Milestone 1 - Mini Reviews

① (0) pretty bad
② (1) good plot
③ (0) not good
④ (1) pretty scenery

["pretty", "good", "bad", "plot", "not", "scenery"]

$W = [0, 0, 0, 0, 0, 0]$, Learning Rate = 0.5

① (0) pretty bad

$y_1 = 0 \longrightarrow \phi x = [1, 0, 1, 0, 0, 0]$

$k = 0+0+0+0+0+0 = 0$

$h = \sigma(0) = 1/1+e^{-0} = 0.5$

$w_1 = w_3 = 0 - 0.5 * 0.5 * 1 = -0.25$

$w_2 = w_4 = w_6 = 0 - 0.5 * 0.5 * 0 = 0$

$\longrightarrow W = [-0.25, 0, -0.25, 0, 0, 0]$ ※

② (1) good plot

$y_2 = 1 \longrightarrow \phi x \ [0, 1, 0, 1, 0, 0]$

$k = 0+0+0+0+0 = 0$

$h = \sigma(W \phi x) = 1/1+e^{-0} = 1/2 = 0.5$

$w_1 = w_3 = -0.25 - 0.5 * (0.5-1) * 0 = -0.25$

$w_2 = w_4 = 0 - 0.5 * (0.5-1) * 1 = 0.25$

$w_5 = w_6 = 0 - 0.5 * (0.5-1) * 0 = 0$

$\longrightarrow W = [-0.25, 0.25, -0.25, 0.25, 0, 0]$ ※

③ (0) not good

$y_3 = 0 \longrightarrow \phi x = [0, 1, 0, 0, 1, 0]$

$k = 0.25$

$h = \sigma(w\phi x) = 1/_{1+e^{-0.25}}$

$w_1 = w_3 = -0.25 - 0 = -0.25$

$w_2 = 0.25 - 0.5 * (1/_{1+e^{-0.25}}) * 1 \fallingdotseq -0.031088$

$w_4 = 0.25 - 0 = 0.25$

$w_5 = 0 - 0.5 * (1/_{1+e^{-0.25}}) * 1 \fallingdotseq -0.281088$

$w_6 = 0 - 0 = 0$

$\longrightarrow w = [-0.25, -0.031088, 0.25, -0.281088, 0]$ ※

---

④ (1) pretty scenery

$y_4 = 1 \longrightarrow \phi x_4 = [1, 0, 0, 0, 0, 1]$

$k = -0.25$

$h = \sigma(-0.25) = 1/_{1+e^{0.25}}$

$w_1 = -0.25 - 0.5 * (1/_{1+e^{0.25}} - 1) * 1 \fallingdotseq 0.031088$

$w_2 = -0.031088$

$w_3 = -0.25$

$w_4 = 0.25$

$w_5 = -0.281087$

$w_6 = 0 - 0.5 * (1/_{1+e^{0.25}} - 1) * 1 \fallingdotseq 0.281088$

$\longrightarrow w = [0.031088, -0.031088, -0.25, 0.25, -0.281088, 0.281088]$ ※

Milestone 2 - Derivatives

請試著推導 **(h-y)Φ(x)** 是怎麼得到的？首先，請推導下列算式之答案

Gradient Descent
for Logistic Regression

※ chain rule

Cost: $J = \frac{1}{m}\sum_{i=1}^{m}\left(-y_i \log h_i + (1-y_i)(\log(1-h_i))\right)$

σ(Predict_val): $h_i = \sigma(k_i) = \frac{1}{1+e^{-k_i}} = \frac{e^{k_i}}{e^{k_i}+1}$

Predict_val: $k_i = \theta x_i$

$$\longrightarrow \frac{dJ}{d\theta} = \frac{dJ}{dh_i}① * \frac{dh_i}{dk_i}② * \frac{dk_i}{d\theta}③$$

①

$$\frac{dJ}{dh_i} = -\frac{1}{m}\sum\left(\frac{d(y_i \log h_i)}{dh_i} + \frac{d(1-y_i)\log(1-h_i)}{dh_i}\right)$$

$$= -\frac{1}{m}\sum\left(\frac{y_i}{h_i} + \frac{(1-y_i)\cdot(-1)}{(1-h_i)}\right)$$

$$= -\frac{1}{m}\sum\left(\frac{y_i}{h_i} - \frac{(1-y_i)}{(1-h_i)}\right) ※$$

②

$$\frac{dh_i}{dk_i} = \frac{d\left(\frac{e^{k_i}}{e^{k_i}+1}\right)}{dk_i}$$

$$= \left(\frac{de^{k_i}}{dk_i}(e^{k_i}+1) - e^{k_i}\left(\frac{d(e^{k_i}+1)}{dk_i}\right)\right) / (e^{k_i}+1)^2$$

$$= e^{k_i}(e^{k_i}+1) - e^{k_i}\left(\frac{de^{k_i}}{dk_i} + \frac{d1}{dk_i}\right) / (e^{k_i}+1)^2$$

$$= e^{k_i} / (e^{k_i}+1)^2 = \frac{e^{k_i}}{e^{k_i}+1} * \frac{1}{e^{k_i}+1}$$

$$= h_i \cdot (1-h_i) ※$$

③

$$\frac{dk_i}{d\theta} = x_i ※$$

$$\longrightarrow \frac{dJ}{d\theta} = ① * ② * ③$$

$$= -\frac{1}{m}\sum\left(\frac{y_i}{h_i} - \frac{(1-y_i)}{(1-h_i)}\right)\cdot\left(h_i\cdot(1-h_i)\right)\cdot x_i$$

$$= -\frac{1}{m}\sum\left(y_i(1-h_i) - (1-y_i)h_i\right)x_i$$

$$= \frac{-1}{m}\sum(y_i - h_i)x_i = \frac{1}{m}\sum(h_i - y_i)x_i ※$$