Q2-C

In order to identify a training problem as concave or convex, we need to derive the Hessian matrix of $\ell(\theta)$.

For data points of type R^n the hessian matrix will be:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

for problem Q2, the data point are in R therefore Taking second deravative would do the work.

the Hessian matrix has only one item.

$$\frac{\partial}{\partial x} \frac{1}{1+e^{ax}} = \frac{-ae^{ax}}{(1+e^{ax})^2}$$

from Q1-a

$$\frac{\partial \ell(\theta)}{\partial \theta} = \sum_{i=1}^{n} \left(y_i - \frac{1}{1+e^{\theta x_i}}\right) x_i$$

$$\frac{\partial^2 \ell(\theta)}{\partial \theta} = \sum_{i=1}^{n} \frac{x_i^T e^{\theta x_i}}{(1+e^{\theta x_i})^2} \times x_i = \sum_{i=1}^{n} \frac{e^{\theta x_i}}{(1+e^{\theta x_i})^2}$$

$$= \sum_{i=1}^{n} \frac{1}{1+e^{\theta x_i}} \cdot \left(\frac{e^{\theta x_i}}{1+e^{\theta x_i}}\right) = \sum_{i=1}^{n} \frac{1}{1+e^{\theta x_i}} \left(\frac{1}{1+e^{\theta x_i}} - 1\right)$$

(*)