Tightness of M-estimators for multiple linear regression in time series

S. Johansen[1] and **B. Nielsen**[2*]

[1] *University of Copenhagen & CREATES; soren.johansen@econ.ku.dk.*
[2] *University of Oxford; bent.nielsen@nuffield.ox.ac.uk*
[*] *Presenting author*

Keywords. *M-estimator; Martingales; Huber-skip; Quantile regression.*

We show tightness for a class of regression M-estimators, where the objective function can be non-monotonic and non-continuous. A prominent example of an estimator is the Huber skip estimator, where each observation contributes to the objective function through a criterion function, which is quadratic in the central part and horizontal otherwise. The tightness result addresses a difficulty which is often met in asymptotic analysis of problems where the objective function is non-convex. A very common solution is to assume that the parameter space is compact. While such an assumption circumvents the problem, it is done through a condition on the unknown parameter and it is therefore rarely satisfactory from an applied viewpoint. Instead, our result only requires an assumption that can be justified by inspecting the observed regressors and the objective function.

We consider the multiple linear regression

$$y_i = \mu + \alpha' x_i + \varepsilon_i, \qquad i = 1, \dots, n, \tag{1}$$

where the innovations ε_i are independent of $\mathcal{F}_{i-1} = \sigma(x_1, \dots, x_i, \varepsilon_1, \dots, \varepsilon_{i-1})$. The regressors x_i have dimension m. They can be deterministic or stochastic, and stationary or stochastically trending. Often we will subsume the intercept in the regressors and use the notation

$$\beta' z_i = \mu + \alpha' x_i.$$

The M-estimator for the parameter β is the minimizer, $\hat{\beta}$, of the objective function

$$R_n(\beta) = \frac{1}{n} \sum_{i=1}^{n} \rho(y_i - z_i'\beta), \tag{2}$$

for some criterion function ρ. M-estimators were originally introduced for location problems by Huber in 1964, but later extended to regression models. The class of M-estimators considered includes the Huber-skip estimator, which has a non-convex criterion function, as well as quantile regression estimators, in particular the least absolute deviation, and least squares estimator, which all have a convex criterion.

The asymptotic theory of the regression M-estimator is well understood for nice criterion functions ρ. **?** provide an asymptotic theory for regression M-estimators and show existence and uniqueness for the case of a convex, differentiable criterion function. **?** give two results on tightness (and consistency) for more general criterion functions. In both cases the criterion function $\rho(u)$ is continuous, non-decreasing in $u > 0$ and non-increasing for $u < 0$. Their Theorem 1 shows tightness when (y_i, z_i') are i.i.d. and $\mathsf{E}\rho(y_i - z_i'\beta)$ has a unique minimum. Their Theorem 4 shows tightness when z_i is deterministic and satisfies a condition on the frequency of small regressors.

In this paper we generalize the result of **?**. We assume ρ is semi-continuous and nonnegative with a minimum at zero and greater then $\rho^* > 0$ for large values of the argument. We also need an extra condition on the expected criterion function $h(v)$, which is assumed to take a value below ρ^* somewhere in the central part of the distribution of the error term. The only condition to the regressors is a condition on the frequency of small regressors, which is weaker than the condition of **?**, albeit stronger than the conditions for the tightness of least square estimators. The latter illustrates the price we pay by leaving the least squares criterion. The condition is related to a condition for deterministic regressor used by **?** for S-estimators. Our condition is, however, formulated in a slightly different way, which seems to be easier to check for particular regressors. Indeed, we check the condition for a few situations. We give a number of examples with deterministic regressors to illustrate the condition. We also show that the condition is satisfied for stationary regressors and for random walk regressors.

It is worth noting that the innovations are neither required to have a zero expectation nor a continuous density. Thus, the results apply both when the innovations follow a non-contaminated reference distribution with density f_0, say, and when the they are contaminated so that they follow a mixture distribution with density $(1-\epsilon)\mathsf{f}_0 + \epsilon\mathsf{f}_1$, say. The proofs use martingale techniques, chaining arguments and the iterated martingale inequality from **?**.

References

Chen, X.R. & Wu, Y.H. (1988). Strong consistency of M-estimates in linear models. *Journal of Multivariate Analysis* **27**, 116–130.

Davies, L. (1990) The asymptotics of S-estimators in the linear regression model. *Annals of Statistics* **18**, 1651–1675.

Johansen, S. & Nielsen, B. (2016) Analysis of the Forward Search using some new results for martingales and empirical processes. *Bernoulli* **22**, 1131-1183.

Maronna, R., Martin, D. & Yohai, V. (2006). Robust Statistics: Theory and Methods. John Wiley & Sons, Chichester.