

ADMM을 이용한 병렬 학습 시스템 최적화

김민우^{0*}, 임환희^{*}, 이병준^{*}, 김경태^{*}, 윤희용^{**}

^{0*}성균관대학교 정보통신대학 전자전기컴퓨터공학과

^{**}성균관대학교 소프트웨어대학 소프트웨어학과

e-mail: {kimmw95, lhh423, byungjun}@skku.edu^{0*}, kyungtaekim76@gmail.com^{*}, youn7147@skku.edu^{**}

Parallel Learning System Optimization using ADMM

Min-Woo Kim⁰, Hwan-Hee Lim^{*}, Byung-Jun Lee^{*}, Kyung-Tae Kim^{*}, Hee-Yong Youn^{**}

^{0*}Dept. of Electrical and Computer Engineering, Sungkyunkwan University

^{**}Dept. of Software, Sungkyunkwan University

● 요약 ●

인공지능의 급격한 발전으로 빅 데이터의 활용이 증가되었지만 이로 인해 머신 러닝에서 일어나는 문제들 또한 해결해야할 과제이다. 본 논문에서는 이에 따라 초래되는 문제들 중 학습 데이터가 많아질 경우의 문제들을 방지하기 위해, 알고리즘의 수정 대신 병렬 처리 기반 시스템을 제안한다. 본 논문에서는 Alternating Direction Method of Multiplier(ADMM) 알고리즘을 소개하고 ADMM 기반의 최적화 기법을 적용하여 병렬 학습 시스템 최적화를 제안하였다.

키워드: 병렬 학습(parallel learning), 분할(segmentation), 머신러닝(machine learning), ADMM

I. Introduction

과거 인간의 뉴런의 정보처리 방식을 참고한 신경망 형태의 병렬 처리기법을 시작으로 딥러닝, 머신러닝 등이 발전하게 되었다. 본 논문에서는 머신러닝에서의 문제에 요구되는 다양한 제약 조건들 중 이를 처리하기 위한 최적화의 방법 중 하나로 Alternating Direction Method of Multiplier(ADMM)를 제시한다[1]. ADMM은 선형 제약조건을 효과적으로 처리 할 수 있으며 병렬 최적화 알고리즘으로도 사용된다. ADMM은 기존의 문제를 보다 최적화하기 쉽게 부분적으로 문제를 나눈 후 다시 결합하여 복잡한 기존의 문제를 해결하는 근사 알고리즘이다[2]. 본 연구에서는 제약이 없는 일반적인 최적화 문제를 해결하기 위한 방법에 대해 이야기 한다.

II. Preliminaries

1. Related works

기존의 연구되어진 ADMM은 머신러닝 뿐만 아니라 딥러닝 분야에 이르기 까지 광범위하게 활용되어지며 범용적인 딥러닝 방법의 분산처리를 위해 활용되고[3], 대표적으로는 Convolutional Neural Network(CNN)방법에서 필터, 채널, 레이어 층의 구조를 조정하기 위해 사용되었다[4]. 또한 분산 뉴런 네트워크 학습과 같은 여러 병렬 학습 접근 방식으로도 사용되었으며 이와 같은 복잡한 노드간의 연결 그래프 구조를 갖추고 있을 적용할 수 있다. 병렬 학습 시스템

최적화 방법에 대한 최근 동향들은 Hadoop 기반의 Skytree와 Mahout 등이 있다.

III. The Proposed Scheme

1. ADMM

1.1 병렬 학습 아키텍처

학습 방법에서 분산 병렬 모두 업무 처리 속도의 증가와 효율성을 높이기 위한다는 공동의 목표를 갖는다. 본 논문에서 제안된 ADMM의 병렬 학습 시스템은 기존의 분산 시스템으로 사용 되었던 목표와 큰 차이가 존재하지는 않다. 하지만 본 논문에서 제안된 병렬 연산은 다른 하위 작업들의 방해 요소가 없기 때문에 차별성을 갖는다. Fig 1은 본 논문의 기본으로 사용되는 병렬 학습 아키텍처이다.

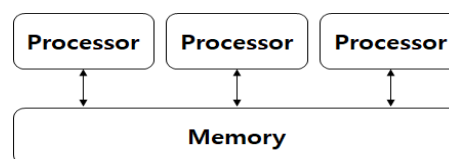


Fig. 1. 병렬 학습 아키텍처

각 프로세서들은 모든 머신러닝 모델의 변수들을 읽고 업데이트를 해야 한다. 이 과정에서의 결과를 통해 변수의 업데이트를 효율적으로 스케줄링 하게 되며 머신러닝의 학습 속도를 향상시킨다.

1.2 기본 원리

먼저 ADMM은 결합 된 제약 조건을 가지고 최적화 문제를 순차적으로 분해하는 고전적인 방법이다. Convex 최적화 방법을 사용하여 증강 라그랑지에 법칙을 기반으로 수식 화 되어 진다. 여기서 ADMM 은 분할을 위한 함수로 f, g 를 사용한다고 가정한다. x 개의 병렬로 나누어진 프로세스들과 z 의 데이터들을 의미하는 다음 식을 최소화하기 위한 변수로 $f(x) + g(z)$ 를 설정하고 $Ax + Bz = c$ 라는 식을 정의 한다. 설정되어진 식은 제약이 있는 최적화 문제와 동일하게 사용되어진다. 아래의 수식은 증강된 라그랑지에 법칙에 의거한 수식이며 변수 값 x 와 z 를 라그랑지에 법칙을 사용하여 동시에 업데이트해야 함을 의미한다.

$$L_{\rho}(x, z, y) = f(x) + g(z) + y^T(Ax + Bz - c) + \frac{1}{2} \| Ax + Bz - c \|_2^2$$

여기서 z 는 고정된 상태에서 x 를 최소화 한 후 최적화하기 때문에 z 를 분할할 수 있다.

IV. Conclusions

본 논문에서 제안된 ADMM 알고리즘을 바탕으로 병렬 학습 패러다임의 초기 단계인 하위 모델의 최적화 문제를 공식화 하였다. 향후 연구로는 병렬 프로그램을 구현시키고 Distributed deep Neural Network(DNN) 훈련에 대해 텍스트 분류와 손으로 쓴 숫자들의 인식, 자연어 처리 등을 기준으로 평가를 진행할 것이다. 실험을 통한 수치 결과는 수렴 속도와 최적화 결과 측면에서 효율성을 보여줄 것이다.

ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 정보통신-방송연구 개발 사업(No. 2016-0-00133, 초연결 IoT 노드의 군집 지능화를 통한 Edge Computing 핵심 기술 연구), SW중심대학지원사업(2015-0-00914), 한국연구재단 기초연구사업(No.2016R1A6A3A11931385, 실시간 공공안전 서비스를 위한 소프트웨어 정의 무선 센서 네트워크 핵심기술 연구, 2017R1A2B2009095, 실시간 스트림 데이터 처리 및 Multi-connectivity를 지원하는 SDN 기반 WSN 핵심 기술 연구), 삼성전자, BK21PLUS 사업의 일환으로 수행되었음.

REFERENCES

- [1] Ermin Wei and Asuman Ozdaglar, “Distributed Alternating Direction Method of Multipliers”, IEEE Conference on Decision and Control , 2012
- [2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers”, Foundations and Trends in Machine Learning, 2011.
- [3] S. Fujishige, “Submodular functions and optimization”, volume 58. Elsevier, 2005.
- [4] D. Povey, X. Zhang, S. Khudanpur, “Parallel training of DNNs with natural gradient and parameter averaging”, International Conference on Learning Representations(ICLR), 2015