INFORMATICA, 1993, Vol.4, No.1-2, 94-110

OFF-LINE ESTIMATION OF DYNAMIC SYSTEMS PARAMETERS IN THE PRESENCE OF OUTLIERS IN AUTOREGRESSIVE NOISE

Rimantas PUPEIKIS

Institute of Mathematics and Informatics,
2600 Vilnius, Akademijos St.4, Lithuania

Abstract. In the previous paper (Pupeikis, 1992) the problem of off-line estimation of dynamic systems parameters in the presence of outliers in observations have been considered, when the filter generating an additive noise has a very special form. The aim of the given paper is the development, in such a case, of classical generalized least squares method (GLSM) algorithms for off-line estimation of unknown parameters of dynamic systems. Two approaches using batch processing of the stored data are worked out. The first approach, is based on the application of S-, H-, W- algorithms used for calculation of M-estimates, and the second one rests on the replacement of the corresponding values of the sample covariance and cross-covariance functions by their robust analogues in respective matrices of GLSM and on a further application of the least squares (LS) parameter estimation algorithms. The results of numerical simulation by IBM PC/AT (Table 1) are given.

Key words: dynamic system, parameter estimation, covariance analysis, outlier, robustness.

1. Statement of the problem. Consider a single input x_k and a single output y_k linear discrete-time system described by the difference equation

$$y_k = -a_1 y_{k-1} - \cdots - a_n y_{k-n} + b_1 x_{k-1} + \cdots + b_n x_{k-n} \quad . \tag{1}$$

Suppose that y_k is observed under additive noise ξ_k^*, i.e.,

$$u_k = y_k + \xi_k^* \tag{2}$$

then

$$u_k = -a_1 u_{k-1} - \cdots - a_n u_{k-n} + b_1 x_{k-1} + \cdots + b_n x_{k-n} + \xi_k^* + a_1 \xi_{k-1}^* + \cdots + a_n \xi_{k-n}^* \tag{3}$$

or

$$u_k = \frac{B(z^{-1})}{1 + A(z^{-1})} x_k + \xi_k^* \tag{4}$$

by introducing the backward shift operator z^{-1} defined by $z^{-1} x_k = x_{k-1}$, where

$$\xi_k = (1 - \gamma_k) v_k + \gamma_k \eta_k \tag{5}$$

is a sequence of independent identically distributed variables with an ε - contaminated distribution of the form

$$p(\xi_k) = (1 - \varepsilon) N(0, \sigma_1^2) + \varepsilon N(0, \sigma_2^2), \tag{6}$$

$p(\xi_k)$ is a probability density distribution of the sequence ξ_k; γ_k is a random variable, taking values 0 or 1 with the probabilities $p(\gamma_k = 1) = \varepsilon$, $p(\gamma_k = 1) = 1 - \varepsilon$; v_k, η_k are sequences of independent Gaussian variables with zero means and variances σ_1^2, σ_2^2 respectively,

$$c^T = (a^T, b^T), \quad a^T = (a_1, \ldots, a_n), \quad b^T = (b_1, \ldots, b_n) \tag{7}$$

$$B(z^{-1}) = \sum_{i=1}^{n} b_i z^{-i}, \quad A(z^{-1}) = \sum_{i=1}^{n} a_i z^{-i}; \tag{8}$$

n is the order of difference equation (1), respectively;

$$\xi_k^* = W(z^{-1}; h) \xi_k, \tag{9}$$

$W(z^{-1}, h)$ is a noise filter transfer function, h is a vector of parameters.

It is assumed that the roots of $A(z^{-1})$ are outside the unit circle of the z^{-1} plane. The true orders of the polynomials $A(z^{-1})$, $B(z^{-1})$

are known. The input signal x_k is persistent excitation of an arbitrary order according to (Åström and Eykhoff,1971).

The aim of the given paper is the development of ordinary GLSM algorithms for the computation of off-line estimates of the unknown parameters of the dynamic system (1) – (9) in the presence of outliers in observations.

2. Parameter estimation in the absence of outliers in observations. Suppose that $\varepsilon = 0$ in equation (6). In this case, as shown in (Åström and Eykhoff, 1971; Isermann (1981); Young (1984); Ljung (1987)) multivariate approaches are worked out to estimate the vector of unknown parameters. It is known that in the case when

$$W(z^{-1};h) = \left[1 + A(z^{-1})\right]^{-1} \tag{10}$$

an ordinary classical LS parameter estimation algorithm is used. On the other hand, it is also known that in a real situation, of course, relationship (10) is hardly satisfied. Therefore the ordinary LS used to estimate unknown parameters of a mathematical model of the dynamic system (1) – (9) is inefficient and that's why the estimates of the above mentioned parameters will be biased.

Clarke (1967) works out the GLSM, which requires that

$$W(z^{-1};h) = \left[1 + G(z^{-1})\right]^{-1}\left[1 + A(z^{-1})\right]^{-1}, \tag{11}$$

where

$$1 + G(z^{-1}) = \sum_{i=1}^{n_g} g_i z^{-i} \tag{12}$$

h is a vector of parameters which correspond to different combinations of $a_1, \ldots, a_n$ and $g_1, \ldots, g_{n_g}$.

This method is more flexible and more useful in practice than LS. In this case the vector $\widehat{c}^T = (\widehat{a}^T, \widehat{b}^T)$ of the estimates $\widehat{a}^T = (\widehat{a}_1, \ldots, \widehat{a}_n)$, $\widehat{b}^T = (\widehat{b}_1, \ldots, \widehat{b}_n)$ of the respective parameters (7) and the vector $\widehat{g}^T = (\widehat{g}_1, \ldots, \widehat{g}_{n_g})$ of the estimates of the parameters $g^T = (g_1, \ldots, g_{n_g})$ are calculated using two classical LS of the form

$$\hat{c}^{(j)} = (\Psi^{T(j)}\Psi^{(j)})^{-1}\Psi^{T(j)}U^{(j)} \ , \tag{13}$$

$$\hat{g}^{(j)} = (Q^{T(j)}Q^{(j)})^{-1}Q^{T(j)}E^{(j)} \ , \tag{14}$$

respectively, where

$$\hat{c}^{T(j)} = (\hat{a}^{T(j)}, \hat{b}^{T(j)}) = (\hat{a}_1^{(j)}, \dots, \hat{a}_n^{(j)}, \hat{b}_1^{(j)}, \dots, \hat{b}_n^{(j)}) \ , \tag{15}$$

$$\hat{g}^{T(j)} = (\hat{g}_1^{(j)}, \dots, \hat{g}_{n_e}^{(j)}) \tag{16}$$

are the estimates of parameters c and g, which are calculated at the j-th iteration using the above mentioned LS algorithms;

$$\Psi^{T(j)}\Psi^{(j)} = \begin{pmatrix} \Psi_{11}^{(j)} & \Psi_{12}^{(j)} \\ \Psi_{21}^{(j)} & \Psi_{22}^{(j)} \end{pmatrix} , \tag{17}$$

$$\Psi_{11}^{(j)} = \begin{bmatrix} R_u{}^{(j)}(0) & R_u{}^{(j)}(1) & \cdots & R_u{}^{(j)}(n-1) \\ & R_u{}^{(j)}(0) & \cdots & R_u{}^{(j)}(n-2) \\ & & \ddots & \vdots \\ & & & R_u{}^{(j)}(0) \end{bmatrix} , \tag{18}$$

$$\Psi_{22}^{(j)} = \begin{bmatrix} R_x{}^{(j)}(0) & R_x{}^{(j)}(1) & \cdots & R_x{}^{(j)}(n-1) \\ & R_x{}^{(j)}(0) & \cdots & R_x{}^{(j)}(n-2) \\ & & \ddots & \vdots \\ & & & R_x{}^{(j)}(0) \end{bmatrix} \tag{19}$$

are $n \times n$ - symmetric submatrices;

$$\Psi_{12}^{(j)} = \Psi_{21}^{(j)} = \\ = \begin{bmatrix} -R_{ux}{}^{(j)}(0) & -R_{xu}{}^{(j)}(1) & \cdots & -R_{xu}{}^{(j)}(n-1) \\ -R_{ux}{}^{(j)}(1) & -R_{ux}{}^{(j)}(0) & \cdots & -R_{xu}{}^{(j)}(n-2) \\ \vdots & \vdots & \cdots & \vdots \\ -R_{ux}{}^{(j)}(n-1) & -R_{ux}{}^{(j)}(n-2) & \cdots & -R_{ux}{}^{(j)}(0) \end{bmatrix} \tag{20}$$

are $n \times n$ submatrices

$$\Psi^T U^{(j)} = (-R_u{}^{(j)}(1) \dots - R_u{}^{(j)}(n) R_{xu}{}^{(j)}(1) \dots R_{xu}{}^{(j)}(n)) \tag{21}$$

is a $2n$ vector;

$$Q^{T(j)}Q^{(j)} = \begin{bmatrix} R_e^{(j)}(0) & R_e^{(j)}(1) & \cdots & R_e^{(j)}(n_s-1) \\ & R_e^{(j)}(0) & \cdots & R_e^{(j)}(n_s-2) \\ & & \ddots & \vdots \\ & & & R_e^{(j)}(0) \end{bmatrix} \tag{22}$$

is an $n_s \times n_s$ symmetric matrix;

$$Q^{T(j)}E^{(j)} = \left(R_e^{(j)}(1) R_e^{(j)}(2) \ldots R_e^{(j)}(n_s)\right)^T \tag{23}$$

is a n_s vector;

$$R_x^{(j)}(i) = \frac{1}{s-i}\sum_{k=1}^{s-i}\left(x_k^{*(j)} - \bar{x}^*\right)\left(x_{k-i}^{*(j)} - \bar{x}^*\right) \qquad i = \overline{0,m} \tag{24}$$

are values of the covariance function of filtered input x_k^*,

$$R_u^{(j)}(i) = \frac{1}{s-i}\sum_{k=1}^{s-i}\left(u_k^{*(j)} - \bar{u}^*\right)\left(u_{k-i}^{*(j)} - \bar{u}^*\right) \qquad i = \overline{0,m} \tag{25}$$

are values of the covariance function of filtered output u_k^*,

$$R_{ux}^{(j)}(i) = \frac{1}{s-i}\sum_{k=1}^{s-i}\left(u_k^{*(j)} - \bar{u}^*\right)\left(x_{k-i}^{*(j)} - \bar{x}^*\right),$$

$$R_{xu}^{(j)}(i) = \frac{1}{s-i}\sum_{k=1}^{s-i}\left(x_k^{*(j)} - \bar{x}^*\right)\left(u_{k-i}^{*(j)} - \bar{u}^*\right), \qquad i = \overline{0,m} \tag{26}$$

are values of cross-covariance functions which are calculated using the filtered sequences $x_k^{*(j)}$ and $u_k^{*(j)}$ of sample size s;

$$R_e^{(j)}(i) = \frac{1}{s-i}\sum_{k=1}^{s-i}\left(e_k^{(j)} - \bar{e}\right)\left(e_{k-i}^{(j)} - \bar{e}\right) \qquad i = \overline{0,m_s} \tag{27}$$

are values of covariance functions of the residuals e_k;

$$\bar{x}^* = s^{-1}\sum_{k=1}^{s} x_k^{*(j)}, \quad \bar{u}^* = s^{-1}\sum_{k=1}^{s} u_k^{*(j)}, \quad \bar{e} = s^{-1}\sum_{k=1}^{s} e_k^{(j)},$$

$$u_k^{*(j)} = \left[1 + \widehat{G}^{(j)}(z^{-1})\right] u_k^{(j)}, \quad x_k^{*(j)} = \left[1 + \widehat{G}^{(j)}(z^{-1})\right] x_k^{(j)}, \tag{28}$$

$$e_k^{(j)} = \left[1 + \widehat{A}^{(j)}(z^{-1})\right] u_k^{(j)} - \widehat{B}^{(j)}(z^{-1}) x_k^{(j)}, \tag{29}$$

$$1 + \widehat{G}^{(j)}(z^{-1}) = 1 + \widehat{g}_1^{(j)} z^{-1} + \ldots + \widehat{g}_{n_g}^{(j)} z^{-n_g},$$

$$1 + \widehat{A}^{(j)}(z^{-1}) = 1 + \widehat{a}_1^{(j)} z^{-1} + \ldots + \widehat{a}_n^{(j)} z^{-n},$$

$$\widehat{B}^{(j)}(z^{-1}) = \widehat{b}_1^{(j)} z^{-1} + \ldots + \widehat{b}_n^{(j)} z^{-n}, \qquad m = n - 1,\ m_g = n_g - 1.$$

The algorithm (13) - (29) of the GLSM is an iterative procedure used for off-line estimation of the above mentioned parameters. Each iteration of the GLSM algorithm consists of six steps. At the first step of the first iteration ($j = 1$) the sequences $x_k^{*(j)}$ and $u_k^{*(j)}$ $k = \overline{1,s}$ are obtained using equation (28), where initial values of the estimates chosen beforehand, are substituted. At the second step the values of the covariance $R_x{}^{(j)}(i)$, $R_u{}^{(j)}(i)$ and cross-covariance $R_{ux}{}^{(j)}(i)$, $R_{xu}{}^{(j)}(i)$ $i = \overline{0,m}$ functions are calculated using filtered sequences $x_k^{*(j)}$, $u_k^{*(j)}$ $k = \overline{1,s}$ and formulas (24) – (26). At the third step the vector of the estimates $\widehat{a}^{T(j)} = (\widehat{a}_1^{(j)}, \ldots, \widehat{a}_n^{(j)})$, $\widehat{b}^{T(j)} = (\widehat{b}_1^{(j)}, \ldots, \widehat{b}_n^{(j)})$ is obtained using the ordinary LS of the form (13). At the fourth step the sequence of the residual $e_k^{(j)}$ is generated using formula (29) and the estimates

$$\widehat{a}^{T(j)} = (\widehat{a}_1^{(j)}, \ldots, \widehat{a}_n^{(j)}), \quad \widehat{b}^{T(j)} = (\widehat{b}_1^{(j)}, \ldots, \widehat{b}_n^{(j)}),$$

which are substituted into the polynomials $\widehat{A}^{(j)}(z^{-1})$ and $\widehat{B}^{(j)}(z^{-1})$. At the fifth step the values of the covariance function $i = \overline{0,m}$ are calculated using an equation of the form (27) and the sequence of $e_k^{(j)}$, generated at the previous step. At the sixth step the vector $\widehat{g}^{T(j)} = (\widehat{g}^{(j)}, \ldots, \widehat{g}_{n_g}^{(j)})$ is obtained using the ordinary LS of the shape (14). Then the iterative stepwise procedure is repeated for

$j = 2, 3, \ldots$ until the respective stopping condition will be satisfied (Eykhoff, 1974).

3. Two approaches for parameter estimation in the presence of outliers in observations. In equation (6) it was assumed that $\varepsilon = 0$. Now let us consider the case when this assumption is invalid. It is known (Novovičova, 1987) that both in this case and for $W(z^{-1}; h)$ of the form (10) M-estimates of unknown parameters $c^T = (a^T, b^T)$ of the linear discrete-time dynamical system (1) – (10) can be calculated using three procedures: the S-algorithm, the H-algorithm, and the W-algorithm. On the other hand, the parameter estimation procedure for $W(z^{-1}; h)$ of the shape (11) and $\varepsilon \neq 0$ isn't worked out up till now. That's why in this section we try to solve this problem using two approaches. By the first approach the two classical LS of the form (13) and (14) are replaced either by two S-algorithms

$$\widehat{c}_M^{(j+1)} = \widehat{c}_M^{(j)} + \widehat{\sigma}\Big[\sum_{t=1}^{o} \psi'(\tilde{e}_t^{(j)}/\widehat{\sigma})\tilde{\varphi}_t^{(j)}\varphi_t^{T(j)}\Big]^{-1} \times \sum_{t=1}^{o} \psi(\tilde{e}_t^{(j)}/\widehat{\sigma})\tilde{\varphi}_t^{(j)}, \tag{30}$$

$$\widehat{g}_M^{(j+1)} = \widehat{g}_M^{(j)} + \widehat{\sigma}\Big[\sum_{t=1}^{o} \psi'(\tilde{e}_t^{(j)}/\widehat{\sigma})\omega_t^{(j)}\omega_t^{T(j)}\Big]^{-1} \times \sum_{t=1}^{o} \psi(\tilde{e}_t^{(j)}/\widehat{\sigma})\omega_t^{(j)}, \tag{31}$$

or by two H-algorithms

$$\widehat{c}_M^{(j+1)} = \widehat{c}_M^{(j)} + \widehat{\sigma}\Big[\sum_{t=1}^{o} \tilde{\varphi}_t^{(j)}\tilde{\varphi}_t^{T(j)}\Big]^{-1} \times \sum_{t=1}^{o} \psi(\tilde{e}_t^{(j)}/\widehat{\sigma})\tilde{\varphi}_t^{(j)}, \tag{32}$$

$$\hat{g}_M^{(j+1)} = \hat{g}_M^{(j)} + \hat{\sigma}\Big[\sum_{t=1}^{o} \omega_t^{(j)} \omega_t^{T(j)}\Big]^{-1} \times \sum_{t=1}^{o} \psi(\tilde{e}_t^{(j)}/\hat{\sigma})\omega_t^{(j)}, \tag{33}$$

or by two W -algorithms

$$\hat{c}_M^{(j+1)} = \hat{c}_M^{(j)} + \hat{\sigma}\Big[\sum_{t=1}^{o} \tilde{w}_t^{(j)} \tilde{\varphi}_t^{(j)} \tilde{\varphi}_t^{T(j)}\Big]^{-1} \times \sum_{t=1}^{o} \psi(\tilde{e}_t^{(j)}/\hat{\sigma})\tilde{\varphi}_t^{(j)}, \tag{34}$$

$$\hat{g}_M^{(j+1)} = \hat{g}_M^{(j)} + \hat{\sigma}\Big[\sum_{t=1}^{o} \tilde{w}_t^{(j)} \omega_t^{(j)} \omega_t^{T(j)}\Big]^{-1} \times \sum_{t=1}^{o} \psi(\tilde{e}_t^{(j)}/\hat{\sigma})\omega_t^{(j)}, \tag{35}$$

respectively.

Here

$$\hat{c}_M^{T(j)} = (\hat{a}_M^T, \hat{b}_M^T) = (\hat{a}_{1M}^{(j)}, \ldots, \hat{a}_{nM}^{(j)}, \hat{b}_{1M}^{(j)}, \ldots, \hat{b}_{nM}^{(j)})$$
$$\hat{g}_M^{T(j)} = (\hat{g}_{1M}^{(j)}, \ldots, \hat{g}_{n_g M}^{(j)})$$

are the estimates of parameter polynomials (8) and (12) which are calculated at the j-th iteration using the above mentioned algorithms; $\hat{\sigma}$ is a scale value of the robust estimate; $\psi(\tilde{e}_t^{(j)}/\hat{\sigma})$, $\psi(e_t^{(j)}/\hat{\sigma})$ are ψ-vectors which can be chosen according to Stockinger and Dutter (1987), Novovičova (1991); whereas $\psi(v)/v$ is non-increasing for $v > 0$ and

$$\lim_{v \to 0} \psi(v)/v = \rho_0'' < \infty;$$

$\psi'(\tilde{e}_t^{(j)}/\hat{\sigma})$, $\psi(e_t^{(j)}/\hat{\sigma})$ are the first order partial derivatives of the $\psi(\tilde{e}_t^{(j)}/\hat{\sigma})$ and $\psi(e_t^{(j)}/\hat{\sigma})$, respectively;

$$e_t^{(j)} = u_t - \varphi_t^{T(j)} \hat{c}^{(j)},$$
$$\tilde{e}_t^{(j)} = u_t^* - \tilde{\varphi}_t^{T(j)} \hat{c}^{(j)}$$

are the errors of a generalized equation and a filtered generalized equation at the jth iteration;

$$\varphi_t^{(j)} = \left(-u_{t-1}, \dots, -u_{t-n}\ x_{t-1}, \dots, x_{t-1} \right)^{T(j)},$$
$$\tilde{\varphi}_t^{(j)} = \left(-u_{t-1}^*, \dots, -u_{t-n}^*\ x_{t-1}^\circ, \dots, x_{t-1}^\circ \right)^{T(j)},$$
$$\omega_t^{(j)} = \left(e_{k-1}, \dots, e_{k-n_e} \right)^{T(j)}$$

are the vectors of observations of the output and input, respectively;

$$\tilde{w}_t^{(j)} = \begin{cases} \hat{\sigma}\psi(\tilde{e}_t^{(j)}/\hat{\sigma})/\tilde{e}_t^{(j)} & \text{for } \tilde{e}_t^{(j)} \neq 0 \\ \rho_o'' & \text{for } \tilde{e}_t^{(j)} = 0 \end{cases},$$

$$w_t^{(j)} = \begin{cases} \hat{\sigma}\psi(e_t^{(j)}/\hat{\sigma})/e_t^{(j)} & \text{for } e_t^{(j)} \neq 0 \\ \rho_o'' & \text{for } e_t^{(j)} = 0 \end{cases}.$$

The second approach is based on the robust covariance analysis and ordinary I/S algorithms for parameter estimation (Pupeikis, 1992). In order to increase their efficiency it is necessary to replace the respective averaging linear operators in matrices (18), (20), (22) and vectors (21), (23) by their nonlinear robust analogues according to Pupeikis (1990). For this purpose in each iteration $j = 1, 2, \dots$ of parameter estimation the values of the sample covariance and cross-covariance functions $R_u^{(j)}(0)$, $R_u^{(j)}(1)$, $\dots$, $R_u^{(j)}(n-1)$, $R_u^{(j)}(n)$, $R_{ux}^{(j)}(0)$, $R_{ux}^{(j)}(1), \dots, R_{ux}^{(j)}(n-1)$, $R_{ux}^{(j)}(n)$, are replaced in respective matrices by their robust analogues, i.e. $r^{(j)}(u_k^2)$, $r^{(j)}(u_k u_{k-1}), \dots, r^{(j)}(u_k u_{k-n+1})$, $r^{(j)}(u_k u_{k-n})$, $r^{(j)}(u_k x_k)$, $r^{(j)}$ $(u_k x_{k-1}), \dots, r^{(j)}(u_k x_{k-n+1})$, $r^{(j)}(u_k x_{k-n})$, $r^{(j)}(u_k^* x_k^*)$, $r^{(j)}(u_k^* x_{k-1}^*)$, $\dots, r^{(j)}(u_k^\circ x_{k-n+1}^\circ)$, $r^{(j)}(u_k^\circ x_k^\circ)$, $r^{(j)}(u_k^\circ x_{k-1}^\circ), \dots, r^{(j)}(u_k^\circ x_{k-n+1}^\circ)$, $r^{(j)}$ $(u_k^\circ x_{k-n}^\circ), r^{(j)}(e_k e_k)$, $r^{(j)}(e_k e_{k-1}), \dots, r^{(j)}(e_k e_{k-n_e+1})$.

Then, in equation (17)

$$\Psi_{11}^{(j)} = \begin{pmatrix} r^{(j)}(u_k^* u_k^*) & r^{(j)}(u_k^* u_{k-1}^*) & \dots & r^{(j)}(u_k^* u_{k-n+1}^*) \\ & r^{(j)}(u_k^* u_k^*) & \dots & r^{(j)}(u_k^* u_{k-n+2}^*) \\ & & \ddots & \vdots \\ & & & r^{(j)}(u_k^* u_k^*) \end{pmatrix},$$

$$\Psi_{12}^{(j)} = \Psi_{21}^{(j)} =$$

$$= \begin{pmatrix} -r^{(j)}(u_k^* x_k^*) & -r^{(j)}(x_k^* u_{k-1}^*) & \dots & -r^{(j)}(x_k^* u_{k-n+1}^*) \\ -r^{(j)}(u_k^* x_{k-1}^*) & -r^{(j)}(u_k^* x_k^*) & \dots & -r^{(j)}(x_k^* u_{k-n+2}^*) \\ \vdots & \vdots & & \vdots \\ -r^{(j)}(u_k^* x_{k-n+1}^*) & -r^{(j)}(u_k^* x_{k-n+2}^*) & \dots & -r^{(j)}(u_k^* x_k^*) \end{pmatrix}$$

and in equation (21)

$$\Psi^{T(j)} U^{(j)} = \begin{pmatrix} -r^{(j)}(u_k^* u_{k-1}^*) \\ \vdots \\ -r^{(j)}(u_k^* u_{k-n}^*) \\ r^{(j)}(x_k^* u_{k-1}^*) \\ \vdots \\ r^{(j)}(x_k^* u_{k-n}^*) \end{pmatrix}.$$

On the other hand, in equation (22) the matrix $Q^{T(j)} Q^{(j)}$ will be of the form

$$Q^{T(j)} Q^{(j)} = \begin{pmatrix} r^{(j)}(e_k e_k) & r^{(j)}(e_k e_{k-1}) & \dots & r^{(j)}(e_k e_{k-n_e+1}) \\ & r^{(j)}(e_k e_k) & \dots & r^{(j)}(e_k e_{k-n_e+2}) \\ & & \ddots & \vdots \\ & & & r^{(j)}(e_k e_k) \end{pmatrix}$$

and the vector $Q^{T(j)} E^{(j)}$ of the form

$$Q^{T(j)} E^{(j)} = \begin{pmatrix} r^{(j)}(e_k e_{k-1}) \\ r^{(j)}(e_k e_{k-2}) \\ \vdots \\ r^{(j)}(e_k e_{k-n_e}) \end{pmatrix}.$$

In this case various robust estimates of the corresponding covariance functions can be used (Gnanadesikan and Kettenring, 1972).

The estimates obtained by the first approach are solutions of the respective nonlinear equations requiring an inversion of the corresponding matrices at each iteration and some initial conditions. The problem of stopping the calculations of M-estimates will arise here too. As shown in (Pupeikis, 1992) the second approach is more helpful and simpler than the first one.

4. Simulation results. As an example we consider the discrete-time object of the form

$$u_k = \frac{z^{-1}}{1+0.7z^{-1}} x_k + \frac{\xi_k}{(1+0.7z^{-1})(1-1.5z^{-1}+0.7z^{-2})} \tag{36}$$

or

$$\begin{aligned} &u_k - 0.8u_{k-1} - 0.35u_{k-2} + 0.49u_{k-3} \\ &\quad = x_{k-1} - 0.8x_{k-2} - 0.35x_{k-3} + 0.49x_{k-4} + \xi_k \,, \end{aligned}$$

where $c^T = (0.7, 1)$ and $g^T = (-1.5, 0.7)$ are real parameters, whose estimates will be obtained using formulas (13), (14), whereas matrices (17), (22) can be rewritten in the forms

$$\Psi^{T(j)}\Psi^{(j)} = s\begin{pmatrix} R_u^{(j)}(0) & -R_{ux}^{(j)}(0) \\ -R_{ux}^{(j)}(0) & R_x^{(j)}(0) \end{pmatrix},$$
$$Q^{T(j)}Q^{(j)} = s\begin{pmatrix} R_e^{(j)}(0) & R_e^{(j)}(1) \\ R_e^{(j)}(1) & R_e^{(j)}(0) \end{pmatrix}$$

and vectors (21), (23) in the forms

$$\Psi^{T(j)}U^{(j)} = s\begin{pmatrix} -R_u^{(j)}(1) \\ R_{xu}^{(j)}(1) \end{pmatrix},$$
$$Q^{T(j)}E^{(j)} = s\begin{pmatrix} R_e^{(j)}(1) \\ R_e^{(j)}(2) \end{pmatrix},$$

respectively.

Then,

$$s\left(\Psi^{T(j)}\Psi^{(j)}\right)^{-1} = q_1^{-1}\begin{pmatrix} R_x^{(j)}(0) & R_{ux}^{(j)}(0) \\ R_{ux}^{(j)}(0) & R_u^{(j)}(0) \end{pmatrix},$$

$$s\left(Q^{T(j)}Q^{(j)}\right)^{-1} = q_2^{-1}\begin{pmatrix} R_e^{(j)}(0) & -R_e^{(j)}(1) \\ -R_e^{(j)}(1) & R_e^{(j)}(0) \end{pmatrix}$$

and

$$\begin{pmatrix} \widehat{a}_1 \\ \widehat{b}_1 \end{pmatrix} = q_1^{-1}\begin{pmatrix} -R_x^{(j)}(0)R_u^{(j)}(1) + R_{ux}^{(j)}(0)R_{xu}^{(j)}(1) \\ -R_{ux}^{(j)}(0)R_u^{(j)}(1) + R_u^{(j)}(0)R_{xu}^{(j)}(1) \end{pmatrix}, \tag{37}$$

$$\begin{pmatrix} \widehat{g}_1 \\ \widehat{g}_2 \end{pmatrix} = q_2^{-1}\begin{pmatrix} R_e^{(j)}(0)R_e^{(j)}(1) - R_e^{(j)}(1)R_e^{(j)}(2) \\ -R_e^{2(j)}(1) + R_e^{(j)}(0)R_e^{(j)}(2) \end{pmatrix}, \tag{38}$$

where

$$q_1 = R_x^{(j)}(0)R_u^{(j)}(0) - R_{ux}^{2(j)}(0),$$
$$q_2 = R_e^{2(j)}(0) - R_e^{2(j)}(1).$$

In order to calculate the robust estimates of a_1, b_1 and g_1, g_2 it is necessary to write the robust analogues in matrices in (37), (38) instead of the respective values of covariance and cross-covariance functions. Then we obtain

$$\begin{pmatrix} \widehat{a}_1 \\ \widehat{b}_1 \end{pmatrix} = q_{1r}^{-1}\begin{pmatrix} R_x^{(j)}(0)r^{(j)}(u_k u_{k-1}) + r^{(j)}(u_k x_k)r^{(j)}(x_k u_{k-1}) \\ -r^{(j)}(u_k x_k)r^{(j)}(u_k u_{k-1}) + r^{(j)}(u_k^2)r^{(j)}(x_k u_{k-1}) \end{pmatrix} \tag{39}$$

$$\begin{pmatrix} \widehat{g}_1 \\ \widehat{g}_2 \end{pmatrix} = q_{2r}^{-1}\begin{pmatrix} r^{(j)}(e_k^2)r^{(j)}(e_k e_{k-1}) - r^{(j)}(e_k e_{k-1})r^{(j)}(e_k e_{k-2}) \\ r^{2(j)}(e_k e_{k-1}) + r^{(j)}(e_k^2)r^{(j)}(e_k e_{k-2}) \end{pmatrix}, \tag{40}$$

where

$$q_{1r} = R_x^{(j)}(0)r^{(j)}(u_k^2) - r^{2(j)}(u_k x_k),$$
$$q_{2r} = r^{2(j)}(e_k^2) - r^{2(j)}(e_k e_{k-1}).$$

As robust analogues of the respective values of covariance and cross-covariance functions for each iteration $j = 1, 2, \ldots$ we choose here

$$r^{(j)}(u_k u_{k-1}) \equiv med^{(j)}(\tilde{u}_k \tilde{u}_{k-1}) = \begin{cases} (\tilde{u}_k \tilde{u}_{k-1})_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(\tilde{u}_k \tilde{u}_{k-1})_{\frac{s}{2}-1} & \text{for even } s \\ \qquad + (\tilde{u}_k \tilde{u}_{k-1})_{\frac{s}{2}+1}\Big] & \end{cases}$$

$$r^{(j)}(x_k u_{k-1}) \equiv med^{(j)}(\dot{x}_k \tilde{u}_{k-1}) = \begin{cases} (\tilde{u}_k \dot{x}_{k-1})_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(\dot{x}_k \tilde{u}_k)_{\frac{s}{2}-1} & \text{for even } s \\ \qquad + (\dot{x}_k \tilde{u}_k)_{\frac{s}{2}+1}\Big] & \end{cases}$$

$$r^{(j)}(u_k^2) \equiv med^{(j)}(\tilde{u}_k^2) = \begin{cases} (\tilde{u}_k^2)_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(\tilde{u}_k^2)_{\frac{s}{2}-1} & \text{for even } s \\ \qquad + (\tilde{u}_k^2)_{\frac{s}{2}+1}\Big] & \end{cases}$$

$$r^{(j)}(e_k^2) \equiv med^{(j)}(\tilde{e}_k^2) = \begin{cases} (\tilde{e}_k^2)_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(\tilde{e}_k^2)_{\frac{s}{2}-1} & \text{for even } s \\ \qquad + (\tilde{e}_k^2)_{\frac{s}{2}+1}\Big] & \end{cases}$$

$$r^{(j)}(u_k x_k) \equiv med^{(j)}(\tilde{u}_k \dot{x}_k) = \begin{cases} (\tilde{u}_k \dot{x}_k)_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(\tilde{u}_k \dot{x}_k)_{\frac{s}{2}-1} & \text{for even } s \\ \qquad + (\tilde{u}_k \dot{x}_k)_{\frac{s}{2}+1}\Big] & \end{cases}$$

$$r^{(j)}(e_k e_{k-1}) \equiv med^{(j)}(\tilde{e}_k \tilde{e}_{k-1}) = \begin{cases} (\tilde{e}_k \tilde{e}_{k-1})_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(\tilde{e}_k \tilde{e}_{k-1})_{\frac{s}{2}-1} & \text{for even } s \\ \qquad + (\tilde{e}_k \tilde{e}_{k-1})_{\frac{s}{2}+1}\Big] & \end{cases}$$

$$r^{(j)}(e_k e_{k-2}) \equiv med^{(j)}(\tilde{e}_k \tilde{e}_{k-2}) = \begin{cases} (\tilde{e}_k \tilde{e}_{k-2})_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(\tilde{e}_k \tilde{e}_{k-2})_{\frac{s}{2}-1} & \text{for even } s \\ \qquad + (\tilde{e}_k \tilde{e}_{k-2})_{\frac{s}{2}+1}\Big] & \end{cases}$$

where

$$\tilde{u}_k = u_k^* - med(u_k^*),$$
$$\dot{x}_k = x_k^* - x^*,$$
$$\tilde{e}_k = e_k - med(e_k),$$

$$med(u_k^*) = \begin{cases} (u_k^*)_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(u_k^*)_{\frac{s}{2}-1} + (u_k^*)_{\frac{s}{2}+1} & \text{for even } s \end{cases}$$

$$med(e_k) = \begin{cases} (e_k)_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}\Big[(e_k)_{\frac{s}{2}-1} + (e_k)_{\frac{s}{2}+1} & \text{for even } s\,. \end{cases}$$

Realizations of independent Gaussian variables ξ_k with zero mean and unitary dispersion and the sequence of the second order AR model of the form

$$x_k = x_{k-1} - 0.5x_{k-2} + \xi_k, \qquad k = \overline{1,100} \tag{41}$$

were used as the input sequence x_k. A realization of the discrete AR process was generated as the additive noise according to equation (11), where $A(z^{-1}) = 0.7z^{-1}$ and $G(z^{-1}) = -1.5z^{-1} + 0.7z^{-2}$. ξ_k is a sequence of independent identically distributed variables of shape (5) with the ε-contaminated distribution (6) and $\sigma_1^2 = 1$, $\sigma_2^2 = 100$. Ten experiments with different realizations of noise ξ_k^* were carried out at the noise level $\lambda = \sigma_{\xi^*}^2/\sigma_y^2 = 0.1$. In each ith experiment the estimates of parameters $a_1 = 0.7$, $b_1 = 1$ and $g_1 = -1.5$, $g_2 = 0.7$ of equation (36) were obtained using formulas (37) - (40) and $s = 100$. In addition, further we replaced the observation u_{50} in the following way

$$\tilde{u}_{50} = u_{50} + 100 \mid u_{50} \mid \tag{42}$$

and processed it together with other observations in formulas (37) - (40). Then we repeated the estimation of the above mentioned parameters once more.

Table 1 illustrates the values of $\bar{b}_1$, $\bar{a}_1$, $\bar{g}_1$ and $\bar{g}_2$ averaged by 10 experiments according to the formula

$$\bar{\kappa} = \frac{1}{10}\sum_{i=1}^{10}\widehat{\kappa}_{(i)} \qquad (43)$$

and their confidence intervals

$$\Delta_\kappa = \pm t_\alpha \frac{\widehat{\sigma}_\kappa}{\sqrt{L}} \ , \qquad (44)$$

where κ is an averaged respective value; $\widehat{\kappa}_{(i)}$ is the estimate of the respective parameter κ obtained after the ith experiment; $\widehat{\sigma}_\kappa$ is the estimate of the variance σ_κ; $\alpha = 0.05$ is the significance level; $t_\alpha = 2.26$ is the $100(1-\alpha)\%$ point of Student's distribution with $v = L - 1$ degree of freedom; $L = 10$ is the number of experiments.

In Table 1 the first and second lines correspond to the estimates, obtained by using formulas (37), (38) and the third and fourth ones – to the estimates, obtained by applying formulas (39), (40). Besides, the first and third lines correspond to the estimates obtained in the case of ε-contaminated distribution (6) of noise and the second and fourth ones – to the estimates, obtained by applying the sequence u_k with damaged $\tilde{u}_{50}$ according to formula (42).

Table 1. Averaged estimates $\bar{b}_1$, $\bar{a}_1$, $\bar{g}_1$, $\bar{g}_2$ and their confidence intervals

$\widehat{b}_1 \pm \Delta_b$	$\widehat{a}_1 \pm \Delta_{a_1}$	$\widehat{g}_1 \pm \Delta_{g_1}$	$\widehat{g}_2 \pm \Delta_{g_2}$
1.05 ± 0.01	0.41 ± 0.01	-2.10 ± 0.02	1.66 ± 0.02
-3.40 ± 0.31	-0.13 ± 0.01	0.12 ± 0.01	0.15 ± 0.01
0.43 ± 0.03	0.37 ± 0.06	-1.54 ± 0.17	1.13 ± 0.16
0.31 ± 0.02	0.49 ± 0.06	-1.08 ± 0.08	0.70 ± 0.08

It follows from the simulation results, presented in Table 1, that in the case of ε-contaminated distribution (6) the accuracy of the averaged estimates b_1, a_1 calculated by formula (37) is higher than that of the same estimates obtained by formula (39). The accuracy of the estimates for the parameters g_1, g_2 is higher in the opposite case. On the other hand, we have always prefered the approach, based on the robust parameter estimation using formulas (39), (40), when the noise, acting on the output of dynamical system (1) has a very large outlier ($\tilde{u}_{50}$ is generated by equation (42)).

5. Conclusions. The results of numerical simulation carried out by computer, prove the efficiency of the robust approach, based on a replacement of the corresponding values of sample covariance and cross-covariance functions by their robust analogues in respective matrices and on a further application of the two ordinary classical LS parameter estimation algorithms. The above mentioned approach can be used instead of the iterative M-procedures in a case of very large outliers in autoregressive noise.

REFERENCES

Åström, K.J., and P.Eykhoff (1971). System identification - a survey. *Automatica*, **7**(2), 123-162.

Clarke, D.W. (1967). Generalized least-squares estimation of the parameters of a dynamic model. *IFAC symp. on identification in automatic control systems*, Paper 3.17. Prague.

Eykhoff, P. (1974). *System identification*. Wiley, New York.

Gnanadesikan, R., and J.R.Kettenring (1972). Robust estimates, residuals and outlier detection with multiresponse data. *Biometrica*, **28**(1), 81–124.

Isermann, R. (1981). *Digital control systems*. Springer Verlag, Berlin.

Ljung, L. (1987). *System identification: Theory for user*. Prentice Hall, Inc.

Novovičova, J. (1987). Recursive computation of M-estimates for the parameters of the linear dynamical system. *Problems of Control and Information Theory*, **16**(1), 19–59.

Pupeikis, R. (1990). Model order robust determination. *Informatica*, **1**(2), 96–109.

Pupeikis, R. (1992). Off-line estimation of dynamic systems parameters in the presence of outliers in observations. *Informatica*, **3**(4), 567–581.

Stockinger, N., and R.Dutter (1987). Robust time-series analysis, an overview. *Kybernetika*, **23**(1–5), 90pp.

Young, P. (1984). *Recursive estimation and time - series analysis*. Springer - Verlag, Berlin.

Received February 1993

R. Pupeikis received the degree of Candidate of Technical Sciences from the Kaunas Polytechnic Institute, Kaunas, Lithuania, 1979. He is a senior researcher worker of the Department of Technological Processes Control at the Institute of Mathematics and Informatics. His research interests include the classical and robust approaches of dynamic system identification as well the technological process control.