Logistic Regression

THEORY

1) It is a machine learning algorithm under supervised learning and classification tasks.
2) $g(z)$ stands for the logistic or sigmoid function & the hypothesis function is given as:

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

$$g'(z) = g(z)(1 - g(z)) \tag{2}$$

3) Our aim is to maximize the likelihood function using the gradient ascent algorithm, given as:

$$L(\theta) = \prod_{i=1}^{n} p(y^{(i)}|x^{(i)}; \theta) \tag{3}$$

Also,

$$p(y|x; \theta) = (h_\theta(x))^y (1 - h_\theta(x))^{1-y} \tag{4}$$

4) We regularly update the parameters using gradient ascent algorithm.

$$\theta_j := \theta_j + \alpha \frac{\partial}{\partial \theta_j} l(\theta) \tag{5}$$

$$\theta_j := \theta_j + \alpha (y^{(i)} - h_\theta(x^{(i)})) x_j^{(i)} \tag{6}$$

QUIZ

1) Logistic regression is a ML algorithm based on:
   a) Regression tasks
   b) Classification tasks
2) What is the resulting method known as when log likelihood function $l(\theta)$ is maximized using Newton's method?
3) Prove that $\nabla_\theta l(\theta) = (y - h_\theta(x)) x_j$