Firstly, we determined how often should we change the learning rate: suppose the learning rate change every T iteration. Then in experiment, we set the value of epsilon to 10^-6, the value of eta to 1, then we get the result in Figure:

We found that MSE does not affected by the value of T, but running time is the lowest when T equals to 10^2, therefore, we set the value of T to 10^2

Then, we determined the value of epsilon, MSE does not change much after epsilon is less than 10^-5. However, the running time is increasing as the value of epsilon decreasing. Therefore, we set the value of epsilon to 10^-5.

Finally, we determined the value of eta, while eta increasing, MSE is decreasing. The maximum eta is 1, therefore we set the value of eta to 1.