

深層学習 day 2 Section 2 学習率最適化手法

Q) モメンタム, AdaGrad, RMSProp, Adam の特徴を
まとめよ。

A) モメンタム

- ・ 局所最適解にはならず、大域的最適解となる。
- ・ 谷間についてから最適値に行くまでの時間が早い。

○ AdaGrad

- ・ 勾配の穏やかな斜面に対して、最適値に近づいた。
- ・ 学習率が徐々に小さくなるので、鞍点問題が
起きる。

○ RMSProp (AdaGradの改善版)

- ・ 局所最適解にはならず、大域的最適解となる。
- ・ ハイパーパラメータの調整が必要ない場合が多い。

○ Adam は、良く利用されており、

- ・ モメンタムの過去の勾配の指数関数的減衰平均
と、RMSProp 1. 2. の
のメリットを孕んだ最適化アルゴリズム