Q1.4 For convienience, here we assume the bias term, is regularized

$$H = \frac{\partial^2 J(\theta)}{\partial\theta_j \cdot \partial\theta_k} = \frac{\partial\left(\frac{1}{m}\left(\sum_{i=1}^{m}\left[h(x^{(i)}) - y^{(i)}\right]\right) \cdot x_j^{(i)} + \frac{\lambda}{m}\theta_j^{(i)}\right)}{\partial\theta_k}$$

$$= \frac{1}{m}\left(\sum_{i=1}^{m} x_j^{(i)} x_k^{(i)} \cdot h(x^{(i)}) \cdot (1 - h(x^{(i)})\right) + \frac{\lambda}{m} \cdot I$$

$$= \frac{1}{m}(X^T \cdot X \cdot S) + \frac{\lambda}{m} I$$

where the scaler terms are combined into S, and

$$S_{ii} = h(x^{(i)}) \cdot (1 - h(x^{(i)}))$$

$$= \frac{1}{m}\left(\underset{d+1 \times m}{X^T}\ \underset{m \times m}{S}\ \underset{m \times d+1}{X} + \lambda \underset{d+1 \times d+1}{I}\right)$$

H is positive definite iff $a^T H a > 0$

$$a^T H a$$

$$= \frac{1}{m} a^T (X^T S X + \lambda I) a$$

$$= \frac{1}{m}(a^T X^T S X a + \lambda a^T a)$$

$$= \frac{1}{m}(a^T X^T S (a^T X^T)^T + \lambda a^T a)$$

$$= \frac{1}{m}(\|a^T S^{\frac{1}{2}} X^T\|^2 + \lambda \|a\|^2)$$

Because $S_{ii} > 0$, X is full rank, all terms > 0

$$a^T H a > 0$$

So H is positive definite