An Adaptive Test on High-dimensional Parameters in Generalized Linear Models

Abstract

Significance testing for high-dimensional generalized linear models (GLMs) has been increasingly needed in various applications, however, existing methods are mainly based on a sum of squares of the score vector and only powerful under certain alternative hypotheses. In practice, depending on whether the true association pattern under an alternative hypothesis is sparse or dense or between, the existing tests may or may not be powerful. In this paper, we propose an adaptive test on a high-dimensional parameter of a GLM (in the presence of a low-dimensional nuisance parameter), which can maintain high power across a wide range of scenarios. To evaluate its p-value, its asymptotic null distribution is derived. We conduct simulations to demonstrate the superior performance of the proposed test. In addition, we apply it and other existing tests to an Alzheimer's Disease Neuroimaging Initiative (ADNI) data set, detecting possible associations between Alzheimer's disease and some gene pathways with a large number of single nucleotide polymorphisms (SNPs). We also implemented the proposed method in R package GLMaSPU that is publicly available on GitHub and CRAN.