

Qualifying Paper Report for Estimation of High Dimensional Mean Regression in the Absence of Symmetry and Light Tail Assumptions

Naitong Chen

November 14, 2022

1 Introduction

Linear regression is an easy-to-use and highly interpretable statistical inference method that has found itself in a wide range of applications. In the fields such as genomics and proteomics, we are often presented with high dimensional problems where the number of predictors is much greater than the number of observations. In such settings, linear regression in its simplest form may not have a unique solution, and so directly applying this method may not yield outputs that are interpretable for statistical inference. While a number of methods have been developed to address this issue, the corresponding theoretical guarantees on the inference quality for most of these methods rely on the assumption that the error distribution is symmetric and light-tailed. However, these assumptions may not be reasonable in practice, thus challenging the reliability of these methods. In this report, we discuss the regression estimator RA-lasso developed in [Fan et al. \(2017\)](#), which is designed to handle high dimensional data whose underlying error distribution may be neither symmetric nor light-tailed.

This report is organized as follows: for the remainder of this section, we more carefully motivate the RA-lasso estimator. In [Section 2](#), we present the proposed method along with some intuition on its theoretical guarantees as well as some practical considerations when applying this method. We comment on their simulation studies in terms of how well they justify the claims made in the paper in [Section 3](#) and reproduce their simulation studies in [Section 4](#). This report is concluded with some final discussions of RA-lasso in [Section 5](#).

2 Proposed method

2.1 Discussion on theoretical results

2.2 Practical implementation

3 Discussion on simulation

4 Reproduced and improved simulation studies

5 Conclusion

References

Jianqing Fan, Qufeng Li, and Yuyan Wang. Estimation of high dimensional mean regression in the absence of symmetry and light tail assumptions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(1):247–265, 2017.