LAHORE UNIVERSITY OF MANAGEMENT SCIENCES
Department of Electrical Engineering

EE514/CS535 Machine Learning
Quiz 05 Solutions

Name: ______________________

Campus ID: ______________________

Total Marks: 10
Time Duration: 12 minutes

Question 1 (4 marks)

True or False:

(a) Lasso can be interpreted as regularized linear regression that regularizes weights with the ℓ_2 norm.(T/F)

Solution: F: Lasso regularizes weights with the ℓ_1 norm.

(b) In polynomial regression, the polynomial degree M quantifies the trade-off between bias and variance. (T/F)

Solution: T

(c) A very large value of the hyper-parameter M in polynomial regression results in the model under-fitting the given data and a smaller mean-squared error on the training data. (T/F)

Solution: F: Overfitting

(d) We introduce a penalty (regularization) term in least-squares linear regression to decrease the training error. (T/F)

Solution: F: Regularization increases the training error.

Question 2 (2 marks)

Given the following information:

$$\boldsymbol{\theta} = \begin{bmatrix} 1.5 \\ -0.5 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 3.6 \\ 1.3 \end{bmatrix}, \theta_0 = 2$$

Consider the columns of the matrix X to be your data points, using this information calculate the mean-squared error (MSE).

Solution: $\tilde{\mathbf{y}} = \mathbf{X}\boldsymbol{\theta} + \theta_0$, MSE $= \frac{1}{2}\|\tilde{\mathbf{y}} - \mathbf{y}\|_2 = \frac{1}{4}$.

Question 3 (4 marks)

For a data-set with feature matrix X, output column vector y and weights column vector θ, the ridge loss $J(\theta)$ is given by:

$$J(\theta) = (y - X\theta)^T(y - X\theta) + \lambda\,\theta^T\theta.$$

Prove that $J(\theta)$ is a convex function in θ using that the fact that X^TX is positive semi-definite.

Solution:

$$J(\theta) = (y^T - \theta^T X^T)(y - X\theta) + \lambda\, \theta^T \theta$$
$$J(\theta) = (y^T y - y^T X\theta - \theta^T X^T y + \theta^T X^T X\theta) + \lambda\, \theta^T \theta$$
$$J(\theta) = (y^T y - 2\theta^T X^T y + \theta^T X^T X\theta) + \lambda\, \theta^T \theta$$
$$\nabla_\theta\, J(\theta) = (-2X^T y + 2X^T X\theta) + \lambda\, \theta$$
$$\nabla_\theta\, J(\theta) = -2X^T (y - X\theta) + \lambda\, \theta$$
$$\nabla_\theta^2\, J(\theta) = 2X^T X + \lambda I$$

X^TX and λI are both positive semi-definite hence $J(\theta)$ is convex in θ.