10 Nonparametric Models and Their Estimation

Göran Kauermann[1]

[1] Faculty of Business Administration and Economics, University of Bielefeld
gkauermann@wiwi.uni-bielefeld.de

Summary: Nonparametric models have become more and more popular over the last two decades. One reason for their popularity is software availability, which easily allows to fit smooth but otherwise unspecified functions to data. A benefit of the models is that the functional shape of a regression function is not prespecified in advance, but determined by the data. Clearly this allows for more insight which can be interpreted on a substance matter level.

This paper gives an overview of available fitting routines, commonly called smoothing procedures. Moreover, a number of extensions to classical scatterplot smoothing are discussed, with examples supporting the advantages of the routines.

10.1 Introduction

Statistics and Econometrics have been dominated by linear or parametric models over decades. A major reason for this were the numerical possibilities which simply forbid to fit highly structured models with functional and dynamic components. These constraints have disappeared in the last 15 to 20 years with new computer technology occurring and with statistical software developed side by side with more flexible statistical models. In particular models with smooth and nonparametrically specified functions became rather popular in the last years and the models are now easily accessible and can be fitted without deep expert knowledge. A milestone for the propagation of the models with smooth components was the introduction to generalized additive models by Hastie and Tibshirani (1990). The presentation of the models was accompanied by its implementation in the software package Splus. In contrast to linear models, in additive models, a response variable y is modelled to depend additively on a number of covariates and in a smooth but otherwise unspecified manner. In this respect, additive models are a flexible way to estimate in a regression setting the influence of a number of covariates x, say, on a response or outcome variable y. Allowing the outcome variable to be non-normally distributed but distributed according to an exponential family (like binomial, Poisson etc.) leads to generalized additive models. The idea of allowing covariates to have nonparamet-

ric influence was extended to varying coefficient models in Hastie and Tibshirani (1993). Here, interaction effects of covariates, particularly between factorial and metrical quantities, are modelled functionally but nonparametrically. Further contributions were proposed including the wide class of semiparametric models. Here, parts of the covariate effects are modelled parametrically while others are included nonparametrically in the model.

Applications of non- or semiparametric models are versatile and found in nearly all scientific fields. Nonetheless, the classical econometric field is interestingly enough still somewhat dominated by classical parametric models and nonparametric models are not common standard. This is, however, changing rapidly now. Nonparametric ideas for time series data have been recently proposed in Fan and Yao (2003) (see also Härdle *et al.*, 1997). For financial data, Ruppert (2004) shows how nonparametric routines can be used to achieve insight beyond the parametric world. We also refer to Pagan and Ullah (1999), Härdle *et al.* (2004) or Akritas and Politis (2003) for further developments of nonparametric routines in econometrics. In this paper we add some further applications in the economic context, primarily though for demonstrational purpose.

The paper is organized as follows. First we give a sketch of the different scatterplot smoothing methods, like local smoothing, spline smoothing and the new proposal of penalized spline smoothing. In Section 3 we present different smoothing models. Data examples are provided as motivation why nonparametric models are worthwhile to be used. A discussion concludes the paper.

10.2 Scatterplot Smoothing

10.2.1 Sketch of Local Smoothing

An early starting point for smoothing was the local estimate formally proposed by Nadaraya (1964) and Watson (1964). The idea is to estimate a regression function $m(x)$, say, locally as weighted mean. Consider data (x_i, y_i), $i = 1, \ldots, n$, with x as (metrically scaled) covariate and y as response. We assume the regression model

$$y_i = m(x_i) + \varepsilon_i, \quad i = 1, \ldots, n \tag{10.1}$$

with ε_i as independent residuals and $m(\cdot)$ as unknown and not further specified regression function. Postulating smoothness for $m(\cdot)$, that is continuity and sufficient differentiability, one can estimate $m(x)$ at a target point x (within the support of x_i, $i = 1, \ldots, n$) by the locally weighted mean

$$\hat{m}(x) = \sum w_{xi} y_i \,. \tag{10.2}$$

Here w_{xi} are weights summing up to 1 mirroring the local estimation. This means that weights w_{xi} take large values if $|x_i - x|$ is small while w_{xi} gets smaller if $|x_i - x|$ increases. A convenient way to construct such weights is to make use of a so called kernel function $K(\cdot)$, where $K(\cdot)$ is typically chosen as positive, symmetric function