RECURSIVE ESTIMATION IN AUTOREGRESSIVE MODELS WITH ADDITIVE OUTLIERS

Tom Cipra, Asuncin Rubio and Jos Luis Canal

The paper deals with recursive robust estimation of the autoregressive models with additive outliers (AO-AR-models). Recursive robust procedures based on the idea of CMM-estimation (Conditional-Mean M-estimation) are suggested that enable to treat the AO-AR-models on-line.

1. INTRODUCTION

This work deals with recursive robust estimation in autoregressive models that are contaminated by additive outliers. The importance of such procedures in applied time series is obvious: (i) the recursive character of the estimation allows to treat time series in real time (on-line) updating previous estimates by means of simple calculations after delivering new observations; (ii) robustness of the estimation procedures makes them insensitive to outlied observations that can distort significantly results of classical non-robust estimation procedures (the additive outliers have also unpleasant consequences for forecasting, see e. g. [10], [14]).

The autoregressive model of the order p with additive outliers denoted as AO-AR(p) has the form

$$y_t = x_t + v_t, \tag{1}$$

where

$$x_t = \varphi_1 \, x_{t-1} + \cdots + \varphi_p \, x_{t-p} + \varepsilon_t \tag{2}$$

is the classical AR(p)-model, i. e. $\varphi_1, \ldots, \varphi_p$ are parameters and $\{\varepsilon_t\}$ is a white noise with variance σ^2. The process $\{x_t\}$ is mostly supposed to be stationary and normal. The process $\{v_t\}$ that is independent of the process $\{x_t\}$ can attain with a small probability very large values contaminating additively according to (1.1) the process $\{x_t\}$. The usual probability model for v_t is the following mixture of independent normal distributions (see e. g. [23])

$$v_t \sim (1-\gamma) \, N(0, \sigma_1^2) + \gamma \, N(0, \sigma_2^2), \tag{3}$$

where the variance σ_2^2 of the contaminating distribution $N(0, \sigma_2^2)$ is much larger than the variance σ_1^2 of the nominal distribution $N(0, \sigma_1^2)$ and the constant $\gamma \in \langle 0, 1)$ is near to zero. Specially, it can be $\sigma_1^2 = 0$ so that

$$P(v_t = 0) = 1 - \gamma. \tag{4}$$

The distribution $N(0, \sigma_2^2)$ in (1.3) ($\sigma_2^2 \gg \sigma_1^2$) can be replaced by a heavy-tailed distribution, e. g. by Cauchy, Laplace, Student, uniform with a large variance or other types of heavy-tailed distributions.

Besides the additive outliers, the robust time series analysis also deals with the innovations outliers (IO) that are not included to the model additively but influence directly the innovation distribution. The corresponding IO-AR(p)-model has the form

$$y_t = \varphi_1 \, y_{t-1} + \cdots + \varphi_p \, y_{t-p} + \varepsilon_t, \tag{5}$$

where the innovations $\{\varepsilon_t\}$ are non-normal with a heavy-tailed distribution or they have the contaminated normal distribution of the type (1.3). The usual robust estimators in the IO-model are the M-estimators. Let us remind briefly the M-estimation for IO-AR(p)-model (1.5) written as

$$y_k = \boldsymbol{\varphi}' \mathbf{z}_k + \varepsilon_k, \quad k = p+1, \ldots, n, \tag{6}$$

where $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_p)'$, $\mathbf{z}_k = (y_{k-1}, \ldots, y_{k-p})'$ and ε_k are i. i. d. random variables with zero mean and $\mathrm{var}(\varepsilon_k) = \sigma^2 > 0$. Let ρ be a robustifying loss function such that its derivative ψ (the so-called psi-function) is monotonous. In this situation the M-estimator of $\boldsymbol{\varphi}$ and σ is obtained by solving the minimization problem

$$\min_{\varphi \in R^p, \, \sigma > 0} \sum_{k=p+1}^{n} \rho \left(\frac{y_k - \boldsymbol{\varphi}' \mathbf{z}_{k-1}}{\sigma} \right) \sigma + c\, \sigma \tag{7}$$

with

$$c = \frac{n-2p-1}{2} \, E[\psi^2(A)] = \frac{n-2p-1}{2} \int \psi^2(x) \, \mathrm{d}\Phi(x), \tag{8}$$

where A is a random variable with distribution $N(0,1)$ (see e. g. [11], [23]).

Although the M-estimation is commonly used in the IO-models due to its advantageous properties in these models (e. g. consistency and asymptotic normality under regularity conditions) it may provide bad results (comparable with the LS-estimation) if it is applied in the AO-models. Namely, the M-estimators in the AO-models may be biased and have substantial variances even in large samples (asymptotically) and lose the advantageous property of the efficiency robustness that can be shown for them in the IO-models (see e. g. [8], [16], [18]). This fact is not surprising since the M-estimators are not capable to cope with additive errors in the regressors $\mathbf{z}_{k-1}$ in (1.7). Therefore in the case of the AO-models, it is necessary to modify the M-estimation in a suitable way.

One of possible modifications is the GM-estimation (General M-estimation, see e. g. [23]) introducing to the minimized expression (1.7) certain weights that reduce the influence of outliers in $\mathbf{z}_k$. If using this method, the minimization problem (1.7) is rewritten to the form

$$\min_{\varphi \in R^p, \, \sigma > 0} \sum_{k=p+1}^{n} u_k \, v_k \, \rho \left(\frac{y_k - \boldsymbol{\varphi}' \mathbf{z}_{k-1}}{u_k \, \sigma} \right) \sigma + d\sigma \tag{9}$$

with

$$d = \frac{n-2p-1}{2}\, E(u_k\, v_k)\, E[\psi^2(A)], \tag{10}$$

where u_k and v_k are weights depending on the size of $\mathbf{z}_{k-1}$ (see e. g. [18], [19], [23, pp. 39–41]). For a Mallows type GM-estimator one uses $u_k = 1$ and $v_k = \psi_1(b_{k-1})/b_{k-1}$, while a Schweppe type GM-estimator has $u_k = v_k = \psi_1(b_{k-1})/b_{k-1}$, where b_{k-1} denotes the size of $\mathbf{z}_{k-1}$ and ψ_1 is a psi-function that can be different from $\psi = \rho'$. The size of $\mathbf{z}_{k-1}$ can be assessed by $b_{k-1} = \left(p^{-1}\, \mathbf{z}_{k-1}{}'\widehat{\mathbf{C}}^{-1}\, \mathbf{z}_{k-1}\right)^{1/2}$, where $\widehat{\mathbf{C}}^{-1}$ is an estimate of the inverse $p \times p$ covariance matrix of the outlier-free process $\{x_t\}$.

This paper concentrates on the other modification that is called the CMM-estimation (Conditional-Mean M-estimation, see [16], [23]). The CMM-estimation replaces the values $\mathbf{z}_{k-1}$ in (1.7) by filtered ones using the ACM-filter (Approximate Conditional-Mean filter, see [17], [23]) which is a robust version of the Kalman filter. The fact that the recursive robust estimation procedures suggested in this paper are based on the CMM-estimation has some advantageous consequences: (1) they provide not only robust parameter estimate but simultaneously smoothed values of the analysed time series and (2) are applicable also to the IO-models (one can compare these procedures with the ones suggested for the IO-models e. g. in [2], [5], [22]).

Section 2 of the paper reminds briefly the ACM-filtering and the CMM-estimation in the context of the AO-AR(p) model. The corresponding recursive robust procedures are suggested in Section 3. Some considerations concerning the convergence of these procedures and results of a simulation study are given in Section 4.

2. ACM–FILTERING AND CMM–ESTIMATION

The ACM-filter is a robust version of the Kalman filter suitable just for the AO-AR-models. The robust Kalman filtering belongs to very up-to-date topics since it enables to treat linear dynamic systems contaminated by outliers, see e. g. references in [5], [6] (there are even attempts of nonlinear robust filtering, see e. g. [3], [7]).

Analogously as the classical Kalman filter in the case without outliers (see e. g. [12]), the ACM-filter enables recursive calculation of the smoothed values

$$\widehat{x}_t^t = E(x_t \,|\, Y^t), \tag{11}$$

where $Y^t = \{y_t, y_{t-1}, \ldots\}$ denotes the values observed in the AO-AR(p)-model (1.1), (1.2) till the current time period t. For simplicity, let be $\sigma_1^2 = 0$ in (1.3) so that (1.4) holds. Then the recursive formula of the ACM-filter has the form

$$\widehat{x}_t^t = \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{t-1}^{t-1} + s_t\, \psi\left(\frac{y_t - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{t-1}^{t-1}}{s_t}\right) \tag{12}$$

where $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_p)'$, $\widehat{\mathbf{x}}_{t-1}^{t-1} = \left(\widehat{x}_{t-1}^{t-1}, \ldots, \widehat{x}_{t-p}^{t-p}\right)'$, ψ is a robustifying psi-function and s_t is a suitable estimate of $\sigma = [\mathrm{var}(\varepsilon_t)]^{1/2}$ constructed at time t. The details concerning the ACM-filter can be found e. g. in [17], [19], [23].

The CMM-estimation combines the principle of M-estimation with the one of ACM-filtering. The CMM-estimates of parameters $\boldsymbol{\varphi}$, σ in the AO-AR(p)-model (1.1), (1.2) are obtained by solving the minimization problem

$$\min_{\varphi \in R^p,\, \sigma > 0} \sum_{k=p+1}^{n} \rho \left(\frac{y_k - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{k-1}^{k-1}}{\sigma} \right) \sigma + c\,\sigma \tag{13}$$

(compare with (1.7)), where c is given in (1.8) and the smoothed values $\widehat{\mathbf{x}}_{k-1}^{k-1}$ are obtained by means of the ACM-filter (2.2). From the practical point of view one recommends to proceed iteratively alternating the ACM-filtering (2.2) and the M-estimation (2.3): the M-estimates $\widehat{\boldsymbol{\varphi}}$, $\widehat{\sigma}$ obtained by (2.3) in a certain step of the recommended iterative procedure are used to obtain the smoothed values $\widehat{x}_t^t$ by (2.2) in the next iteration. The estimate s_t of σ in (2.2) at time t can be calculated by means of Kalman filter scheme, or it is possible to use directly the estimate $\widehat{\sigma}$ if one replaces the ACM-filter formula (2.2) by the simplified form

$$\widehat{x}_t^t = \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{t-1}^{t-1} + s\,\psi \left(\frac{y_t - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{t-1}^{t-1}}{s} \right) \tag{14}$$

(see e. g. [4], [16], [23]). Some evaluations of the properties of the CMM-estimates in the framework of the spectral autoregressive estimation are given in [13], [20].

3. RECURSIVE ROBUST ESTIMATION

The recursive robust estimation procedures for the AO-AR(p)-model suggested in this section are based on the idea that we replace the iteration scheme by the recursive one in the CMM-estimation (see Section 2). In principle it is possible to do it in two different ways:

In the first suggested procedure one combines the ACM-filter (2.2) with the recursive least squares (RLS) estimation (the RLS-estimation is described e. g. in [9], [15]) solving the minimization problem

$$\min_{\varphi \in R^p} \sum_{k=p+1}^{n} \left(\widehat{x}_k^k - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{k-1}^{k-1} \right)^2 \tag{15}$$

(the application of the classical non-robustified LS-criterion (3.1) is justified due to the replacement of the contaminated values y_k and $\mathbf{z}_{k-1}$ in (1.7) by the smoothed values $\widehat{x}_k^k$ and $\widehat{\mathbf{x}}_{k-1}^{k-1}$, respectively). Moreover, the recursive estimation of the parameter σ^2 can be obtained if one rewrites the estimation formula

$$\widehat{\sigma}^2 = \frac{1}{n-2p-1} \sum_{k=p+1}^{n} \left(\widehat{x}_k^k - \widehat{\boldsymbol{\varphi}}'\widehat{\mathbf{x}}_{k-1}^{k-1} \right)^2 \tag{16}$$

in a recursive way. Thus the first suggested recursive procedure can be summarized as

$$\widehat{x}_t^t = \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1} + \widehat{\sigma}_{t-1}\,\psi\left(\frac{y_t - \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1}}{\widehat{\sigma}_{t-1}}\right), \tag{17}$$

$$\widehat{\boldsymbol{\varphi}}_t = \widehat{\boldsymbol{\varphi}}_{t-1} + \frac{\mathbf{V}_{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}}{1+\widehat{\mathbf{x}}_{t-1}^{t-1}{}'\mathbf{V}_{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}}\left(\widehat{x}_t^t - \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1}\right), \tag{18}$$

$$\mathbf{V}_t = \mathbf{V}_{t-1} - \frac{\mathbf{V}_{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}{}'\mathbf{V}_{t-1}}{1+\widehat{\mathbf{x}}_{t-1}^{t-1}{}'\mathbf{V}_{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}}, \tag{19}$$

$$\widehat{\sigma}_t^2 = \frac{1}{t-2p-1}\left[(t-2p-2)\,\widehat{\sigma}_{t-1}^2 + (\widehat{x}_t^t - \widehat{\boldsymbol{\varphi}}_t{}'\widehat{\mathbf{x}}_{t-1}^{t-1})^2\right], \tag{20}$$

$$\widehat{\mathbf{x}}_t^t = \left(\widehat{x}_t^t,\, \widehat{x}_{t-1}^{t-1}, \ldots,\, \widehat{x}_{t-p+1}^{t-p+1}\right)'. \tag{21}$$

The initial values that must be prechosen for the first recursive procedure (3.3) – (3.7) are $\widehat{\mathbf{x}}_p^p$, $\widehat{\boldsymbol{\varphi}}_p$, $\mathbf{V}_p$, $\widehat{\sigma}_p$.

The second procedure combines the ACM-filter (2.2) with the recursive weighted least squares (RWLS) estimation solving the minimization problem (2.3). The RWLS-estimation is motivated by the iterated weighted least squares (IWLS) estimation described e. g. in [11], [23].

Let us remind briefly the IWLS-scheme in the context of the problem (2.3). The corresponding normal equations obtained by deriving the minimized expression (2.3) with respect to $\boldsymbol{\varphi}$ and σ have the form (see e. g. [23])

$$\sum_{k=p+1}^{n} \psi\left(\frac{y_k - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{k-1}^{k-1}}{\sigma}\right)\widehat{\mathbf{x}}_{k-1}^{k-1} = \mathbf{0}, \tag{22}$$

$$\frac{1}{n-2p-1}\sum_{k=p+1}^{n} \chi\left(\frac{y_k - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{k-1}^{k-1}}{\sigma}\right) = d \tag{23}$$

with

$$\chi(x) = x\,\psi(x) - \rho(x), \tag{24}$$

$$d = E\left[\psi^2(A)\right]/\,2, \tag{25}$$

where A is a random variable with distribution $N(0,1)$ similarly as in (1.8). Let us denote

$$w_k^{(m-1)} = \psi\left(\frac{y_k - \widehat{\boldsymbol{\varphi}}^{(m-1)}{}'\widehat{\mathbf{x}}_{k-1}^{k-1}}{\widehat{\sigma}^{(m-1)}}\right)\Bigg/\frac{y_k - \widehat{\boldsymbol{\varphi}}^{(m-1)}{}'\widehat{\mathbf{x}}_{k-1}^{k-1}}{\widehat{\sigma}^{(m-1)}}, \tag{26}$$

where $\widehat{\boldsymbol{\varphi}}^{(m-1)}$ and $\widehat{\sigma}^{(m-1)}$ are estimates obtained in the iteration $m-1$. Then the equation (3.8) is approximated by the WLS-form

$$\sum_{k=p+1}^{n} w_k^{(m-1)}\left(y_k - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{k-1}^{k-1}\right)\widehat{\mathbf{x}}_{k-1}^{k-1} = \mathbf{0}. \tag{27}$$

This approximation, that is recommended in the robust statistical literature (see e. g. [23, pp. 26–27]), consists in replacing the expression

$$\psi\left(\frac{y_k - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{k-1}^{k-1}}{\sigma}\right) \Bigg/ \frac{y_k - \boldsymbol{\varphi}'\widehat{\mathbf{x}}_{k-1}^{k-1}}{\sigma}$$

by its estimate $w_k^{(m-1)}$ (from the previous iterative step) in (3.8).

The resulting IWLS-scheme gives the estimates $\widehat{\boldsymbol{\varphi}}^{(m)}$ and $\widehat{\sigma}^{(m)}$ of the iteration m in the following form

$$\widehat{\boldsymbol{\varphi}}^{(m)} = \left[\sum_{k=p+1}^{n} w_k^{(m-1)}\,\widehat{\mathbf{x}}_{k-1}^{k-1}\,\widehat{\mathbf{x}}_{k-1}^{k-1\prime}\right]^{-1} \sum_{k=p+1}^{n} w_k^{(m-1)}\,\widehat{\mathbf{x}}_{k-1}^{k-1}\,y_k, \tag{28}$$

$$\left[\widehat{\sigma}^{(m)}\right]^2 = \frac{[\widehat{\sigma}^{(m-1)}]^2}{(n-2p-1)\,d} \sum_{k=p+1}^{n} \chi\left(\frac{y_k - \widehat{\boldsymbol{\varphi}}^{(m-1)\prime}\,\widehat{\mathbf{x}}_{k-1}^{k-1}}{\widehat{\sigma}^{(m-1)}}\right). \tag{29}$$

If one rearranges (3.14), (3.15) recursively using the estimates from the previous recursive step instead of the ones from the previous iterative step then one obtains the second recursive procedure that can be summarized in the following way:

$$\widehat{\sigma}_t^2 = \frac{\widehat{\sigma}_{t-1}^2}{t-2p-1}\left[t-2p-2+\frac{1}{d}\,\chi\left(\frac{y_t - \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1}}{\widehat{\sigma}_{t-1}}\right)\right], \tag{30}$$

$$w_t = \psi\left(\frac{y_t - \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1}}{\widehat{\sigma}_t}\right) \Bigg/ \frac{y_t - \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1}}{\widehat{\sigma}_t}, \tag{31}$$

$$\widehat{\boldsymbol{\varphi}}_t = \widehat{\boldsymbol{\varphi}}_{t-1} + \frac{\mathbf{V}_{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}}{w_t^{-1} + \widehat{\mathbf{x}}_{t-1}^{t-1}{}'\mathbf{V}_{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}}\left(y_t - \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1}\right), \tag{32}$$

$$\mathbf{V}_t = \mathbf{V}_{t-1} - \frac{\mathbf{V}_{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}\,\widehat{\mathbf{x}}_{t-1}^{t-1}{}'\mathbf{V}_{t-1}}{w_t^{-1} + \widehat{\mathbf{x}}_{t-1}^{t-1}{}'\mathbf{V}_{t-1}\,\widehat{x}_{t-1}^{t-1}}, \tag{33}$$

$$\widehat{x}_t^t = \widehat{\boldsymbol{\varphi}}_t{}'\widehat{\mathbf{x}}_{t-1}^{t-1} + \widehat{\sigma}_t\,\psi\left(\frac{y_t - \widehat{\boldsymbol{\varphi}}_t{}'\widehat{\mathbf{x}}_{t-1}^{t-1}}{\widehat{\sigma}_t}\right), \tag{34}$$

$$\widehat{\mathbf{x}}_t^t = \left(\widehat{x}_t^t,\,\widehat{x}_{t-1}^{t-1},\ldots,\widehat{x}_{t-p+1}^{t-p+1}\right)'. \tag{35}$$

The initial values for the second recursive procedure (3.16) – (3.21) are again $\widehat{\mathbf{x}}_p^p$, $\widehat{\boldsymbol{\varphi}}_p$, $\mathbf{V}_p$, $\widehat{\sigma}_p$.

Remark 3.1. The procedure (3.16) – (3.21) can be simplified in various ways. For example, the complicated formula (3.16) for recursive calculation of $\widehat{\sigma}_t$ can be replaced by the following "ad hoc" one

$$\widehat{\sigma}_t = 1.25\,\nu\,\widehat{\sigma}_{t-1}\,\psi\left(\frac{|y_t - \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1}|}{\widehat{\sigma}_{t-1}}\right) + (1-\nu)\,\widehat{\sigma}_{t-1}, \tag{36}$$

where ν $(0 < \nu < 1)$ is a constant near to zero. Namely, the formulas of the type $\widehat{\sigma}_t = 1.25\,\nu\,\left|y_t - \widehat{\boldsymbol{\varphi}}_{t-1}{}'\widehat{\mathbf{x}}_{t-1}^{t-1}\right| + (1-\nu)\,\widehat{\sigma}_{t-1}$ are popular in the framework of the exponential smoothing in time series analysis (see e. g. [21]). In (3.22), this formula is robustified using the psi-function ψ.

4. SOME REMARKS ON SUGGESTED PROCEDURES

Although the suggested procedures seem to give good numerical results it is very difficult to investigate them theoretically (specially, it concerns the proof of their convergence). Such a lack of theoretical results appears also in other works dealing with recursive time series procedures (see e. g. [9], [22]).

On the other hand, some related results under certain simplifying assumptions can serve as a partial justification of the suggested algorithms from the theoretical point of view. For example, if one replaces the estimated values $\widehat{x}_t^t$ by the theoretical values x_t in the procedure (3.3) – (3.7) for $p = 1$ then the following convergence result holds.

Lemma. Let the model (1.2) with $p = 1$ fulfil the following assumptions

(i) $|\varphi_1| < 1$;

(ii) $\varepsilon_t \sim$ iid with a distribution F that is symmetric and increasing in zero (i. e. $F(-\eta) < F(\eta)$ for each $\eta > 0$).

Let the estimate $\widehat{\varphi}_{1t}$ of the parameter φ_1 at time t be given by means of the recursive formulas

$$\widehat{\varphi}_{1t} = \widehat{\varphi}_{1,t-1} + \frac{x_{t-1}\,V_{t-1}}{1 + x_{t-1}^2\,V_{t-1}}\,\psi_H\left(x_t - \widehat{\varphi}_{1,t-1}\,x_{t-1}\right), \tag{37}$$

$$V_t = V_{t-1} - \frac{x_{t-1}^2\,V_{t-1}^2}{1 + x_{t-1}^2\,V_{t-1}} = \frac{V_{t-1}}{1 + x_{t-1}^2\,V_{t-1}}, \tag{38}$$

$$E(\widehat{\varphi}_{1,0}) < \infty,\ V_0 > 0 \text{ a.s.},\ \widehat{\varphi}_{1,0},\ V_0,\ \varepsilon_t \text{ are independent}, \tag{39}$$

where ψ_H is the Huber's psi-function

$$\psi_H(x) = \begin{cases} x, & |x| \le c, \\ c\,\mathrm{sgn}\,(x), & |x| > c \end{cases} \tag{40}$$

with an arbitrary positive constant c. Then

$$\widehat{\varphi}_{1t} \longrightarrow \varphi_1 \quad \text{a.s.} \tag{41}$$

The proof is similar as for Theorem in [5]. □

Remark 4.1. Obviously, the formula (4.1) in comparison with the original formula (3.4) contains the values x_t, x_{t-1} instead of $\widehat{x}_t^t$, $\widehat{x}_{t-1}^{t-1}$. Since the practical numerical experiences with the performance of the ACM-filter (3.3) are favourable this Lemma gives some hints concerning the convergence of the recursive procedure (3.3) – (3.7). On the other hand, one can see that it is necessary from the asymptotic point of view to trim off large residuals $x_t - \widehat{\varphi}_{1,t-1}\, x_{t-1}$. Similar hints can be obtained from the convergence results by [2].

In the literature it is recommended (see e. g. [15, p. 136]) that theoretical analyses of recursive time series methods should be complemented by simulation studies to judge some practical aspects (finite-sample properties, the influence of choice of initial values). As an example let us present numerical results of a simulation study evaluating the influence of choice of initial values. A process $\{y_t\}$ $(t = 1, \ldots, 100)$ is generated as

$$y_t = x_t + v_t,$$

where $\{x_t\}$ is the autoregressive process of the first order

$$x_t = 0.5\, x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, 1)$$

(i. e. $\varphi_1 = 0.5,\ \sigma = 1$) and $\{v_t\}$ is the contaminating process

$$v_t = \begin{cases} 10, & t = 20,\ 40,\ 60,\ 80,\ 100, \\ 0, & \text{otherwise.} \end{cases}$$

The recursive formulas (3.17) – (3.22) are applied (see the second recursive procedure in Section 3 and Remark 3.1) with $p = 1$ and $\nu = 0.1$. The Huber's psi-function (4.4) is used in (3.17), (3.20) and (3.22) with $c = 1.645$ (this choice of the constant c corresponds to the 5

Table 1 contains the values $\widehat{\varphi}_{1,t}$ and $\widehat{\sigma}_t$ for $t = 10,\ 20,\ 30, \ldots,\ 100$ corresponding to various choices of $\widehat{\varphi}_{1,1}$ and $\widehat{\sigma}_1$ (it is $\widehat{x}_1^1 = y_1$ and $V_1 = 1/y_1^2$ in all cases). One can see that $\widehat{\varphi}_{1t}$ and $\widehat{\sigma}_t$ quickly become acceptable estimates of the true values $\varphi_1 = 0.5$ and $\sigma = 1$ even if the initial values are not chosen near to the true values.

Table 1. Simulation example of the recursive robust estimation for the process A0-AR(1) with $\varphi_1 = 0.5,\ \sigma = 1$ and artificial outliers $v_{20} = v_{40} = v_{60} = v_{100} = 10$ for various choices of initial values.

	$\widehat{\varphi}_{1,1} = -0.8$	$\widehat{\sigma}_1 = 3$	$\widehat{\varphi}_{1,1} = -0.5$	$\widehat{\sigma}_1 = 2$	$\widehat{\varphi}_{1,1} = 0.3$	$\widehat{\sigma}_1 = 0.5$	$\widehat{\varphi}_{1,1} = 0.8$	$\widehat{\sigma}_1 = 3$
t	$\widehat{\varphi}_{1t}$	$\widehat{\sigma}_t$	$\widehat{\varphi}_{1t}$	$\widehat{\sigma}_t$	$\widehat{\varphi}_{1t}$	$\widehat{\sigma}_t$	$\widehat{\varphi}_{1t}$	$\widehat{\sigma}_t$
10	0.41	2.04	0.42	1.70	0.69	0.67	0.67	1.74
20	0.66	1.52	0.67	1.38	0.74	0.82	0.74	1.41
30	0.46	1.36	0.48	1.30	0.58	0.97	0.54	1.33
40	0.45	1.24	0.47	1.21	0.56	1.05	0.52	1.25
50	0.42	1.14	0.44	1.13	0.50	1.08	0.48	1.14
60	0.43	1.19	0.44	1.18	0.50	1.16	0.48	1.19
70	0.37	0.89	0.38	0.89	0.43	0.88	0.42	0.90
80	0.46	1.53	0.47	1.53	0.51	1.56	0.50	1.56
90	0.45	1.13	0.46	1.13	0.49	1.16	0.49	1.16
100	0.48	1.03	0.48	1.03	0.52	1.02	0.51	1.03

Table 2. Simulation example of the recursive robust estimation and the GM-estimation for the process A0-AR(1) with $\varphi_1 = 0.5$, $\sigma = 1$ and artificial outliers $v_{20} = v_{40} = v_{60} = v_{100} = 10$.

	Recursive procedure (3.17) – (3.22)		GM-estimation	
t	$\widehat{\varphi}_{1t}$	$\widehat{\sigma}_t$	$\widehat{\varphi}_{1t}$	$\widehat{\sigma}_t$
5	0.470	7.394	-0.010	1.211
10	0.491	4.638	0.500	1.149
15	0.560	3.219	0.443	1.268
20	0.593	2.671	0.379	1.157
25	0.364	2.122	0.089	1.057
30	0.381	1.612	0.078	0.981
35	0.369	1.430	0.048	0.982
40	0.431	1.289	0.072	0.970
45	0.459	1.291	0.061	1.009
50	0.506	1.295	0.115	1.085
55	0.506	1.072	0.188	1.139
60	0.509	1.136	0.234	1.184
65	0.545	1.025	0.233	1.193
70	0.543	0.880	0.257	1.189
75	0.544	0.948	0.274	1.191
80	0.559	0.960	0.329	1.254
85	0.528	1.243	0.273	1.282
90	0.531	1.143	0.284	1.271
95	0.550	1.069	0.328	1.321
100	0.578	1.192	0.350	1.330

Table 3. Means and MSE's of the estimated parameters over 100 simulations of the same type as the one in Table 2.

	Recursive procedure (3.17) – (3.22)				GM-estimation			
	$\widehat{\varphi}_{1t}$		$\widehat{\sigma}_t$		$\widehat{\varphi}_{1t}$		$\widehat{\sigma}_t$	
t	Mean	MSE	Mean	MSE	Mean	MSE	Mean	MSE
20	0.39	0.15	4.86	2.05	0.22	0.32	1.22	0.47
40	0.60	0.09	1.20	0.44	0.34	0.21	0.96	0.21
60	0.49	0.09	1.17	0.25	0.29	0.18	1.16	0.11
80	0.51	0.06	0.92	0.18	0.38	0.19	1.19	0.09
100	0.53	0.06	1.13	0.20	0.41	0.12	1.25	0.09

Tables 2 and 3 compare the suggested methodology with the GM-estimation for the simulated AO-AR(1) process of the same type as in Table 1. The initial values $\widehat{\varphi}_{1,1} = 0$ and $\widehat{\sigma}_1 = 10$ for the recursive formulas (3.17) – (3.22) with $p = 1$, $\nu = 0.1$, $\psi = \psi_H$ and $c = 1.645$ (see the second recursive procedure in Section 3 and Remark 3.1) have been chosen far from the true values $\varphi_1 = 0.5$ and $\sigma = 1$ (the remaining initial values are $\widehat{x}_1^1 = y_1$ and $V_1 = 1/y_1^2$). The GM-estimates have been obtained using "Fit Autoregression Using Robust GM-Estimates" code of the system S-PLUS (Version 3.0 from September 1991, Statistical Sciences, Inc., Seattle, Washington) installed at the Department of Statistics, University of British Columbia, Vancouver. The recommended defaults of this code have been respected replacing only the constant $c = 1.5$ of the Huber's psi-function by its previous value $c = 1.645$. Table 2 contains the estimates $\widehat{\varphi}_{1t}$ and $\widehat{\sigma}_t$ ($t = 5, 10, 15, \ldots, 100$) for a simulation of the previous type. It is necessary to keep in mind that the GM-estimation used is a non-recursive procedure so that e. g. the GM-estimates $\widehat{\varphi}_{1,50}$ and $\widehat{\sigma}_{50}$ are obtained in a non-recursive way using all observations $y_1, \ldots, y_{50}$. Table 3 contains the means and MSE's of the estimated parameters ($t = 20, 40, 60, 80, 100$) over 100 simulations of the same type as the one in Table 2. This comparison indicates that the GM-estimation requires much more observations than the suggested recursive procedure even if the initial parameter values for the recursive procedure are chosen far away from the true ones.

(Received December 12, 1991.)

REFERENCES

[1] D. F. Andrews *et al.*: Robust Estimates of Location: Survey and Advances. Princeton University Press, Princeton 1972.

[2] K. Campbel: Recursive computation of M-estimates for the parameters of a finite autoregressive process. Ann. Statist. *10* (1982), 442–453.

[3] T. Cipra: Note on approximate non-Gaussian filtering with nonlinear observation relation. Comment. Math. Univ. Carolinae *31* (1990), 601–605.

[4] T. Cipra and J. Hanousek: Robust Filtering and Estimation in Autoregressive Models (in Czech). Research Report, Institute of Hygiene and Epidemiology, Prague 1989.

[5] T. Cipra and R. Romera: Robust Kalman filter and its application in time series analysis. Kybernetika *27* (1991), 481–494.

[6] T. Cipra and R. Romera: Recursive time series methods in L_1-norm. In: L_1-Statistical Analysis and Related Methods (Y. Dodge, ed.), North Holland, Amsterdam 1992, pp. 233–243.

[7] T. Cipra and A. Rubio: Kalman filter with a non-linear non-Gaussian observation relation. Trabajos de Estadstica *6* (1991), 111–119.

[8] L. Denby and R. D. Martin: Robust estimation of the first-order autoregressive parameter. J. Amer. Statist. Assoc. *74* (1979), 140–146.

[9] L. Fahrmeier: Rekursive Algorithmen fr Zeitreihenmodelle. Vandenhoeck und Ruprecht, Gttingen 1981.

[10] S. Hillmer: Monitoring and adjusting forecasts in the presence of additive outliers. J. Forecasting *3* (1984), 205–215.

[11] P. J. Huber: Robust Statistics. J. Wiley, New York 1981.

[12] A. Jazwinski: Stochastic Processes and Filtering Theory. Academic Press, New York 1970.

[13] B. Kleiner, R. D. Martin and D. J. Thomson: Robust estimation of power spectra. J. Roy. Statist. Soc. Ser. B *41* (1979), 313–351.

[14] J. Ledolter: The effect of additive outliers on the forecasts from ARIMA models. Internat. J. Forecasting *5* (1989), 1–10.

[15] L. Ljung and T. Sderstrm: Theory and Practice of Recursive Identification. MIT Press, Cambridge, Massachusetts 1983.

[16] R. D. Martin: Robust estimation for time series autoregressions. In: Robustness in Statistics (R. L. Launer and G. N. Wilkinson, eds.), Academic Press, New York 1979, pp. 147–176.

[17] R. D. Martin: Approximate conditional-mean type smoothers and interpolators. In: Smoothing Techniques for Curve Estimation (T. Gasser and M. Rosenblatt, eds.), Springer, New York 1979, pp. 117–143.

[18] R. D. Martin: Robust estimation of autoregressive models. In: Directions in Time Series (D. R. Brillinger and G. C. Tiao, eds.), Inst. Math. Statist. Publications, Hayward, CA 1980, pp. 228–254.

[19] R. D. Martin: Robust methods for time series. In: Applied Time Series II (D. F. Findley, ed.), Academic Press, New York 1981, pp. 683–759.

[20] R. D. Martin and D. J. Thomson: Robust-resistant spectrum estimation. Proc. IEEE *70* (1982), 1097–1114.

[21] D. C. Montgomery and L. A. Johnson: Forecasting and Time Series Analysis. McGraw-Hill, New York 1976.

[22] K. Sejling, H. Madsen, J. Holst, U. Holst and J.-E. England: A method for recursive robust estimation of AR-parameters. Preprint, Technical University Lyngby, Denmark and University of Lund, Sweden 1990.

[23] N. Stockinger and R. Dutter: Robust Time Series and Analysis: A Survey. Supplement to Kybernetika *23* (1987).

Doc. RNDr. Tom Cipra, CSc., matematicko-fyzikln fakulta UK (Faculty of Mathematics and Physics – Charles University), Sokolovsk 83, 186 00 Praha 8. Czech Republic.

Prof. Dr. Asuncin Rubio, Departamento de Matemticas, Facultad de Veterinaria, Universidad de Extremadura, 10071 Cceres. Spain.

Prof. Dr. Jos Luis Canal, Departamento de Matemticas, Escuela Politcnica, Universidad de Extremadura, 10071 Cceres. Spain.