Solution HW2

1. Find the error of each row $(\log e)$

row 1 ≈ 0.438 ✓

row 2 ≈ 1.038 ✓

row 3 ≈ 0.313 ✓

EX

$$\text{Row 1} = \underline{y^{(i)}(-\log(h(x)))} + \underline{(1-y^{(i)})(-\log(1-h(x)))}$$

$$0.19 = 1\times\left(-\log_e\left(\sigma\left(0.1+(0.1\times2)+(0.1\times3)\right)\right)\right) + 0$$

$$\sigma(0.1+0.2+0.3)$$

$$\sigma(0.6)$$

$$= \quad -\log(0.645)$$

$$-(-0.438)$$

2) Let SGD = $\theta_j = \theta_j - \alpha \left(h(x)^{(i)} - y^{(i)} \right) x_j^{(i)}$

$\alpha = 0.01, \theta_0 = \theta_1 = \theta_2 = 0.01$

<u>iter 1</u>   assume select row 2

$h(x^{(2)}) = \dfrac{1}{1 + e^{(-0.6)}} = 0.646$

$\theta_0 = 0.1 - 0.01 (0.646 - 0) (1) = 0.094$

$\theta_1 = 0.1 - 0.01 (0.646 - 0)(1) = 0.094$

$\theta_2 = 0.1 - 0.01(0.646 - 0)(4) = 0.074$

<u>iter 2</u>   Assume select row 3

use $\theta_0 = 0.94, \theta_1 = 0.94, \theta_2 = 0.074$

$h(x^{(3)}) = \dfrac{1}{1 + e^{(-0.84)}} = 0.698$

$\theta_0 = 0.094 - 0.01 (0.698 - 1)(1) = 0.097$

$\theta_1 = 0.094 - 0.01 (0.698 - 1)(4) = 0.106$

$\theta_2 = 0.074 - 0.01(0.698 - 1)(5) = 0.089$

3) Work again on 2) with ridge & Lasso

let $\theta_0 = \theta_1 = \theta_2 = 0.1$

ridge regression $= \theta_j = \theta_j - \alpha[(h(x)-y)x_j^i + \lambda\theta_j]$

iter1 $\alpha = 0.01$, $\lambda = 10$, select row $\underline{\underline{2}}$

$h_\theta(x^{(2)}) = 0.646$

$\theta_0 = 0.1 - 0.01((0.646-0)(1) + 10(0.1))$

$\theta_1 = 0.1 - 0.01((0.646-0)(1) + 10(0.1))$

$\theta_2 = 0.1 - 0.01((0.646-0)(4) + 10(0.1))$

$\theta_0 = 0.084$, $\theta_1 = 0.084$, $\theta_2 = 0.064$

$\qquad\qquad\qquad\qquad 0.084 - 0.01(0.517)$

iter 2 select row 3, $h_\theta(x^3) = 0.677$

$\theta_0 = 0.084 - 0.01[(0.677-1)(1) + 10(0.084)]$

$\qquad\qquad\qquad\qquad -0.452$

$\theta_1 = 0.084 - 0.01[(0.677-1)(4) + 10(0.084)]$

$\qquad\qquad\qquad\qquad -1.292 \uparrow 0.84$

$\theta_2 = 0.064 - 0.01[(0.677-1)(5) + 10(0.064)]$

$\theta_0 = 0.078$, $\theta_1 = 0.088$ $\theta_2 = 0.073$

3) Work again on 2) with ridge & Lasso

let $\theta_0 = \theta_1 = \theta_2 = 0.1$

Lasso regression $= \theta_j = \theta_j - \alpha\left[(h(x)-y)x_j^i + \lambda \, \text{sign}(\theta)\right]$

iter 1 $\alpha = 0.01, \quad \lambda = 0.2, \quad \text{select row } \underline{\underline{2}}$

$h_\theta(x^{(2)}) = 0.646$

$\theta_0 = 0.1 - 0.01\left((0.646-0)(1) + 0.2(1)\right)$

$\theta_1 = 0.1 - 0.01\left((0.646-0)(1) + 0.2(1)\right)$

$\theta_2 = 0.1 - 0.01\left((0.646-0)(4) + 0.2(1)\right)$

$\underline{\theta_0 = 0.092}, \quad \underline{\theta_1 = 0.092} \quad \underline{\theta_2 = 0.072}$

---

iter 2 select row 3, $h_\theta(x^3) = 0.694$

$\theta_0 = 0.092 - 0.01\left[(0.694-1)(1) + 0.2(1)\right]$

$\theta_1 = 0.092 - 0.01\left[(0.694-1)(4) + 0.2(1)\right]$

$\theta_2 = 0.072 - 0.01\left[(0.694-1)(5) + 0.2(1)\right]$

$\theta_0 = 0.093, \quad \theta_1 = 0.102, \quad \theta_2 = 0.085$