Exercise 1
For all answers y, x and $h_\theta\left(x^{(i)}\right)$ denote a vector and multiplications are vector multiplications.

a) The vectorized expression for the hypothesis function is :

$$h_\theta\left(x^{(i)}\right) = \frac{1}{1 + e^{-\theta^{(i)\,T} \cdot x^{(i)}}}$$

This function is also used in the assignment code.

b) The vectorized expression of the cost function with summation is:

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

This function includes both cases if y=0 and y=1.
This function is also used in the assignment code.

c) The vectorized expression of the gradient function with summation is:

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

This function is also used in the assignment code.

d) The vectorized update expression of the theta update rule in the gradient procedure is:

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

Where j denotes the single instace in the training set that that specific theta belongs to. J is a row in x with all its features in the columns.

e) The matrix vector multiplication for this vectorization problem is:

$$\theta = \frac{a}{m}X^T(h_\theta(X) - y^{->})$$

Exercise 2
Initially I would say this can be done by just calculating the mean and standard deviation, but because a hint of the derivative is given I would not use this. The only other option I could come up with would be by doing least squares.

$$\theta = (X^T X)^{-1} - X^T y$$