

Vertical federative learning for distributed networks

Pyotr Lisov
Ivan Toropin

Table of contents

- Intro
 - Motivation
 - Problem Statement
- Related works
- Plans and expectations
- Bibliography

Problem statement

Задача распределенной оптимизации в общем виде:

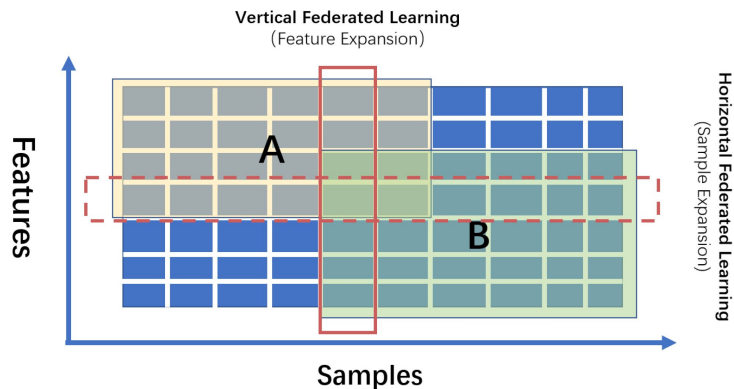
$$\min_{\Theta} \ell(\Theta; \mathcal{D}) \triangleq \frac{1}{N} \sum_{i=1}^N f(\Theta; \mathbf{x}_i, y_i)$$

\mathbf{x}_i, y_i - датасет. Θ - некая обобщенная модель

Motivation

Вертикальное обучение - подход к обучению моделей, при котором мы делим исходные данные на признаки. Это позволяет эффективно использовать данные, находящиеся в различных узлах сети, к тому же, такой подход является более безопасным.

(пример)



Related work

В [2] задача ставится так:

$$\min_{x \in \mathbb{R}^d} \left[f(x) := \frac{1}{s} \left\| \sum_{i=1}^n A_i x_i - b \right\|^2 \right]$$

Каждая нода имеет свой собственный набор фич A_i , x_i и вычисляет градиент независимо с помощью L-Katyusha, а затем передает его с помощью функции AllReduce[3].

Related Work

Работаем в 3 предположениях

- 1) μ сильная выпуклость
- 2) L гладкость
- 3) Собственные значения $AA^T \in [\mu, L]$

Related works

В данных предположениях достигается такая оценка на число итераций.

$$\mathcal{O} \left(\frac{s}{b_s} \left(\sqrt{\frac{L}{\mu}} \sqrt{\frac{1}{\beta} \left(\frac{1}{s} + (w-1) \frac{sL^2}{\mu^2} \right)} + 1 \right) \log \frac{1}{\varepsilon} \right)$$

Эта статья примечательна прежде всего тем, что она является одной из первых, которая дала гарантии теоретической сходимости в VFL, а не только в численных тестах

Plans and Expectations

Как можно улучшить результаты статьи?

- Доказать сходимость метода для нелинейной функции

В условиях общей задачи распределенной оптимизации авторами [2], в предположении несмещенности градиента не была доказана сходимость для нелинейного случая

Подробнее можно посмотреть в доказательстве леммы C.1 в [2]

Bibliography

- [1] Allen-Zhu, Z. (2016). Katyusha: The first direct acceleration of stochastic gradient methods
- [2] Anonymous Authors (2024). Accelerated Methods with Compression for Horizontal and Vertical Federated Learning
- [3] Kovalev, D., Horvath, S., and Richtarik, P. (2019). Don't Jump Through Hoops and Remove Those Loops: SVRG and Katyusha are Better Without the Outer Loop
- [4] Ernie Chan, Marcel Heimlich, Avi Purkayastha and Robert van de Geijn (2006). Collective communication: theory, practice, and experience. Wiley InterScience.