ACTA SOCIETATIS MATHEMATICAE LATVIENSIS
Abstracts of the 6th Latvian Mathematical Conference, April 7–8, 2006, Liepāja, Latvia

ON ESTIMATION OF STATISTICAL DATA BY SMOOTHING SPLINES[1]

NATALJA BUDKINA

Riga Technical University

Meža iela 1, LV-1048, Rīga, Latvia

Institute of Mathematics of Latvian Academy of Sciences and University of Latvia

Akadēmijas laukums 1, LV-1524, Rīga, Latvia

E-mail: budkinanat@navigator.lv

Extensive research demonstrated splines utility for solving a lot of statistical estimation problems. In this report we will observe some developments of spline smoothing methods for applying them to the nonparametric regression estimation. We will review the method for the estimation of a given histogramm and will suggest one more way of solving of this problem.

Suppose that we have data (x_i, y_i) on a mesh $\triangle_n : a = x_0 < x_1 < \ldots < x_n = b$, the regression relationship between an explanatory variable x and a response variable y can be modelled as

$$y_i = g(x_i) + \varepsilon_i, \ i = 0, \ldots, n,$$

where g is a smooth regression function equal to the conditional mean of y_i given x_i and ε_i are the observation errors (independent identically distributed normal variates $\varepsilon_i \sim N(0, \sigma^2)$).

An estimation of regression function g could be obtained as the solution of the following minimization problems in Sobolev space $W_2^q[a,b]$

$$\min\{\sum_{i=0}^{n}(y_i - f(x_i))^2 + \lambda \int_a^b (f(x)^{(q)})^2 dx \ : \ f \in W_2^q[a,b]\},$$

or

$$\min\{\int_a^b (f(x)^{(q)})^2 dx \ : \ f \in W_2^q[a,b], \quad \sum_{i=0}^{n}(y_i - f(x_i))^2 \leq \Lambda\},$$

for some smoothing parameters $\lambda > 0$ and $\Lambda > 0$. Some of the known methods for solving nonparametric regression estimation problem and the problem of the choice of smoothing parameters will be observed in the report in a more detailed way (general cross-validation method, the regression splines).

This aproach was applied for solving the problem of approximation of a histogram $F = \{f_1, \ldots, f_n\}$ given on a mesh $\triangle_n$ by a function from Sobolev space. We consider two smoothing problems

$$\min\{\sum_{i=1}^{n}(\int_{x_{i-1}}^{x_i} f(x)dx - f_i(x_i - x_{i-1}))^2 + \lambda \int_a^b (f(x)^{(q)})^2 dx \ : \ f \in W_2^q[a,b]\}$$

and

$$\min\{\int_a^b (f(x)^{(q)})^2 dx \ : \ f \in W_2^q[a,b], \quad \sum_{i=1}^{n}(\int_{x_{i-1}}^{x_i} f(x)dx - f_i(x_i - x_{i-1}))^2 \leq \Lambda\}.$$

We investigate the connection between these problems and justify the choice of the smoothing parameters λ and Λ.

[1]This work was partially supported by The European Social Fund (ESF).