

周报

刘精昌

2016 年 12 月 16 日

本周工作

1. 看完《Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers》(ADMM), 看了一点《convex optimization》内容。
2. 看了 Bach 等在 NIPS16 报告的 tutorial《Stochastic optimization: Beyond stochastic gradients and convexity》, 主要是对最近优化工作的总结, 有较大启发。目前的优化关注于 4 个问题:
 - (a) 有限次求和问题, 主要工作是各种 SGD 的推广: SAG、SDCA、SARG、SVRG、ADAM 等
 - (b) 有限次求和加正则化项, 主要是 prox 相关算法。
 - (c) 非凸优化问题, 和凸优化问题比较本质的区别是证明收敛性等时, 凸优化中的一些性质不能直接使用, 且证明目标和凸优化有一定区别。
 - (d) 分布式优化问题, 近年来有很多工作围绕 SARG、SVRG 等的分布式展开。
3. 看了关于两篇 SGD 变形的 papers: 《A Stochastic Gradient Method with an Exponential Convergence Rate for Finite Training Sets》(SAG) nips12 和《Accelerating Stochastic Gradient Descent using Predictive Variance Reduction》(SVRG) nips13, SAG 和 SVRG 提升了 SGD 的收敛率。比较经典的两篇文章, 把 SGD 的次线性收敛率提升到线性, SVRG 相对于 SAG 的改进在于不需要再记录下之前的梯度, 论文里说它更适合神经网络训练。

4. 准备本周组会报告。

想了下《Distribution multi-task learning》AISTST16，认为该工作解决 MTL 问题有一定的针对性，不 general，而且不能做到隐私保护。

下周计划

- 看其他关于 SGD 改进文章。
- 考虑分布式 MTL 问题，考虑请教师兄，听听师兄看法。
- 花时间应对考试
- 发周报前突然看见《Distribution multi-task learning》作者挂在 Arxiv 上的文章 Distributed Multi-Task Learning with Shared Representation》，里面提到了很多方法，包括 ADMM，周五仔细看看。