INFORMATICA, 1994, Vol. 5, No. 1–2, 189–210

ON-LINE ESTIMATION OF DYNAMIC SYSTEMS PARAMETERS IN THE PRESENCE OF OUTLIERS IN OBSERVATIONS

Rimantas PUPEIKIS

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

Abstract. In the previous papers (Novovičova, 1987; Pupeikis 1991) the problem of recursive least square (RLS) estimation of dynamic systems parameters in the presence of outliers in observations has been considered, when the filter, generating an additive noise, has a transfer function of a particular form, see Fig. 1, 2. The aim of the given paper is the development of well-known classical techniques for robust on-line estimation of unknown parameters of linear dynamic systems in the case of additive noises with different transfer functions. In this connection various ordinary recursive procedures, see Fig. 2–6, are worked out when systems' output is corrupted by the correlated noise containing outliers. The results of numerical simulation by IBM PC/AT (Table 1) are given.

Key words: dynamic system, parameter estimation, recursive algorithm, outlier, robustness.

1. Statement of the problem. By parameter estimation of real objects it is often assumed that the additive noise acting on the output of a dynamic system has a Gaussian distribution. However, data sets, for which the Gaussian model is often assumed, sometimes contain a small fraction of outliers (Stockinger and Dutter, 1987). That is why the recursive classical algorithms applied in the on-line estimation of unknown parameters, appeared to be inefficient. In this case it is necessary to work out robust on-line algorithms or to improve the ordinary ones.

Consider a single input x_k and single output y_k linear discrete–time system, which is shown in Fig. 1, described by the difference equation

$$y_k = -a_1 y_{k-1} - \cdots - a_n y_{k-n} + b_0 x_{k-\tau} + \cdots + b_m x_{k-m-\tau}. \quad (1)$$

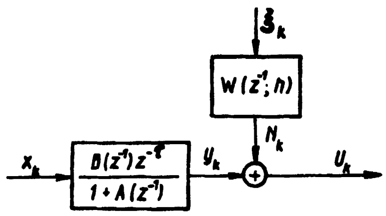

Fig. 1. Dynamic system with correlated noise.

Suppose that y_k is observed with an additive noise N_k, i.e.,

$$u_k = y_k + N_k. \quad (2)$$

Then

$$\begin{aligned} u_k = &- a_1 u_{k-1} - \cdots - a_n u_{k-n} + b_0 x_{x-\tau} + \ldots + b_m x_{k-\tau} \\ &+ N_k + a_1 N_{k-1} + \cdots + a_n N_{k-n}, \end{aligned} \quad (3)$$

or

$$u_k = \frac{B(z^{-1})z^{-\tau}}{1 + A(z^{-1})} x_k + W(z^{-1}; \mathbf{h})\xi_k, \quad (4)$$

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

$$\xi_k = (1 - \gamma_k) v_k + \gamma_k \eta_k \quad (5)$$

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 probabilities $p(\gamma_k = 0) = 1 - \varepsilon$, v_k, η_k are sequences of independent Gaussian variables with zero means and variances σ_1^2, σ_2^2 respectively;

$$\mathbf{c}^{\mathrm{T}} = (\mathbf{a}^{\mathrm{T}}, \mathbf{b}^{\mathrm{T}}), \quad \mathbf{a}^{\mathrm{T}} = (a_1, \ldots, a_n), \quad \mathbf{b}^{\mathrm{T}} = (b_0, \ldots, b_m), \tag{7}$$

$$B(z^{-1}) = \sum_{i=0}^{m} b_i z^{-i}, \quad A(z^{-1}) = \sum_{i=1}^{n} a_i z^{-i}, \tag{8}$$

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

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

is a noise filter transfer function; $\mathbf{h}$ is a parameter vector, τ is the time delay.

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 exitation of arbitrary order according to Åströ́m and Eykhoff (1971).

2. Recursive parameter estimation in the absence of outliers in observations. Suppose that $\tau = 0$ in equation (1) and $\varepsilon = 0$ in equation (6), therefore $p(\xi_k) = N(0, \sigma_1^2)$. In this case, as shown in Åströ́m and Eykhoff (1971), to estimate the vector of unknown parameters $\mathbf{c}^{\mathrm{T}} = (\mathbf{a}^{\mathrm{T}}, \mathbf{b}^{\mathrm{T}})$ multivariate on-line approaches and algorithms are worked out. On the other hand, it is known that in the case when

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

an ordinary classical recursive least square (RLS) parameter estimation algorithm of the shape

$$\widehat{\mathbf{c}}_{k+1} = \widehat{\mathbf{c}}_k + \mathbf{K}_{k+1} e_{k+1}, \tag{11}$$

$$\mathbf{K}_{k+1} = \frac{\boldsymbol{\Gamma}_k \nabla_{\mathbf{c}}\, e_{k+1}}{\lambda_{k+1} + \nabla_{\mathbf{c}}^{\mathrm{T}}\, e_{k+1} \boldsymbol{\Gamma}_k \nabla_{\mathbf{c}}\, e_{k+1}}, \tag{12}$$

$$\boldsymbol{\Gamma}_{k+1} = \left(\boldsymbol{\Gamma}_k - \frac{\boldsymbol{\Gamma}_k \nabla_{\mathbf{c}}\, e_{k+1} \nabla_{\mathbf{c}}^{\mathrm{T}}\, e_{k+1} \boldsymbol{\Gamma}_k}{\lambda_{k+1} + \nabla_{\mathbf{c}}^{\mathrm{T}}\, e_{k+1} \boldsymbol{\Gamma}_k \nabla_{\mathbf{c}}\, e_{k+1}}\right) \lambda_{k+1}^{-1}, \tag{13}$$

$$e_{k+1} = u_{k+1} - \nabla_{\mathbf{c}}^{\mathrm{T}}\, e_{k+1} \widehat{\mathbf{c}}_k, \tag{14}$$

$$\boldsymbol{\Gamma}_0 = \alpha I, \quad \alpha \gg 1$$

is used, see Fig. 2. Here

$$\widehat{\mathbf{c}}_{k+1} = \left(\widehat{\mathbf{a}}^{\mathrm{T}}, \widehat{\mathbf{b}}^{\mathrm{T}}\right)_{k+1} = \left(\widehat{a}_1, \ldots, \widehat{a}_n, \widehat{b}_0, \ldots, \widehat{b}_m\right)_{k+1} \tag{15}$$

is the vector of unknown parameter estimates after processing $k+1$ samples;

$$\nabla_{\mathbf{c}}\, e_{k+1} = \left(-u_k, \ldots, -u_{k+1-n}, x_{k+1}, \ldots, x_{k+1-m}\right)^{\mathrm{T}} \tag{16}$$

is the vector of n and $m+1$ most recent observations of input x_k and output u_k; $0.95 \leqslant \lambda_{k+1} \leqslant 1$ is a time–varying weighing factor.

It is known (Ljung, 1977) that under the above mentioned and some other conditions the RLS is going to have the maximal rate of convergence.

In practice, the assumption (10) is invalid as a rule, and the classical RLS is of little use. Therefore a multivariate set of recursive algorithms is worked out. In the case when

$$W\left(z^{-1};\ \mathbf{h}\right) \equiv 1 \tag{17}$$

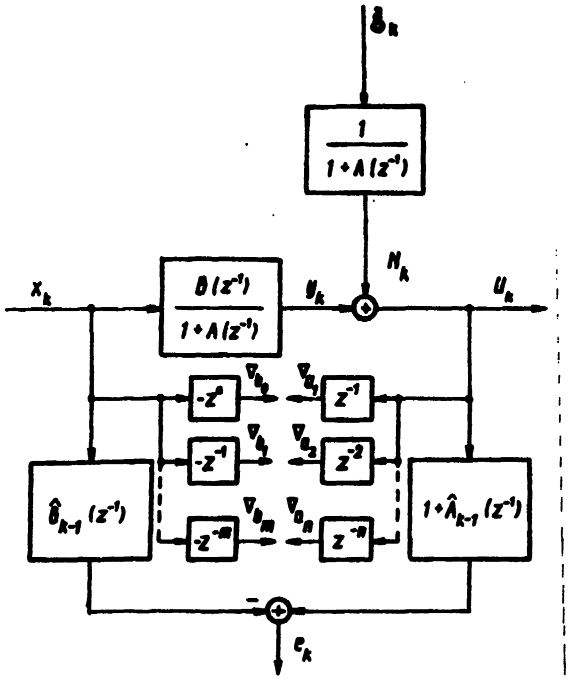

Fig. 2. Recursive parameter estimation by the RLS.

the recursive algorithm, based on the technique, which is developed in Steiglitz and McBride (1965), may be used, see Fig. 3. Then, in formulas (11)–(14) the vector $\nabla_{\mathrm{c}}\, e_{k+1}$ and equation error e_{k+1} must be replaced by

$$\nabla_{\mathrm{c}}\, e_{k+1} = \big(-u_k^*, \dots, -u_{k+1-n}^*, x_{k+1}^*, \dots, x_{k+1-m}^*\big)^{\mathrm{T}}, \quad (18)$$

and

$$e_{k+1}^* = u_{k+1}^* - \nabla_{\mathrm{c}}^{\mathrm{T}}\, e_{k+1}^* \widehat{\mathrm{c}}_k, \quad (19)$$

respectively. Here

$$u_k^* = \big[1 + \widehat{A}_{k-2}(z^{-1})\big]^{-1} u_k,$$
$$x_k^* = \big[1 + \widehat{A}_{k-2}(z^{-1})\big]^{-1} x_k, \quad (20)$$

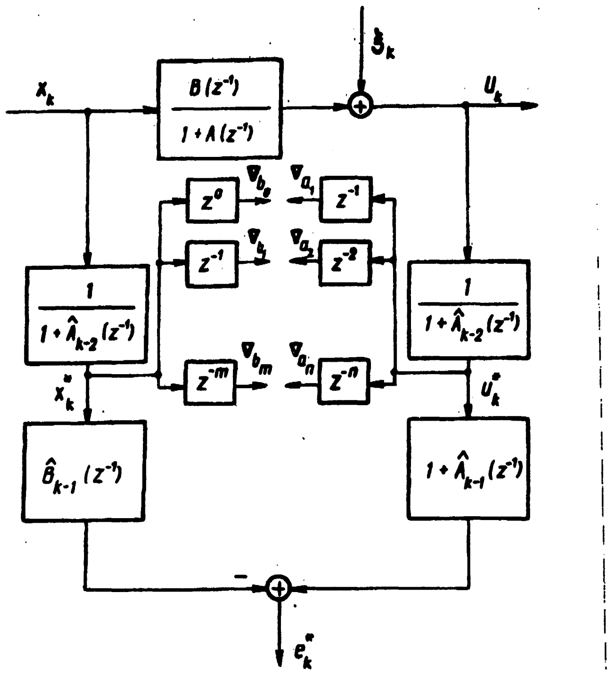

Fig. 3. Recursive parameter estimation by the Steiglitz and McBridge method.

are filtered observations after k samples;

$$\widehat{A}_{k-2}(z^{-1}) = \widehat{a}_{1_{k-2}} z^{-1} + \cdots + \widehat{a}_{n_{k-2}} z^{-n}. \tag{21}$$

The characteristics and convergence conditions of the above mentioned algorithm were investigated by Stoica and Sőderstrőm (1981).

Let us assume that

$$W(z^{-1};\ \mathbf{h}) = \left[1 + G(z^{-1})\right]^{-1}\left[1 + A(z^{-1})\right]^{-1}, \tag{22}$$

where

$$G(z^{-1}) = g_1 z^{-1} + \cdots + g_q z^{-q}. \tag{23}$$

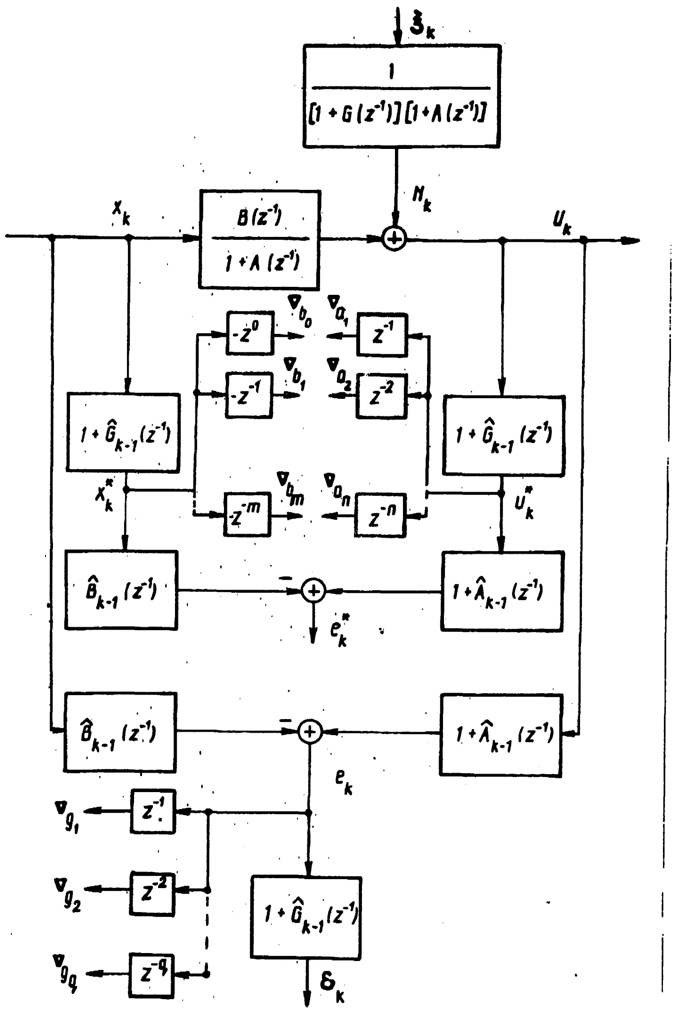

Fig. 4. Recursive parameter estimation by the GRLS.

In such a case the generalized RLS (GRLS) algorithm, see Fig. 4, consisting of two RLS algorithms, is used for estimating the vectors of unknown parameters and $\mathbf{g}^{\mathrm{T}} = (g_1, \ldots, g_q)$ according to (Hastings–James R., Sage M.W., 1969). The first RLS algorithm calculates the vector of the estimates $\widehat{\mathbf{c}}_{k+1}^{\mathrm{T}} =$

$(\widehat{\mathbf{a}}^{\mathrm{T}}, \widehat{\mathbf{b}}^{\mathrm{T}})_{k+1}$ according to formulas (11)–(13), where

$$e_{k+1} = u^*_{k+1} - \nabla^{\mathrm{T}}_{\mathbf{c}}\, e^*_{k+1}\widehat{c}_k, \tag{24}$$

$$\nabla_{\mathbf{c}}\, e^*_{k+1} = \big(- u^*_k, \ldots, -u^*_{k+1-n}, x^*_{k+1}, \ldots, x^*_{k+1-m}\big)^{\mathrm{T}}, \tag{25}$$

$$\begin{aligned} u^*_{k+1} &= \big[1 + \widehat{G}_k(z^{-1})\big] u_{k+1}, \\ x^*_{k+1} &= \big[1 + \widehat{G}_k(z^{-1})\big] x_{k+1}, \end{aligned} \tag{26}$$

$$\widehat{G}_k(z^{-1}) = \widehat{g}_{1_k} z^{-1} + \cdots + \widehat{g}_q z^{-q}. \tag{27}$$

The second algorithm calculates the vector of the estimates $\widehat{\mathbf{g}}^{\mathrm{T}}_{k+1} = (\widehat{g}_1, \ldots, \widehat{g}_q)_{k+1}$ using recursive equations of the form

$$\widehat{\mathbf{g}}_{k+1} = \widehat{\mathbf{g}}_k + \Lambda_{k+1} \nabla_{\mathbf{g}}\, \varepsilon_{k+1} \varepsilon_{k+1}, \tag{28}$$

$$\Lambda_{k+1} = \Big(\Lambda_k - \frac{\Lambda_k \nabla_{\mathbf{g}}\, \varepsilon_{k+1} \nabla^{\mathrm{T}}_{\mathbf{g}}\, \varepsilon_{k+1} \Lambda_k}{\lambda_{k+1} + \nabla^{\mathrm{T}}_{\mathbf{g}}\, \varepsilon_{k+1} \Lambda_k \nabla_{\mathbf{g}}\, \varepsilon_{k+1}}\Big) \lambda^{-1}_{k+1}, \tag{29}$$

$$\varepsilon_{k+1} = \big[1 + G_k(z^{-1})\big] e_{k+1}, \tag{30}$$

where

$$\nabla_{\mathbf{g}}\, \varepsilon_{k+1} = (e_{k+1}, \ldots, e_{k-q})^{\mathrm{T}},$$

and e_{k+1} is of the shape (14).

The initial conditions for RGLS can be chosen according to (Clarce, 1967).

Further, suppose that

$$W(z^{-1}; \mathbf{h}) = \big[1 + F(z^{-1})\big] \big[1 + A(z^{-1})\big]^{-1}, \tag{31}$$

where

$$F(z^{-1}) = f_1 z^{-1} + \cdots + f_\varkappa z^{-\varkappa}, \tag{32}$$

Then, the estimates of vectors of the parameters $\mathbf{c}$ and $\mathbf{f}^{\mathrm{T}} = (f_1, \dots, f_\varkappa)$ can be calculated using the maximum likelihood method (ML) or its on-line version consisting of two recursive algorithms, described in (Åström, Bohlin, 1965), see Fig. 5. The first recursive ML algorithm (RML) calculates the vector of the estimates $\widehat{\mathbf{c}}_k^{\mathrm{T}} = (\widehat{\mathbf{a}}^{\mathrm{T}}, \widehat{\mathbf{b}}^{\mathrm{T}})_k$ according to formulas (11)–(13), where e_{k+1} and $\nabla_{\mathbf{c}}\ e^*_{k+1}$ are of the shapes (24), (25) respectively, with the exeption of

$$u^*_{k+1} = \left[1 + \widehat{F}_k z(^{-1})\right]^{-1} u_{k+1},$$
$$x^*_{k+1} = \left[1 + \widehat{F}_k(z^{-1})\right]^{-1} x_{k+1}, \tag{33}$$

$$\widehat{F}_k(z^{-1}) = \widehat{f}_{1_k} z^{-1} + \dots + \widehat{f}_{\varkappa_k} z^{-\varkappa}. \tag{34}$$

The second algorithm calculates the vector of the estimates $\widehat{\mathbf{f}}_k^{\mathrm{T}} = (\widehat{f}_1, \dots, \widehat{f}_\varkappa)_k$ using recursive equations of the, form

$$\widehat{\mathbf{f}}_{k+1} = \widehat{\mathbf{f}}_k + \boldsymbol{\Pi}_{k+1} \nabla_{\mathbf{f}}\ \varepsilon_{k+1} \varepsilon_{k+1}, \tag{35}$$

$$\boldsymbol{\Pi}_{k+1} = \left(\boldsymbol{\Pi}_k - \frac{\boldsymbol{\Pi}_k \nabla_{\mathbf{f}}\ \varepsilon_{k+1} \nabla_{\mathbf{f}}^{\mathrm{T}} \varepsilon_{k+1} \boldsymbol{\Pi}_k}{\lambda_{k+1} + \nabla_{\mathbf{f}}^{\mathrm{T}} \varepsilon_{k+1} \boldsymbol{\Pi}_k \nabla_{\mathbf{f}}\ \varepsilon_{k+1}}\right) \lambda_{k+1}^{-1}, \tag{36}$$

$$\varepsilon_{k+1} = \left[1 + \widehat{F}_k(z^{-1})\right]^{-1} e_{k+1}, \tag{37}$$

where

$$\nabla_{\mathbf{f}}\ \varepsilon_{k+1} = (-\varepsilon^*_{k+1}, \dots, -\varepsilon^*_{k-\varkappa})^{\mathrm{T}}, \tag{38}$$

$$\varepsilon^*_{k+1} = \left[1 + \widehat{F}_k(z^{-1})\right]^{-1} e_{k+1}, \tag{39}$$

e_{k+1} is of the shape (14).

The initial conditions for RML can be chosen according to (Eykhoff, 1975; Isermann, 1974).

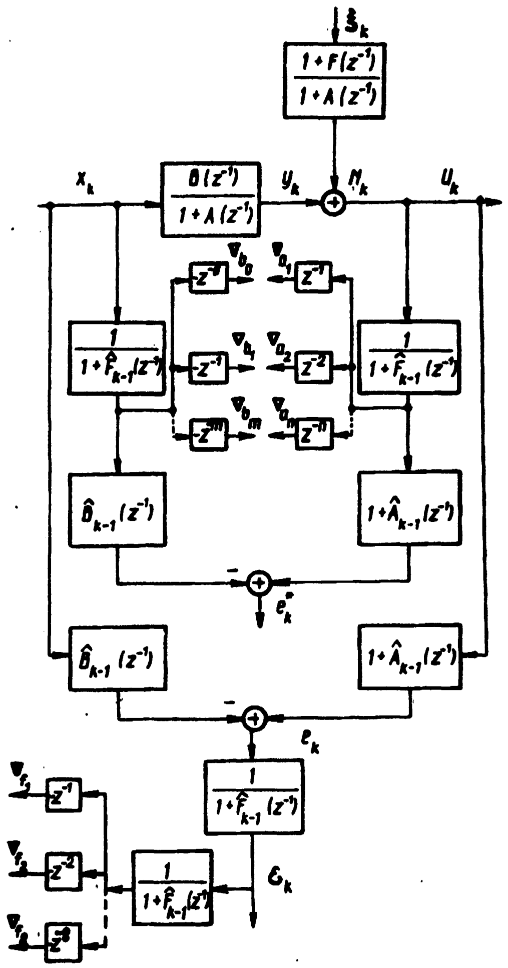

Fig. 5. Recursive parameter estimation by the ML.

It is known (Van den Boom; 1981), that in a more general case, i.e., when

$$W(z^{-1};\mathbf{h}) = \left[1+F(z^{-1})\right]\left[1+G(z^{-1})\right]^{-1}\left[1+A(z^{-1})\right]^{-1}, \quad (40)$$

the recursive algorithm of an extended least square method can be used, which for $F(z^{-1}) \equiv 0$ turned into GRLS and for

$G(z^{-1}) \equiv 0$ – into the RML. There also exist the algorithms of stochastic approximation which usually are recursive procedures with a scalar step (Saridis, 1974). The convergence of the recursive algorithms on the basis of a united approach is investigated in (Ljung, 1977). The simulation results of the above mentioned algorithms are given in (Isermann and coworkers, 1974).

3. Recursive parameter estimation in the presence of outliers in observations. It has been already earlier assumed that $\varepsilon = 0$. Now let us consider the situation when this assumption is not satisfied. It is known (Novovičova, 1987) that in both such cases, i.e., $\varepsilon = 0$ and $W(z^{-1};\ \mathbf{h})$ of the form (10) M-estimates of unknown parameters of linear dynamic systems (1)–(9) can be calculated using three recursive algorithms:

1) the S-algorithm

$$\widehat{\mathbf{c}}_{k+1} = \widehat{\mathbf{c}}_k + \frac{\mathbf{P}_k \varphi_{k+1} \widehat{\boldsymbol{\sigma}} \psi(r_{k+1}^{(k)}/\widehat{\boldsymbol{\sigma}})}{\left[\psi'(r_{k+1}^{(k)}/\widehat{\boldsymbol{\sigma}})\right]^{-1} + \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k \varphi_{k+1}}, \tag{41}$$

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \frac{\mathbf{P}_k \varphi_{k+1} \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k}{\left[\psi'(r_{k+1}^{(k)}/\widehat{\boldsymbol{\sigma}})\right]^{-1} + \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k \varphi_{k+1}}, \tag{42}$$

2) the H-algorithm

$$\widehat{\mathbf{c}}_{k+1} = \widehat{\mathbf{c}}_k + \frac{\mathbf{P}_k \varphi_{k+1} \widehat{\boldsymbol{\sigma}} \psi(r_{k+1}^{(k)}/\widehat{\boldsymbol{\sigma}})}{1 + \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k \varphi_{k+1}}, \tag{43}$$

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \frac{\mathbf{P}_k \varphi_{k+1} \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k}{1 + \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k \varphi_{k+1}}, \tag{44}$$

3) and the W-algorithm

$$\widehat{\mathbf{c}}_{k+1} = \widehat{\mathbf{c}}_k + \frac{\mathbf{P}_k \varphi_{k+1} \widehat{\boldsymbol{\sigma}} r_{k+1}^{(k)}}{\left[w_{k+1}^{(k)}\right]^{-1} + \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k \varphi_{k+1}}, \tag{45}$$

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \frac{\mathbf{P}_k \varphi_{k+1} \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k}{\left[w_{k+1}^{(k)}\right]^{-1} + \varphi_{k+1}^{\mathrm{T}} \mathbf{P}_k \varphi_{k+1}}, \tag{46}$$

$$w_{k+1} = \begin{cases} \widehat{\boldsymbol{\sigma}} \psi(r_{k+1}^{(k)}/\widehat{\boldsymbol{\sigma}}) r_{k+1}^{(k)} & \text{for } r_{k+1}^{(k)} \neq 0, \\ \rho_0'' & \text{for } r_{k+1}^{(k)} = 0, \end{cases} \tag{47}$$

$$r_{k+1}^{(k)} = u_{k+1} - \varphi_{k+1}^{\mathrm{T}} \widehat{\mathbf{c}}_k, \tag{48}$$

generating current M-estimates by means of minimizing sums

$$\sum_{i=1}^{t} \rho\left(\frac{e_i}{\boldsymbol{\sigma}}\right) = \min, \tag{49}$$

or by solving the system of nonlinear equations

$$\sum_{i=1}^{t} \psi\left(\frac{e_i}{\boldsymbol{\sigma}}\right) \varphi_i = 0, \tag{50}$$

if the derivatives are taken with respect to $\mathbf{c}$. Here $\rho(\cdot)$ is a symmetric robustifying loss function, $\psi = \rho'(\cdot)$ is a psi-function which can be chosen according to (Huber, 1984; Stockinger and Dutter, 1987; Pupeikis, 1991), $\widehat{\boldsymbol{\sigma}}$ denotes an estimate of the innovation scale and may be obtained simultaneously (Novovičova, 1987).

On the other hand, robust recursive methods used for parameter estimation of dynamic systems (1)–(9), when assumption (10) is invalid and other assumptions, i.e., (17) or (22) or (31) are satisfied, are not developed up till now. Therefore, in this paper we try to work out such recursive algorithms that will be efficient in the above mentioned case. We choose here the H-algorithm of the shape (43)–(44) as initial, because it is the simplest one of all recursive algorithms used for calculation of M-estimates (it requires only to insert the "winsorization" step into the RLS).

Thus, it is not difficult to show that if $W(z^{-1};\mathbf{h})$ is of the form (17), then in formulas (43)–(44) the vector φ_{k+1} must be replaced by (18) and $r_{k+1}^{(k)}$ by e_{k+1} of the shape (19), where u_k^* and x_k^* are observations filtered according to (20).

Suppose that $W(z^{-1};\ \mathbf{h})$ is of the form (22). Then the first RLS, calculating the vector of the estimates $\widehat{\mathbf{c}}_k$ according to formulas (11)–(13), can be replaced by the H-algorithm of the shape (43)–(44), where the vector φ_{k+1} must be replaced by (25) and $r_{k+1}^{(k)}$ by e_{k+1} of the shape (24). The observations (u_{k+1}^*, x_{k+1}^*) in e_{k+1} and φ_{k+1}^* are filtered according to (26). The same preparations must be done when $W(z^{-1};\ \mathbf{h})$ is of the form (31) and the RML is used. The difference is that the observations (u_{k+1}^*, x_{k+1}^*) in e_{k+1} and φ_{k+1}^* are filtered according to (33).

It should be noted in respect of the estimates $\widehat{\mathbf{g}}_k$ and $\widehat{\mathbf{f}}_k$ calculated by (28)–(30) or by (35)–(39) for GRLS and RML, respectively, that the recursive equations, mentioned above can be left unchanged when the detection and correction of outliers in observations are used according to Pupeikis (1991).

As initial values for the above algorithms the least square (LS) estimates, obtained for small data sets, can be used. The simulation results of the H-algorithm with Huber's ψ-function and an adaptive ψ-function are given in (Pupeikis, 1991).

It can be mentioned that the recursive algorithm with quasilinearization of nonlinear expressions for estimating the parameter $\mathbf{c}$ when $W(z^{-1};\ \mathbf{h}) \equiv 1$ or $W(z^{-1};\mathbf{h})$ is of the shape

$$W(z^{-1};\ \mathbf{h}) = \frac{1 + P(z^{-1})}{1 + R(z^{-1})} \tag{51}$$

is worked out in (Kaminskas, Pupeikis, 1974; Kaminskas, Pupeikis, 1975), see Fig. 6.

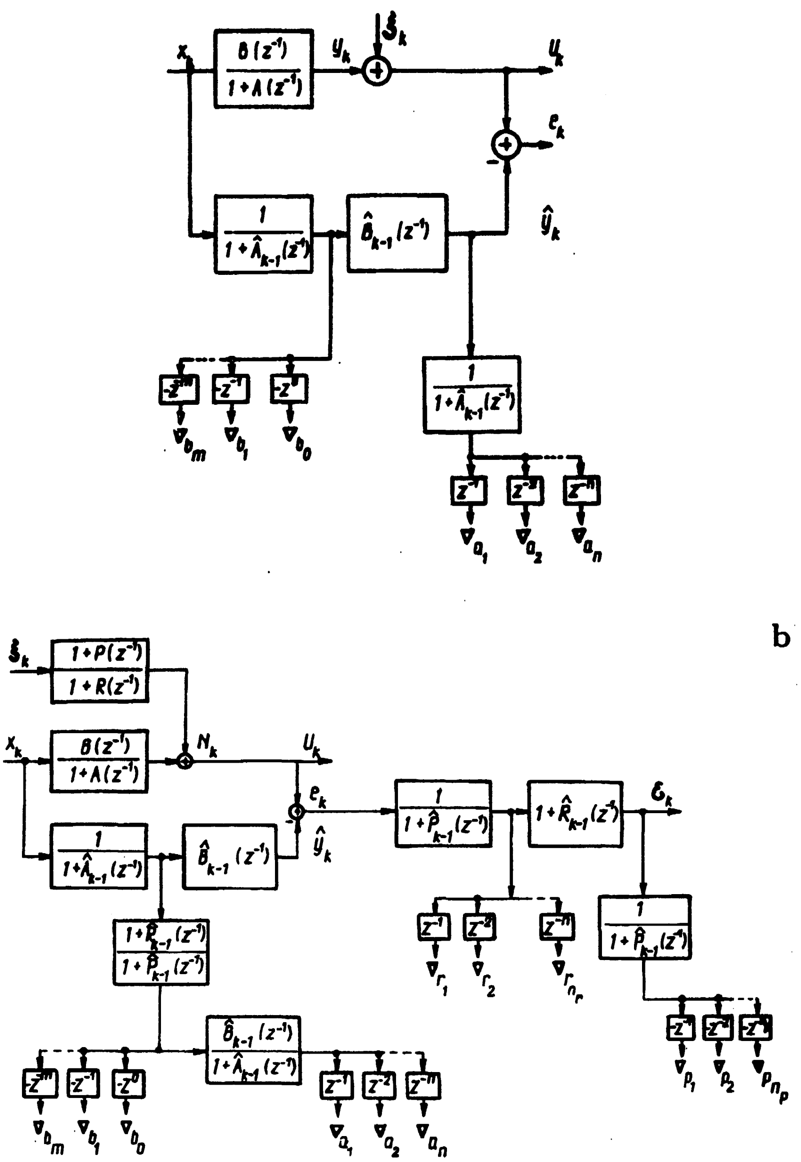

Fig. 6. Recursive parameter estimation by the quazilinearization algorithm. noise: a – independent, b – correlated.

Here

$$\left.\begin{aligned} R(z^{-1}) &= \sum_{i=1}^{n_r} r_i z^{-i} \qquad P(z^{-1}) = \sum_{i=1}^{n_p} p_i z^{-i}, \\ \mathbf{h}^{\mathrm{T}} &= (\mathbf{p}^{\mathrm{T}}, \mathbf{r}^{\mathrm{T}}) = (p_1, \ldots, p_{n_p}, r_1, \ldots, r_{n_r}). \end{aligned}\right\} \tag{52}$$

The current estimates $\widehat{\boldsymbol{\beta}}_k^{\mathrm{T}} = (\widehat{\mathbf{h}}_k^{\mathrm{T}}, \widehat{\mathbf{c}}_k^{\mathrm{T}})$ are calculated according to

$$\widehat{\boldsymbol{\beta}}_{k+1} = \widehat{\boldsymbol{\beta}}_k - \rho_k \boldsymbol{\Lambda}_k \nabla_{\beta} \varepsilon_{k+1} \varepsilon_{k+1}, \tag{53}$$

$$\boldsymbol{\Lambda}_k = \frac{\boldsymbol{\Gamma}_k}{\lambda_{k+1} + \nabla_{\boldsymbol{\beta}}^{\mathrm{T}} \varepsilon_{k+1} \boldsymbol{\Gamma}_k \nabla_{\beta} \varepsilon_{k+1}}, \tag{54}$$

$$\boldsymbol{\Gamma}_{k+1} = \left(\boldsymbol{\Gamma}_k - \frac{\boldsymbol{\Gamma}_k \nabla_{\beta} \varepsilon_{k+1} \nabla_{\beta}^{\mathrm{T}} \varepsilon_{k+1} \boldsymbol{\Gamma}_k}{\lambda_{k+1} + \nabla_{\beta}^{\mathrm{T}} \varepsilon_{k+1} \boldsymbol{\Gamma}_k \nabla_{\beta} \varepsilon_{k+1}}\right) \lambda_{k+1}^{-1}, \tag{55}$$

where

$$\left.\begin{aligned} \widehat{\beta}_k^{\mathrm{T}} &= (\widehat{\mathbf{h}}_k^{\mathrm{T}}, \widehat{\mathbf{c}}_k^{\mathrm{T}}), \\ \mathbf{h}_k^{\mathrm{T}} &= (\widehat{p}_1, \ldots, \widehat{p}_{n_p}, \widehat{r}_1, \ldots, \widehat{r}_{n_r})_k \end{aligned}\right\} \tag{56}$$

are the vectors of unknown parameter estimates after processing k samples;

$$\left.\begin{aligned} \nabla_{\beta}^{\mathrm{T}} e_{k+1} &= \left(\nabla_{\mathbf{p}}^{\mathrm{T}}, \nabla_{\mathbf{r}}^{\mathrm{T}}, \nabla_{\mathbf{a}}^{\mathrm{T}}, \nabla_{b}^{\mathrm{T}}\right), \\ \nabla_{\mathbf{p}}^{\mathrm{T}} &= -\frac{z^{-i}}{1 + \widehat{P}_k(z^{-1})} \varepsilon_{k+1}, \quad i = \overline{1, n_p}, \\ \nabla_{\mathbf{r}}^{\mathrm{T}} &= \frac{z^{-i}}{1 + \widehat{P}_k(z^{-1})} e_{k+1}, \quad i = \overline{1, n_r}, \\ \nabla_{\mathbf{a}}^{\mathrm{T}} &= \frac{1 + \widehat{R}_k(z^{-1})}{1 + \widehat{P}_k(z^{-1})} \frac{z^{-i}}{1 + \widehat{A}_k(z^{-1})} \widehat{y}_{k+1}, \quad i = \overline{1, n}, \\ \nabla_{\mathbf{b}}^{\mathrm{T}} &= -\frac{1 + \widehat{R}_k(z^{-1})}{1 + \widehat{P}_k(z^{-1})} \frac{z^{-i}}{1 + \widehat{A}_k(z^{-1})} x_{k+1}, \quad i = \overline{0, m}, \end{aligned}\right. \tag{57}$$

is the operator of the first partial derivatives, obtained by using the current residual of the shape

$$\varepsilon_{k+1} = \frac{1+\widehat{R}_k(z^{-1})}{1+\widehat{P}_k(z^{-1})}\, e_{k+1}, \tag{58}$$

where

$$e_{k+1} = u_{k+1} - \widehat{y}_{k+1} = u_{k+1} - \frac{\widehat{B}_k(z^{-1})}{1+\widehat{A}_k(z^{-1})}\, x_{k+1} \tag{59}$$

is a current residual nonlinear with respect to parameters; ρ_k is the scalar multiplier choosen according to

$$\rho_k = \begin{cases} 1, & \text{if } \rho^k_{\max} \geqslant 1, \\ \rho^k_{\max} - \delta, & \text{if } \rho^k_{\max} < 1, \end{cases} \tag{60}$$

δ is some positive quantity; $\rho^k_{\max}$ is the respective step in the direction of

$$\Delta\beta_{k+1} = \Lambda_{k+1}\nabla_\beta \varepsilon_{k+1}\varepsilon_{k+1}, \tag{61}$$

as far as the boundary of the area D in which the stability conditions for the parameters of transfer functions $W(z^{-1};\mathbf{c})$ and $W(z^{-1};\mathbf{h})$ are satisfied.

Thus, on each iteration the stability of the current system and noise filter models is guaranteed by ρ_k. It is necessary in the case of arbitrary initial conditions for the recursive algorithm (53)–(61) and in a case of outliers in observations to be processed.

It is shown (in Kaminskas, Pupeikis, 1974) that if $W(z^{-1};\mathbf{h}) \equiv 1$, then the recursive algorithm (53)–(55) appeares to be optimal in the sense of minimization of

$$M(\widehat{\mathbf{c}}_{k+1} - \mathbf{c})^{\mathrm{T}}(\widehat{\mathbf{c}}_{k+1} - \mathbf{c})/\sigma^2_\xi, \quad \mathbf{c}^{\mathrm{T}} = (\mathbf{a}^{\mathrm{T}}, \mathbf{b}^{\mathrm{T}})$$

on each current iteration, see Fig. 6a. In this case the operator of the first partial derivatives turns into

$$\left.\begin{aligned} \nabla_{\mathbf{c}}^{\mathrm{T}} e_{k+1} &= (\nabla_{\mathbf{a}}^{\mathrm{T}}, \nabla_{\mathbf{b}}^{\mathrm{T}}), \\ \nabla_{\mathbf{a}}^{\mathrm{T}} &= \frac{z^{-i}}{1+\widehat{A}_k(z^{-1})}\,\widehat{y}_{k+1}, \quad i=\overline{1,n}, \\ \nabla_{\mathbf{b}}^{\mathrm{T}} &= \frac{z^{-i}}{1+\widehat{A}_k(z^{-1})}\,x_{k+1}, \quad i=\overline{0,m}. \end{aligned}\right\} \tag{62}$$

It ought to be noted that the interesting estimation results, obtained by robust filtering algorithms in discrete–time dynamic systems with abrupt changes are presented in (Nemura, Kliokys, 1988).

4. Simulation results. The efficiency of above mentioned recursive algorithms was investigated by numerical simulation by means of IBM PC/AT. The noiseless sequence y_k was generated by the equation from the paper (Åström and Eykhoff, 1971)

$$y_k = \frac{z^{-1}+0.5z^{-2}}{1-1.5z^{-1}+0.7z^{-2}}\,x_k, \quad k=\overline{1,1000}. \tag{63}$$

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

$$x_k = x_{k-1} - 0.5x_{k-2} + \nu_k, \quad k=\overline{1,1000} \tag{64}$$

were used as the input sequence x_k. A realization of the discrete AR process was generated as the additive noise according to equation (9), where $W(z^{-1};\mathbf{h})$ is of the shape (22), $G(z^{-1}) = -z^{-1} + 0.4z^{-2}$ and $A(z^{-1}) = -1.5z^{-1} + 0.7z^{-2}$. Ten experiments with the different realizations of the noise N_k at the noise level $\sigma_N^2/\sigma_y^2 = 0.5$ were carried out. In each experiment we replaced only the observation u_{500} in the following way

$$u_{500} = 1000, \tag{65}$$

and processed it together with other observations by the GRLS. The initial conditions for the GRLS were obtained using off-line GLS and the sequences x_k, u_k of simple size $s = 200$. In each experiment the estimates of the parameters b_0, b_1, b_2, a_1, a_2, g_1, g_2 were obtained using ordinary GRLS and robust one based on the two H-algorithms with adaptive Huber's ψ-function (Pupeikis, 1991).

In Table 1 the estimates $\widehat{b}_0$, $\widehat{b}_1$, $\widehat{b}_2$, $\widehat{a}_1$, $\widehat{a}_2$, $\widehat{g}_1$, $\widehat{g}_2$ of the true parameters $b_0 = 0$, $b_1 = 1$, $b_2 = 0.5$, $a_1 = 1.5$, $a_2 = -0.7$, $g_1 = 1$, $g_2 = -0.4$ respectively and the variables

$$\delta_1^{(1)} = \frac{\sum_{j=1}^{5}(\theta_j - \widehat{\theta}_j)^2}{\sum_{j=1}^{5}\theta_j^2}\, 100\%, \tag{66}$$

$$\delta_2^{(1)} = \frac{\sum_{j=6}^{7}(\theta_j - \widehat{\theta}_j)^2}{\sum_{j=6}^{7}\theta_j^2}\, 100\%, \tag{67}$$

calculated after processing the 1000 values of the sequences (x_k, u_k) in the first experiment are given. Here $\theta_1 = b_0$, $\theta_2 = b_1$, $\theta_3 = b_2$, $\theta_4 = a_1$, $\theta_5 = a_2$, $\theta_6 = g_1$, $\theta_7 = g_2$, $\widehat{\theta}_1 = \widehat{b}_0$, $\widehat{\theta}_2 = \widehat{b}_1$, $\widehat{\theta}_3 = \widehat{b}_2$, $\widehat{\theta}_4 = \widehat{a}_1$, $\widehat{\theta}_5 = \widehat{a}_2$, $\widehat{\theta}_6 = \widehat{g}_1$, $\theta_7 = \widehat{g}_2$. It should be noted that in Table 1 the first line for different inputs corresponds to the estimates, obtained using ordinary GRLS and the second one – to the estimates, obtained by applying robust GRLS. Further we calculate the averaged by 10 experiments variables

$$\overline{\delta}_1 = \frac{1}{10}\sum_{i=1}^{10}\delta_1^{(i)}, \tag{68}$$

$$\bar{\delta}_2 = \frac{1}{10}\sum_{i=1}^{10}\delta_2^{(i)}, \tag{69}$$

with their confidence intervals Δ, using classical formulas (Bendat and Piersol, 1971). There are obtained such results:

$$\bar{\delta}_1 \pm \Delta = 167 \pm 8.75; \quad \bar{\delta}_2 \pm \Delta = 967 \pm 13.64;$$
$$\bar{\delta}_1 \pm \Delta = 1.56 \pm 0.28; \quad \bar{\delta}_2 \pm \Delta = 2.27 \pm 1.32;$$

for Gaussian input and

$$\bar{\delta}_1 \pm \Delta = 289 \pm 14.73; \quad \bar{\delta}_2 \pm \Delta = 146 \pm 11.23;$$
$$\bar{\delta}_1 \pm \Delta = 5.06 \pm 1.12 \quad \bar{\delta}_2 \pm \Delta = 4.26 \pm 1.78;$$

for AR process. The first line of each input corresponds to the estimates wich were calculated using ordinary GRLS and the second line – to the estimates wich were obtained by applying robust GRLS.

Table 1. Estimates of the parameters and variables (66), (67) in the presence of outlier in observations

b_0	b_1	b_2	a_1	a_2	g_1	g_2	$\delta_1\%$	$\delta_2\%$
0	1	0.5	1.5	–0.7	1	–0.4		
Input – Gaussian process								
1.206	2.855	–0.280	1.947	–1.385	1.444	2.839	155	921
0.028	1.030	0.423	1.602	–0.812	0.853	–0.375	0.8	1.9
Input – AR process								
1.139	2.082	–2.271	2.093	–1.331	1.439	0.782	273	137
0.040	1.105	0.143	1.635	–0.837	0.806	–0.365	4.4	3.3

It follows from the simulation and estimation results, presented here, that in the presence of outlier in observations the accuracy of the estimates $\widehat{b}_0$, $\widehat{b}_1$, $\widehat{b}_2$, $\widehat{a}_1$, $\widehat{a}_2$, $\widehat{g}_1$, $\widehat{g}_2$ obtained using robust GRLS is more higher than that of the same estimates calculated by ordinary GRLS.

5. Conclusions. In spite of a great variety recursive algorithms, worked out for the estimation of parameters of the dynamic systems, described by difference equation (4), it is possible to use only several of them, e.g. the algorithm with quasilinearization, in the presence of outliers in observations. In this case the ordinary recursive techniques are inefficient. However, there also exists an approach, based on the replacement of ordinary RLS in classical recursive schemes by the H-algorithm, and on a further substitution into its formulas of the corresponding vectors and residuals according to the estimation method to be used.

REFERENCES

Åström, K.J., and T. Bohlin. (1965). Numerical identification of linear dynamic systems from normal operating records. *IFAC Symp. on Theory of Selfadaptive Control Systems,* Teddington, N.Y.: Plenum Press, 96–106.

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

Bendat, J.S., and A.G. Piersol (1971). *Measurements and Analysis of Random Data.* Mir, Moscow. 408pp. (in Russian).

Clarce, 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. (1975). *System Identification. Parameter and State Estimation.* Mir, Moscow. 683pp. (in Russian).

Hastings–James R., and M.W. Sage (1969). Recursive generalized least–squares procedure for on-line identification of process parameters.

Proc. IEE, **116**, 2057–2062.

Huber, P. (1984). *Robust statistics.* Mir, Moscow. 303pp. (in Russian).

Isermann, R (1974). *Prozessidentifikation.* Springer Verlag, Berlin. 188pp.

Isermann, R., U. Baur, W. Bamberger, P. Kneppo, and Siebert (1974). Comparison of six on-line identification and parameter estimation methods. *Automatica,* **10**(1), 81–103.

Kaminskas, V., and R. Pupeikis (1974). On sequential parameter estimation of linear dynamic systems. *Works of the Academy of Sciences the Lithuanian SSR,* Ser.B, **6**(85), 123–129 (in Russian).

Kaminskas, V., and R. Pupeikis. (1975). On successive parameter estimation of linear dynamic systems with a noise of rational fraction spectral density observed in the output. *Works of the Academy of Sciences the Lithuanian SSR,* Ser.B, **3**(88), 163–168. (in Russian).

Ljung, L. (1977). Analysis of recursive stochastic algorithms. *IEE Trans. AC.,* **22**(4), 551–575.

Nemura, A., and E. Kliokys (1988). *Parameter and State Estimation in Systems. Systems with Abruptly Changing Properties.* Mokslas, Vilnius. 184pp. (in Russian).

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. (1991). Recursive robust estimation of dynamic systems parameters. *Informatica,* **2**(4), 579–592.

Saridis, G. (1974). Stochastic approximation methods for identification and control – a survey. *IEEE Trans. AC.,* **19**(6), 798–809.

Steiglitz, K., and L.E. McBride (1965). A technique for the identification of the linear systems. *IEEE Trans. AC.* , **10**(4), 461–464.

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

Stoica, P., and T. Sőderstrőm (1981). The Steiglitz–McBride identification algorithm revisited–convergence analysis and accuracy aspects. *IEEE Trans.AC.,* **26**(3), 712–717.

Van den Boom, A.J.W. (1981). *Doctoral Thesis.* Technishe Hogeshool Eindhofen. 238pp.

Received March 1994

R. Pupeikis received the Degree of Candidate Technical Science from the Kaunas Polytechnic Institute, Kaunas, Lithuania, 1979. He is a senior research worker of the Group of the Technological Process Control at the Institute of Mathematics and Informatics. His research interest include the classical and robust approaches of dynamic system identification as well technological process control.