Summary: In high-dimensional longitudinal data with multinomial response, the number of covariates is always much larger than the number of subjects and when modelling such data, variable selection is always an important issue. In this study, we developed the penalized generalized estimating equation for multinomial responses for identifying important variables and estimation of their regression coefficients simultaneously. An iterative algorithm is used to solve the penalized estimating equation by combining the Fisher-scoring algorithm and minorization-maximization algorithm. We used a penalty term to regularize the slope part only because category-specific intercept terms should be included in the multinomial model. We conducted a simulation study to investigate the performance of the proposed method and demonstrated its performance using real dataset.

MSC:

62J12 Generalized linear models (logistic models)
62H12 Estimation in multivariate analysis
62P10 Applications of statistics to biology and medical sciences; meta analysis

Keywords: high-dimensional data; longitudinal data; multinomial response; minorization-maximization algorithm; smoothly clipped absolute deviation penalty; minimax concave penalty; variable selection

Software:
HandTill2001; Fahrmeir; SimCorMultRes

Full Text: DOI