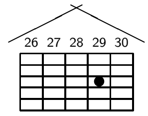

Poster2

An evolutionary algorithm for LTS-Regression: A comparative study

Morell, Oliver; Bernholt, Thorsten; Fried, Roland; Kunert, Joachim; Nunkesser, Robin

Least Trimmed Squares (LTS) Regression (Rousseeuw and Leroy, 1987) is a robust regression technique, which is becoming increasingly popular. This is due to its optimal breakdown properties and its better large-sample efficiency in comparison to e.g. the Least Median of Squares. LTS is defined by the non-convex minimization problem of finding a subset of predetermined size for which the least squares distance to a regression hyperplane is minimized. However, because of its NP-Hardness (Bernholt, 2005) the exact computation of LTS is very demanding, particularly in the case of a large number of explanatory variables. As an alternative to exact algorithms, procedures like the Fast-LTS (Rousseeuw and van Driessen, 2006) have been suggested, which calculate good solutions by restricting the calculations to a subset of all possible subsamples.

We use evolutionary algorithms (EA) to compute LTS Regression and introduce a new problem specific search operator. Since there are many design parameters of the algorithm that need to be chosen, we search for an optimal setting of these parameters. This search can be done with design of experiments. We compare the suitability of factorial and spacefilling designs for finding an optimal setting. In addition, the performance of the evolutionary algorithm is compared to that of other LTS-algorithms like the *lqs* and *ltsReg* implemented in R, both of which are based on the principle of the Fast-LTS. The performance is measured by the value of the LTS-criterion and the computation time, which are optimized jointly with the help of desirability indices. It turns out that the EA requires a higher computation time, but can deliver considerably better values of the LTS-criterion. M

References:

Bernholt, T. (2006), Robust Estimators are Hard to Compute, Technischer Report, Universität Dortmund.

Rousseeuw, P.J., Leroy, A.M. (1987), Robust Regression and Outlier Detection, John Wiley & Sons, New York.

Rousseeuw, P.J., van Driessen, K. (2006), Computing LTS Regression for Large Data Sets, Data Mining and Knowledge Discovery, 12, 29-45.

Submitting author: Morell, Oliver – University of Dortmund.