CIS 419/519: Homework 2

Jiatong Sun

Although the solutions are entirely my own, I consulted with the following people and sources while working on this homework: $JunfanPan$
$https://machinelearningmastery.com/understand-the-dynamics-of-learning-rate-on-deep-learning-neural-networks/$
$https://en.wikipedia.org/wiki/Learning_rate$
$https://machinelearningmastery.com/how-to-tune-algorithm-parameters-with-scikit-learn/$.

1 Gradient Descent

a. The implication of the learning rate α_k is to control how big a step should be taken in the gradient descent direction towards the minimum, where a too small α_k may result in a long training time and a too large α_k may lead to an overshooting training process.

b. The implications of setting α_k as a function of k is to select an adaptive learning rate based on the training process, since the best step to take can vary as the the training goes gradually towards the minimum and a preset constant α_k may not work well in the whole process.

2 Linear Regression [CIS 519 ONLY]

3 Polynomial Regression

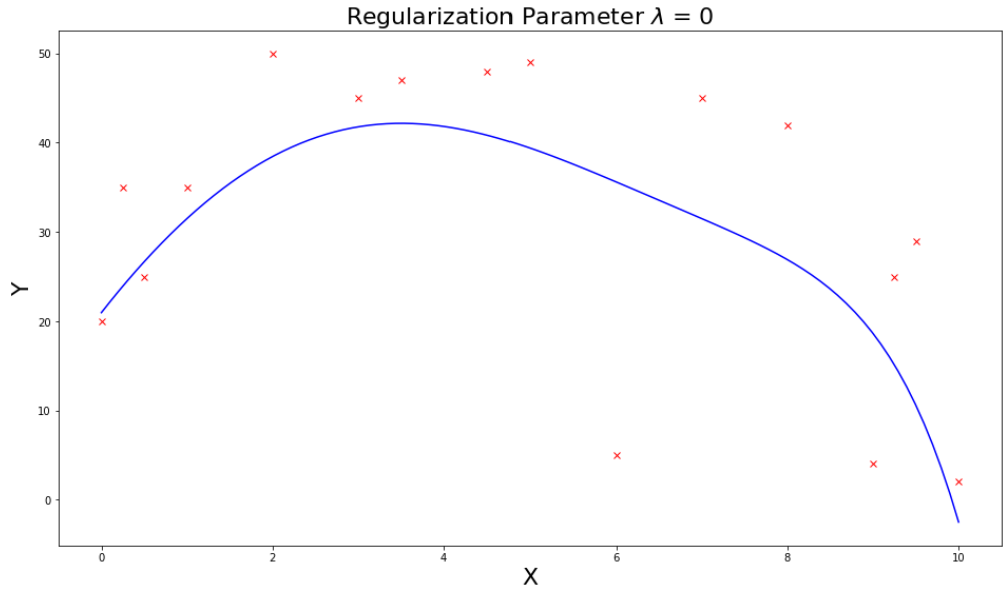

Figure 1: $\lambda = 0$

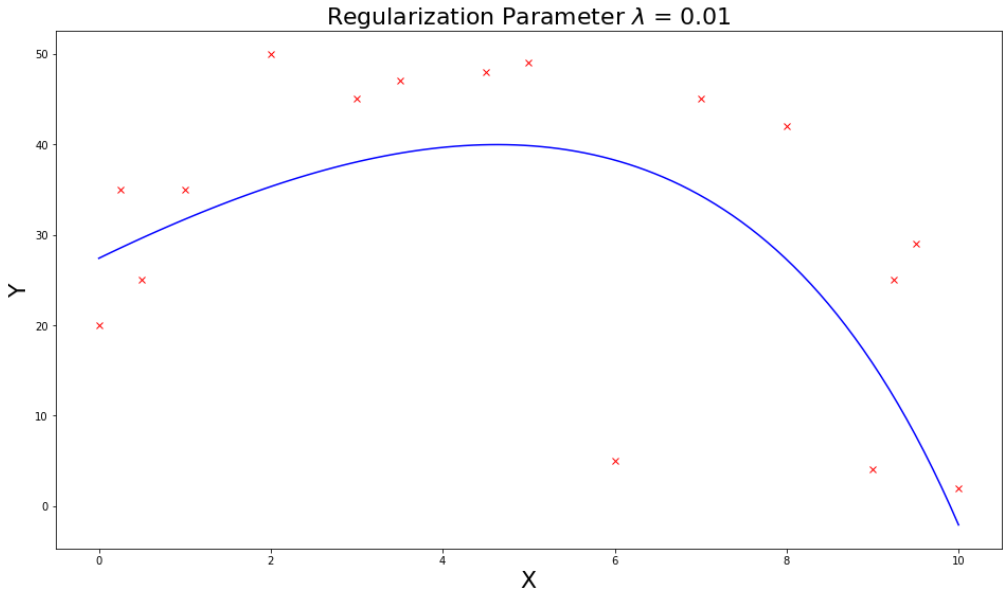

Figure 2: $\lambda = 0.01$