# Problem Set 1

DuannYu

January 5, 2021

## 1 Logistic Regression

### (a)

*Proof.* We denote $x^{(k)}$ as $k$-th sample in dataset $X$, use $i$ and $j$ to index the dimensions of $x$. Using chain rule,

$$\begin{aligned}
\frac{\partial J(\theta)}{\partial \theta_i} &= -\frac{1}{m}\sum_{k=1}^{m}\frac{1}{g(z)}g(z)[1-g(z)]\frac{\partial z}{\partial \theta_i} \\
&= \frac{1}{m}\sum_{k=1}^{m}[g(z)-1]y^{(k)}x_i^{(k)}
\end{aligned}$$

Therefore, Hessian Matrix can be derived

$$\begin{aligned}
H_{ij} &= \frac{\partial^2 J(\theta)}{\partial \theta_i \partial \theta_j} \\
&= \frac{1}{m}\sum_{k=1}^{m}\frac{\partial g(z)}{\partial \theta_j}y^{(k)}x_i^{(k)} \\
&= \frac{1}{m}\sum_{k=1}^{m}g(z)[1-g(z)]y^{(k)}x_j^{(k)}y^{(k)}x_i^{(k)}
\end{aligned}$$

Because $y^{(k)}y^{(k)}$ always equals to 1, there is

$$H_{ij} = \frac{1}{m}\sum_{k=1}^{m}g(z)[1-g(z)]x_i^{(k)}x_j^{(k)}$$

Lastly,

$$\begin{aligned}
z^T H z &= \sum_i \sum_j z_i H_{ij} z_j \\
&= \frac{1}{m}\sum_{k=1}^{m}\sum_i \sum_j z_i g(z)[1-g(z)]x_i^{(k)}x_j^{(k)}z_j \\
&= \frac{1}{m}\sum_{k=1}^{m}g(z)[1-g(z)]\sum_i \sum_j z_i z_j x_i^{(k)}x_j^{(k)}
\end{aligned}$$

Given that $g(z) \in (0,1)$ and $\sum_i \sum_j z_i z_j x_i^{(k)} x_j^{(k)} = (x^T z)^2 \geq 0$, So $z^T H z \geq 0$ $\square$

question **(b)(c)**s' codes can be see in this file.