Lab 3 - Optimise it

Abhilash Pulickal Scaria
aps1n21@soton.ac.uk

Introduction

The performance of different optimisers like SGD, SGD with momentum, Adam and Adagrad on Rastrigin function optimisation is observed. In the second part a soft margin SVM is applied to iris dataset and its parameters are optimised using SGD and Adam. Observations in both cases are noted.

1 Exploring optimisation of analytic functions

1.1 Rastrigin

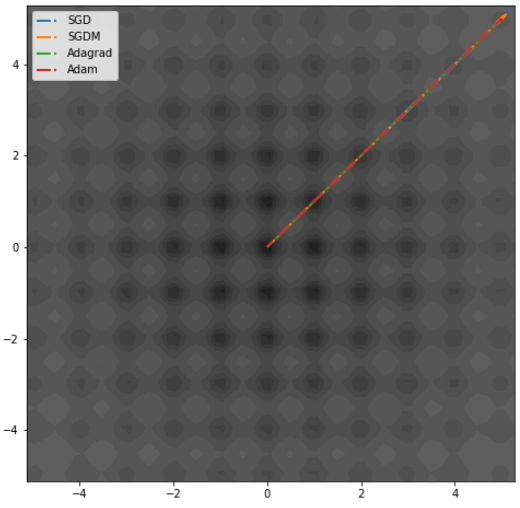
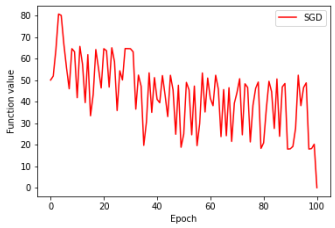
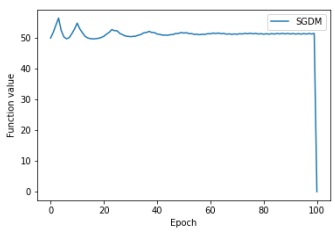


Figure 1: Trajectory of different optimisers.

From figure 1 we can observe that all the optimisers converge to (0,0) which is the global minima for Rastrigin function.

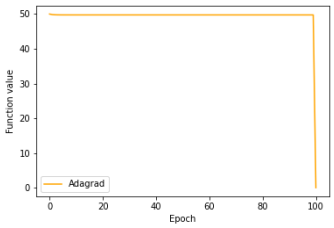


SGD Loss plot

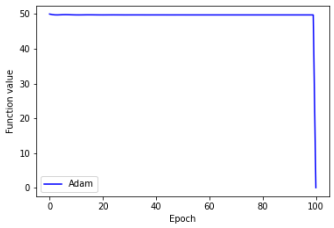


SGD + Momentum Loss plot

Figure 2: Loss plot of SGD and SGD with momentum



Adagrad Loss plot



Adam Loss plot

Figure 3: Loss plot of Adagrad and Adam

From the Loss plots we can observe that all of them converges to minima at similar epochs. Adam and Adagrad are stable throughout the epochs. SGD and SGD + Momentum oscilates a lot during the epochs. Taking stability into account we can say Adam and Adagrad are the best choices for optimising Rastrigin function

2 Optimisation of a SVM on real data

2.1 Iris SVM

The mean validation accuracy obtained after training a soft margin linear SVM on Iris dataset using Adam and repeating it 20 times was 87.6. For SGD the mean validation accuracy was 90. Adam works based on a combination of SGD with momentum and RMSP algorithm. It is supposed to be an improved version of SGD but still it was observed to have less validation accuracy than SGD.