One Fourth Labs

Summary and what next

What have we learned this chapter

1. We have seen two new update rules, namely Momentum based gradient descent and Nesterov Accelerated Gradient Descent
2. These each mitigate some of the shortcomings of vanilla gradient descent
3. MGD allows for faster movement at plateau regions, thereby saving a lot of time/epochs
4. However, MGD can be a bit wasteful as it approaches the minima valley and oscillates till it stops
5. This flaw was remedied using NAG, whereby the oscillations near the minima valley are drastically reduced
6. NAG offers a good improvement to MGD