A generalized estimating equation approach to multivariate adaptive regression splines.

Summary: Multivariate adaptive regression splines (MARS) is a popular nonparametric regression tool often used for prediction and for uncovering important data patterns between the response and predictor variables. The standard MARS algorithm assumes responses are normally distributed and independent, but in this article we relax both of these assumptions by extending MARS to generalized estimating equations. We refer to this MARS-for-GEEs algorithm as “MARGE.” Our algorithm makes use of fast forward selection techniques, such that in the univariate case, MARGE has similar computation speed to a standard MARS implementation. Through simulation we show that the proposed algorithm has improved predictive performance than the original MARS algorithm when using correlated and/or nonnormal response data. MARGE is also competitive with alternatives in the literature, especially for problems with multiple interacting predictors. We apply MARGE to various ecological examples with different data types. Supplementary material for this article is available online.

MSC:
62-XX Statistics

Keywords:
GEE; GLM; MARS; model selection; negative binomial regression; nonparametric smoothing

Software:
gamair; BayesTree; SemiPar; Glmulti; mvabund; STEPCEE; R; earth

Full Text: DOI

References:
[9] Chipman, H.; McCulloch, R., BayesTree: Bayesian Additive Regression Trees, R Package Version 0.3-1.2 (2015)