

Generalized additive latent variable modeling

Øystein Sørensen

*Center for Lifespan Change in Brain and Cognition, Department of Psychology,
University of Oslo, Norway, oystein.sorensen@psykologi.uio.no*

Generalized linear mixed models (GLMMs) are an essential tool when analyzing clustered data. However, both GLMMs and their nonlinear extensions require the parametric form of nonlinear effects to be specified a priori, and this is often not practical. For example, cognitive abilities across various domains follow distinctive trajectories across the lifespan, which are not easily parametrized. Generalized additive mixed models (GAMMs) are an excellent alternative in these cases, able to flexibly adapt to the underlying nonlinear shape. With multivariate response data, however, GLMMs and GAMMs are not able to project the responses onto a lower-dimensional space of latent variables. Structural equation models are ideal for this type of modeling, but also have important limitations in terms of incorporating explanatory variables, modeling multilevel data, and analyzing repeated measures data with irregular time intervals. Generalized linear latent and mixed models (GLLAMMs)[2] combine the best of both worlds by allowing latent variable modeling combined with the flexibility GLMMs. While GLLAMMs allow nonlinear effects in the explanatory variables, they also require the parametric form of the relationships to be explicitly formulated. We have therefore developed generalized additive latent and mixed models (GALAMMs), an extension of GLLAMMs in which both the linear predictor and latent variables may depend smoothly on observed explanatory variables, as in GAMMs. Utilizing the mixed model view of smoothing, we show that any GALAMM can be represented as a GLLAMM, with smoothing parameters estimated by maximum likelihood. This allows fitting GALAMMs using a profile likelihood approach developed for GLLAMMs[1]. We further show how standard errors and confidence bands of the estimated smooth functions can be computed, extending upon existing methods for GAMMs. The application motivating the development concerns how level and change of human cognitive function is correlated across cognitive domains, and how environmental and genetic factors affect the lifespan trajectories, and we show results of analyses using GALAMMs on a dataset with repeated high-dimensional measures relating to various cognitive domains in a dataset with 1850 participants between 6 and 93 years. The methods are implemented in the R package 'galamm', which will also briefly be demonstrated.

References

- [1] Jeon, M. and Rabe-Hesketh, S. (2012). Profile-Likelihood Approach for Estimating Generalized Linear Mixed Models With Factor Structures. *Journal of Educational and Behavioral Statistics*, **37**(4):518–542.
- [2] Rabe-Hesketh, S., Skrondal, A., and Pickles, A. (2004). Generalized multilevel structural equation modeling. *Psychometrika*, **69**(2):167–190.