S-weighted estimators

J.Á. Víšek[1*]

[1] *Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague*
[*] *Presenting author*

Keywords. *Robustness; Weighting the order statistics of the squared residuals plugged into a ρ-function; Consistency of S-weighted estimator.*

After an unsuccessful pursuit for a robust estimator with 50% breakdown point the task became at the end of seventies a nightmare of statisticians. Feasible versions of such estimators - the *least median of squares* and the *least trimmed squares* - fulfilled the desire but the discontinuity of objective function and the presence of order statistics of the squared residuals in their definitions were not favorable for studying the properties of estimators in question. All after, it caused that the consistency of the *least trimmed squares* was in full generality proved more than 20 years after the proposal of $\hat{\beta}^{(LTS,n,h)}$, Víšek (2006). *S-estimators*, defined as

$$\hat{\beta}^{(S,n,\rho)} = \arg\min_{\beta \in R^p} \left\{ \sigma \in R^+ \, : \, \sum_{i=1}^{n} \rho\left(\frac{r_i(\beta)}{\sigma}\right) = b \right\} \tag{1}$$

see Rousseeuw & Yohai (1984), removed both these snags simultaneously preserving the high breakdown point. The extraordinary virtue of all these estimators was their "innate" *scale-* and *regression-equivariance* in contrast to *M-estimators* which require special studentization of residuals, see Bickel (1975). However, the requirement on high breakdown point (seemingly) yields the high sensitivity to a small shift of "in-liers", see Hettmansperger & Sheather(1992). Although, their results were - due to a bad algorithm - a bit biased, they opened a discussion on the sensitivity of (robust) estimators to "inliers". The *least weighted squares* (LWS)

$$\hat{\beta}^{(LWS,n,w)} = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} w\left(\frac{i-1}{n}\right) r^2_{(i)}(\beta)$$

where

$$r^2_{(1)}(\beta) \le r^2_{(2)}(\beta) \le \; ... \; \le r^2_{(n)}(\beta),$$

see Víšek (2000), offered a chance to cope with this drawback. and represented an alternative to S-estimators (it is easy to see that $\hat{\beta}^{(LWS,n,w)}$ is not a special case of $\hat{\beta}^{(S,\rho,n)}$ and vice versa). The high speed of modern computational means allowed to select the weight function w just tailored to the level and even to the character of contamination. Utilization of a generalized version of Kolmogorov-Smirnov result on the convergence of empirical distribution functions to the underlying one - generalized for the regression framework, see Víšek (2011) - then simplified the proofs of consistency, $\sqrt{n}$-consistency, etc . Finally, recently proposed *S-weighted estimator*

$$\hat{\beta}^{(SW,n,w,\rho)} = \arg\min_{\beta \in R^p} \left\{ \sigma(\beta) \in R^+ \; : \; \frac{1}{n} \sum_{i=1}^{n} w\left(\frac{i-1}{n}\right) \rho\left(\frac{r^2_{(i)}(\beta)}{\sigma^2}\right) = b \right\}, \qquad (2)$$

see Víšek (2015), where $b = \mathbb{E}\rho(\frac{\varepsilon_1^2}{\sigma_0^2})$, $\rho : (0,\infty) \;\rightarrow\; (0,\infty)$ and nondecraesing on $(0,\infty)$, inherited plausible properties of *S-estimators* as well as of the *least weighted squares*, allowing for a wide range of objective (unbounded) functions and simultaneously offering to adjust the estimator to the level and to the character of contamination. Notice that in (**??**) in contrast to (**??**) we need the order statistics of squared residuals. Fortunately, employing a trick from *rank statistics*, see Hájek *&* Šidák (1967), we can rid of the technical problems with order statistics and then to employ Kolmogorov-Smirnov result recalled above. The paper summarizes the ideas of proving the consistency and $\sqrt{n}$-consistency under heteroscedasticity of error terms. It offers also a few patterns of results of numerical studies of their behavior for moderate sample size. These patterns demonstrate that a widely spread idea that the leverage points cause much more serious problems than the outliers need not have the absolute validity.

References

Bickel, P. J. (1975): One-step Huber estimates in the linear model. *J. Amer. Statist. Assoc. 70, 428–433.*

Hájek, J. *&* Šidák, Z. (1967): *Theory of Rank Test.* Academic Press, New York.

Hettmansperger, T. P. *&* Sheather, S. J. (1992): A Cautionary Note on the Method of Least Median Squares. *The American Statistician 46, 79–83.*

Rousseeuw, P. J. *&* Yohai, V. (1984): Robust regressiom by means of S-estimators. In: *Robust and Nonlinear Time Series Analysis. eds. J. Franke, W. Hárdle and R. D. Martin, Lecture Notes in Statistics 26, Springer Verlag, New York, 256-272.*

Víšek, J. Á. (2000): Regression with high breakdown point. *Robust 2000 (eds. Jaromír Antoch & Gejza Dohnal, published by Union of Czech Mathematicians and Physicists), 2001, 324 - 356.*

Víšek, J. Á. (2006): The least trimmed squares. Consistency. $\sqrt{n}$-consistency. Asymptotic normality and Bahadur representation. *Kybernetika 42, 1 - 36, 181 - 202, 203 - 224.*

Víšek, J. Á. (2011): Empirical distribution function under heteroscedasticity. *Statistics 45, 497-508.*

Víšek, J. Á. (2015): S-weighted estimators. *Proc. of the 16th Conference on the Applied Stochastic Models, Data Analysis and Demographics 2015, 1031 - 1042.*