CMPS242 Machine Learning

Spring 18 Warmuth

handed out Tu 4-3-18

Homework Assignment 1

due Tu 4-10-18 beginning of class

The goal of this homework is to teach you about over fitting and cross validation.

You are to implement regularized polynomial curve fitting.
Examples $(x_n, t_n),\ n = 1..N$
w_i is the weight for x_n^i, $(0 \le i \le 9)$

- $$\mathbf{w}^* \quad = \quad \text{argmin}_{\mathbf{w}} \left(\sum_n (\underbrace{\sum_{i=0}^{9} w_i\, x_n^i}_{y_n} - t_n)^2 + \lambda \sum_{i=0}^{9} w_i^2 \right)$$

  Let's rewrite the above in matrix notation: $\mathbf{X}$ is a $10 \times N$ matrix, where N is the number of examples and the nth column of $\mathbf{X}$ is $(1, x_n, x_n^2, \dots, x_n^9)^\top$. $\mathbf{t}$ is the vector of outputs $\mathbf{t} = (t_1, t_2, \dots, t_n)^\top$.

  $$\mathbf{w}^* = \text{argmin}_{\mathbf{w}} \left(\|\mathbf{X}^\top \mathbf{w} - \mathbf{t}\|^2 + \lambda \|\mathbf{w}\|^2 \right)$$

  This is solved by differentiation:

  $$\begin{aligned} \mathbf{X}(\mathbf{X}^\top \mathbf{w}^* - \mathbf{t}) + \lambda \mathbf{w}^* &= \mathbf{0} \\ (\mathbf{X}\mathbf{X}^\top + \lambda \mathbf{I})\mathbf{w}^* - \mathbf{X}\mathbf{t} &= \mathbf{0} \\ \mathbf{w}^* &= (\mathbf{X}\mathbf{X}^\top + \lambda \mathbf{I})^{-1} \mathbf{X}\mathbf{t} \end{aligned}$$

- We will provide a training set (train.txt) and a test set (test.txt) Each row contains one value of x and the corresponding value of t, separated by space. Use 10-fold cross validation to find the best choice of λ and report the loss on the test set.
- Write a roughly 3 page summary of what you did:
  - at least one plot
  - report the best value of λ and the test error.
- Note that you need to find a suitable way to discretize λ.

We encourage you to work in groups of up to three.
Every group is to do their own work. Same score for all members of a group.

Hint: First implement a simple split of the train.txt into train and validation set. After that implement 10-fold cross validation.

Extra credit:

- Modify the above solution so that the bias term/weight is not regularized. Justify your answer.

- Show variance bars for the losses of the different holdouts.
- Implement leave one out cross validation and show that in some sense this is better than 10-fold cross validation.

Remember you to look at the test score only once to report results!
Strong suggestion: Do this homework in Matlab, Octave or R.