INFORMATICA, 1995, Vol. 6, No. 3, 299–312

ON ADAPTIVE THRESHOLD INTERVALS FOR STOPPING RECURSIVE LEAST SQUARES IN THE SPACE OF PARAMETERS

Rimantas PUPEIKIS

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

Abstract. In the papers (Kaminskas, 1972; Kaminskas and Nemura, 1975; Yin, 1989) the stopping rules of recursive least squares (RLS) are worked out using the ellipsoidal confidence region for the respective parameter vector of a linear dynamic system. The aim of the given paper is the development of the technique for calculating threshold intervals of respective criterions, used in a stopping rule, which are presented in Kaminskas (1972). In this connection adaptive threshold intervals based on the Cramer–Rao lower bound are proposed here. The results of numerical simulation by IBM PC/AT are given.

Key words: dynamic system, recursive algorithm, threshold, least squares.

1. Stopping rule of RLS. Consider a single input $\mathbf{x}_k$ and output U_k linear discrete-time system, described by the difference equation

$$u_k = -a_1 u_{k-1} - \ldots - a_p u_{k-p} + b_0 x_k + \ldots + b_q x_{k-q} + N_k, \qquad (1)$$

where a_i, $i = \overline{1,p}$, and b_j, $j = \overline{0,q}$, are unknown parameters to be estimated by processing some input–output observations x_k and u_k; p, q are known positive integers; N_k is a sequence of independent Gaussian variables with zero mean and σ_N^2.

To calculate the estimate $\widehat{\mathbf{c}}_{s+1}$ of the parameter vector $\mathbf{c}^T = (a_1, \ldots, a_p, b_0, \ldots, b_q)$ we use ordinary RLS of the shape

$$\widehat{\mathbf{c}}_{s+1} = \widehat{\mathbf{c}}_s + \boldsymbol{\Gamma}_{s+1} \nabla_c e_{s+1} e_{s+1}, \qquad (2)$$

$$\boldsymbol{\Gamma}_{s+1} = \boldsymbol{\Gamma}_s - \frac{\boldsymbol{\Gamma}_s \nabla_{\mathbf{c}} e_{s+1} \nabla_{\mathbf{c}}^T e_{s+1} \boldsymbol{\Gamma}_s}{1 + \nabla_{\mathbf{c}}^T e_{s+1} \boldsymbol{\Gamma}_s \nabla_{\mathbf{c}} e_{s+1}}, \qquad (3)$$

$$e_{s+1} = u_{s+1} - \nabla_{\mathbf{c}}^T e_{s+1} \hat{\mathbf{c}}_s, \tag{4}$$

$$\boldsymbol{\Gamma}_o = \gamma \mathbf{I}, \quad \gamma \gg 1.$$

Here

$$\hat{\mathbf{c}}_{s+1} = (\hat{\mathbf{a}}^T, \hat{\mathbf{b}}^T)_{s+1} = (\hat{a}_1, \ldots, \hat{a}_p, \hat{b}_0, \ldots, \hat{b}_q)_{s+1}^T \tag{5}$$

is a vector of unknown parameter estimates obtained by recursive processing of $s+1$ samples of x_k and u_k, $k = 1, 2, \ldots, s, s+1$;

$$\nabla_{\mathbf{c}} e_{s+1} = (-u_s, \ldots, -u_{s+1-p}, x_{s+1}, \ldots, x_{s+1-q})^T \tag{6}$$

is a vector of p and $q+1$ most recent observations of input $\mathbf{x}_k$ and output U_k; $\boldsymbol{\Gamma}_s$ is an $m \times m$ positive definite matrix; $\mathbf{I}$ is the $m \times m$ unit matrix; $m = p + q + 1$.

According to Kaminskas (1972) the stopping rule for RLS is based on the appropriate ellipsoidal m-dimensional confidence region

$$p\{(\hat{\mathbf{c}}_s - \mathbf{c})^T \mathbf{K}_s (\hat{\mathbf{c}}_s - \mathbf{c}) \leqslant 1\} = 1 - \alpha, \tag{7}$$

$$\mathbf{K}_s = \frac{\boldsymbol{\Gamma}_s^{-1}}{m \hat{\sigma}_{N_s}^2 F_\alpha}, \tag{8}$$

with the centre at the point $\mathbf{c} = \hat{\mathbf{c}}_s$.

Here $\hat{\sigma}_{N_s}^2$ is the estimate of variance of σ_N^2, that can be calculated recursively; F_α is such that

$$p\{F \leqslant F_\alpha\} = 1 - \alpha, \quad F \sim F_{m,s-m}, \tag{9}$$

and it is tabulated; $p\{\cdot\}$ is a probability; α is a significance level; $F_{m,s-m}$ denotes Fisher's distribution with m and $s - m$ degrees of freedom.

Then the criterions

$$\mu_s = Tr\{\mathbf{K}_s^{-1}\}, \tag{10}$$

$$\mu_s = \max_i \lambda_i\{\mathbf{K}_s^{-1}\}, \quad i = \overline{1, m}, \tag{11}$$

$$\mu_s = \det\{\mathbf{K}_s^{-1}\} \tag{12}$$

are elaborated in the above mentioned paper.

In (10) – (12)

$$\mathbf{K}_s^{-1} = m\hat{\sigma}_{N_s}^2 F_\alpha \mathbf{\Gamma}_s. \tag{13}$$

$Tr\{\cdot\}, \max_i \lambda_i\{\cdot\}$ $i = \overline{1,m}$ and $\det\{\cdot\}$ denote the trace, the maximal eigenvalue and the determinant of matrix $\mathbf{K}_s^{-1}$, respectively.

Criterions (10), (11) define the sum of squares of the main semiaxles and the square of the maximal semiaxle of the confidence ellipsoid, respectively. The criterion (12) is proportional to the square of the mentioned ellipsoid volume.

Recursive calculations by RLS are stopped when

$$\mu_s \leqslant \mu_0, \tag{14}$$

where μ_0 is the threshold to be chosen beforehand.

2. Calculation of adaptive thresholds. There exist *three* main uncertainties while using such a stopping rule. *First*, for different measures (10) – (12) the same threshold μ_0 is chosen; *second*, there are no suggestions whatsoever as to the choise of μ_0; *third*, there is not clearly shown the efficiency of different μ_s. In Yin (1989) it is recommended to solve this problem using $\mu_0 = \mu_\alpha$ so that $p\{\chi_{p+q}^2 \geqslant \mu_\alpha\} = \alpha$, where χ_{p+q}^2 denotes the chi-square distribution with $p+q$ degrees of freedom. We try to obtain here the threshold values using the Cramer–Rao lower bound (Rao, 1968; Pupeikis, 1988).

It is known (Ljung, 1977; Cypkin, 1984) that under some conditions the RLS technique has the maximal rate of convergence. Then according to Cypkin (1984) for the asymptotically optimal algorithm (2) – (4) the asymptotic covariance matrix of errors (ACME) is

$$\mathbf{V} = \lim_{k\to\infty} k\mathbf{V}(\hat{\mathbf{c}}) = \mathbf{I}^{-1}(p_0)A^{-1}(\mathbf{c}, \sigma^2(p_0)), \tag{15}$$

and the equalities

$$\mathbf{V}_{jj} = \tilde{\mathbf{I}}_{jj}^{-1}(\mathbf{c}), \quad j = \overline{1,m}, \tag{16}$$

$$Tr\{\mathbf{V}\} = Tr\left\{\tilde{\mathbf{I}}^{-1}(\mathbf{c})\right\}, \tag{17}$$

$$\max_i \lambda_i\{\mathbf{V}\} = \max_i \lambda_i\left\{\tilde{\mathbf{I}}^{-1}(\mathbf{c})\right\}, \tag{18}$$

$$\det\{\mathbf{V}\} = \det\left\{\tilde{\mathbf{I}}^{-1}(\mathbf{c})\right\} \tag{19}$$

are satisfied, because in such a case the Cramer–Rao inequality

$$\mathbf{V} \geqslant \mathbf{I}^{-1}(p_0)A^{-1}(\mathbf{c}, \sigma^2(p_0)) \tag{20}$$

turns out into an equality.

Here

$$\mathbf{V}(\hat{\mathbf{c}}) = M\{(\hat{\mathbf{c}}_s - \mathbf{c})(\hat{\mathbf{c}}_s - \mathbf{c})^T\} \tag{21}$$

is the covariance matrix of errors (CME);

$$\mathbf{I}(p_0) = M\left\{\left(\frac{p_0'(N)}{p_0(N)}\right)\right\} \tag{22}$$

is Fisher's information; $p_0(N)$ and $p_0'(N)$ are the probability density function and its first derivative, respectively; $A(\mathbf{c}, \sigma^2(p_0))$ is the normed information matrix NIM); $\sigma^2(p_0)$ is the variance of noise N_k;

$$\tilde{\mathbf{I}}(c) = I(p_0)A(c, \sigma^2(p_0)); \tag{23}$$

$\mathbf{V}_{jj}$ and $\hat{\mathbf{I}}_{jj}(\mathbf{c})$, $j = \overline{1,m}$, are diagonal elements of matrices (15) and (23), respectively; $M\{\cdot\}$ is a mean value.

The estimate of ACME can be calculated by

$$\hat{\mathbf{V}}_s = s\hat{\sigma}_{e_s}^2(\boldsymbol{\Phi}_s^T\boldsymbol{\Phi}_s)^{-1} \tag{24}$$

in off-line operation and

$$\hat{\mathbf{V}}_k = k\sigma_{e_k}^2\boldsymbol{\Gamma}_k, \quad k = 1, 2, \ldots, s, s+1, \ldots \tag{25}$$

in on-line one.

Here s is the sample size; $\hat{\sigma}_{e_k}^2$ is the estimate of variance of σ_e^2 after processing s pairs of input–output observations; $\mathbf{e} = (e_1, \ldots, e_s)^T$ is the vector of residuals (4);

$$\boldsymbol{\Phi}_s = \begin{bmatrix} -u_i & \ldots & -u_{i-p+1} & x_{i+1} & \ldots & x_{i-q+1} \\ -u_{i+1} & \ldots & -u_{i-p+2} & x_{i+2} & \ldots & x_{i-q+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -u_{s-2} & \ldots & -u_{s-p-1} & x_{s-1} & \ldots & x_{s-m-1} \\ -u_{s-1} & \ldots & -u_{s-1} & x_s & \ldots & x_{s-m} \end{bmatrix} \tag{26}$$

is the matrix of input–output observations; $i = \max(p, q)$.

Taking into account (25) equalities (16) – (19) can be rewritten in such a way

$$k\hat{\sigma}^2_{e_k}(\Gamma_k)_{jj} \approx \widetilde{\mathbf{I}}^{-1}_{jj}(\hat{\mathbf{c}}_k), \quad j = \overline{1, m}, \tag{27}$$

$$k\hat{\sigma}^2_{e_k} Tr\{\Gamma_k\} \approx Tr\left\{\widetilde{\mathbf{I}}^{-1}(\hat{\mathbf{c}}_k)\right\}, \tag{28}$$

$$k\hat{\sigma}^2_{e_k} \max_i \lambda_i\{\Gamma_k\} \approx \max_i \lambda_i\left\{\widetilde{\mathbf{I}}^{-1}(\hat{\mathbf{c}}_k)\right\}, \quad i = \overline{1, m}, \tag{29}$$

$$k\hat{\sigma}^2_{e_k} \det\{\Gamma_k\} \approx \det\left\{\widetilde{\mathbf{I}}^{-1}(\hat{\mathbf{c}}_k)\right\}. \tag{30}$$

Hence, taking into account (13) it follows for minimal values of thresholds that

$$\mu^{(0)}_{k\,\min} = k^{-1} m F_\alpha \varkappa_k \widetilde{\mathbf{I}}^{-1}_{jj}(\hat{\mathbf{c}}_k), \quad j = \overline{1, m}, \tag{31}$$

$$\mu^{(1)}_{k\,\min} = k^{-1} m F_\alpha \varkappa_k Tr\left\{\widetilde{\mathbf{I}}^{-1}(\hat{\mathbf{c}}_k)\right\}, \tag{32}$$

$$\mu^{(2)}_{k\,\min} = k^{-1} m F_\alpha \varkappa_k \max_i \lambda_i\left\{\widetilde{\mathbf{I}}^{-1}(\hat{\mathbf{c}}_k)\right\}, \quad i = \overline{1, m}, \tag{33}$$

$$\mu^{(3)}_{k\,\min} = k^{-1} m F_\alpha \varkappa_k \det\left\{\widetilde{\mathbf{I}}^{-1}(\hat{\mathbf{c}}_k)\right\}, \tag{34}$$

where $\varkappa_k = \hat{\sigma}^2_{N_k}/\hat{\sigma}^2_{e_k}$, $k = 1, 2, \ldots, s, s+1 \ldots$. The maximal values of thresholds $\mu^{(0)}_{s\,\max}$, $\mu^{(1)}_{s\,\max}$, $\mu^{(2)}_{s\,\max}$, $\mu^{(3)}_{s\,\max}$ could be calculated by the same formulas supposing $k = s$ in (31) – (34).

It might be mentioned that the respective minimal and the maximal value of thresholds are time varying not only because of current k and $\varkappa_k$ but also because of the meanings of some parameters whose current estimates ought to be substituted into $\widetilde{\mathbf{I}}(\mathbf{c})$ of the shape (23).

Then, recursive calculations by RLS of the shape (2) – (4) are stopped if one of the conditions

$$\mu^{(0)}_{s\,\max} \geqslant \mu^{(0)}_k \geqslant \mu^{(0)}_{k\,\min}, \quad j = \overline{1, m}, \tag{35}$$

$$\mu^{(1)}_{s\,\max} \geqslant \mu^{(1)}_k \geqslant \mu^{(1)}_{k\,\min}, \tag{36}$$

$$\mu^{(2)}_{s\,\max} \geqslant \mu^{(2)}_k \geqslant \mu^{(2)}_{k\,\min}, \tag{37}$$

$$\mu^{(3)}_{s\,\max} \geqslant \mu^{(3)}_k \geqslant \mu^{(3)}_{k\,\min}, \quad k = s+1, s+2, \ldots \tag{38}$$

is satisfied or two conditions or three conditions or even all the four conditions are satisfied at the same time.

Here

$$\mu_{k,j}^{(0)} = (\mathbf{K}_s^{-1})_{jj}, \quad j = \overline{1,m}, \tag{39}$$

$$\mu_k^{(1)} = Tr\{\mathbf{K}_s^{-1}\}, \tag{40}$$

$$\mu_k^{(2)} = \max_i \lambda_i\{\mathbf{K}_s^{-1}\}, \quad i = \overline{1,m}, \tag{41}$$

$$\mu_k^{(3)} = \det\{\mathbf{K}_s^{-1}\}, \tag{42}$$

where $\mathbf{K}_s$ is of the shape (13).

3. Adaptive threshold for maximale length of the confidence interval. Values (10)–(12) are more general characteristics of the accuracy of estimates $\hat{\mathbf{c}}_s$. Sometimes it suffices (Kaminskas and Nemura, 1975) to calculate particular characteristics, e.g., the maximal length of the confidence interval for a separate coordinate c_i of the parameter vector $\mathbf{c}$. According to Rao (1968), Kaminskas and Nemura (1975) we use here the statistics

$$t = \frac{\hat{c}_{i,s} - c_i}{\sqrt{\hat{\sigma}_{N_s}^2 (\mathbf{\Gamma}_s)_{ii}}}, \quad i = \overline{1,\nu}, \tag{43}$$

where t denotes the t distribution with $s - \nu$ degrees of freedom.

The confidence interval for the i-th coordinate c_i is

$$\hat{c}_{i,s} - t_\alpha \hat{\sigma}_{N_s} \sqrt{(\mathbf{\Gamma}_s)_{ii}} < c_i < \hat{c}_{i,s} + t_\alpha \hat{\sigma}_{N_s} \sqrt{(\mathbf{\Gamma}_s)_{ii}}, \quad i = \overline{1,\nu}, \tag{44}$$

where t_α is such that

$$p\{t \leqslant t_\alpha\} = 1 - \alpha, \quad t \sim t_{s-m}, \tag{45}$$

and it is tabulated.

In Kaminskas and Nemura (1975) a criterion

$$\mu_s^{(4)} = \max_i \left\{ 2t_\alpha \hat{\sigma}_{N_s} \sqrt{(\mathbf{\Gamma}_s)_{ii}} \right\}, \quad i = \overline{1,\nu}, \tag{46}$$

is proposed which corresponds to the maximal length of the confidence interval (44). Therefore the recursive calculations by RLS could be stopped if the condition

$$\mu_{i_{\max}}^{(4)} \geqslant \mu_k^{(4)} \geqslant \mu_{k_{\min}}^{(4)}, \quad k = s+1, s+2, \ldots \tag{47}$$

with

$$\mu^{(4)}_{\substack{k,j\\ \min}} = k^{-1} \max_i \left\{ 2t_\alpha m F_\alpha \sqrt{\varkappa_k k \widetilde{\mathbf{I}}^{-1}_{ii}(\widehat{\mathbf{c}}_k)} \right\}, \tag{48}$$

$$\mu^{(4)}_{\substack{s,j\\ \max}} = s^{-1} \max_i \left\{ 2t_\alpha m F_\alpha \sqrt{\varkappa_s s \widetilde{\mathbf{I}}^{-1}_{ii}(\widehat{\mathbf{c}}_s)} \right\}, \quad i = \overline{1,\nu}, \tag{49}$$

is satisfied.

4. Time varying threshold intervals for the first order object. As an example we consider here a discrete-time object of the shape

$$u_k + a u_{k-1} = b_0 x_k + N_k, \tag{50}$$

where a and b_0 are the coefficients of difference equation (50).

In such a case ACME and NIM are

$$V = \sigma_e^2 \frac{\begin{bmatrix} K_x(0) & K_{ux}(1) \\ K_{ux}(1) & K_u(0) \end{bmatrix}}{K_x(0)\, K_u(0) - K_{ux}^2(1)}, \tag{51}$$

and

$$A^{-1}(\mathbf{c}, \sigma_N^2) = \begin{bmatrix} \dfrac{1-a^2}{\sigma_N^2 + b_0 \sigma_x^2} & 0 \\ 0 & 1/\sigma_x^2 \end{bmatrix}, \tag{52}$$

respectively, where

$$e_k = u_k - b_0 x_k + a u_{k-1} \tag{53}$$

is residual; $K_x(\cdot)$, $K_u(\cdot)$, $K_{ux}(\cdot)$ are the meanings of input–output autocovariance and crosscovariance function values, respectively; $\sigma^2_{e_k}$ and $\sigma^2_{x_k}$ are variances of residuals and input signal, respectively.

Then inequality (20) can be rewritten in such a way

$$\sigma_e^2 \frac{\begin{bmatrix} K_x(0) & K_{ux}(1) \\ K_{ux}(1) & K_u(0) \end{bmatrix}}{K_x(0) K_u(0) - K_{ux}^2(1)} \geqslant \sigma^2_{N_k} \begin{bmatrix} \dfrac{1-a^2}{\sigma_N^2 + b_0 \sigma^2_{x_k}} & 0 \\ 0 & \dfrac{1}{\sigma^2_{x_k}} \end{bmatrix} \tag{54}$$

since Fisher's information

$$\boldsymbol{I}(p_0) = 1/\sigma_N^2. \tag{55}$$

In view of the mentioned expressions time varying thresholds can be obtained for object (50) by the formulas:

$$\mu^{(0)}_{\substack{k,1\\ \min}} = k^{-1} m F_\alpha \varkappa_k \frac{1-\hat{a}_k^2}{1+\hat{b}_{0_k}\hat{\sigma}^2_{x_k}/\hat{\sigma}^2_{N_k}}, \tag{56}$$

$$\mu^{(0)}_{\substack{s,1\\ \max}} = s^{-1} m F_\alpha \varkappa_s \frac{1-\hat{a}_s^2}{1+\hat{b}_{0_s}\hat{\sigma}^2_{x_s}/\hat{\sigma}^2_{N_s}}, \tag{57}$$

$$\mu^{(0)}_{\substack{k,2\\ \min}} = k^{-1} m F_\alpha \varkappa_k \hat{\sigma}^2_{N_k}/\hat{\sigma}^2_{x_k}, \tag{58}$$

$$\mu^{(0)}_{\substack{s,2\\ \max}} = s^{-1} m F_\alpha \varkappa_s \hat{\sigma}^2_{N_s}/\hat{\sigma}^2_{x_s}, \tag{59}$$

$$\mu^{(1)}_{\substack{k\\ \min}} = k^{-1} \frac{1-\hat{a}_k^2}{1+\hat{b}_{0_k}\frac{\hat{\sigma}^2_{x_k}}{\hat{\sigma}^2_{N_s}}} + \frac{\hat{\sigma}^2_{N_k}}{\hat{\sigma}^2_{x_k}} m F_\alpha \varkappa_k, \tag{60}$$

$$\mu^{(1)}_{\substack{s\\ \max}} = s^{-1} \frac{1-\hat{a}_s^2}{1+\hat{b}_{0_s}\frac{\hat{\sigma}^2_{x_s}}{\hat{\sigma}^2_{N_s}}} + \frac{\hat{\sigma}^2_{N_s}}{\hat{\sigma}^2_{x_s}} m F_\alpha \varkappa_s, \tag{61}$$

$$\mu^{(2)}_{\substack{k\\ \min}} = k^{-1} \begin{cases} m F_\alpha \varkappa_k w_{1_k}, & \text{if } w_{1_k} > w_{2_k}, \\ m F_\alpha \varkappa_k w_{2_k}, & \text{if } w_{1_k} < w_{2_k}, \end{cases} \tag{62}$$

$$\mu^{(2)}_{\substack{s\\ \max}} = s^{-1} \begin{cases} m F_\alpha \varkappa_s w_{1_s}, & \text{if } w_{1_s} > w_{2_s}, \\ m F_\alpha \varkappa_s w_{2_s}, & \text{if } w_{1_s} < w_{2_s}, \end{cases} \tag{63}$$

$$\mu^{(3)}_{\substack{k\\ \min}} = k^{-1} \frac{\hat{\sigma}^2_{N_k}\left(1-\hat{a}_k^2\right)}{\hat{\sigma}^2_x\left(\hat{\sigma}^2_N + \hat{b}^2_{0_k}\hat{\sigma}^2_{x_k}\right)} m F_\alpha \varkappa_k, \tag{64}$$

$$\mu^{(3)}_{\substack{s\\ \max}} = s^{-1} \frac{\hat{\sigma}^2_{N_s}\left(1-\hat{a}_s^2\right)}{\hat{\sigma}^2_x\left(\hat{\sigma}^2_N + \hat{b}^2_{0_s}\hat{\sigma}^2_{x_s}\right)} m F_\alpha \varkappa_s, \tag{65}$$

$$\mu^{(4)}_{\substack{k,j\\ \min}} = k^{-1} \begin{cases} 2t_\alpha m F_\alpha \sqrt{\varkappa_k k w_{1_k}}, & \text{if } w_{1_k} > w_{2_k}, \\ 2t_\alpha m F_\alpha \sqrt{\varkappa_k k w_{2_k}}, & \text{if } w_{1_k} < w_{2_k}, \end{cases} \tag{66}$$

$$\mu^{(4)}_{s,j\ \max} = s^{-1}\begin{cases} 2t_\alpha m F_\alpha \sqrt{\varkappa_s s w_{1_s}}, & \text{if } w_{1_s} > w_{2_s}, \\ 2t_\alpha m F_\alpha \sqrt{\varkappa_s s w_{2_s}}, & \text{if } w_{1_s} < w_{2_s}, \end{cases} \tag{67}$$

$$j = \overline{1,2},$$

where

$$w_{1_k} = \frac{1-\hat{a}_k^2}{1+\hat{b}_{0_k}\dfrac{\hat{\sigma}^2_{x_k}}{\sigma_{N_k}}}, \quad w_{2_k} = \sigma_{N_k}/\sigma_{x_k},$$

$$w_{1_s} = \frac{1-\hat{a}_k^2}{1+\hat{b}_{0_s}\dfrac{\hat{\sigma}^2_{x_s}}{\sigma_{N_s}}}, \quad w_{2_s} = \sigma_{N_s}/\sigma_{x_s}.$$

Equations (56) – (67) could be realized in on-line operation if the estimates $\hat{a}_k$, $\hat{b}_{0_k}$, $\hat{K}_x(0,k)$, $\hat{K}_u(0,k)$, $\hat{K}_{ux}(\tau,k)$, $\tau = \overline{0,1}$, σ_{x_k}, σ_{e_k}, σ_{N_k} are substituted into the above mentioned expresions instead of their unknown values, respectively. They may be calculated in such a way:

$$\begin{bmatrix} \hat{a}_{1_k} \\ \hat{b}_{0_k} \end{bmatrix} = \frac{\begin{bmatrix} -\hat{K}_x(0,k) & \hat{K}_u(1,k)+\hat{K}_{ux}(1,k) & \hat{K}_{xu}(0,k) \\ -\hat{K}_{ux}(1,k) & \hat{K}_u(1,k)+\hat{K}_u(0,k) & \hat{K}_{ux}(0,k) \end{bmatrix}}{\hat{K}_x(0,k)\hat{K}_u(0,k) - \hat{K}^2_{ux}(1,k)}, \tag{68}$$

$$\hat{K}_x(0,k) = \hat{K}_x(0,k-1) + \frac{1}{1+k}\left[x_k x_k - \hat{K}_x(0,k-1)\right], \quad \tau = \overline{0,1},$$

$$\hat{K}_u(\tau,k) = \hat{K}_u(\tau,k-1) + \frac{1}{1+k}\left[u_{k-\tau}u_k - \hat{K}_u(\tau,k-1)\right],$$

$$\hat{K}_{ux}(\tau,k) = \hat{K}_{ux}(\tau,k-1) + \frac{1}{1+k}\left[u_{k-\tau}x_k - \hat{K}_{ux}(\tau,k-1)\right],$$

$$\sigma_{x_k} = \sigma_{x_{k-1}} + \frac{1}{k-1}\left[(x_k - \overline{x}_k)^2 - \sigma_{x_{k-1}}\right],$$

$$\overline{x}_k = x_{k-1} + \frac{1}{k}(x_k - \overline{x}_{k-1}),$$

$$\sigma_{e_k} = \sigma_{e_{k-1}} + \frac{1}{k-1}\frac{\left(u_k - \hat{b}_{0_k}x_k + \hat{a}_k u_{k-1}\right)^2}{1+\eta_k},$$

$$\eta_k = \frac{\hat{K}_x(0,k-1)u^2_{k-1} - \hat{K}_{ux}(1,k-1)x_k u_{k-1}}{(k-1)\left[\hat{K}_x(0,k)\hat{K}_u(0,k) - \hat{K}^2_{ux}(1,k)\right]}$$

$$+\frac{\widehat{K}_{ux}(1,k-1)u_{k-1}^2+\widehat{K}_u(0,k-1)x_k^2}{(k-1)\left[\widehat{K}_x(0,k)\widehat{K}_u(0,k)-\widehat{K}_{ux}^2(1,k)\right]},$$

$$\sigma_{N_k}=\sigma_{N_{k-1}}+\frac{1}{k-1}\left[\left(\widehat{N}_k-\overline{N}_k\right)^2-\sigma_{N_{k-1}}\right],$$

$$\overline{N}_k=\overline{N}_{k-1}+\frac{1}{k}\left(\widehat{N}_k-\overline{N}_{k-1}\right),$$

$$\widehat{N}_k=\widehat{v}_k+\widehat{a}_k\widehat{v}_{k-1},$$

$$\widehat{v}_k=u_k-\widehat{y}_k=u_k-\widehat{b}_{0_k}x_k+\widehat{a}_k\widehat{y}_{k-1}.$$

For higher order objects the stopping conditions are considerably more complicated, but not so much that their determination were imposible. Recommendations referring to the information matrix can be found in Cypkin (1984); Klein and Melard (1994).

5. Simulation results. The stopping rule in (10)–(12) with adaptive thresholds (56)–(67) for a discrete-time object with $a = 0.7$ and $b_0 = 1$ in (50) was investigated by numerical simulation by means of IBM PC/AT. Realizations of independent Gaussian variables ξ_k with zero mean and unitary variance and a sequence of the second order model of the form

$$x_k = x_{k-1} - 0.5x_{k-2} + \xi_k, \quad k = \overline{1,500}, \tag{69}$$

were used as an input sequence $\mathbf{x}_k$. Ten experiments with different realizations of noise N_k at the noise level $\sigma_N^2/\sigma_y^2 = 0.5$ were carried out. In each i-th experiment, first, the estimates of parameters $a = 0.7$ and $b_0 = 1$ of Eq. 50, the criterions $\mu_k^{(1)}$, $\mu_k^{(3)}$ of the shape (40), (42), respectively, and their minimal and maximal threshold values were obtained by formulas (60), (61), (64), (65). Afterwards, the same values were calculated recursively, using the above mentioned on-line procedure.

In Table 1 the estimates, averaged by 10 experiments,

$$\overline{\widehat{a}} = \frac{1}{10}\sum_{i=1}^{10}\widehat{a}^{(i)}, \tag{70}$$

$$\overline{\widehat{b}} = \frac{1}{10}\sum_{i=1}^{10}\widehat{b}^{(i)}, \tag{71}$$

the criterions

$$\overline{\mu}_k^{(1)} = \frac{1}{10}\sum_{i=1}^{10}\mu_{k,i}^{(1)}, \tag{72}$$

$$\overline{\mu}_k^{(3)} = \frac{1}{10}\sum_{i=1}^{10}\mu_{k,i}^{(3)}, \tag{73}$$

and their maximal threshold values

$$\overline{\mu}_{s\ \max}^{(1)} = \frac{1}{10}\sum_{i=1}^{10}\mu_{s,i\ \max}^{(1)}, \tag{74}$$

$$\overline{\mu}_{s\ \max}^{(3)} = \frac{1}{10}\sum_{i=1}^{10}\mu_{s,i\ \max}^{(3)}. \tag{75}$$

with their confidence intervals Δ, calculated in each i-th experiment after processing 15 values of observations (x_k, u_k) are given.

Table 1. Estimates (70), (71), criterions (72), (73) and their maximal threshold values (74), (75), averaged by 10 experiments, with confidence intervals after processing 15 values of observations

$\widehat{\overline{a}}$	$\widehat{\overline{b}}$	$\overline{\mu}_{k\ \max}^{(1)}$	$\overline{\mu}_k^{(1)}$	$\overline{\mu}_{k\ \max}^{(3)}$	$\overline{\mu}_k^{(3)}$
Input – Gaussian process					
0.39 ± 0.03	0.84 ± 0.07	2.26 ± 0.45	3.33 ± 0.17	0.89 ± 0.20	1.06 ± 0.01
Input – AR process					
0.24 ± 0.04	1.17 ± 0.12	0.66 ± 0.22	1.89 ± 0.19	0.15 ± 0.05	1.57 ± 0.16

Table 2 presents the same estimates, respective criterions and their minimal threshold values

$$\overline{\mu}_{k\ \min}^{(1)} = \frac{1}{10}\sum_{i=1}^{10}\mu_{k,i\ \min}^{(1)}, \tag{76}$$

$$\overline{\mu}_{k\ \min}^{(3)} = \frac{1}{10}\sum_{i=1}^{10}\mu_{k,i\ \min}^{(3)}, \tag{77}$$

Table 2. Estimates (70), (71), criterions (72), (73) and their minimal threshold values (76), (77), averaged by 10 experiments, with confidence intervals depending on k

$\widehat{\widehat{a}}$	$\widehat{\widehat{b}}$	$\overline{\mu}^{(1)}_{k\ \min}$	$\overline{\mu}^{(1)}_{k}$	$\overline{\mu}^{(3)}_{k\ \min}$	$\overline{\mu}^{(3)}_{k}$
		$k = 100$			
0.69 ± 0.01	0.99 ± 0.01	1.52 ± 0.08	1.99 ± 0.02	0.49 ± 0.03	1.01 ± 0.01
0.70 ± 0.01	1.00 ± 0.01	0.70 ± 0.04	0.91 ± 0.01	0.16 ± 0.01	1.08 ± 0.01
		$k = 200$			
0.70 ± 0.01	0.99 ± 0.01	1.52 ± 0.05	1.84 ± 0.01	0.49 ± 0.02	1.01 ± 0.01
0.69 ± 0.01	1.00 ± 0.01	0.63 ± 0.02	0.77 ± 0.00	0.14 ± 0.01	1.04 ± 0.01
		$k = 300$			
0.70 ± 0.01	0.99 ± 0.01	1.27 ± 0.07	1.66 ± 0.01	0.39 ± 0.03	1.01 ± 0.01
0.70 ± 0.01	1.01 ± 0.02	0.47 ± 0.03	0.67 ± 0.00	0.10 ± 0.01	1.05 ± 0.00
		$k = 400$			
0.70 ± 0.01	1.00 ± 0.01	1.16 ± 0.04	1.56 ± 0.01	0.35 ± 0.02	1.00 ± 0.00
0.70 ± 0.01	0.99 ± 0.01	0.38 ± 0.02	0.60 ± 0.00	0.07 ± 0.00	1.04 ± 0.00
		$k = 500$			
0.70 ± 0.01	0.99 ± 0.01	1.13 ± 0.04	1.48 ± 0.01	0.33 ± 0.01	1.00 ± 0.00
0.70 ± 0.01	0.98 ± 0.01	0.33 ± 0.01	0.55 ± 0.00	0.06 ± 0.00	1.04 ± 0.00

averaged by 10 experiments and calculated in each experiment after processing different number of observations (x_k, u_k). The first line of each k corresponds to the meanings which were calculated using a Gaussian process as input and the second line – to the meanings obtained by applying sequence of the shape (69) as input. It follows from the simulation and estimation results, presented here, that condition (36) for averaged measure (72) will be satisfied even for $k = 100$, if x_k is a Gaussian process and for $k = 400$, if x_k is AR sequence. On the other hand the condition (38) will be not satisfied for both inputs even at

$k = 500$. It can be mentioned that a decrease in $\overline{\mu}_k^{(3)}$ is negligible for increased number of k.

6. Conclusions. The results of numerical simulations carried out by computer prove the applicability of adaptive threshold intervals if the right criterion is chosen for recursive least squares stopping. Otherwise, the proposed here approach will be inefficient. That is why it is important recursively to calculate different criterions and their minimal and maximal time varying thresholds in order to choose the right ones and to use them for stopping LS while estimating unknown parameters.

REFERENCES

Cypkin, Ya.Z. (1984). *Principles of Information Theory of Identification.* Nauka, Moscow. 320 pp. (in Russian).

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

Kaminskas, V.A. (1972). On stopping rule of some successive identification algorithms. *Automatics and Computing Technique,* **4**, 39–42. (in Russian).

Kaminskas, V.A., and A. Nemura (1975). *Statistical Methods in Dynamic Systems Identification.* Mintis, Vilnius. 197 pp. (in Russian).

Klein, A., and G. Melard (1994). Computation of the Fisher information matrix for SISO models. *IEEE Trans. on SP.,* **42**(3), 684–688.

Pupeikis, R. (1988). Optimal recursive identification methods (2. Stopping of algorithms). *Transactions of the Academy of Sciences of the Lithuanian SSR. Ser. B,* **4**(167), 109–116. (in Russian).

Rao, C.R. (1968). *Linear Statistical Inference and its Applications.* Nauka, Moscow. 548 pp. (in Russian).

Yin, G. (1989). A stopping rule for least-squares identification. *IEEE Trans. AC.,* **34**(6), 659–662.

Received September 1995

R. Pupeikis received Ph.D. degree from Kaunas Polytechnic Institute, Kaunas, Lithuania, 1979. He is a senior researcher 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.

APIE ADAPTYVIUS SLENKSČIŲ INTERVALUS, STABDANT REKURENTINĮ MAŽIAUSIŲJŲ KVADRATŲ ALGORITMĄ PARAMETRŲ ERDVĖJE

Rimantas PUPEIKIS

Analitinio tyrimo būdu, taikant Kramerio–Rao nelygybę, sudaryti adaptyvūs slenksčių intervalai prof. V. Kaminsko ir prof. A. Nemuros kriterijams, jų pasiūlytiems parametrų įverčių skaičiavimams stabdyti, pasiekus pageidaujamą šių įverčių tikslumą. Darbe pateikti net ir riboto pobūdžio skaitinio modeliavimo rezultatai (Lentelės 1, 2) parodo, kad ne visi minėtų autorių kriterijai gali būti panaudoti rekurentinio mažiausiųjų kvadratų algoritmo stabdymui, taikant adaptyvius slenksčius.