

Guest Editorial: Regularisation Methods in Regression and Classification

Gerhard Tutz

Received: 24 January 2010 / Accepted: 1 February 2010 / Published online: 2 March 2010
© Springer Science+Business Media, LLC 2010

Regularisation has become an important tool in statistical modelling. In particular the challenge of high dimensional data boosted the fitting of more and more complex models that can not be fitted without appropriate regularisation. The need for regularisation, however, is not restricted to the modelling of high dimensional data, it is mainly driven by the complexity of the model. When the model includes nonparametric function estimation regularisation restricts the class of functions that are fitted. In regression and classification complexity of the model is usually determined by the structuring of the predictor. Regularisation helps to identify the relevant parts, which can consist of simple linear terms, functions, parametric or nonparametric interaction terms, or complex spatially and temporally structured terms. Regularisation can be made explicit by using penalty terms that restrict estimates or can be implicitly determined by the algorithm as for example in boosting methods.

This issue of Statistics and Computing collects ten papers that focus on regularisation methods in different areas and with different methodology. Two papers are devoted to the extension of boosting techniques. In the paper “*Twin Boosting: Improved Feature Selection and Prediction*” P. Bühlmann and T. Hothorn propose a boosting method that consists of two rounds of common boosting with the second boosting process being forced to resemble the first round of boosting. The method shows much better feature selection behaviour than common boosting, in particular as far as false positives are concerned. The paper “*Estimation and Regularization Techniques for Regression Models with Multidimensional Prediction Functions*”

by M. Schmid, S. Potapov, A. Pfahlberg, T. Hothorn extends boosting to applications where the model is determined by more than one component, as for example in zero-inflated Poisson models where a model for the counts and a model for the inflation of zeros are combined.

Several papers focus on specific classes of models. Survival models are the topic in the papers “*Sparse regression techniques in low-dimensional survival data settings*” by C. Porzelius, M. Schumacher and H. Binder and “*Rank-based variable selection with censored data*” by J. Xu. In the first paper various methods of variable selection by regularisation are examined, including SCAD, lasso, elastic net and boosting methods. The second paper proposes regularisation methods for simultaneous estimation and variable selection in the accelerated failure time model. The modelling of two-dimensional predictors is considered in the papers “*Least squares and shrinkage estimation under bi-monotonicity*” by R. Beran and L. Dümbgen and “*Bilinear Modulation Models for Seasonal Tables of Counts*” by B. Marx, P. Eilers, J. Gampe, R. Rau. Beran and Dümbgen show how to estimate smooth functions under general order constraint, a method that also applies in image denoising. Marx et al. fit two-dimensional varying coefficient surfaces to age-time tables of seasonal counts by using a parsimonious representation of time modulated effects.

The Bayesian approach to regularisation in structured additive regression is outlined in the survey paper “*Bayesian Regularisation in Structured Additive Regression*” by L. Fahrmeir, T. Kneib and S. Konrath. The authors develop a unifying perspective on shrinkage, smoothing and predictor selection based on conditionally Gaussian priors. The method works for additive regression in a rather general sense: the predictor can include linear, nonlinear smooth, interaction surfaces, spatial and random effects as well as varying coefficients. The second paper that advo-

G. Tutz (✉)
Ludwig-Maximilians-Universität München, Akademiestraße 1,
80799 Munich, Germany
e-mail: tutz@stat.uni-muenchen.de

cates the Bayesian approach is “*Model uncertainty and variable selection in Bayesian lasso regression*” by C. Hans. He proposes and investigates analytic and computational approaches for handling uncertainty in the special case of Bayesian Lasso.

Finally two papers explicitly deal with prediction issues. G. Obozinski, B. Taskar and M. Jordan address the problem of recovering a common set of covariates that are relevant simultaneously to several classification problems. The methods are extended to the more broader problem of subspace selection which is closely related to sufficient dimension reduction.

In the paper “*Sparse Conformal Predictors*” M. Hebiri shows how to construct prediction intervals for the response variable in multivariate linear models that are sparse. The method combines the principle of conformal prediction and lasso type penalties.

I want to express my gratitude to the authors and the referees who helped to improve the papers. Moreover, I want to thank the current editor-in chief, Gilles Celeux, for the opportunity of publishing the collected papers. I hope that the readers will enjoy this issue of Statistics and Computing.