Machine Learning Assignment 1

許晉 314653002

September 9, 2025

Question 1

Consider stochastic gradient descent method to learn the house price model

$$h(x_1, x_2) = \sigma(b + w_1 x_1 + w_2 x_2).$$

where σ is the sigmoid function.

Given one signle data point $(x_1, x_2, y) = (1, 2, 3)$, and assuming that the current parameter is $\theta^0 = (b, w_1, w_2) = (4, 5, 6)$, evaluate θ^1.

(Just write the expression and substitute the numbers; no need to simplify or evaluate.)

Solution

Consider the gradient descent method, we have

$$\theta^{n+1} = \theta^n + \frac{2\alpha}{m}\sum_{i=1}^{m}(y^i - h(x_1^i, x_2^i))\nabla_\theta h(x_1^i, x_2^i)$$

When in stochastic case, which is $m = 1$.

Thus, we have $\theta^1 = \begin{bmatrix}4\\5\\6\end{bmatrix} + 2\alpha(3-\sigma(21))\begin{bmatrix}\sigma'(21)\\\sigma'(21)\cdot 1\\\sigma'(21)\cdot 2\end{bmatrix} = \begin{bmatrix}4\\5\\6\end{bmatrix} + 2\alpha(3-\sigma(21))\begin{bmatrix}\sigma(21)(1-\sigma(21))\\\sigma(21)(1-\sigma(21))\\2\sigma(21)(1-\sigma(21))\end{bmatrix}$

Question 2

(a) Find the expression of $\frac{d^k}{dx^k}\sigma(x)$ in terms of $\sigma(x)$ for $k = 1, \cdots, 3$ where σ is the sigmoid function.

(b) Find the relation between sigmoid function and hyberbolic functions.

Solution

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

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

$$\frac{d^3}{dx^3}\sigma(x) = \sigma''(x)(1-2\sigma(x)) - 2(\sigma'(x))^2$$

(b) As we know that $\sigma(x) = \dfrac{1}{1+e^{-x}} \Rightarrow 1-\sigma(x) = \dfrac{e^{-x}}{1+e^{-x}}$. Then we have $e^{-x} = \dfrac{1-\sigma(x)}{\sigma(x)}$.

Thus, $\sinh(x) = \dfrac{\dfrac{1-\sigma(-x)}{\sigma(-x)} - \dfrac{1-\sigma(x)}{\sigma(x)}}{2} = \dfrac{\sigma(x)(1-\sigma(-x)) - \sigma(-x)(1-\sigma(x))}{2\sigma(x)\sigma(-x)} = \dfrac{\sigma(x)-\sigma(-x)}{2\sigma(x)\sigma(-x)}$

$\cosh(x) = \dfrac{\dfrac{1-\sigma(-x)}{\sigma(-x)} + \dfrac{1-\sigma(x)}{\sigma(x)}}{2} = \dfrac{\sigma(x)(1-\sigma(-x)) + \sigma(-x)(1-\sigma(x))}{2\sigma(x)\sigma(-x)} = \dfrac{\sigma(x)+\sigma(-x)}{2\sigma(x)\sigma(-x)} - 1$

$\tanh(x) = \dfrac{e^x - e^{-x}}{e^x + e^{-x}} = \dfrac{e^{2x}-1}{e^{2x}+1} = \dfrac{1+e^{2x}-2}{1+e^{2x}} = 1 - 2\sigma(-2x)$