Robust CUSUM-M test in the presence of long-memory disturbances [1]

by

Philipp Sibbertsen
Fachbereich Statistik, Universität Dortmund, D-44221 Dortmund, Germany

Version March 2000

Abstract

We derive the limiting null distribution of the robust CUSUM-M test and the recursive CUSUM-M test for structural change of the coefficients of a linear regression model with long-memory disturbances. It turns out that the asymptotic null distribution of the CUSUM-M statistic is a fractional Brownian Bridge and the asymptotic null distribution of the recursive CUSUM-M statistic is fractional Brownian motion.

KEY WORDS: CUSUM test; robust regression; long range dependence

1 Introduction

Consider the linear regression model

$$y_i = \beta^\top x_i + \varepsilon_i, \qquad i = 1, \dots T, \tag{1}$$

where y_i is the dependent variable, x_i is a p - dimensional vector of fixed regressors, β is the p - dimensional parameter vector and ε_i is an error process. Here we consider the case where ε_i is a long-memory stationary process.

[1] Research supported by "Deutsche Forschungsgemeinschaft" (DFG).

This paper considers tests of the null hypothesis that the parameter vector β is constant over time:

$$H_0 : \beta^{(1)} = \ldots = \beta^{(T)} = \beta, \tag{2}$$

β unknown, versus

$$H_1 : \beta^{(1)} = \ldots = \beta^{(m)} \neq \beta^{(m+1)} = \ldots = \beta^{(T)} \tag{3}$$

for some m, where $m\,(1 \leq m \leq T)$ is unknown.

The most important tests dealing with this problem are the standard CUSUM test introduced by Brown, Durbin, Evans(1975) based on recursive OLS-residuals and the OLS-based CUSUM test by McCabe, Harrison(1980) based on standard OLS-residuals. The respective asymptotic null distributions are known for the case of independent disturbances (Brown, Durbin, Evans(1975), Sen(1982), Krämer, Ploberger, Alt(1988) or Ploberger, Krämer(1992)). These null distributions are no longer valid in the case of long-memory error terms. Neglecting long range dependencies in the disturbances leads to a rejection of the null hypothesis with probability one (see for example Wright(1998)).

Defining $R(k) := Cov(X_i, X_{i+k})$ long-memory time series can be modeled as stationary processes satisfying

$$\frac{R(k)}{L(k)|k|^{2H-2}} \to 1, \qquad k \to \infty,$$

where $\frac{1}{2} < H < 1$, $L(k)$ is a slowly varying function. In the special case of a fractional Gaussian noise introduced by Mandelbrot and van Ness(1968) $R(k)$ has the form

$$R(k) = \frac{\sigma_X^2(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H})}{2}.$$

The equation $\mathrm{Var}(\bar{X}_n) = \sigma_X^2 n^{2H-2}$ also holds so that we have the convergence $n^{2-2H} L_{\mathrm{Var}}^{-1}(n)\mathrm{Var}(\bar{X}_n)$ to 1, where $L_{\mathrm{Var}}(n) = L(n)/(H(2H-1))$.

So the main property of long-memory processes is the slow decay of the correlations. For a more detailed discussion of long range dependence see Beran(1994) or Sibbertsen(1999a).

Hidalgo, Robinson(1996) determine the asymptotic null distribution for tests for structural change with long-memory errors when the breakpoint is known. Wright(1998) derives the asymptotic null distribution of the OLS-based CUSUM test for a polynomial design.

The standard CUSUM test and the OLS-based CUSUM test are based on ordinary or recursive least squares residuals. Since the least squares estimator is not robust against outliers, the standard CUSUM and the OLS-based CUSUM test are not robust against outliers either. Both tests reject the null hypotheses of no structural break if there is only one outlier with probability one even if there is no structural break in the data. Sen(1984) therefore introduced the robust CUSUM-M test and the recursive CUSUM-M test by replacing the OLS-residuals and recursive residuals by M-residuals and recursive M-residuals respectively.

The aim of this paper is to determine the asymptotic null distribution of the CUSUM-M test and the recursive CUSUM-M test in the case of long-memory disturbances.

There is an extensive literature about the behaviour of robust estimates in the context of long-memory processes. For example Beran(1991) considers M-estimates for location and Sibbertsen(1999a,b,c) considers S-estimates for the linear and nonlinear regression model with long-memory disturbances.

The outline of the paper is as follows. In the next chapter we consider the CUSUM-M test and in chapter 3 we derive the asymptotic limit distribution of the recursive CUSUM-M test in the presence of long-memory.

2 The CUSUM-M Test

The M-estimator of the regression parameter β in model (1) is given as the solution of

$$\sum_{i=1}^{T} \psi(y_i - x_i^T \beta) = 0, \tag{4}$$

where the function ψ meets the following restrictions:

- ψ is skew-symmetric, that is $\psi(x) + \psi(-x) = 0$
- ψ is nondecreasing
- ψ is almost everywhere differentiable
- $E[\psi'] \neq 0$
- $g_\varepsilon(y) := \sup_{\delta \leq \varepsilon} |\psi'(y + \delta) - \psi'(y)| \leq c$ almost sure for some $\varepsilon > 0$ and $c > 0$.

Here and in the rest of the paper the expectation is taken with respect to the standard normal distribution.

One example for such a ψ function is Huber's ψ given by

$$\psi_H(x) = \min(c, \max(x, -c)).$$

The test statistic of the CUSUM-M test is

$$\sup_{0 \leq \lambda \leq 1} |\frac{1}{\hat{\sigma}\sqrt{T}} W_{CM}^{(T)}(\lambda)|, \tag{5}$$

where

$$W_{CM}^{(T)}(\lambda) = \sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \hat{\beta}),$$

and $\hat{\beta}$ is a M-estimator derived from (4).

In what follows the regressor matrix $X(T)$ ($T \times p$ with columns $x_i(T)$) is assumed to be fixed, with

$$\sum_{i=1}^{T} x_i(T)x_i(T)^T = TI_p, \tag{6}$$

where I_p is the p-dimensional identity matrix. This implies that

$$\sum_{i=1}^{T} \|x_i(T)\|^2 = Tp. \tag{7}$$

This involves no loss in generality as such a regressor matrix can always be obtained by an appropriate transformation of the $x_i(T)$. In addition, the following conditions are assumed to hold:

A1 There are positive numbers η_1, η_2 and T_0 such that

$\sum_{i=1}^{T} x_i(T)x_i(T)^T 1_{\{\|x_i(T)\|<\eta_1\}} - T\eta_2 I_p$ is positiv definite for all $T \geq n_0$.

A2 $\lim_{T\to\infty} \frac{1}{T}\sum_{i=1}^{T} \|x_i(T)\|^2 1_{\{\|x_i(T)\|>\delta\sqrt{T}\}} = 0 \qquad \forall\delta > 0.$

From this condition we have a sequence $[\delta(T)]_{T=1}^{\infty}$ with $\lim_{T\to\infty}\delta(T) = 0$ and

$$\lim_{T\to\infty} \frac{1}{\delta(T)^2 T} \sum_{i=1}^{T} \|x_i(T)\|^2 1_{\{\|x_i(T)\|\geq\delta(T)\sqrt{T}\}} = 0, \tag{8}$$

consequently

$$\lim_{T\to\infty} \sum_{\{\|x_i(T)\|>\delta(T)\sqrt{T}\}} 1 = 0. \tag{9}$$

These are quite mild regularity conditions for the regressors allowing in particular for trends. The dependence on the observation size T will not be mentioned in the following any more.

The proof of the asymptotic properties of the CUSUM-M test statistic requires the Hermite Rank of a function.

Definition (Hermite rank)

Let Z be a standard normal random variable. A function $G : \mathbb{R} \to \mathbb{R}$ with $E[G(Z)] = 0$ and $E[G^2(Z)] < \infty$, is said to have Hermite rank m, if $E[G(Z)P_q(Z)] = 0$ for all Hermite polynomials $P_q, q = 1, \ldots, m-1$ and $E[G(Z)P_m(Z)] := J_G(m) \neq 0$.

The function $J_G(l)$ is defined by

$$J_G(l) := E[G(Z)P_l(Z)], \qquad l \in \mathbb{N}. \tag{10}$$

Theorem 1 *Under the above assumptions we have*

$$T^{-1/2-d} W_{CM}^{(T)}(\lambda) \xrightarrow{d} B_d(\lambda) - \xi, \tag{11}$$

where $B_d(\lambda)$ is a fractional Brownian motion with self-similarity parameter d and ξ is a normal random vector with zero mean and variance $E\psi^2 I_p$ and I_p is the p-dimensional identity matrix.

Proof: From (4) we have

$$\sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \hat{\beta}) = 0.$$

A Taylor expansion around the true parameter vector β_0 gives

$$\sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \hat{\beta}) = \sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \beta_0) - \sum_{i=1}^{[T\lambda]} (\hat{\beta} - \beta_0)^\top \psi'(y_i - x_i^\top \beta_0) x_i + o(1). \tag{12}$$

Denote by $L_{\mathrm{Var}}(T)$ a slowly varying function and remember that ψ is skew-symmetric and therefore has Hermite Rank one. So theorem 5.1 in Taqqu(1975) in conjunction with Marmol(1995) gives

$$T^{-1/2-d} L_{\mathrm{Var}}^{-1/2}(T) J_\psi(1)^{-1} \sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \beta_0) \xrightarrow{d} B_d(\lambda). \tag{13}$$

Sibbertsen(1999a, b) shows that

$$T^{1/2-d} L_{\mathrm{Var}}^{-1/2}(T) J_\psi(1)^{-1}(\hat{\beta}-\beta_0) \xrightarrow{d} N(0, \frac{E\psi^2}{(E\psi')^2} I_p). \tag{14}$$

To see this, keep in mind that the asymptotic distribution of S-estimators is equivalent to that of M-estimators.

We also have

$$\frac{1}{T}\sum_{i=1}^{[T\lambda]} \psi'(y_i - x_i^\top \beta_0)x_i \xrightarrow{d} E[\psi']I_p. \tag{15}$$

Remember that assumption (A1) gives $\frac{1}{T}\sum_{i=1}^{[T\lambda]} x_i x_i^\top = I_p$.

Equation (12) to (15) gives the assertion. $\diamondsuit$

Remark 1: For $d = 0$ and $\psi(x) = x$ the limit distribution is the standard Brownian Bridge as in the case of independent errors. $B_0(\lambda)$ is standard Brownian motion.

We can also obtain the limiting distribution of the non-recursive CUSUM-M test in the case of short-memory disturbances, that is $-1/2 < d < 0$. The correlations $R(k)$ of a short-memory process are summable.

Theorem 2 *In the case $-1/2 < d < 0$ we have*

$$T^{-1/2-d} W_{CM}^{(T)}(\lambda) \xrightarrow{d} B_d(\lambda) - \eta, \tag{16}$$

where $B_d(\lambda)$ is a standard Brownian motion and η is a Gaussian random vector with mean zero and variance $E\psi^2 I_p$.

Proof: The proof of this theorem is similar to the proof of Theorem 1. We have (see also Breuer and Major(1983)) that

$$T^{-1/2-d} L_{\mathrm{Var}}^{-1/2}(T) \sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \beta_0) \xrightarrow{d} B(\lambda). \tag{17}$$

From Sibbertsen(1999b)

$$T^{1/2-d}L_{\text{Var}}^{-1/2}(T)(\hat{\beta}-\beta_0) \xrightarrow{d} N(0, \frac{E\psi^2}{(E\psi')^2}I_p). \tag{18}$$

Again we have

$$\frac{1}{T}\sum_{i=1}^{[T\lambda]}\psi'(y_i - x_i^\top\beta_0)x_i \xrightarrow{d} E[\psi']I_p. \tag{19}$$

Equations (17) to (19) prove the theorem. $\diamondsuit$

Remark 2:

a) A generalization of fractional Brownian motion to short-memory processes can be found in Taqqu(1977).

b) For $d = 0$ we obtain the classical ARMA processes and the well known rate of convergence of $T^{1/2}$.

3 Recursive CUSUM-M Test

This section considers the asymptotic behaviour of the recursive CUSUM-M test in the case of long-memory disturbances. The idea of the recursive CUSUM-M test is to replace the M-residuals in the non-recursive version by recursive M-residuals. The k-th recursive M-residual is thereby given as

$$r_k = y_k - x_k^\top\hat{\beta}_{k-1},$$

where $\hat{\beta}_{k-1}$ is the M-estimator based on the first $k-1$ observations. The test statistic of the recursive CUSUM-M test is defined by

$$\sup_{0\leq\lambda\leq1}|\frac{1}{\hat{\sigma}\sqrt{T}}W_{RCM}^{(T)}(\lambda)|,$$

with

$$W_{RCM}^{(T)}(\lambda) = \sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \hat{\beta}_{(i-1)}).$$

and $\hat{\beta}_{(i-1)}$ is the M-estimator based on the first $i-1$ observations.

Theorem 3 *Under the above assumptions we have*

$$T^{-1/2-d} W_{RCM}^{(T)}(\lambda) \xrightarrow{d} B_d(\lambda), \tag{20}$$

where $B_d(\lambda)$ is a fractional Brownian motion with self-similarity parameter d.

Proof: Using Taylor expansion around the true parameter vector β_0 we have

$$\begin{aligned} \sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \hat{\beta}_{(i-1)}) &= \sum_{i=1}^{[T\lambda]} \psi(y_i - x_i^\top \beta_0) \\ &- \sum_{i=1}^{[T\lambda]} (\hat{\beta}_{(i-1)} - \beta_0)^\top \psi'(y_i - x_i^\top \beta_0) x_i + o(1). \end{aligned} \tag{21}$$

First we show that

$$\max_{1 \le i \le T} \|(\hat{\beta}_i - \beta_0) - \frac{1}{E\psi'} \sum_{j=1}^{i} (\psi(y_j - x_j^\top \beta_0) x_j) Q_j^{-1}\| = o_P(T^{(d+1/2)} (\log\log T)^{1/2}), \tag{22}$$

where

$$\frac{1}{T} \sum_{i=1}^{j} x_i x_i^\top =: C_j.$$

We have $C_T = I_p$ because of assumption (A1).

Denote

$$\sum_{j=1}^{i} (\psi(y_j - x_j^\top \beta_0) x_j) =: S_i.$$

The law of the iterated logarithm for sums of non-linear functions of Gaussian random variables with long-memory gives us for S_i (see Taqqu(1977))

$$\max_{1\leq i\leq T}\left[\frac{S_i}{\left(\frac{2J_\psi(1)^2}{(d+1/2)(2(d+1/2)-1)}T^{2(d+1/2)}L_{\mathrm{Var}}(T)\log\log T\right)^{1/2}}\right] = o_P(1). \tag{23}$$

Using (23) equation (22) follows from Lemma 3.1 in Jureckova, Sen(1984).

Combining (21) with (22) gives

$$\begin{aligned}\sum_{i=1}^{[T\lambda]}\psi(y_i - x_i^\top\hat{\beta}_{(i-1)}) &= \sum_{i=1}^{[T\lambda]}\psi(y_i - x_i^\top\beta_0)\\ &- \frac{1}{E\psi'}\sum_{i=1}^{[T\lambda]}\left(\sum_{j=1}^{i-1}\psi(y_j - x_j^\top\beta_0)x_j^\top Q_j^{-1}\right)\\ &\quad\psi'(y_i - x_i^\top\beta_0)x_i + o(T^{(d+1/2)}(\log\log T)^{1/2})\\ &= \sum_{i=1}^{[T\lambda]}\left(\sum_{j=1}^{i}c_{ij}\psi(y_j - x_j^\top\beta_0) + o(T^{(d+1/2)}(\log\log T)^{1/2}\right),\end{aligned} \tag{24}$$

where $c_{ij} = -x_j^\top Q_{i-1}^{-1}x_i$ for $j < i$, $c_{ii} = 1$ and $c_{ij} = 0$ for $j > i$.

Now from Sen(1984) we obtain

$$\sum_{j\leq i}c_{ij}^2 = 1 + o(i^{-1}). \tag{25}$$

Applying theorem 5.1 of Taqqu(1975) and Marmol(1995) to the last equation of (24) gives the assertion. $\diamondsuit$

As for the non-recursive CUSUM-M test a slight modifications of the above proof gives the null distribution in the short-memory case also for the recursive test.

Theorem 4 *Under the above assumptions and for $-1/2 < d < 0$ we have*

$$T^{-1/2-d} W_{RCM}^{(T)}(\lambda) \xrightarrow{d} B_d(\lambda), \tag{26}$$

where $B_d(\lambda)$ is a fractional Brownian motion with self-similarity parameter d.

Proof: The proof of this theorem is again similar to the proof of theorem 3, but instead of using the law of the iterated logarithm for long-memory Gaussian processes we need in (23) the similar law for short-memory Gaussian processes (see Taqqu(1977)). Applying the limit theorem of Breuer and Major(1983) to (24) gives the assertion. $\diamondsuit$

References

Beran, J. (1991): "M estimators of location for Gaussian and related processes with slowly decaying serial correlations." *J. Amer. Statist. Assoc.* 68, 704 - 707.

Beran, J. (1994): *Statistics for long-memory processes.* Chapman & Hall, New York.

Breuer, P. and Major, P. (1983): "Central limit theorems for non-linear functionals of Gaussian fields." *Journal of Multivariate Analysis* 13, 425 - 441.

Brown, R. L.; Durbin, J. and Evans, J. M. (1975): "Techniques for testing the constancy of regression relationships over time." *Journal of the Royal Statistical Society B* 37, 149 - 163.

Hidalgo, J. and Robinson, P.M. (1996): "Testing for structural change in a long–memory environment." *Journal of Econometrics* 70, 159 – 174.

Jureckova, J. and Sen, P. K. (1984): "On adaptive scale-equivariant M-estimators in linear models." *Statistics & Decision* 2, 31 - 46.

Krämer, W.; Ploberger, W. and Alt, R. (1988): "Testing for structural change in dynamic models." *Econometrica* 56, 1355 – 1369.

Mandelbrot, B. and van Ness, J. (1968): "Fractional brownian motions, fractional noises and applications." *SIAM Reviews* 10, 422 - 437.

Marmol, F. (1995): "Spurious Regression Between I(d)Processes." *Journal of Time Series Analysis* 16, 313 – 321.

Ploberger, W. and Krämer, W. (1992): "The CUSUM–test with OLS–residuals." *Econometrica* 60, 271 – 286.

Sen, P. K. (1982): "Invariance principles for recursive residuals." *The Annals of Statistics* 10, 307 - 312.

Sen, P. K. (1984): " Recursive M–tests for the constancy of multivariate regression relationships over time." *Sequential Ananlysis* 3, 191 - 211.

Sibbertsen, P. (1999a): *Robuste Parameterschaetzung im linearen Regressionsmodell bei Fehlertermen mit langem Gedaechtnis.* Verlag fuer Wissenschaft und Forschung, Berlin (in German).

Sibbertsen, P. (1999b): "S-estimators in the linear regression model with long-memory error terms under trend." *Journal of Time Series Analysis* to appear.

Sibbertsen, P. (1999c): "S-estimators in the nonlinear regression model with long-memory error terms." *Technical Report 36/99, SFB 475, University of Dortmund.*

Taqqu, M. S. (1975): "Weak convergence to fractional Brownian Motion and to the Rosenblatt Process." *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 31, 287 - 302.

Taqqu, M. S. (1977): "Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibits a long range dependence." *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 40, 203 - 238.

Wright (1998): "Testing for a structural break at unknown date with long-memory disturbances." *Journal of Time Series Analysis* 19, 369 – 379.