

# Assignment 1

(1) by question,  $h(x_1, x_2) = \sigma(b + w_1 x_1 + w_2 x_2)$

$$\sigma \text{ is the sigmoid} \Rightarrow \sigma(x) = \frac{1}{1+e^{-x}}$$

$$\theta_{n+1} = \theta_n - \alpha \nabla_{\theta} \text{Loss} \Rightarrow \theta_1 = \theta_0 - \alpha \nabla L(\theta_0) \quad (\alpha: \text{學習率})$$

$$L(\theta) = \frac{1}{2} (h(x_1, x_2) - y)^2$$

$$\text{single data, general form is } L(\theta) = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|y_i - h(x_i^i, x_2^i)\|^2$$

$$\sigma'(x) = \frac{d}{dx} (1+e^{-x})^{-1} = -(1+e^{-x})^{-2} \cdot -e^{-x} = e^{-x} (1+e^{-x})^{-2}$$

$$\Rightarrow \sigma'(x) = \left( \frac{e^{-x}}{1+e^{-x}} \right) \left( \frac{1}{1+e^{-x}} \right) = (1-\sigma(x))\sigma(x)$$

$$\text{Let } z = b + w_1 x_1 + w_2 x_2 \Rightarrow L(\theta) = \frac{1}{2} (\sigma(z) - y)^2$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \sigma} \cdot \frac{\partial \sigma}{\partial z} \cdot \frac{\partial z}{\partial b} = (\sigma(z) - y)(1-\sigma(z))\sigma(z) \cdot 1$$

same way to get

$$\frac{\partial L}{\partial w_1} = (\sigma(z) - y)(1-\sigma(z))\sigma(z) \cdot x_1$$

$$\frac{\partial L}{\partial w_2} = (\sigma(z) - y)(1-\sigma(z))\sigma(z) \cdot x_2$$

$$z = b + w_1 x_1 + w_2 x_2 = 4 + 5 - 1 + 6 \times 2 = 21$$

$$b' = b_0 - \frac{\partial L}{\partial b} = 4 - \alpha (\sigma(21) - 3)(1-\sigma(21))\sigma(21)$$

$$w_1' = 5 - \alpha (\sigma(21) - 3)(1-\sigma(21))\sigma(21)$$

$$w_2' = 6 - 2\alpha (\sigma(21) - 3)(1-\sigma(21))\sigma(21)$$

※

(2)

(a) by question 1 blue part

we know when  $k=1 \Rightarrow \sigma'(x) = (1-\sigma(x))\sigma(x)$

$$k=2 \Rightarrow \frac{d}{dx} \left[ \frac{(1-\sigma(x))\sigma(x)}{u} \right]$$

$$\frac{du}{dx} = -\sigma(x)(1-\sigma(x)), \quad \frac{dv}{dx} = \sigma(x)(1-\sigma(x))$$

$$\begin{aligned} \frac{du}{dx}v + \frac{dv}{dx}u &= \sigma(x)(1-\sigma(x)) \left[ -\sigma(x) + (1-\sigma(x)) \right] \\ &= \sigma(x)(1-\sigma(x))(1-2\sigma(x)) = \sigma''(x) \end{aligned}$$

$$\frac{d}{dx} \sigma''(x) = \frac{d}{dx} \left[ \frac{\sigma'(x)}{u} \frac{(1-2\sigma(x))}{v} \right]$$

$$\frac{du}{dx} = \sigma'(x)(1-2\sigma(x)), \quad \frac{dv}{dx} = -2\sigma'(x)$$

$$\frac{du}{dx}v + \frac{dv}{dx}u = \sigma'(x) \left[ (1-2\sigma(x))^2 + (-2\sigma'(x)) \right]$$

$$= \sigma'(x)(1-\sigma(x)) \left[ (1-2\sigma(x))^2 - 2\sigma(x)(1-\sigma(x)) \right] = \sigma'''(x)$$

(b)

$$\sinh(x) = \frac{e^x - e^{-x}}{2}, \quad \cosh(x) = \frac{e^x + e^{-x}}{2}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$\sigma(x) = \frac{1}{1+e^{-x}}, \quad \text{try to find same denominator}$$

$$\tanh\left(\frac{x}{2}\right) = \frac{e^{\frac{x}{2}} - e^{-\frac{x}{2}}}{e^{\frac{x}{2}} + e^{-\frac{x}{2}}} = \frac{1-e^{-x}}{1+e^{-x}} \Rightarrow 1+\tanh\left(\frac{x}{2}\right) = \frac{2}{1+e^{-x}} \Rightarrow \sigma(x) = \frac{1+\tanh\left(\frac{x}{2}\right)}{2}$$

\*

(3)

(1) 上課有提到房價模型有可能會有負的房價

老師有提到可能可以用  $h(x_1, x_2) = \sigma(b + x_1 w_1 + x_2 w_2)$

$$\sigma(x) = \frac{1}{1+e^{-x}}, \text{ 假設 } x_1 = 0, x_2 = 0$$

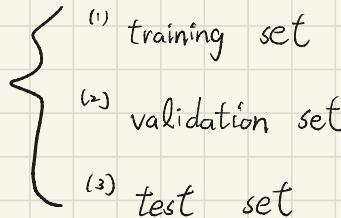
↑ 屋齡 ↑ 地坪

可能房價還是  $> 0$ , 不知道  $b + w_1 x_1 + w_2 x_2$

重新設計能不能解決這問題

(2)

data 需要分



不知道三者佔比的比例不同，會不會

導致 training 的模型結果有很大差距？