Efficient robust estimation of regression models

by Pavel Čížek

This paper introduces a new class of regression estimators robust to outliers, measurement errors, and other data irregularities. The proposed two-step least weighted squares (2S-LWS) estimator employs weights determined adaptively using the empirical distribution function of regression residuals obtained from an initial robust fit. Similarly to other methods based on an initial robust estimate, such as reweighted least squares, the proposed 2S-LWS estimator preserves the main robust properties of the initial robust estimate. However contrary to existing methods, the first-order asymptotic behavior is independent of the initial robust fit under mild conditions; most importantly, the initial estimator does not have to be $\sqrt{n}$ consistent. Moreover, the proposed method is proved to be asymptotically normal under weak β-mixing conditions and to be asymptotically efficient if unobservable errors follow a normal distribution. Additional features of the proposed method are that it does not have to fully reject any observation and that it does not have to rely on auxiliary tuning parameters. A simulation study documents that 2S-LWS outperform in most situations both least squares and existing robust estimators in finite samples. In particular, the relative efficiency of 2S-LWS reaches 85–90% in samples of several tens of observations under various distributional models.