta)$는 $\mathcal{L}^{(i)}(\theta)$와 마찬가지로 θ에 대한 2차식이고 convex하다.