

Lucas Resende

IMPA (Instituto de Matemática Pura e Aplicada), Brazil
lucas.resende@impa.br

It is well-known that trimmed sample means are robust against heavy tails and data contamination. This poster presents the results in [1], where Oliveira and I analyzed the performance of trimmed means and related methods in two novel contexts. The first one consists of estimating expectations of functions in a given family, with uniform error bounds; this is closely related to the problem of estimating the mean of a random vector under a general norm. The second problem considered is that of regression with quadratic loss. In both cases, trimmed-mean-based estimators are the first to obtain optimal dependence on the (adversarial) contamination level. Moreover, they also match or improve upon the state of the art in terms of heavy tails. Experiments with synthetic data show that a natural “trimmed mean linear regression” method often performs better than both ordinary least squares and alternative methods based on median-of-means.

[1] Oliveira, Roberto I. and Lucas Resende. “Trimmed sample means for robust uniform mean estimation and regression.” (2023).

Joint work with Roberto Imbuzeiro Oliveira (IMPA).