A General Framework for Functional Regression Modelling

S. Greven¹, F. Scheipl¹

sonja.greven@stat.uni-muenchen.de, fabian.scheipl@stat.uni-muenchen.de

¹Department of Statistics, Ludwig-Maximilians-Universität München, Germany

Recent technological advances generate an increasing amount of functional data, data where each observation represents a curve or an image (Ramsay and Silverman, 2005). Examples of technologies that generate functional data include imaging techniques, accelerometers, spectroscopy and spectrometry. Any kind of measurement collected over time - data usually referred to as longitudinal - can also be viewed as potentially sparsely observed functional data.

Researchers are increasingly interested in regression models for functional data to relate functional observations to other variables of interest. We will discuss a comprehensive framework for additive (mixed) models for functional responses and/or functional covariates. The guiding principle is to reframe functional regression in terms of corresponding models for scalar data, allowing the adaptation of a large body of existing methods for these novel tasks. The framework encompasses many existing as well as new models. It includes regression for “generalized” functional data, mean regression, quantile regression as well as generalized additive models for location, shape and scale (GAMLSS) for functional data. It admits many flexible linear, smooth or interaction terms of scalar and functional covariates as well as (functional) random effects and allows flexible choices of bases - in particular splines and functional principal components - and corresponding penalties for each term. It covers functional data observed on common (dense) or curve-specific (sparse) grids. Penalized likelihood based and gradient-boosting based inference for these models are implemented in R packages refund and FDboost, respectively. We also discuss identifiability and computational complexity for the functional regression models covered.

A running example on a longitudinal multiple sclerosis imaging study serves to illustrate the flexibility and utility of the proposed model class. Reproducible code for this case study is also available online with the recent discussion paper Greven and Scheipl (2017) this talk is based on.

Keywords: functional additive mixed model, functional data, functional principal components, GAMLSS, gradient boosting, penalized splines.

References
