Statistica Sinica **35** (2025), 1-19
doi:https://doi.org/10.5705/ss.202023.0022

REGULARIZED ADAPTIVE HUBER MATRIX REGRESSION AND DISTRIBUTED LEARNING

Yue Wang[1], Wenqi Lu[2], Lei Wang[2], Zhongyi Zhu[3],
Hongmei Lin*[4] and Heng Lian[1,5]

[1] *City University of Hong Kong,* [2] *Nankai University,* [3] *Fudan University,*
[4] *Shanghai University of International Business and Economics*
and [5] *City University of Hong Kong Shenzhen Research Institute*

Abstract: Matrix regression provides a powerful technique for analyzing matrix-type data, as exemplified by many contemporary applications. Despite the rapid advance, distributed learning for robust matrix regression to deal with heavy-tailed noises in the big data regime still remains untouched. In this paper, we first consider adaptive Huber matrix regression with a nuclear norm penalty, which enjoys insensitivity to heavy-tailed noises without losing the statistical accuracy. To further enhance the scalability in massive data applications, we employ the communication-efficient surrogate likelihood framework to develop distributed robust matrix regression, which can be efficiently implemented through the ADMM algorithms. Under only bounded $(1+\delta)$-th moment on the noise for some $\delta \in (0, 1]$, we provide upper bounds for the estimation error of the central estimator and the distributed estimator, and prove they can achieve the same rate as established with sub-Gaussian tails when only the second moment of noise exists. Numerical studies verify the advantage of the proposed method over existing methods in heavy-tailed noise settings.

Key words and phrases: Big data, communication-efficient, Huber loss, nuclear norm, robust matrix regression.

1. Introduction

Advances of modern technologies have made matrix-type data increasingly frequent in various applications, including image processing in computer vision, microarray gene study in medicine and asset allocation in economics (Rohde and Tsybakov, 2011; Senneret et al., 2016; Yang et al., 2016; Fan, Wang and Zhu, 2021). Although one intuitive idea is to reshape the matrix into a vector and apply popular vector-based regression methods, this may incur ultrahigh dimensionality and also destroy the inherent structure of matrix data such as the correlation between rows and columns. When considering matrix estimation, the rank plays an important part in constraining the model complexity, and the nuclear norm is a convex surrogate for rank (Candès and Tao, 2010). Indeed, the

*Corresponding author. E-mail: hongmeilin66@outlook.com