When do S- and CM-estimates for regression coincide?

Olcay Arslan
Department of Mathematics
University of Cukurova, 01330 Balcali
Adana Turkey

Ove Edlund and Håkan Ekblom
Department of Mathematics
Luleå University of Technology
97331 Luleå, Sweden

Robust estimation methods for regression have been developed for many years. Well-known examples are M-estimates (Huber(1981)) and S-estimates (Rousseeuw and Yohai (1984)). A more recent approach is "Constrained M-estimates", or CM-estimates for short (Mendes and Tyler (1995)). We can formulate CM-estimation for regression the following way.

Consider the linear model $y_i = x_i^T\beta + e_i, i = 1, 2, ..., n$, where $y = (y_1, y_2, ..., y_n)^T$ is the response vector, x_i is the i:th row in the $(n \times p)$ design matrix X, β a p-dimensional vector of unknowns and $e = (e_1, e_2, ..., e_n)^T$ the error vector. Define the residuals as $r_i = y_i - x_i^T\beta$, $i = 1, 2, ..., n$. Using the notation "ave" for the arithmetic average, the CM-estimation problem is to find the global minimum of

$$L(\beta, \sigma) = c^2\, ave\{\rho(r_i/c\sigma)\} + \log(\sigma)$$

over $\beta \in R^p$ and $\sigma \in R^+$ subject to the constraint

$$ave\{\rho(r_i/c\sigma)\} \leq \varepsilon\rho(\infty) \tag{1}$$

Here $\rho(t)$ is a bounded, nondecreasing function of $t \geq 0$ with tuning parameter $c > 0$. If strict inequality holds in the constraint (1) we get the redescending M-estimating equations for β and σ. To find the S-estimate, we minimize L with respect to σ. This implies equality in the constraint (1).

The CM-estimates are regression and affine equivariant, and possess, at the same time, the good local properties of the M-estimates for regression and good global robustness properties of the regression S-estimates. They are consistent, asymptotically normal and very efficient estimators. They will have low gross error sensitivity and high asymptotic efficiency. Their asymptotic breakdown point is $\min(\varepsilon, 1 - \varepsilon)$, so that for $\varepsilon = 1/2$ the asymptotic breakdown point is 1/2. One can see Mendes and Tyler (1995) for the robustness and the asymptotic properties of the CM-estimates for regression.

From the definition of the S- and CM-estimates, it is appearant that they will in many cases coincide. An interval for c when this happens for the case that ρ is the Tukey biweight is derived in Arslan, Edlund and Ekblom (2001). We will give intervals when this happens for two other ρ functions. We have

$$\frac{\partial L(\beta, \sigma)}{\partial \sigma} = c^2 ave\{\rho'(\frac{r_i}{c\sigma})(-\frac{r_i}{c\sigma^2})\} + \frac{1}{\sigma} = \frac{c^2}{\sigma}\left[\frac{1}{c^2} - ave\{\rho'(\frac{r_i}{c\sigma})\frac{r_i}{c\sigma}\}\right]$$

Now consider the function $g(t) = \rho'(t)t$. For the Welsh function $\rho(t) = \frac{1}{2}(1 - e^{-t^2})$, giving $g(t) = t^2 e^{-t^2}$ and $g'(t) = 2t(1-t^2)e^{-t^2}$. Thus $g(t)$ has its maximum $= e^{-1}$ for $t = \pm 1$. This implies that $\frac{\partial L(\beta,\sigma)}{\partial \sigma} > 0$ if $\frac{1}{c^2} > \frac{1}{e}$, i.e. if $c < \sqrt{e} = 1.772$. For these c-values, σ should be choosen as small as possible. The left hand side of (1) increases when σ gets smaller. This means that σ is limited by the equality case in the constraint , i.e. we have the S-estimate.

For the Geman and McClur function $\rho(t) = t^2(1+t^2)^{-1}$, giving $g(t) = 2t^2(1+t^2)^{-2}$ and $g'(t) = 4t(1-t^2)(1+t^2)^{-1}$. Thus the maximum of $g(t)$ is $\frac{1}{2}$, which gives the interval $c < \sqrt{2} = 1.414$. We summarize the result in the following table,

Name	ρ function	Interval where S- and CM-estimates always coincide
Tukey	$\rho(t) = \begin{cases} \frac{x^2}{2} - \frac{x^4}{2} + \frac{x^6}{6} & if \quad \|x\| \le 1 \\ \frac{1}{6} & if \quad \|x\| > 1 \end{cases}$	$0 < c < 2.598$
Welsh	$\rho(t) = \frac{1}{2}(1 - e^{-t^2})$	$0 < c < 1.772$
Geman and McClur	$\rho(t) = t^2(1+t^2)^{-1}$	$0 < c < 1.414$

REFERENCES

Arslan,O., Edlund,O. and Ekblom,H.(2001) "Algorithms to compute CM- and S-estimates for regression." Tech.report, Dept of Mathematics, Luleå University of Technology , Sweden (to be published).

Huber, P.J. (1981) Robust Statistics. New York. Wiley.

Kent, J.T. and Tyler, D.E.(1996) "Constrained M-estimation for multivariate location and scatter." The Annals of Statistics, 24(3), 1346-1370.

Mendes, B. and Tyler, D.E.(1995) "Constrained M estimates for regression." In Robust Statistics; Data Analysis, and Computer Intensive Methods. Lecture Notes In Statistics 109, Springer, New York, 299-320.

Rousseeuw P.J. and Yohai V.J. (1984), Robust regression by means of S-estimators. In Robust and Nonlinear Time Series Analysis, eds: J.Frank,W.Härdle and R.D.Martin (Lecture Notes in Statistics), New York Springer-Verlag, pp.256-272.

RESUME

Intervals for the tuning parameter, where the regression S- and CM-estimates coincide, are derived.