EE798L: Machine Learning in Wireless Communications **EE Dept. IITK**

MATLAB Assignment-1

Linear modelling - least squares and maximum likelihood approach (5 marks) *Jan. 2024*

In this tutorial, we will generate various plots given in the book. You will have to submit this assignment. Also, MATLAB codes for these questions (except Q3) are available online. You can learn from the online code, but write your own code.

1. (Fig. 1.9 of FCML) *Non-linear response from a linear model*: For the following model:

$$t = w_0 + w_1 x + w_2 x^2 + n,$$

the values of parameters are $w_0 = 1$, $w_1 = -2$ and $w_2 = 0.5$. Random noise n is distributed as $\mathcal{N}(0, 1)$.

   (a) *Data set*: Generate $N = 200$ values of x which are uniformly distributed between -5 and 5. Calculate t for all the x values.

   (b) Fit a linear and quadratic model to this data set using least squares approach.

2. (Fig. 1.9 of FCML) *Regularized least squares*: Consider the following model

$$t = 2x - 3 + n.$$

Random noise n is distributed as $\mathcal{N}(0, 3)$.

   (a) *Data set*: Generate six values of x which are uniformly spread between 0 and 1. Calculate t corresponding to all the x values.

   (b) Fit a fifth-order polynomial to this data set using regularized least squares approach for $\lambda = 0, 10^{-6}, 0.01$ and 0.1.

   (c) Verify that the behavior matches with the one discussed in the class.

3. (Prob 1.12 of FCML) Using K-fold cross-validation, and the value of λ that gives the best predictive performance on the Olympic men's 100m data for

   (a) a first-order polynomial (i.e. the standard linear model).

   (b) a fourth-order polynomial.

   Use the Olympic data given in Table 1.2 of the book for this problem. You could copy and save this data in a .mat file. *(Learn how to create .mat file in MATLAB from the web.)*

4. (Fig. 2.8 of FCML) *pdf of two-dimensional Gaussian random vector*: Plot the pdf expression with the following specifications:

   (a) $\boldsymbol{\mu} = [2, 1]^T, \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

(b) $\boldsymbol{\mu} = [2, 1]^T, \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$.

Generate the contour plots also. *Read about contour plots from the web.*

5. (Fig. 2.17 of FCML) *Predictive variance*: Consider the following model

$$t = 5x^3 - x^2 + x + n.$$

Random noise n is distributed as $\mathcal{N}(0, 300)$.

(a) *Data set*: Generate $N = 100$ values of x which are uniformly distributed between -5 and 5. Calculate t for all the x values.

(b) Fit linear, cubic and sixth-order models to this data set using maximum-likelihood approach.

(c) Plot the predictive error bars for all these models.

Please follow these Coding instructions:

- Properly comment your code.
- The code should execute and generate the desired output.
- Your submission should be self-contained (should include all the files required for running it).
- Avoid hard-coding the values of the variables for specific configurations. The code should be generic.

Please follow these submission instructions

- Deadline is 4th of Feb, 11:59 pm.
- All codes should be in one .zip folder. Please do not submit separate files.
- Upload your properly commented in drive link which will be provided to you. Name your code as rollno.zip.
- Please submit one final zip file.
- Please do not mail your file to me.

Please also read this carefully.

- Each one of you have to individually do all the MATLAB assignments. You can discuss with your friends but you will have to completely write your own code.
- Copying also means sharing your code with some else for them to copy. We will not differentiate between the two acts, and both such cases will be awarded zero. Our decision will be final.