

Sign operator for (L0, L1) smooth optimization

Tuesday, 20 May 2025 13:18 (12 minutes)

In Machine Learning, the non-smoothness of optimization problems, the high cost of communicating gradients between workers, and severely corrupted data during training necessitate further research of optimization methods under broader assumptions. This paper explores the efficacy of sign-based methods, which address slow transmission by communicating only the sign of each stochastic gradient. We investigate these methods for (L_0, L_1) -smooth problems, which encompass a wider range of problems than the L -smoothness assumption. To address the problem of data accuracy, we introduce the convergence bounds for (L_0, L_1) -SignSGD and -M-SignSGD under heavy-tailed noise, defined as noise with bounded κ -th moment $\kappa \in (1, 2]$.

Primary author: ИКОННИКОВ, Марк (МФТИ)

Co-authors: БЕЗНОСИКОВ, Александр (МФТИ); КОРНИЛОВ, Никита (МИРТ)

Presenter: ИКОННИКОВ, Марк (МФТИ)

Session Classification: 20-Математическая оптимизация

Track Classification: Математическая оптимизация