ML 2017/2018: Graded Assignment 1 (5 points) Deadline: October 10

Miriam Riefel - 11332549

1. **3 points This question is about vectorization, i.e. writing expressions in matrixvector form. The goal is to vectorize the update rule for multivariate linear regression.**
   (a) **Let θ be the parameter vector θ = (θ0 θ1 ···θn)T and let the i-th data vector be: x(i) = (x0 x1 ···xn)T where x0 = 1. What is the vectorial expression for the hypothesis function hθ(x)?**

   $g(z) = \frac{1}{1+e^{-z}}$

   $h_\theta(x) = g(\theta^T x)$

   (b) **What is the vectorized expression for the cost function: J(θ) (still using the explicit summation over all training examples).**

   $J\theta = \frac{1}{m}\sum_{i=1}^{m}([y_{(i)}\log(h_\theta(x_{(i)})) + (1 - y_{(i)})\log(1 - h_\theta(x_{(i)}))]$

   (c) **What is the vectorized expression for the gradient of the cost function?**

   $J\theta = \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x(i)) - y_{(i)})\mathrm{x}_{j(i)}$

   (d) **What is the vectorized expression for the θ update rule in the gradient descent procedure?**

   $\theta_j := \theta_j - \frac{\alpha}{m}\sum_{i=1}^{m}(h_\theta(x(i)) - y_{(i)})\mathrm{x}_{j(i)}$

2. **2 points Derive an equation that can be used to find the optimal value of the parameter θ1 for univariate linear regression without doing gradient descent. This can be done by setting the value of the derivative equal to 0. You may assume that the value of θ0 is fixed.**

$J\theta = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x(i)) - y_{(i)})^{2}$

$2m * J\theta = \sum_{i=1}^{m}(h_\theta(x(i)) - y_{(i)})^{2}$

Derivation

$-2\sum_{i=1}^{m}(h_\theta(x(i)) - y_{(i)})$

Comparison to 0

$-2\sum_{i=1}^{m}(h_\theta(x(i)) - y_{(i)}) = 0$

$\sum_{i=1}^{m}(h_\theta(x(i)) - y_{(i)}) = 0$

$\sum_{i=1}^{m}(h_\theta(x(i)) - \sum_{i=1}^{m}(y(i)) = 0$

$\sum_{i=1}^{m}(h\theta(x(i)) = \sum_{i=1}^{m}(y(i))$

$h\theta(x(i))$ = y(i)