

On Realizing Alternating Minimization Algorithm with TensorFlow

Hsin-Yu Chen, Chao-Hsiang Hung, Kan-Lin Hsiung

Abstract—In this brief note, an implementation of alternating minimization algorithm (AMA) using TensorFlow is reported.

Keywords—distributed optimization, AMA, TensorFlow

I. Introduction

Optimization algorithms solve problems in a wide variety of application domains. However, a lot of problems of interest are too large to be solved with traditional interior point methods. TensorFlow, developed by Google with the purpose of conducting machine learning and deep neural networks research, is general enough to be applied for solving various large-scale distributed optimization problems. In [4], a realization of AMA over Apache Spark framework has been reported. In this note, we briefly report an implementation of AMA using TensorFlow 2.0 [1].

II. Alternating Minimization Algorithm

Consider optimization problems of the following form with optimization variables u and v [2].

$$\begin{aligned} & \text{minimize} && H(u) + G(v) \\ & \text{subject to} && Au + Bv = b \end{aligned} \quad (1)$$

Note that both H and G are assumed to be (closed) convex functions. In many applications, optimization problems of the form (1) are frequently encountered. If H and G have an exploitable structure, alternating direction methods, which solve the *coupled* problem (1) by using *uncoupled* steps of addressing H and G alternatively, are often easy to implement and computationally efficient.

One of the popular alternating direction methods is Tseng's alternating minimization algorithm (AMA) [3], which assumes strong convexity for one of the objectives and alternately minimizes over the Lagrangian with respect to u and then the augmented Lagrangian with respect to v .

Alternating Minimization Algorithm

Require: $\lambda_0, \tau > 0$

- 1: **for** $k = 0, 1, 2, \dots$ **do**
 - 2: $u_{k+1} = \operatorname{argmin}_u H(u) + \langle \lambda_k, -Au \rangle$
 - 3: $v_{k+1} = \operatorname{argmin}_v G(v) + \langle \lambda_k, -Bv \rangle + \frac{\tau}{2} \|b - Au_{k+1} - Bv\|^2$
 - 4: $\lambda_{k+1} = \lambda_k + \tau(b - Au_{k+1} - Bv_{k+1})$
 - 5: **end for**
-

III. Numerical Results

Our AMA solver implemented using TensorFlow was tested over an AWS cluster. Five “c5.4xlarge” instances were used as workers with 16 vCPUs and 32 Gb of memory each. Dense Lasso problems [5] of the following standard form

$$\text{minimize} \quad \|Ax - b\|_2^2 + \mu \|x\|_1$$

with various problem sizes were randomly generated. In each problem, the matrix A had 10^3 columns and we varied the number of rows from 10^3 to 10^7 . In our preliminary numerical experiments, the computation time for solving each randomly generated problem ranged from a few seconds to 1 minute or so.

References

- [1] E. Bisong, “TensorFlow 2.0 and Keras,” *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, pp. 347–399, 2019.
- [2] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, “Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers,” *Foundation and Trends in Machine Learning*, 3(1), pp. 1–122, 2011.
- [3] P. Tseng, “Applications of Splitting Algorithm to Decomposition in Convex Programming and Variational Inequalities,” *SIAM Journal on Control and Optimization*, vol. 29, pp. 119–138, 1991.
- [4] H.-Y. Chen, K.-M. Huang, K.-L. Hsiung, “On Realizing Alternating Minimization Algorithm with Apache Spark,” in *Proceedings of the International Conference on Big Data, Internet of Things, and Zero-Size Intelligence (BIZ2015)*, pp. 1–2, 2015.
- [5] R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1996.

H.-Y. Chen, C.-H. Hung, K.-L. Hsiung, Department of Electrical Engineering, Yuan Ze University, TAIWAN.