

Distributed optimization method based on L-Katusha and Gossip-mix

Ivan Toropin¹, Petr Lisov¹, Alexander Beznosikov¹

¹ MIPT

1 Abstract

Distributed optimization algorithms have emerged as a superior approaches for solving machine learning problems. To accommodate the diverse ways in which data can be stored across devices, these methods must be adaptable to a wide range of situations. As a result, two orthogonal regimes of distributed algorithms are distinguished: horizontal and vertical. During parallel training, communication between nodes can become a critical bottleneck, particularly for high-dimensional and over-parameterized models. Therefore, it is crucial to enhance current methods with strategies that minimize the amount of data transmitted during training while still achieving a model of similar quality. This paper introduces a new accelerated algorithm, working in the regime of vertical data division. By utilizing a momentum and variance reduction technique from the Loopless-Katyusha algorithm and Gossip procedure for communications, we provide one of the first theoretical convergence guarantees for the vertical regime.

2 Problem Statement

This paper considers optimization problems of the form

$$\min_{x \in \mathbb{R}^d} \left\{ f(x) := \frac{1}{n} \sum_{i=1}^n f_i(x) \right\},$$

3 Theoretical result

In our research we showed convergence with Lipschitz constant

4 Experiments

More information on Github