Machine Learning, 2024 Spring
Assignment 4

Notice

Plagiarizer will get 0 points.
LaTeXis highly recommended. Otherwise you should write as legibly as possible.

Problem 1 For problem 3 in assignment 3, change your GD code to SGD and complete the tasks below:

- Present your code.
- How to choose (mini -) batch size?
- How to choose learning rate?
- How to terminate?
- Demonstrate the impact of different learning rates on the accuracy of the solution. In other words, your program should output an image similar to the image on page 34 of the Lecture 6-SGD PPT.

Solution:
0. Data preparation: The data is the same as the data in assignment 3.: The data is divided into training set and validation set. The training set is used to train the model, and the validation set is used to evaluate the model. The data is normalized to make the data have zero mean and unit variance. The data is shuffled to make the data have no order.

1. The code and the method to run the code are all in the folder 'code'.

2. The mini-batch size is set to be 1 as we are applying the stochastic gradient descent.

3. The learning rate η is tried with the following values: $0.1, 0.01, 0.001, 0.0001$. The learning rate for each step is similarly with the GD method, i.e. set $\eta_t = \eta\|\nabla e_t\|$. The learning rate is chosen by the learning rate schedule. From the results show in Figure **??**, we can see that the learning rate $\eta = 0.1$ is too large, which makes the loss diverge. The learning rate $\eta = 0.01$ is also too large, which makes the loss oscillate. The learning rate $\eta = 0.001$ is a good choice, which makes the loss converge. The learning rate $\eta = 0.0001$ is too small, which makes the loss converge slowly. So we finally choose the learning rate $\eta = 0.001$. The accuracy in the validation set.

4. The termination condition is chosen by the number of iterations. The number of iterations should be large enough to make the loss converge.

5. The impact of different learning rates on the accuracy of the solution is shown in Figure **??**.