Andrey A. Kudryavtsev, St. Petersburg State University, Russia
Mirbulat B. Sikhov, Al Farabi Kazakh National University, Kazakhstan

Some aspects of nonparametric regression applications to rate making

The idea of rate making is to estimate of rate $P(\mathbf{y}_j)$ on the base of loss data x_j and risk factors information $\mathbf{y}_j = (y_{1j},..., y_{mj})$, $j = 1, 2,..., n$. Although a construction of $P(\mathbf{y}_j)$ is based on a parametric (regression) model, nonparametric regression could also be used.

The basic idea of the latter approach is to use the Rosenblatt density estimator that is modified by an adjustment for risk factors:

$$\hat{f}(x|\mathbf{y}) = \frac{1}{nh_n}\sum_{j=1}^{n} K\left(\frac{x_j - x}{h_n}\right)\mathbf{1}_{\{\mathbf{y}_j \in A(\mathbf{y})\}},$$

where $\mathbf{1}_{\{\cdot\}}$ is an indicator function being equal to one if the condition is true or to zero otherwise and $A(\mathbf{y})$ is an appropriate partition of the risk factor set. Then the aimed rate estimate is usually as follows: $P(\mathbf{y}) = (1+\lambda)\,\mathrm{E}[X|\mathbf{y}]$ with a relative risk loading $\lambda > 0$. In other words, the key aspects of this approach are not only choosing kernel function $K(\cdot)$ but also constructing an appropriate partition. The latter constitutes actually a tariff classification.

The partitioning estimator (that based on an indicator function) is simplest and gives a compromise between these two aspects. However, its result is too elementary as the estimator is a sample mean. Other kernel functions lead to more consistent estimators, but the special attention to possible contradiction between above mentioned aspects should be paid. Some recommendations for choosing kernel functions and partitioning risk factor set are given in the report.

Keywords: nonparametric regression, rate making, Rosenblatt density estimator, partitioning risk factor set