

Assignment 2

EECS 4404/5327, Fall '20

The assignment is due on Monday, October 26, by the end of the day.

In this assignment you will implement linear least squares regression. For the data provided, set up the design matrix (using non-linear basis functions for polynomial curve fitting) and solve for the parameter vector \mathbf{w} as discussed in class. Print out your code and include it in your submission.

Step 1 - load the data

The data is stored in two files, `dataset1_inputs.txt` and `dataset1_outputs.txt` which contain the input values (i.e., values x_i) and the target values (i.e., values t_i) respectively. These files are simple text files which can be loaded with the load function in Matlab/Octave. Plot the outputs as a function of the inputs (ie plot the datapoints, not a curve) and include this plot in your write-up.

Step 2 - ERM

For degrees $W = 1, \dots, 30$, fit a polynomial of degree W to the data using (unregularized) least squares regression. For each learned function, compute the empirical square loss on the data and plot it as a function of W . Include this plot in your report. Given the curve, which value of W do you think would be suitable?

Step 3 - RLM

Repeat the previous step using regularized least squares polynomial regression. Each time train polynomial of degree 20 for regularization parameters λ so that $\ln(\lambda) = -1, -2, \dots, -20$. This time plot (and include) the empirical loss as a function of i . Compare and discuss the two curves you get for ERM and RLM.

Step 4 - cross validation

Implement 10-fold cross validation for ERM. That is, randomly divide that data into 10 chunks of equal size. Then train a model on 9 chunks and test on the 10th that was not used for training. For each model you train, average the 10 test scores you got and plot these again as a function of W . Which value of W do you think would be suitable?

Step 5 - visualization

For the degrees $W = 1, 5, 10, 20, 30$ plot the data along with the ERM learned models. Do the same for models learned with RLM with a fixed regularization parameter $\lambda = 0.0025$ (while varying the degree as for ERM). Discuss the plots. Which degree seems most suitable? What is the effect of adding the regularizer here?

(4 + 4 + 4 + 4 + 4 marks)