CPSC-340 Machine Learning and Data Mining 2004

Homework # 4

Due Friday, Nov 5 at 4pm.

NAME:______________________________

Signature:______________________________

STD. NUM: ______________________________

Email: ______________________________

General guidelines for homeworks:

You are encouraged to discuss the problems with others in the class, but all write-ups are to be done on your own.

Homework grades will be based not only on getting the "correct answer," but also on good writing style and clear presentation of your solution. It is your responsibility to make sure that the graders can easily follow your line of reasoning.

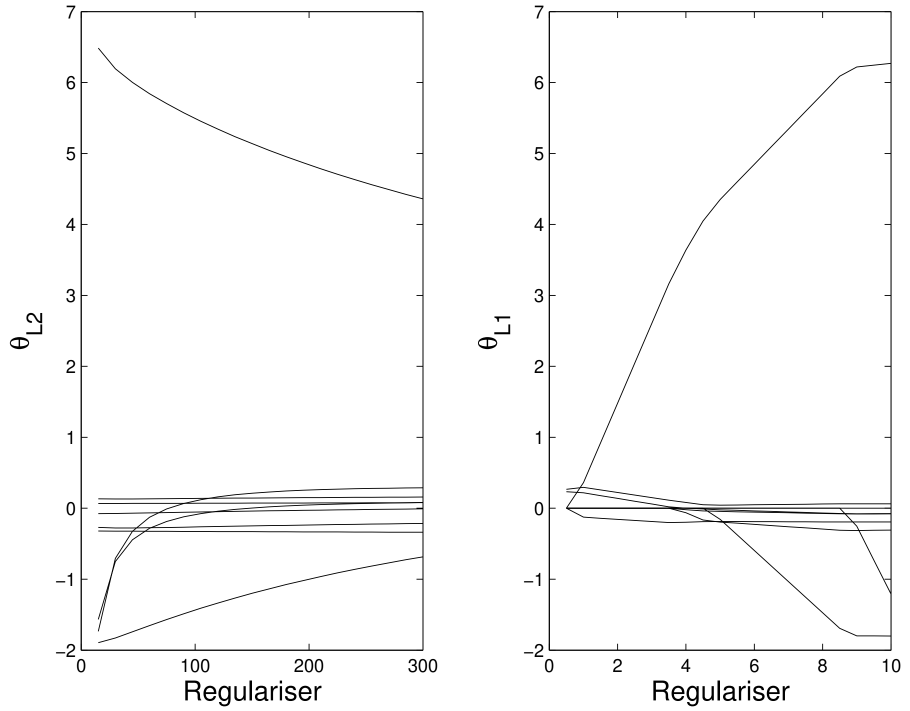
Try every problem. Even if you can't solve the problem, you will receive partial credit for explaining why you got stuck on a promising line of attack. More importantly, you will get valuable feedback that will help you learn the material.

Please acknowledge the people with whom you discussed the problems and what sources you used to help you solve the problem (e.g. websites, books from the library). This won't affect your grade but is important as academic honesty.

When dealing with Matlab exercises, please attach a printout with all your code and show your results clearly.

1. **(Linear supervised learning)**

   Download the Boston housing dataset from the course website. Repeat the experiment of homework 3, but this time instead of least squares, you should use both ridge regression and Lasso. You should repeat the experiment for different values of the regularisers δ^2 and t. Your code should be able to generate the following plots or θ against these regularisers.



   You should also use cross-validation to choose a good value of δ^2 and t. Hand in the plots of training and test set errors, your code, the figure as the one above, and the values of the regularisers that provide the best results in a min-max sense.

2. **(Nonlinear supervised learning)**

   Repeat the experiment using a nonlinear kernel and ridge regression or Lasso (choose your favorite). To guide you, you should download the files kerneldemo.m and so on. You should hand in the code and the plots of the train and test set errors.