



Knot deletion for robust penalized spline regression

Steffen Liebscher*

Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany - steffen.liebscher@wiwi.uni-halle.de

Thomas Kirschstein

Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany - thomas.kirschstein@wiwi.uni-halle.de

Penalized spline regression is a well-known method to fit a smooth spline function to a set of observations. This is done by fitting to the data a system of basis functions using standard regression techniques. Regression coefficients are determined by minimizing a combination of a classical RSS criteria and a roughness penalty (in order to achieve smoothness of the resulting spline function). In using a RSS criteria the approach, however, becomes sensitive towards outliers. Therefore, *robust* penalized spline regression approaches replace the standard RSS measure by an appropriate robust alternative.

The underlying system of basis functions is defined piecewise based on knots/breakpoints placed along the sample space. For positioning of these knots, two rather vague strategies are commonly used. Either the knots are uniformly distributed along the sample space, or they are placed according to the empirical density (i.e. by using quantiles). We propose a new placement strategy which builds upon the idea of identifying contaminated regions within the sample space. By reducing the number of knots within contaminated regions, the effect of outliers on the regression estimates can be further reduced (beyond the effect that is already achieved by robust estimators). The efficacy of the proposed procedure is shown in terms of a simulation study.

Keywords: penalized spline regression; robust statistics; knot deletion.