DISCUSSION

Comments on: Inference and computation with Generalized Additive Models and their extensions

Thomas Kneib[1]

It is my pleasure to contribute to the discussion of Simon's paper on "Inference and computation with Generalized Additive Models and their extensions" which provides an excellent overview of the current state of the art of the class of Generalized Additive Models in a broad sense, i.e., including several modern developments such as functional effects, interaction surfaces or distributional regression. I particularly enjoyed the brief yet very informative summaries of inferential results and statistical computing where Simon takes a delightful pragmatic perspective by focusing on the applied and computational pros and cons of approaches such as penalized likelihood, Markov chain Monte Carlo simulations, integrated nested Laplace approximations or functional gradient descent boosting. I believe that such a pragmatic perspective is indeed required to bring recent advances concerning statistical modeling to applied scientists utilizing these modeling techniques.

Another necessity for the future success of extended generalized additive models, from my perspective, is considerably more work focusing on interpretation, visualization or uncertainty quantification for such models if these should be routinely used by applied researchers. In fact, already simple generalized linear models pose considerable challenges concerning interpretation. While in some cases ceteris paribus-type interpretations are still conceivable, these are usually restricted either to transformations of the expectation of the response (e.g., log odds in logistic regression or log expectations in Poisson regression) or to relative effects (e.g., on odds in logistic regression or the expectation in Poisson regression). While such relations are certainly relevant and can be interpreted correctly with enough care, they can also easily lead to misleading conclusions. For example, a significant multiplicative and therefore relative effect on the odds in logistic regression does not necessarily lead to a relevant effect on the actual probability for observing the event of interest, depending, for example, on the value of the intercept or the values of the other covariates consid-

✉ Thomas Kneib
tkneib@uni-goettingen.de

1 Chair of Statistics, Georg-August-Universität Göttingen, Göttingen, Germany

ered. For specific types of covariate combinations, this may easily lead to the situation that a significant relative effect leads to basically no absolute change in the success probability.

While this is certainly well known for generalized linear models, the situation gets considerably more complex in case of generalized additive models and their extensions. For example, it is a common practice (admittedly also by the author of this comment) to show the nonlinear additive effects $f_j(x_j)$ of a generalized additive model only on the predictor scale and centered around zero. It is then very tempting to identify regions where the corresponding covariate has a "positive" and a "negative" effect although this indeed depends very much on the values of the intercept and all other covariates. In fact, all additive components can only be interpreted ceteris paribus in terms of differences $f_j(x_{j1}) - f_j(x_{j1})$ where x_{j1} and x_{j2} are pre-chosen values of the covariate.

In models that comprise more than one predictor such as the "several smooth linear predictors" models discussed in Section 3.4 of Simon's paper, the situation gets even more complicated since the same covariate may impact several of the distributional parameters of the response distribution. As a consequence, it is rather difficult to judge the actual effect of a given covariate on the response distribution since differences in a covariate of interest may easily compensate or reinforce each other due to their effects on the different distributional parameters.

Given these issues, I believe that future applied research on generalized additive models and their extensions will have to develop appropriate visualization tools assisting the user in interpreting the effect of covariates on the response distribution and in checking the adequacy of the model. Furthermore, measures of effect relevance and possibilities to quantify uncertainty for such derived measures (or other complex functionals of the original model output) will certainly deserve more attention. From my perspective, Bayesian inference based on Markov chain Monte Carlo simulations will prove particularly useful at this point due to its ease in performing finite sample uncertainty assessments via sampling-based inference. A similar, yet asymptotic approach is to perform a parametric bootstrap based on the asymptotic normality of the regression coefficients where the bootstrap samples can also be plugged into complex transformations to achieve sample-based measures of uncertainty.

Two other aspects that I would like to comment on concern the posterior consistency of Bayesian quantile regression with asymmetric Laplace likelihood and the potential of other inferential approaches for generalized additive models and their extensions. For the former, Simon states that the consideration of the asymmetric Laplace distribution as a working likelihood "is invalid since the asymmetric Laplace is mis-specified as a probability model, and this mis-specification tends to become extreme as we move away from the median quantile." This is in contrast to the work of Sriram et al. (2013) who showed posterior consistency of Bayesian inference even under this mis-specification. While this is certainly only an asymptotic argument and only concerns concentration of the posterior around the true value and therefore does not cover uncertainty quantification, it still indicates that some sensible conclusions can be drawn from models estimated under the mis-specified asymmetric Laplace likelihood.

Under the topic of other inferential approaches, I would particularly be interested in hearing Simon's opinion on the ability of variational approximations for estimating complex generalized additive models. Waldmann and Kneib (2015) have utilized these for inference in Gaussian mean regression and quantile regression (again based on the working likelihood of the asymmetric Laplace distribution) where similar schemes as with Gibbs sampling in Markov chain Monte Carlo simulations can be derived. On the other hand, they also found that uncertainty quantification tends to be complicated using simple forms of variational approximations.

References

Sriram K, Ramamoorthi R, Ghosh P (2013) Posterior consistency of Bayesian quantile regression based on the misspecified asymmetric Laplace density. Bayesian Anal 8(2):479–504

Waldmann E, Kneib T (2015) Variational approximations in geoadditive latent Gaussian regression: mean and quantile regression. Stat Comput 25:1247–1263