

Lab 3 Exercise - Optimise it!

Amishapriya Singh
Student id: 33333904
as6u21@soton.ac.uk

1. Exploring optimisation of analytic functions

1.1. Rastrigin

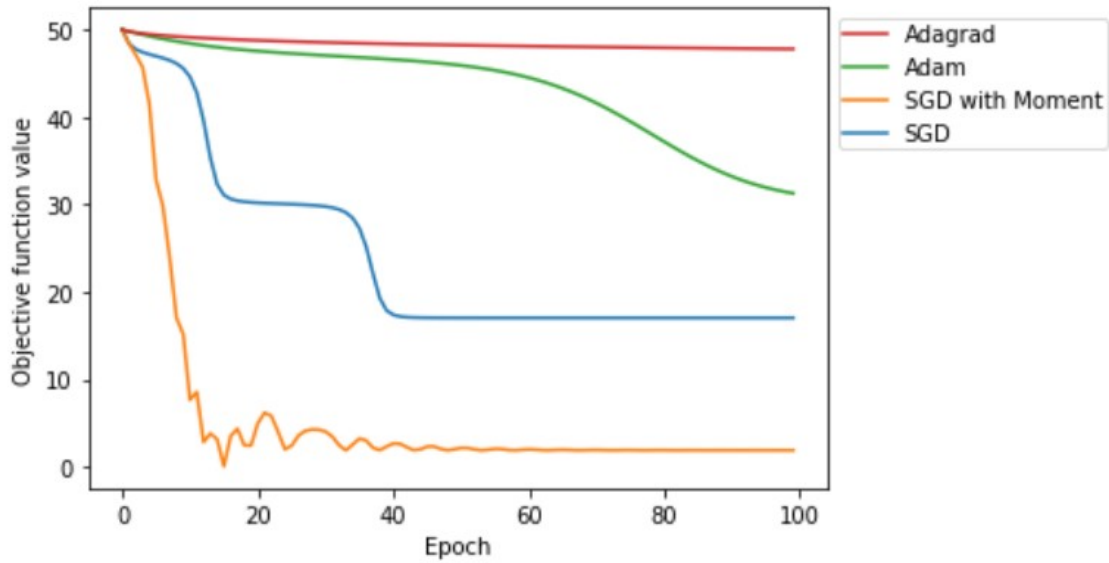


Figure 1. Loss plot of function value at each epoch

We can see from Fig. 1 that of the four optimisers, SGD with momentum reaches the lowest minimum function value and therefore gives the best performance. This could be explained as the momentum helping accelerate gradients vectors in the right directions, thus leading to faster convergence. Also, SGD with momentum is likely to skip over local minima as it would not stop at minima but rather move through it if the direction of momentum does not change while passing through the critical point.

Adagrad	Adam	SGD with moment	SGD
47.81	31.30	1.90	17.05

Table 1. Starting at $[5.0, 5.0]$, minimum value reached in 100 iterations by each optimiser.

2. Optimisation of a SVM on real data

2.1. Iris SVM

Validation accuracy for Adam: 0.88

Validation accuracy for SGD: 0.92

If we run the two optimisers 20 times, we see that the two methods give the same validation accuracy at times (in different iterations). If we take expected value over these 20 iterations, the accuracy for SGD comes as 0.91 and that of Adam is 0.89, indicating that SGD generalises better than Adam.