id: 23341130

(a)

Hypothesis function,

$$h_\theta(x) = g(\theta^T x)$$

$$= \frac{1}{1+e^{-\theta^T x}}$$

cost function:

$$J_i(\theta) = -\left[y \log(h_\theta(x)) + (1-y)\log(1-h_\theta(x))\right]$$

$$J_i(\theta) = \begin{cases} -\log(h_\theta(x)) & \text{if } y=1 \\ -\log(1-h_\theta(x)) & \text{if } y=0 \end{cases}$$

$$\therefore J(\theta) = \frac{1}{n}\sum_{i=0}^{n-1} J_i(\theta)$$

(b) for (1)

$$\theta^T x = \begin{bmatrix} 0 & 0.1 & 0.008 & 0.001 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 10 & 5 & 3 \\ 124 & 398 & 312 \\ -139 & -112 & -172 \end{bmatrix}$$

$$\Rightarrow \theta^T x = \begin{bmatrix} 1.853 & 3.5 & 2.624 \end{bmatrix}$$

$\cancel{\theta_a \leftarrow J_\theta(x_1) =}$

$$J_1(\theta) = -\log(1 - 1.853)$$

$$J_2(\theta) = -\log(3.5)$$

$$J_3(\theta) = -\log(1 - 2.624)$$

for ② $\rightarrow$

$$\theta^T x = [0 \quad 0.5 \quad 0.008 \quad 0.002] \begin{bmatrix} 1 & 1 & 1 \\ 10 & 5 & -139 \\ 124 & 398 & -112 \\ -139 & 312 & -172 \end{bmatrix}$$

$$\theta^T x = [5.714 \quad 6.308 \quad -70.74]$$

$$J_1(\theta) = -\log(1 - 5.714)$$

$$J_2(\theta) = -\log(6.308)$$

$$J_3(\theta) = -\log(1 + 70.74)$$

for ① $\rightarrow$

$$h_\theta^{(1)}(x_1) = \frac{1}{1+e^{-1.853}}$$

$$= 0.86$$

$$h_\theta^{(2)}(x_2) = \frac{1}{1+e^{-3.5}}$$

$$= 0.97$$

$$h_\theta^{(3)}(x) = \frac{1}{1+e^{-2.624}}$$

$$= 0.93$$

$$\therefore J(\theta)_{(1)} = -\log(0.86) - \log(0.97) - \log(1-0.93)$$

for ② $\rightarrow$

$$h^1(x) = \frac{1}{1+e^{-5.714}}$$

$$= 0.996$$

$$h^2(x) = \frac{1}{1+e^{-6.308}}$$

$$= 0.998$$

$$h^3(x) = \frac{1}{1+e^{70.74}}$$

$$= 1.896 \times 10^{-31}$$

$$\therefore J(\theta)_{(2)} = -\log(1-0.996) - \log(0.998)$$

$$- \log(1-1.896\times10^{-31})$$

$$= 5.523$$

As $J(\theta)_{(1)} < J(\theta)_{(2)}$

So 1st set of θ on model is better.