Recursive identification of smoothing spline ANOVA models

Marco Ratto, Andrea Pagano *
European Commission, Joint Research Centre, Ispra, ITALY

February 25, 2009

Keywords: Smoothing spline ANOVA models, recursive algorithms, backfitting, sensitivity analysis.

In this paper we present a unified discussion of different approaches to identification of smoothing spline ANOVA models. The 'classical' approach to smoothing spline ANOVA models can be referred to in the line of Wahba (1990) and Gu (2002). Recently, Storlie et al. presented 'a new regularization method for simultaneous model fitting and variable selection in nonparametric regression models in the framework of smoothing spline ANOVA'. This method is an improvement on the COSSO (Lin and Zhang, 2006), penalizing the sum of component norms, instead of the squared norm employed in the traditional smoothing spline method. Storlie et al. introduce an adaptive weight to be used in the COSSO penalty which allows for more flexibility to estimate important functional components while giving heavier penalty to unimportant functional components.

In a 'parallel' stream of research, using the the so-called State-Dependent Regression (SDR) approach of Young (2001), Ratto et al. (2007) have developed a non-parametric approach which is very similar to smoothing splines and kernel regression approaches, but which is based on recursive filtering and smoothing estimation (the Kalman Filter combined with Fixed Interval Smoothing). Such a recursive least-squares implementation has some fundamental advantages: (a) it is couched with optimal Maximum Likelihood estimation, thus allowing for an objective estimation of the smoothing

* *Address for correspondence:* Marco Ratto, Joint Research Centre, European Commission, TP361 IPSC, via Fermi 1, 21020 Ispra (VA), Italy . E-mail: marco.ratto@jrc.it

hyper-parameters , and (b) it allows for greater flexibility in adapting to local discontinuities, heavy non-linearity and heteroscedastic error terms.

The purposes of this paper are:

1. develop a formal comparison and demonstrate equivalences between the 'classical' tensor product cubic spline approach with reproducing Kernel Hilbert space algebra (RKHS) and the SDR approach;

2. discuss advantages of disadvantages of these approaches;

3. propose a unified approach to smoothing spline ANOVA models that combines the best of the discussed methods, in particular the use of the recursive algorithms can be very effective in detecting the important functional components, adding valuable information in the ACOSSO framework.

References

C. Gu. *Smoothing Spline* ANOVA *Models*. Springer-Verlag, 2002.

Y. Lin and H. Zhang. Component selection and smoothing in smoothing spline analysis of variance models. *Annals of Statistics*, 34:2272–2297, 2006.

M. Ratto, A. Pagano, and P. C. Young. State dependent parameter meta-modelling and sensitivity analysis. *Computer Physics Communications*, 177:863–876, 2007.

C.B. Storlie, H. Bondell, B. Reich, and H.H. Zhang. The adaptive cosso for nonparametric surface estimation and model selection. *The annals of statistics*. submitted.

G. Wahba. *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics., 1990.

P. C. Young. The identification and estimation of nonlinear stochastic systems. In A. I. et al. Mees, editor, *Nonlinear Dynamics and Statistics*. Birkhauser, Boston, 2001.