CS 229 Machine Learning, spring 2019

Homework 1:

Regression

Due Monday February 11, 11:59pm

Submit by the **blackboard system**

***Requirements:***

1. Discussion is encouraged, but independent write-up and submission in your own hand is required for credit.
2. The answers copied from textbook or internet will be graded as a score of 0 point.
3. The work should be written in a \*clear\* way.
   1. The process of your solution must be unambiguously outlined or explicitly shown.
   2. The explanation and discussion must be well-founded by examples, tables, figures, and literatures and so on.
   3. The work should be submitted electronically in pdf/doc format. The scan of hand-written answers is acceptable.
4. Work submitted after the due date will be graded for correction, but not credited.

The goal of this homework is to become familiar with the gradient descent algorithms and maximum likelihood algorithm.

**Code:**

Write your code by any programming languages and submit your results together with the programs.

**Data:**

Generate a set of data samples (x,t) by yourself, e.g., t=f(x)+noise. x is a variable with real value in one dimension, and t is the target variable with real value in one dimension too. Function f(x) can be any non-linear function, e.g., sin(x), cos(x), or more complex ones, and random noises should be added.

Randomly choose 80% of the data for learning the regression function. Use the remaining 20% for evaluating the learned function.

**Regression basis function:**

Choose any kind of basis function, e.g., polynomial function. Try an appropriate number of basis functions.

**Task:**

1. (33 points) implement the ***batch gradient descent*** algorithm.
2. Show the decreasing of error function with the increasing of iteration numbers.
3. Give the results of obtained coefficient, **w**
4. Show the predicted ***f*(*x*)** when applying the learned regression function on testing data ***x***, and compare it with the corresponding actual target ***t*** on the same figure;
5. What is the Root-Mean-Square Error on test set?
6. (33 points) implement the ***stochastic gradient descent*** algorithm.
7. Show the decreasing of error function with the increasing number of used data points.
8. Give the results of obtained coefficient, **w**
9. Show the predicted ***f*(*x*)** when applying the learned regression function on testing data ***x***, and compare it with the corresponding actual target ***t*** on the same figure;
10. What is the Root-Mean-Square Error on test set?
11. (34 points) implement the ***maximum likelihood*** algorithm.
12. Give the results of obtained coefficient, **w**
13. Compare the value of **w** obtained at (1) and (2) by gradient descent algorithm, and (3) by maximum likelihood
14. Show the predicted ***f*(*x*)** when applying the learned regression function on testing data ***x***, and compare it with the corresponding actual target ***t*** on the same figure;
15. What is the Root-Mean-Square Error on test set? Which method has the best performance? Batch gradient descent, stochastic gradient descent or maximum likelihood?