Problem Set #2

2.(a)

$$\frac{\sum_{i\in I_{a,b}} P(y^{(i)}=1|x^{(i)};\theta)}{|\{i\in I_{a,b}\}|} = \frac{\sum_{i\in I_{a,b}} \mathbb{I}\{y^{(i)}=1\}}{|\{i\in I_{a,b}\}|}$$

$$P(y=1|x,\theta) = h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

$I_{a,b} = \{i \mid i\in\{1,\dots,m\}, h_\theta(x^{(i)})\in(a,b)\}$ ← index set of training examples

$|S|$: size of the set S.

Show that the above property holds true for the described logistic regression model over the range of $(a,b)=(0,1)$

$\{x^{(i)}, y^{(i)}\}_{i=1}^{m}$

$x_0^{(i)}=1$ $\quad x^{(i)}\in\mathbb{R}^{n+1}$ $\quad y^{(i)}\in\{0,1\}$ $\quad \theta\in\mathbb{R}^{n+1}$

$$\frac{\sum_{i\in I_{a,b}} P(y^{(i)}=1|x^{(i)};\theta)}{|\{i\in I_{a,b}\}|} = \frac{\sum_{i\in I_{a,b}} h_\theta(x)}{|\{i\in I_{a,b}\}|} = \frac{\sum_{i\in I_{a,b}} \frac{1}{1+e^{-\theta^T x^{(i)}}}}{|\{i\in I_{a,b}\}|} \qquad |\{i\in I_{a,b}\}| = |\{i\in I_{0,1}\}| = m$$

$$= \frac{\sum_{i\in I_{a,b}} \frac{1}{1+e^{-\theta^T x^{(i)}}}}{m} = \frac{1}{m}\sum_{i\in I_{a,b}} \frac{1}{1+e^{-\theta^T x^{(i)}}} = \frac{1}{m}\sum_{i\in I_{a,b}} \frac{1}{1+e^{-(\theta_0+\theta'^T x'^{(i)})}} = \frac{1}{m}\left\{\sum_{i=1}^{m} \frac{1}{1+e^{-(\theta_0+\theta^T x^{(i)})}}\right\}$$

$$\theta = \begin{bmatrix}\theta_0\\ \theta_1\\ \vdots\\ \theta_n\end{bmatrix} \quad x^{(i)} = \begin{bmatrix}x_0^{(i)}=1\\ x_1^{(i)}\\ \vdots\\ x_n^{(i)}\end{bmatrix} \quad \theta' = \begin{bmatrix}\theta_1\\ \vdots\\ \theta_n\end{bmatrix} \quad x'^{(i)} = \begin{bmatrix}x_1^{(i)}\\ x_2^{(i)}\\ \vdots\\ x_n^{(i)}\end{bmatrix} \quad \theta_0 x_0^{(i)} = \theta_0 = \text{bias term}$$

$$\frac{1}{m}\left\{\sum_{i=1}^{m} \frac{1}{1+e^{-(\theta_0+\theta^T x^{(i)})}}\right\} = \frac{\sum_{i\in I_{a,b}} \mathbb{I}\{y^{(i)}=1\}}{|\{i\in I_{a,b}\}|} = \frac{1}{m}\sum_{i\in I_{a,b}} \mathbb{I}\{y^{(i)}=1\}$$

$$\sum_{i=1}^{m} \frac{1}{1+e^{-(\theta_0+\theta^T x^{(i)})}} = \prod_{i=1}^{m} \{y^{(i)}=1\} = \sum_{i=1}^{m} y^{(i)}$$

negative log likelihood

$$l(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log h_\theta(x^{(i)}) + (1-y^{(i)})\log(1-h_\theta(x^{(i)}))\right]$$

$$h_\theta(x^{(i)}) = \frac{1}{1+e^{-(\theta^T x+\theta_0)}} = \sigma(x^{(i)})$$

(i) when $j=0$

$$\frac{\partial}{\partial\theta_0} l(\theta) = \frac{\partial}{\partial\theta_0} -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log h_\theta(x^{(i)}) + (1-y^{(i)})\log(1-h_\theta(x^{(i)}))\right]$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)} \times \frac{1}{h_\theta(x^{(i)})} \times \sigma(x)(1-\sigma(x)) \times 1 + (1-y^{(i)}) \times \frac{1}{1-h_\theta(x^{(i)})} \times (-\sigma(x))(1-\sigma(x)) \times 1\right]$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}(1-\sigma(x^{(i)})) + (1-y^{(i)})(-\sigma(x^{(i)}))\right]$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}(1-h_\theta(x^{(i)})) - h_\theta(x^{(i)}) + y^{(i)}h_\theta(x^{(i)})\right]$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)} - h_\theta(x^{(i)})\right] = 0$$

$$\therefore \quad \sum_{i=1}^{m} y^{(i)} = \sum_{i=1}^{m} h_\theta(x^{(i)})$$

(ii) when $j \neq 0$

$$\frac{\partial}{\partial \theta_j} l(\theta) = \frac{\partial}{\partial \theta_j} -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)} \log h_\theta(x^{(i)}) + (1-y^{(i)}) \log(1-h_\theta(x^{(i)}))\right] = 0$$

$$h_\theta(x^{(i)}) = \frac{1}{1+e^{-(\theta^T x^{(i)}+\theta_0)}}$$

$$\frac{\partial}{\partial \theta_j} l(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)} \times \frac{1}{h_\theta(x^{(i)})} \times \sigma(x^{(i)})(1-\sigma(x^{(i)}))x_j^{(i)} + (1-y^{(i)})\frac{1}{1-h_\theta(x^{(i)})} \times (-1)\sigma(x^{(i)})(1-\sigma(x^{(i)}))x_j^{(i)}\right]$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}(1-h_\theta(x^{(i)}))x_j^{(i)} + (1-y^{(i)})(-1)\sigma(x^{(i)})x_j^{(i)}\right] = 0$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}x_j^{(i)} - y^{(i)}h_\theta(x^{(i)})x_j^{(i)} - (1-y^{(i)})h_\theta(x^{(i)})x_j^{(i)}\right] = 0$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}x_j^{(i)} - y^{(i)}h_\theta(x^{(i)})x_j^{(i)} - h_\theta(x^{(i)})x_j^{(i)} + y^{(i)}h_\theta(x^{(i)})x_j^{(i)}\right] = 0$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}x_j^{(i)} - h_\theta(x^{(i)})x_j^{(i)}\right] = 0$$

$$\sum_{i=1}^{m} y^{(i)}x_j^{(i)} = \sum_{i=1}^{m} h_\theta(x^{(i)})x_j^{(i)}$$

$$y^T x_j = h_\theta(x)^T x_j$$

$$\therefore \sum_{i=1}^{m} y^{(i)} = \sum_{i=1}^{m} h_\theta(x^{(i)})$$

Therefore,

$$\frac{\sum_{i\in I_{a,b}} h_\theta(x^{(i)})}{|\{i \in I_{a,b}\}|} = \frac{\sum_{i\in I_{a,b}} y^{(i)}}{|\{i \in I_{a,b}\}|} \qquad (a,b) = (0,1)$$