

MODEL ORDER ROBUST DETERMINATION

Rimantas PUPEIKIS

Institute of Mathematics and Informatics,
Lithuanian Academy of Sciences,
232600 Vilnius, Akademijos St.4, Lithuania

Abstract. In the papers (Pupeikis, 1988a, b; 1989a, b, c) the problems of efficiency determination, stopping and increase of the effectiveness of asymptotically optimal recursive algorithms are considered respectively by means of estimating time delay in an object and also introducing their robust analogues, stable to outliers in observations. The aim of the given paper is the development of the robust method for a determination of the model order on the basis of determinant ratio. The three methods forming the initial moment matrices are considered. By the first method the elements of the matrix, being the corresponding values of the sample covariance and cross-covariance functions, are calculated by classical formulas. In the case of the second method the same elements are substituted by their robust analogues. The third method is based on an application of auxiliary variables. The results of numerical simulation on a computer (Table 1) indicate the advisability to apply the robust method for determining the model order in the presence of outliers.

Key words: model order, outliers, robustness.

Statement of the problem. Suppose that the recursive least squares algorithm used for the current estimation of the

unknown parameters of a mathematical model of the dynamic object, described by the linear difference equation

$$[1 + A(z^{-1})]u_k = B(z^{-1})x_k + \xi_k, \quad (1)$$

appeared to be non-effective. Let the reason for this be a disagreement between the model order and the object order.

In equation (1)

$$\begin{aligned} A(z^{-1}) &= a_1 z^{-1} + \dots + a_n z^{-n}, \\ B(z^{-1}) &= b_1 z^{-1} + \dots + b_n z^{-n}, \end{aligned} \quad (2)$$

$a^T = (a_1, \dots, a_n)$, $b^T = (b_1, \dots, b_n)$ are object parameters, subject to estimating; z^{-1} is a backward shift operator; x_k , $u_k = y_k + \xi_k^*$ are input and output sequences of the object; y_k is a noise free sequence of the object,

$$\xi_k = (1 - \gamma_k)v_k + \gamma_k\eta_k \quad (3)$$

is the sequence of independent identically distributed variables with ε -contaminated distribution of the shape

$$p(\xi_k) = (1 - \varepsilon)N(0, \sigma_1^2) + \varepsilon N(0, \sigma_2^2), \quad (4)$$

$p(\xi_k)$ is a probability density distribution of the sequence ξ_k ; γ_k is a random variable, taking the values of 0 and 1 with the probabilities $p(\gamma_k = 1) = \varepsilon$, $p(\gamma_k = 0) = 1 - \varepsilon$; v_k , η_k are the sequences of independent Gaussian variables with the zero means and σ_1^2 , σ_2^2 , respectively; n is the order of difference equation (1), further called the model order; $\xi_k^* = [1 + A(z^{-1})]^{-1}\xi_k$.

It is possible to divide the existing multivariate methods of determination of the model order into two groups depending on the determination of n , i.e. before or after the parametric identification of an object. To the first group the

methods presented in the papers (Söderström, 1977; Unbehauen and Göhring, 1974; Van den Boom and Van den Eindhoven, 1974) are referred, which are used after an estimation of the parameters a^T , b^T when n is arbitrarily defined. On the other hand the methods in the second group, presented in the papers (Lee, 1966; Setsuo Sagara, Hiromu Gotanda, Kiyoshi Wada, 1982; Unbehauen and Göhring, 1974; Van den Boom and Van den Eindhoven, 1974; Wellstead and Rojas, 1982) allow to determine the model order before the beginning of the parametric identification of an object what is their essential advantage over the methods of the first group. It should be noted that until now their efficiency has not been investigated in the presence of the outliers in observations, and the ways to increase their effectiveness have not been searched out yet. The present paper is devoted to the solution of these problems.

Model order determination in the absence of outliers in observations. Suppose that in equation (4) $\varepsilon = 0$, therefore $p(\xi_k) = N(0, \sigma_1^2)$. In this case to determine model order test statistics of the shape

$$\beta_{m+1} = \left| \frac{\det \Phi^T(m) \Phi(m)}{\det \Phi^T(m+1) \Phi(m+1)} \right| \quad m = 1, 2, \dots, \quad (5)$$

is applied, where

$$\Phi^T(m) \Phi(m) = s \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \quad (6)$$

$$\Phi^T(m+1) \Phi(m+1) = s \begin{pmatrix} \Phi_{11}^* & \Phi_{12}^* \\ \Phi_{21}^* & \Phi_{22}^* \end{pmatrix} \quad (7)$$

are $-2m \times 2m$ and $2(m+1) \times 2(m+1)$ symmetric covariance matrixes, respectively:

$$\Phi_{11} = \begin{pmatrix} R_u(0) & R_u(1) & \dots & R_u(m-1) \\ & R_u(0) & \dots & R_u(m-2) \\ & & \ddots & \vdots \\ & & & R_u(0) \end{pmatrix}$$

$$\Phi_{12} = \begin{pmatrix} -R_{ux}(0) & -R_{xu}(1) & \dots & -R_{xu}(m-1) \\ -R_{ux}(1) & -R_{ux}(0) & \dots & -R_{xu}(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ -R_{ux}(m-1) & -R_{ux}(m-2) & \dots & -R_{ux}(0) \end{pmatrix}$$

$$\Phi_{22} = \begin{pmatrix} R_x(0) & R_x(1) & \dots & R_x(m-1) \\ & R_x(0) & \dots & R_x(m-2) \\ & & \ddots & \vdots \\ & & & R_x(0) \end{pmatrix}$$

are the $m \times m$ symmetric and non-symmetric submatrices, respectively, where $\Phi_{21}^T = \Phi_{12}$ and

$$\Phi_{11}^* = \begin{pmatrix} R_u(0) & R_u(1) & \dots & R_u(m) \\ & R_u(0) & \dots & R_u(m-1) \\ & & \ddots & \vdots \\ & & & R_u(0) \end{pmatrix}$$

$$\Phi_{12}^* = \begin{pmatrix} -R_{ux}(0) & -R_{xu}(1) & \dots & -R_{xu}(m) \\ -R_{ux}(1) & -R_{ux}(0) & \dots & -R_{xu}(m-1) \\ \vdots & \vdots & \ddots & \vdots \\ -R_{ux}(m) & -R_{ux}(m-1) & \dots & -R_{ux}(0) \end{pmatrix}$$

$$\Phi_{22}^* = \begin{pmatrix} R_x(0) & R_x(1) & \dots & R_x(m) \\ & R_x(0) & \dots & R_x(m-1) \\ & & \ddots & \vdots \\ & & & R_x(0) \end{pmatrix}$$

are the $(m+1) \times (m+1)$ symmetric and non-symmetric submatrices, respectively, where $\Phi_{21}^{*T} = \Phi_{12}^*$.

$$\begin{aligned}
R_x(i) &= \frac{1}{s-i} \sum_{k=1}^{s-i} x_k x_{k+i}, \\
R_x(u) &= \frac{1}{s-i} \sum_{k=1}^{s-i} u_k u_{k+i}, \\
R_{ux}(i) &= \frac{1}{s-i} \sum_{k=1}^{s-i} u_k x_{k+i}, \\
R_{xu}(i) &= \frac{1}{s-i} \sum_{k=1}^{s-i} x_k u_{k+i} \quad (i = \overline{0, m}) \quad (8)
\end{aligned}$$

are the values of covariance and cross-covariance functions, which are calculated by the sequences x_k and u_k of a sample size s ; \det is a determinant symbol of the corresponding matrix; $|\cdot|$ is an absolute value of the corresponding variable.

In the capacity of the order n such a value $n = m$ is selected, under which β_{m+1} sharply grows, i.e. $\beta_{m+1} \gg \beta_m$. It occurs due to the fact that the matrix $\Phi^T(m+1)\Phi(m+1)$ is close to the degenerated matrix. In this connection $\det \Phi^T(i)\Phi(i) \forall i = m+1, m+2, \dots$ takes the values significantly not differing from the zero value.

In the case of independent additive noise according to the paper (Woodside, 1971) it is necessary to have a noise matrix Γ at the output of an object, moreover $M\{\Gamma^T\Gamma\} = \sigma^2 G$, where G is a given positively determined matrix, σ^2 is an unknown dispersion of noises.

Then in equation (5) instead of the matrices $\Phi(\cdot)$, the matrices $D(\cdot)$ are substituted, which are of the shape

$$D(\cdot) = \Phi(\cdot) - \sigma^2 G. \quad (9)$$

The model order determination method, based on statistics (5), loses its efficiency if additive noise is a sequence, correlated in time. For an increase of its effectiveness in the paper (Young, Jakeman and McMartrie, 1980) it is suggested to change the noise observations u_k in the matrices $\Phi(\cdot)$ by the corresponding values of the sequence of the auxiliary variable h_k . In this case the second, third and fourth equations in the systems of equations (8) will be written in the following way:

$$\begin{aligned} R_h(i) &= \frac{1}{s-i} \sum_{k=1}^{s-i} h_k h_{k+i}, \\ R_{hx}(i) &= \frac{1}{s-i} \sum_{k=1}^{s-i} h_k x_{k+i}, \\ R_{xh}(i) &= \frac{1}{s-i} \sum_{k=1}^{s-i} x_k h_{k+i} \quad (i = \overline{0, m}). \end{aligned} \quad (10)$$

The values of the covariance and cross-covariance functions, calculated according to (10), are substituted instead of the corresponding values $R_u(i)$, $R_{ux}(i)$ and $R_{xu}(i)$ in matrices (6), (7).

The efficiency of statistics (5) depends on the way of choice of h_k , therefore in the capacity of the values of an auxiliary variable it is proposed to use the corresponding values of an input sequence x_k introduced with some known lagging. There also exist some other variants of h_k selection (Söderström and Stoica, 1983).

Model order determination in the presence of outliers in observations. It was assumed earlier that in equation (4) $\varepsilon = 0$. Now let us consider such a case when this assumption is not valid. It is known (Gnanadesikan and Kettenring, 1972; Hampel et al., 1989; Huber, 1984) that then equations (8), (10) give strongly biased estimates of the sample

covariance functions and therefore test statistics (5) becomes of little use. In order to increase its efficiency it is necessary in equations (8), (10) to substitute the averaging linear operators by their non-linear robust analogues according to the formulas

$$\begin{aligned}
 R_x(i) &= \frac{1}{s-i} \widetilde{\sum}_\alpha \psi(x_k - \tilde{x}) \psi(x_{k+i} - \tilde{x}), \\
 R_u(i) &= \frac{1}{s-i} \widetilde{\sum}_\alpha \psi(u_k - \tilde{u}) \psi(u_{k+i} - \tilde{u}), \\
 R_{ux}(i) &= \frac{1}{s-i} \widetilde{\sum}_\alpha \psi(u_k - \tilde{u}) \psi(x_{k+i} - \tilde{x}), \\
 R_{xu}(i) &= \frac{1}{s-i} \widetilde{\sum}_\alpha \psi(x_k - \tilde{x}) \psi(u_{k+i} - \tilde{u}) \quad (11)
 \end{aligned}$$

for the additive independent noise or

$$\begin{aligned}
 R_x(i) &= \frac{1}{s-i} \widetilde{\sum}_\alpha \psi(x_k - \tilde{x}) \psi(x_{k+i} - \tilde{x}), \\
 R_h(i) &= \frac{1}{s-i} \widetilde{\sum}_\alpha \psi(h_k - \tilde{h}) \psi(h_{k+i} - \tilde{h}), \\
 R_{hx}(i) &= \frac{1}{s-i} \widetilde{\sum}_\alpha \psi(h_k - \tilde{h}) \psi(x_{k+i} - \tilde{x}), \\
 R_{xh}(i) &= \frac{1}{s-i} \widetilde{\sum}_\alpha \psi(x_k - \tilde{x}) \psi(h_{k+i} - \tilde{h}) \quad (12)
 \end{aligned}$$

for the additive correlated noise at the output or at the input of an object, where ($i = \overline{0, m-1}$).

Here $\psi = \psi(\cdot)$ is a monotone function, \tilde{x} , \tilde{u} , \tilde{h} are the robust analogues of the corresponding means, e.g. sample medians, calculated by the equations

$$\tilde{x} = m(x_i) = \begin{cases} x_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}(x_{\frac{s}{2}-1} + x_{\frac{s}{2}+1}) & \text{for even } s \end{cases} \quad (13)$$

$$\tilde{u} = m(u_i) = \begin{cases} u_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}(u_{\frac{s}{2}-1} + u_{\frac{s}{2}+1}) & \text{for even } s \end{cases} \quad (14)$$

$$\tilde{h} = m(h_i) = \begin{cases} h_{\frac{s+1}{2}} & \text{for odd } s \\ \frac{1}{2}(h_{\frac{s}{2}-1} + h_{\frac{s}{2}+1}) & \text{for even } s \end{cases} \quad (15)$$

respectively, where $x_1 \leq x_2 \leq \dots \leq x_s$, $u_1 \leq u_2 \leq \dots \leq u_s$, $h_1 \leq h_2 \leq \dots \leq h_s$; $\widetilde{\sum}_\alpha$ is a non-linear analogue of the sum, which is based on a rejection of lateral terms of the variational series and summing up of the remaining ones by the equation

$$\widetilde{\sum}_\alpha u_i = s(s - 2r)^{-1} \sum_{i=r+1}^{s-r} u_{(i)}, \quad r = [\alpha s]$$

$u_{(i)}$ is the i -th order statistics of the sample $\{u_i\}$ ($i = \overline{1, s}$); $m(\cdot)$ is a sample median; $[\alpha s]$ is an integer part of the variable αs ; $1 > \alpha > 0$.

Since for $\alpha = 0.5$ the operation of robust summing up converges to the use of the sample median, in matrices (6) and (7) the values of the sample covariance functions $R_u(0)$, $R_u(1)$, \dots , $R_u(m - 1)$, $R_u(m)$, $R_{ux}(0)$, $R_{ux}(1)$, \dots , $R_{ux}(m - 1)$, $R_{ux}(m)$, $R_x(0)$, $R_x(1)$, \dots , $R_x(m - 1)$, $R_x(m)$ are substituted correspondingly by $m(u_k^2)$, $m(u_k u_{k-1})$, \dots , $m(u_k u_{k-m+1})$, $m(u_k u_{k-m})$, $m(u_k x_k)$, $m(u_k x_{k+1})$, \dots , $m(u_k x_{k+m-1})$, $m(u_k x_{k+m})$, $m(x_k^2)$, $m(x_k x_{k-1})$, \dots , $m(x_k x_{k-m+1})$, $m(x_k x_{k-m})$. For $\alpha = 0$ the operations an ordinary and robust summing up coincide.

It is also possible to obtain the robust estimates of the covariance functions by the equations from the paper (Gnanadesikan and Kettenring, 1972) of the shape

$$R_x(i) = \rho(x_k, x_{k+i}) \widehat{\sigma}_x^2,$$

$$R_u(i) = \rho(u_k, u_{k+i}) \widehat{\sigma}_u^2,$$

$$\begin{aligned}
R_{ux}(i) &= \rho(u_k, x_{k-i}) \hat{\sigma}_x \hat{\sigma}_u, \\
R_{xu}(i) &= \rho(x_k, u_{k+i}) \hat{\sigma}_u \hat{\sigma}_x \quad (i = \overline{0, m-1}) \quad (16)
\end{aligned}$$

for the additive independent noise or

$$\begin{aligned}
R_x(i) &= \rho(x_k, x_{k+i}) \hat{\sigma}_x^2, \\
R_h(i) &= \rho(h_k, h_{k+i}) \hat{\sigma}_h^2, \\
R_{hx}(i) &= \rho(h_k, x_{k+i}) \hat{\sigma}_x \hat{\sigma}_h, \\
R_{xh}(i) &= \rho(x_k, h_{k+i}) \hat{\sigma}_h \hat{\sigma}_x \quad (i = \overline{0, m-1}) \quad (17)
\end{aligned}$$

for the additive correlated noise at the output of an object.

Here

$$\hat{\sigma}_q = m(|q_i - m(q_i)|)/0.6745 \quad (18)$$

is a scale value the robust estimate;

$$\rho(\omega_k, \theta_k) = \frac{\hat{\sigma}^2(\omega_k + \theta_{k+i}) - \hat{\sigma}^2(\omega_k - \theta_{k+i})}{\hat{\sigma}^2(\omega_k + \theta_{k+i}) + \hat{\sigma}^2(\omega_k - \theta_{k+i})} \quad (19)$$

is a robust analogue of the correlation coefficient between the sequences ω_k and θ_k .

The estimates of the sample covariance and cross-covariance functions may be calculated by equations (8), (10), if before this a rejection of the outliers was carried out or a robust smoothing of observations was realized.

Simulation results. The efficiency of test statistics (5) was investigated by numerical simulation by means of a computer. The noiseless sequence y_k was generated by the equation from the paper (Åström and Eykhoff, 1971)

$$y_k = \frac{z^{-1} + 0.5z^{-2}}{1 - 1.5z^{-1} + 0.7z^{-2}} \quad (k = \overline{1, 500}). \quad (20)$$

In the capacity of the input sequence x_k the realization of the sequence of independent Gaussian variables with the zero mean and unitary dispersion was used. In the capacity an additive noise ξ_k^* the realization of a discrete process of autoregression – moving average was generated by the equation from the paper (Talmon and Van den Boom, 1973)

$$\xi_k^* = \frac{1 + 0.3z^{-1}}{1 - 0.5z^{-1}} \xi_k, \quad (21)$$

where ξ_k is a sequence of independent identically distributed variables of shape (3) with ε – contaminated distribution of shape (4), where $\sigma_1^2 = 1$, $\sigma_2^2 = 100$, and $\sigma_{\xi^*}^2 / \sigma_y^2 = 0.5$.

In Table 1 the values of determinants of the matrices $\Phi^T(m) \Phi(m)$ are presented. They are calculated for various m by a pair (x_k, u_k) . In this connection the first line of each m corresponds to the values of determinants of these matrices, the elements of which were calculated by equations (8). The second line corresponds to the values of determinants the elements of which were substituted by sample medians according to the operation of robust summing up for $\alpha = 0.5$. The third line is obtained in the same way as the first one in the table, though with such a difference that in equation (8) auxiliary variables h_k were substituted instead of u_k . In the capacity of h_k the values x_k were used, introduced with a lagging $\tau = 1$. From the simulation results, presented in Table 1, it follows that for $m > 2$ only the values of determinants, given for each m in the second line of the table do not significantly differ from the zero values. Therefore further test statistics (5) was calculated on the basis of determinants, presented in this line. For $m = (\overline{1, 5})$ 10 experiments with different realizations of the additive noise ξ_k^* at the noise level $\sigma_{\xi^*}^2 / \sigma_y^2 = 0.5$ were carried out. In each i -th experiment the test statistics of shape (5) as well as the determinants of the matrices $\Phi^T(m) \Phi(m)$ were calculated. While simulating it was assumed that in expression (4) $\varepsilon = 0.25$.

Table 1. Determinant values of the matrices $\Phi^T(m)\Phi(m)$, test statistics β_m values and their confidence intervals depending on m

m	Determinant values	$\beta_m \pm \delta\beta_m$
1	25.687	42.469±79.926
	4.835	
	1.038	
2	20.162	16.750±17.144
	0.269	
	1.0576	
3	9.180	18.432±23.483
	$1.4 \cdot 10^{-2}$	
	1.798	
4	3.992	26.724±28.387
	$7.8 \cdot 10^{-4}$	
	2.156	
5	1.565	43.971±64.273
	$3.023 \cdot 10^{-5}$	
	2.632	
6	0.567	
	$1.07 \cdot 10^{-6}$	
	3.244	

In Table 1 also the averaged by 10 experiments variables

$$\bar{\beta}_m = \frac{1}{10} \sum_{i=1}^{10} \beta_m^i \quad (22)$$

and their confidence intervals for an object (20), (21) are presented. From the simulation results, presented in Table 1, it follows that for $m = 2$ the variable β_m takes a minimal value. For $m > 2$ the values of this variable increase. On the basis of the results of numerical experiments it is possible to make a conclusion that test statistics (5) may be applied to determine the model order if classical estimates of the sample covariance functions are substituted by their robust analogues. Comparatively large values of the confidence intervals are evidently explained by such a circumstance that classical formulas, developed in the absence of outliers in observations, have been applied for their calculation. It is advisable to continue the investigations in the determination of a threshold value for the decision rule of solution on the model order in the further works in this direction.

Conclusions. The results of numerical simulation, carried out by computer prove the efficiency of determinant ratio test (5) calculated on the basis of robust analogues of sample covariance and cross-covariance functions. It is possible to apply the developed method of robust determination of the model order as an alternative approach to the classical one in the case of outliers in observations.

REFERENCES

- Åström, K.J., and P.Eykhoff (1971). System identification a survey. *Automatica*, **7**(2), 123–162.
- Gnanadesikan, R., and J.R.Kettenring (1972). Robust estimates, residuals and outlier detection with multiresponse data. *Biometrika*, **28**(1), 81–124.
- Hampel, F.R., E.M.Ronchetti, P.J.Rousseeuw and W.A.Stahel (1989). Robust statistics. *The Approach Based on Influence Functions*. Mir, Moscow. 512pp. (in Russian).

- Huber, P. (1984). *Robust statistics*. Mir, Moscow. 303pp. (in Russian).
- Lee, R.C.K. (1966). *Optimal estimation, identification, and control*. Nauka, Moscow. 176pp. (in Russian).
- Pupeikis, R. (1988). Optimal recursive identification methods (1. Algorithm efficiency determination). *Works of the Academy of Sciences the Lithuanian SSR*, Ser. B, **3**(166), 132–140 (in Russian).
- Pupeikis, R. (1988). Optimal recursive identification methods (2. Stopping of algorithms). *Works of the Academy of Sciences the Lithuanian SSR*, Ser. B, **4**(167), 109–116 (in Russian).
- Pupeikis, R. (1989). Optimal recursive identification methods (3. Algorithms efficiency increasing by joint estimation of parameters and time delay in object). *Works of the Academy of Sciences the Lithuanian SSR*, Ser. B, **4**(173), 105–115 (in Russian).
- Pupeikis, R. (1989). Optimal recursive identification methods (4. Algorithms adaptive stabilization in case of outliers). *Works of the Academy of Sciences the Lithuanian SSR*, Ser. B, **5**(174), 134–142 (in Russian).
- Pupeikis, R. (1989). Optimal recursive identification methods (5. Recursive M-estimation of parameter and time delay of linear discrete time systems). In C.Paulauskas (Ed.), *Statisticheskie Problemy Upravleniya*. Vol.84. Inst. Math. Cybern. Lith. Acad. Sci., Vilnius (in Russian).
- Setsuo Sagara, Hiromu Gotanda and Kiyoshi Wada (1982). Dimensionally recursive order determination of linear discrete system. *Int.J.Control*, **35**(4), 637–651.
- Söderström, T. (1977). On model structure testing in system identification. *Int.J.Control*, **26**(1), 1–18.
- Söderström, T. and P.G.Stoica (1983). *Instrumental Variable Methods for System Identification*. Springer, Berlin. 243pp.
- Talmon, J.L., and A.J.W.Van den Boom (1973). On the estimation of the transfer function parameters of process and noise dynam-

- ics using a single – stage estimator. In *Proc. of the 3rd IFAC Conf. on Identif. Syst. Param. Estim.*, Hague. pp. 711–720.
- Unbehauen, H., and B.Göhring (1974). Tests for determining model order in parameter estimation. *Automatica*, **10**, 233–244.
- Van den Boom, A.J.W and A.W.M.Van den Enden (1974). The determination of the orders of process–and noise dynamics. *Automatica*, **10**, 245–256.
- Wellstead, P.E., and R.A.Rojas (1982). Instrumental product moment model-order testing: extensions and application. *Int. J. Control*, **35**(6), 1013–1027.
- Woodside, C.H. (1971). Estimation of the order of linear systems. *Automatica*, **7**, 727–733.
- Young, P., A.Jakeman and R.Mc Martrie (1980). An instrumental variable method for model order identification. *Automatica*, **16**, 281–294.

Received June 1990

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