

周报

刘精昌

2016 年 12 月 22 日

本周工作

1. 看了《Distributed Multi-Task Learning with Shared Representation》，主要关于分布式多任务学习。《An Empirical Study of ADMM for Nonconvex Problems》NIPS16 opt workshop，较短，主要是 ADMM 相关实验结果。
2. 课程作业、复习备考相关。

对于《Distributed Multi-Task Learning with Shared Representation》这篇 paper，比较了一些用来解决 MTL 分布式优化问题，如图1 所示。

Samples 这一列表示为了达到 ϵ 的精度，每个 worker 需要的采样数目，牵扯到 learning theory 中的 PAC 学习。Communication 这一列表示迭代的每轮中，通信次数。ERM 表示 workers 在该算法中做经验风险最小化的工作，Gradient Comp 表示 worker 做的是计算梯度的工作。

Approach	Samples	Rounds	Communication	Worker Comp.	Master Comp.
Local	$\frac{A^2}{\epsilon^2}$	1	0	ERM	0
Centralize	$\frac{A^2}{\epsilon^2} \left(\frac{r}{m} + \frac{r}{p} \right)$	1	$\frac{A^2}{\epsilon^2} \left(\frac{r}{m} + \frac{r}{p} \right)$	0	Nuclear Norm Minimization
ProxGD	$\frac{A^2}{\epsilon^2} \left(\frac{r}{m} + \frac{r}{p} \right)$	$\frac{mHA^2}{\epsilon}$	$2 \cdot p$	Gradient Comp.	SV Shrinkage
AccProxGD	$\frac{A^2}{\epsilon^2} \left(\frac{r}{m} + \frac{r}{p} \right)$	$\sqrt{\frac{mHA^2}{\epsilon}}$	$2 \cdot p$	Gradient Comp.	SV Shrinkage
ADMM	$\frac{A^2}{\epsilon^2} \left(\frac{r}{m} + \frac{r}{p} \right)$	$\frac{mA^2}{\epsilon}$	$3 \cdot p$	ERM	SV Shrinkage
DFW	$\frac{A^2}{\epsilon^2} \left(\frac{r}{m} + \frac{r}{p} \right)$	$\frac{mHA^2}{\epsilon}$	$2 \cdot p$	Gradient Comp.	Leading SV Comp.
DGSP	—	$\frac{mHA^2}{\epsilon}$	$2 \cdot p$	ERM	Leading SV Comp.
DNSP	—	—	$2 \cdot p$	ERM	Leading SV Comp.

图 1: MTL 分布式优化问题的比较

下周计划

- 看其他关于 SGD 改进文章。
- 看看周师姐今天报告的第二篇 paper，考虑能不能在 View learning 中，考虑多个 View 的情况。以及放松对 views 之间条件独立的 assumption，对相关设置个容忍度。
- 花时间应对考试。