A Robust Multifactor Dimensionality Reduction Method Using Residuals of the Survival Phenotype for Detecting Gene-gene Interactions

Seungyeoun Lee
leesy@sejong.ac.kr
Department of Mathematics and Statistics, Sejong University, Korea

Genome-wide association studies have successfully found genetic susceptibility to complex diseases such as hypertension, diabetes and cancer, but there is still a missing heritability problem since only a few proportion of genetic heritability could be explained. One of the methods for solving the missing heritability problem is to consider gene-gene interaction associated with the phenotype of interest since most of genome-wide association studies have been based on the single gene approach due to high-dimensional problem. The MDR method has been proposed to identify gene-gene interaction by reducing a high-dimensional genetic level into one dimensional binary level such as high and low risk groups for the case-control study. The main component of MDR is a classifier for distinguishing high and low risk groups. Originally the case-control ratio of each genetic combination is compared with the total ratio of case-control for the case-control study. However, a new classifier has been proposed for the quantitative trait under the generalized linear model by Lou et al. (2007) in which a case-control ratio is replaced by a residual score to classify multi-level genetic combination into high and low risk groups. Recently, the Cox-MDR method has been proposed for the survival phenotype by Lee et al. (2012), in which the martingale residual of a Cox model is used as a new classifier. In this paper, we propose a simple and robust approach based on the residuals of the survival phenotype. In Cox-MDR, each genotype is classified as high risk if the sum of the martingale residuals for those who have the corresponding genotype is positive, and as low risk, otherwise. However, the new approach transforms each individual into either a case or a control, depending on the sign of its martingale residual. Once all individuals are transformed into either case or control, the rest of the procedure is the same as that of MDR to identify gene-gene interaction. For the parametric regression model in survival analysis, the standardized residual score can be used for transforming all individuals into either case or control. In the proposed approach, the residual plays a role of transforming the survival phenotype into a binary phenotype. Since the residual is obtained from a regression model with adjusting covariates, this approach also has the same advantage as Cox-MDR. Simulation results show that the proposed approach performs similarly as Cox-MDR and is robust even when the censoring fraction is moderate. The proposed method for the survival phenotype can be easily implemented using the algorithm of MDR with adjusting for covariates and more robust to censoring fraction.