LAHORE UNIVERSITY OF MANAGEMENT SCIENCES
Department of Electrical Engineering

EE 514 (CS 535) Machine Learning
Quiz 5

Name: _____

Campus ID: _____

Total Marks: 10
Time Duration: 15 minutes

We have gone through the details of Ridge Regression and have studied gradient descent algorithm. In this quiz we will reinforce the same understanding. To keep things really simple – Since you must have completed your first homework, let us redo the Problem 2 (exactly the same ... or is it ?) from your homework.

Question 1 (7 marks)

You need to compute the gradient $\nabla J(\theta)$ and derive the gradient descent update rule $\theta^{(t+1)}$ for the regularized objective:

$$J(\theta) = \frac{1}{2n}\|y - A\theta\|^2 + \frac{\lambda}{2}\|\theta\|^2$$

.

Question 2 (1 mark)

Explain the impact of λ (very high / very low values) in terms of model under-fitting and over-fitting.

Question 3 (2 marks)

In gradient descent, the cost per update for SGD is $O(1)$ which is much lower then $O(m)$ of Mini-Batch GD. In general SGD takes large number of epoch to converge as compared Mini-Batch Gradient descent. (e.g 200 Vs 100). Based on this information, justify which one is more computationally **expensive** ?