Summary: We study a sparse negative binomial regression (NBR) for count data by showing the non-asymptotic advantages of using the elastic-net estimator. Two types of oracle inequalities are derived for the NBR’s elastic-net estimates by using the Compatibility Factor Condition and the Stabil Condition. The second type of oracle inequality is for the random design and can be extended to many \( \ell_1 + \ell_2 \) regularized M-estimations, with the corresponding empirical process having stochastic Lipschitz properties. We derive the concentration inequality for the suprema empirical processes for the weighted sum of negative binomial variables to show some high-probability events. We apply the method by showing the sign consistency, provided that the nonzero components in the true sparse vector are larger than a proper choice of the weakest signal detection threshold. In the second application, we show the grouping effect inequality with high probability. Third, under some assumptions for a design matrix, we can recover the true variable set with a high probability if the weakest signal detection threshold is larger than the turning parameter up to a known constant. Lastly, we briefly discuss the de-biased elastic-net estimator, and numerical studies are given to support the proposal.

MSC:

62-XX Statistics

Keywords:
de-biased elastic-net; empirical processes; high-dimensional count data regressions; oracle inequalities; sign consistency; stochastic Lipschitz condition

Software:
catdata; logcondiscr; COUNT

Full Text: DOI arXiv

References:

[39] Huiming Zhang School of Mathematical Sciences and Center for Statistical Science, Peking University, Beijing 100871, P.R. China. Department of Mathematics, Faculty of Science and Technology, University of Macau, Taipa, Macau, P.R. China.
[40] Jinzhu Jia School of Public Health and Center for Statistical Science, Peking University, Beijing 100191, P.R. China.

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.