EE2211 Tutorial 6

(Polynomial Regression, 1D data)
Question 1:

Given the following data pairs for training

$$\{x = -10\} \rightarrow \{y = 5\}$$
$$\{x = -8\} \rightarrow \{y = 5\}$$
$$\{x = -3\} \rightarrow \{y = 4\}$$
$$\{x = -1\} \rightarrow \{y = 3\}$$
$$\{\ x = 2\ \} \rightarrow \{y = 2\}$$
$$\{\ x = 8\ \} \rightarrow \{y = 2\}$$

(a) Perform a 3rd-order polynomial regression and sketch the result of line fitting.

(b) Given a test point $\{\ x = 9\ \}$ predict y using the polynomial model.

(c) Compare this prediction with that of a linear regression.

(Polynomial Regression, 3D data, Python)
Question 2:

(a) Write down the expression for a 3rd order polynomial model having a 3-dimensional input.

(b) Write down the **P** matrix for this polynomial given $\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$.

(c) Given $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, can a unique solution be obtained in dual form? If so, proceed to solve it.

(d) Given $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, can the primal ridge regression be applied to obtain a unique solution? If so, proceed to solve it.

(e)

(Binary Classification, Python)
Question 3:
Given the training data:

$$\{x = -1\} \rightarrow \{\ y = class1\ \}$$
$$\{x = \quad 0\ \} \rightarrow \{\ y = class1\ \}$$
$$\{x = 0.5\} \rightarrow \{\ y = class2\ \}$$
$$\{x = 0.3\} \rightarrow \{\ y = class1\ \}$$
$$\{x = 0.8\} \rightarrow \{\ y = class2\ \}$$

Predict the class label for $\{x = -0.1\}$ and $\{x = 0.4\}$ using linear regression with signum discrimination.

(Multi-Category Classification, Python)
Question 4:
Given the training data:

$$\{x = -1\} \rightarrow \{\ y = class1\ \}$$
$$\{x = \quad 0\ \} \rightarrow \{\ y = class1\ \}$$

$$\{x = 0.5\} \rightarrow \{\, y = class2\,\}$$
$$\{x = 0.3\} \rightarrow \{\, y = class3\,\}$$
$$\{x = 0.8\} \rightarrow \{\, y = class2\,\}$$

(a) Predict the class label for $\{x = -0.1\}$ and $\{x = 0.4\}$ based on linear regression towards a one-hot encoded target.

(b) Predict the class label for $\{x = -0.1\}$ and $\{x = 0.4\}$ using a polynomial model of 5^{th} order and a one-hot encoded target.

(Multi-Category Classification, Python)
Question 5 (continued from Q3 of Tutorial 2):
Get the data set "`from sklearn.datasets import load_iris`". Use Python to perform the following tasks.

(a) Split the database into two sets: 74% of samples for training, and 26% of samples for testing. Hint: you might want to utilize `from sklearn.model_selection import train_test_split` for the splitting.

(b) Construct the target output using one-hot encoding.

(c) Perform a linear regression for classification (without inclusion of ridge, utilizing one-hot encoding for the learning target) and compute the number of test samples that are classified correctly.

(d) Using the same training and test sets as in above, perform a 2^{nd} order polynomial regression for classification (again, without inclusion of ridge, utilizing one-hot encoding for the learning target) and compute the number of test samples that are classified correctly. Hint: you might want to use `from sklearn.preprocessing import PolynomialFeatures` for generation of the polynomial matrix.

Question 6

MCQ: there could be more than one answer. Given three samples of two-dimensional data points $\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 3 & 3 \end{bmatrix}$ with corresponding target vector $\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. Suppose you want to use a full third-order polynomial model to fit these data. Which of the following is/are true?

a) The polynomials model has 10 parameters to learn
b) The polynomial learning system is an under-determined one
c) The learning of the polynomial model has infinite number of solutions
d) The input matrix **X** has linearly dependent samples
e) None of the above

Question 7

MCQ: there could be more than one answer. Which of the following is/are true?

a) The polynomial model can be used to solve problems with nonlinear decision boundary.

b) The ridge regression cannot be applied to multi-target regression.

c) The solution for learning feature $\mathbf{X}$ with target $\mathbf{y}$ based on linear ridge regression can be written as $\hat{\mathbf{w}} = (\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}$ for $\lambda > 0$. As λ increases, $\hat{\mathbf{w}}^T\hat{\mathbf{w}}$ decreases.

d) If there are four data samples with two input features each, the full second-order polynomial model is an over-determined system.

EE2211: Spring 2023
Tutorial 6 (Additional Questions)

8. Show that the primal and dual forms of the ridge regression solution are the same, i.e., for any $\lambda > 0$,

$$(\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}_{d+1})^{-1} \mathbf{X}^\top \mathbf{y} = \mathbf{X}^\top (\mathbf{X}\mathbf{X}^\top + \lambda \mathbf{I}_m)^{-1} \mathbf{y}.$$

   You may use the Woodbury identity

$$(\mathbf{I} + \mathbf{U}\mathbf{V})^{-1} = \mathbf{I} - \mathbf{U}(\mathbf{I} + \mathbf{V}\mathbf{U})^{-1}\mathbf{V}.$$

9. Derive the solution to the following weighted ridge regression problem

$$\min_{\mathbf{w}} \quad (\mathbf{y} - \mathbf{X}\mathbf{w})^\top (\mathbf{y} - \mathbf{X}\mathbf{w}) + \lambda \|\mathbf{w}\|_{\mathbf{q}}^2,$$

   where $\|\mathbf{w}\|_{\mathbf{q}}^2 = \sum_{i=1}^d q_i w_i^2$ where $q_i > 0$ for all i.

10. Derive the solution to the following shifted ridge regression problem

$$\min_{\mathbf{w}} \quad (\mathbf{y} - \mathbf{X}\mathbf{w})^\top (\mathbf{y} - \mathbf{X}\mathbf{w}) + \lambda \|\mathbf{w} - \mathbf{v}\|^2,$$

   where $\mathbf{v} \in \mathbb{R}^d$ is a fixed vector.