EE655000 Machine learning HW2

TA: Kang-yu Liu, Chao-yi Peng, Jun-Rei Wu

asdfghj49888@gmail.com
sky135410@yahoo.com.tw
heathentw@gmail.com

Grading policy

- In the handwriting assignment, you need to provide detailed derivations. Partial points will be credited when a wrong answer is accompanied by correct reasoning.

- In the programing assignment, the code, test data and report should be compressed into a **ZIP** file and upload to iLMS website. Also, please write a Readme file to explain how to run your code and discuss characteristics in your report. The Report format is not limited.

- The programming language that can be used on this assignment includes Python and Matlab.

- Discussions are encouraged, **but plagiarism is strictly prohibited.**

Part1. Handwriting homework assignment

You can find the corresponding problems from the textbook.

1. (30 points)
   Exercise 4.7
2. (40 points)
   Exercise 4.13
3. (30 points)
   Exercise 4.14

Part2. Computer assignment

In this problem, you need to apply the Maximum Likelihood (ML) and Bayesian linear regression methods to train a linear model in order to predict the chance of being admitted to graduate admissions.

Data

Data, contained in the two csv files **Training_set.csv** and **Testing_set.csv**. In this problem is the prediction of being admitted to graduate admissions.
More detailed descriptions are given below:

- In the Training_set.csv total have 400 pieces of data
  Collum1: GRE Scores
  Collum2: TOEFL Scores
  Collum3: Research Experience (either 0 or 1)
  Collum4: Chance of Admit (ranging from 0 to 1)

- In the Testing_set.csv total have 100 pieces of data
  Collum1: GRE Scores
  Collum2: TOEFL Scores
  Collum3: Research Experience (either 0 or 1)
  Collum4: Chance of Admit (ranging from 0 to 1)

Feature Vector

In this problem, we utilize the Gaussian basis function and the Research Experience to form the feature vector, denoted as

$$\phi(\mathrm{x}) = [\phi_1(\mathrm{x}), \phi_2(\mathrm{x}), \ldots, \phi_P(\mathrm{x}), \phi_{P+1}(\mathrm{x}), \phi_{P+2}(\mathrm{x})]^T$$

where we place P Gaussian basis functions uniformly over the spatial domain with $P = O_1 \times O_2$, $x = (x_1, x_2, x_3)$ is the input data (the scores together with the Research Experience), and O_1 and O_2 denote the number of locations along the horizontal and vertical directions, respectively, that you choose for your model in the prediction. (That is, you need to discuss the impact of different choices of O_1 and O_2.)

More specifically, for $1 \leq k \leq P$, the Gaussian basis function is defined as

$$\phi_k(\mathrm{x}) = \exp\left\{-\frac{(x_1-\mu_i)^2}{2s_1^2} - \frac{(x_2-\mu_j)^2}{2s_2^2}\right\}, \quad \text{for} \quad 1 \leq i \leq O_1, 1 \leq j \leq O_2,$$

Where

$$k = O_2 \times (i-1) + j,$$

$$\mu_i = \left(\frac{x_{1\_max} - x_{1\_min}}{O_1 - 1}\right) \times (i-1), \quad \mu_j = \left(\frac{x_{2\_max} - x_{2\_min}}{O_2 - 1}\right) \times (j-1)$$

$$s_1 = \frac{x_{1\_max} - x_{1\_min}}{O_1 - 1}, \quad s_2 = \frac{x_{2\_max} - x_{2\_min}}{O_2 - 1}.$$

Finally, the last two components of the feature vector are $\varphi_{P+1}(x) = x_3$(Research Experience) and $\varphi_{P+2}(x) = 1$ (bias).

Problem

1) (80 points)

Please employ the linear model

$$y(\mathrm{x}, \mathrm{w}) = \sum_{j=1}^{P+2} w_j \phi_j(\mathrm{x})$$

to predict the chance of being admitted to graduate admissions given in the testing_set.csv.

Please use Maximum Likelihood to train the model. Then, use your trained linear model to predict the chance of admit and compute the squared error ($(y(\mathrm{x}) - t(\mathrm{x}))^2$) for each data in testing_set.

2) (20 points)

Please discuss the impact of different choices of O_1 and O_2 and results in your report.