Model Averaging for Linear Mixed Models via Augmented Lagrangian

René-Marcel Kruse and Benjamin Säfken*
Chair of Statistics, University of Göttingen

March 30, 2020

Abstract

Model selection for linear mixed models has been a focus of recent research in statistics. The method of model averaging, however, has been not in the centre of attention. This paper presents a weight finding method for model averaging of linear mixed models, as well as its implementation. However, the optimization of that criterion is a non-trivial task due to its non-linear characteristics and the constraints imposed. Therefore, this paper proposes to use the augmented Lagrangian optimization method which is a non-typical approach for the field of statistics. The characteristics of the approach as well as its implementation are discussed in detail. Additionally, we present a stable, robust and fast implementation of the algorithm, which, together with the weight-finding methodology provided, is released as an extension of the R-package **cAIC4**. Furthermore, the influence of the weight finding criterion on the resulting model averaging estimator will be discussed through simulation studies as well as an application to real data.

Keywords: Optimization, Augmented Lagrangian, **cAIC4**, Model Averaging, Linear Mixed Models, conditional AIC

*The authors gratefully acknowledge *please remember to list all relevant funding sources in the unblinded version*

introduction

1 motivation

2 Introduction

The class of linear mixed models (Henderson, 1950) is a very powerful and flexible analytic tool, that enjoys popularity especially for the analysis of clustered or longitudinal data (Laird and Ware, 1982; Verbeke and Molenberghs, 2009), for splines smoothing (Ruppert et al., 2003; Wager et al., 2007) and for functional data analysis (Di et al., 2009; Cederbaum et al., 2016). Due to the flexibility and therefore possible complexity of the models, the question of valid model selection procedures comes to the centre of attention. of the included random effects, this methodology suffers from limitations as methods such as likelihood-ratio tests can encounter potential boundary problems (Crainiceanu and Ruppert, 2004; Wood, 2013). Furthermore, the deviations from the regularity conditions of the linear mixed model pose a serious problem with the use of information criteria such as the commonly used Akaike Information Criterion (Akaike, 1973) for model choice (Wager et al., 2007). However, deviations from the regularity of the linear mixed model pose a serious problem with the use of information criteria such as the commonly used Akaike Information Criterion (Akaike, 1973) for model choice (Wager et al., 2007). Furthermore, evaluating the suitability of the included random effects of models with nested or clustered structures suffers from limitations as methods such as the likelihood-ratio test encounter boundary problems (Crainiceanu and Ruppert, 2004; Wood, 2013).

Vaida and Blanchard (2005), however, showed that it is also possible to derive an AIC from the conditional formulation, which in turn is particularly suitable for accounting for possible shrinkage in random effects. Liang et al. (2008) suggest a corrected version of the conditional AIC which incorporates the effect of estimating the variance parameter. Nevertheless, this proposed version is computationally intensive as it relies on numerical approximation. Greven and Kneib (2010) demonstrate that an analytical solution can be derived and thus minimize the computational intensity of the corrected version of the conditional AIC. Another approach to addressing model uncertainty is model averaging. Instead of choosing a single model from a list of candidate models based on information criteria such as AIC or BIC, an weighted average of the considered models is calculated and then used for analysis. Selecting the underlying weights for the averaged models is an important factor when dealing with model averaging. Different proposals have been brought forward, the most prominent one being the approach of information criteria based weights (Buckland et al., 1997). Yet a majority of proposals aim at classical linear models and do not encounter difficulties when applied to the model framework of linear mixed models. A proposal by Zhang et al. (2014) demonstrates that it is possible to construct an asymptotically optimal weight finding criterion for model averaging of linear mixed models based on the conditional AIC. One issue not addressed by the authors of the proposed weighting criterion is a computationally stable and fast minimization of the underlying target function. The non-linear form of the criterion itself, as well as the nature of the constraints in the form of simultaneous equality and inequality conditions, makes it necessary to resort to complex, advanced optimization methods that are not part of the basic version of the programming language R.

In this paper, we present an implemented version of the proposed weight finding criterion by Zhang et al. (2014) in the statistical programming language R (R Core Team, 2019). Furthermore, we describe the special non-linear optimization under equality and inequality constraints of the underlying problem. We illustrate the approach of solving the problem by applying the augmented Lagrangian method (Hestenes, 1969; Powell, 1969) and present our implementation of the solver which is a uniquely customized version of the augmented Lagrangian optimization method for optimization of our weight finding problem.

This paper is structured as follows: Section 2 introduces the theory and formulations of linear mixed models, as well as the estimation and the application of linear mixed models for spline smoothing. Section 3 presents the concept of the conditional AIC and its corrections, in addition this section also induces the concept of model averaging and the weight finding criterion proposed by Zhang et al. (2014), which is the most one important for our implementation. The following section 4 gives an introduction to the newly implemented functions of the **cAIC4** R-package (Säfken et al., 2018) and the underlying mathematical concept of the augmented Lagrangian method. Section 5 analyzises the properties of the implemented methods by applying them to three different simulation studies. The last section 6, gives a summary of the findings of the previous sections and also gives an outlook of further work concerning model selection and averaging for linear mixed models. linearmixedmodels

3 motivation

modelchoice

4 motivation

practicalma

5 motivation

simulations

6 motivation

applications

7 motivation

conclusions

8 motivation

appendix

9 motivation

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, proceedings of the 2nd international symposium on information. *Czaki, Akademiai Kiado, Budapest*.

Buckland, S. T., Burnham, K. P., and Augustin, N. H. (1997). Model selection: an integral part of inference. *Biometrics*, pages 603–618.

Cederbaum, J., Pouplier, M., Hoole, P., and Greven, S. (2016). Functional linear mixed models for irregularly or sparsely sampled data. *Statistical Modelling*, 16(1):67–88.

Crainiceanu, C. M. and Ruppert, D. (2004). Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(1):165–185.

Di, C.-Z., Crainiceanu, C. M., Caffo, B. S., and Punjabi, N. M. (2009). Multilevel functional principal component analysis. *The annals of applied statistics*, 3(1):458.

Greven, S. and Kneib, T. (2010). On the behaviour of marginal and conditional aic in linear mixed models. *Biometrika*, 97(4):773–789.

Henderson, C. R. (1950). Estimation of genetic parameters. In *Biometrics*, volume 6, pages 186–187. International Biometric Soc 1441 I ST, NW, Suite 700, Washington, DC 20005-2210.

Hestenes, M. R. (1969). Multiplier and gradient methods. *Journal of optimization theory and applications*, 4(5):303–320.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, pages 963–974.

Liang, H., Wu, H., and Zou, G. (2008). A note on conditional aic for linear mixed-effects models. *Biometrika*, 95(3):773–778.

Powell, M. J. (1969). A method for nonlinear constraints in minimization problems. *Optimization*, pages 283–298.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric regression*. Number 12. Cambridge university press.

Säfken, B., Rügamer, D., Kneib, T., and Greven, S. (2018). Conditional model selection in mixed-effects models with caic4.

Vaida, F. and Blanchard, S. (2005). Conditional akaike information for mixed-effects models. *Biometrika*, 92(2):351–370.

Verbeke, G. and Molenberghs, G. (2009). *Linear mixed models for longitudinal data*. Springer Science & Business Media.

Wager, C., Vaida, F., and Kauermann, G. (2007). Model selection for penalized spline smoothing using akaike information criteria. *Australian & New Zealand Journal of Statistics*, 49(2):173–190.

Wood, S. N. (2013). A simple test for random effects in regression models. *Biometrika*, 100(4):1005–1010.

Zhang, X., Zou, G., and Liang, H. (2014). Model averaging and weight choice in linear mixed-effects models. *Biometrika*, 101(1):205–218.