

周报

刘精昌

2017 年 2 月 10 日

本周工作

1. 结合提出 SAGA、SVRG 的经典论文，继续阅读 Bottou 等人的《Optimization Methods for Large-Scale Machine Learning》Noise Reduction Methods、Other Popular Methods 两部分。SGD 方法每次迭代用单个梯度替代全梯度，因为单个梯度并不是全梯度，也就是单个梯度替代全梯度有一定的 noise。这种 noise 的存在，导致 SGD 的收敛慢，于是可以考虑 reduce noise。Paper 中总结了三种 noise reduce 的方法：Dynamic Sample Size Methods、Gradient Aggregation、Iterate Averaging Methods。其中 SAGA、SVRG 这些经典的 SGD 变种就属于从 Gradient Aggregation 角度 reduce noise。
2. 和亦铨师兄联系。浏览了他提供的两篇 paper。一篇主要是解决 model 异步并行的 $f(x) + g(x)$ 优化问题。下面会继续看。

后续计划

- 亦铨师兄推荐的那两篇文章相关。