The optimal estimation of parameters of models of controlled stochastic systems based on the experiment design

V I Denisov¹, V M Chubich¹, E V Filippova¹

¹ Novosibirsk State Technical University, 20, Karla Markska ave., Novosibirsk, 630073, Russia
E-mail: videnis@nstu.ru, chubich@ami.nstu.ru, e.filippova@corp.nstu.ru

Abstract. The procedure of active identification, which is resistant to the appearance of anomalous observations, includes robust estimation of unknown parameters and optimal design of input signals for models of non-stationary linear discrete systems is proposed. The general case of the entry of unknown parameters into the equations of state and observation, the initial condition and the covariance noise matrix of the system is considered. Using the developed software, the efficiency of this procedure is demonstrated by the example of a direct current motor control system.

1. Introduction

Active identification of dynamic systems [1-6], consisting in a combination of methods of parametric evaluation and experiment design, is a powerful and effective means of building high-quality mathematical models used in many applied fields of science and technology. Traditionally, the estimation of unknown parameters is carried out based on the classical Kalman filter, which makes it possible to find estimates of the state vector and corresponding covariance matrices under the assumption of the normality of the noise distribution of the system and measurements. When this assumption is violated, the specified filter can lead to biased estimates or even diverge. In this regard, it seems expedient and promising to develop methods of active identification based on robust estimation of unknown parameters resistant to the appearance of anomalous emissions.

Recently, there have been many works devoted to the development of robust filters, including the so-called correntropy, allowing solving the problem of parametric identification even in the presence of noise other than Gaussian. In [7] a comparative analysis of the most promising in the opinion of the authors of filters for models of non-stationary linear discrete systems was carried out. According to the research results, the most effective of them were chosen. They were Izanloo – Fakoorian – Yazdi – Simon filter and Särkkä – Nummenmaa filter. The first one relates to correntropy filters, the weak point of which is the choice of the size of the Gaussian kernel. The second filter is devoid of this drawback, and therefore will be used in this work as a robust filter in solving the problem of active identification.

2. Problem definition

Let us consider the following controlled, observed and identifiable model of stochastic linear discrete system in the state space:
\[ x(t_{k+1}) = \Phi(t_k)x(t_k) + \Psi(t_k)u(t_k) + \Gamma(t_k)w(t_k) , \]
\[ y(t_{k+1}) = H(t_{k+1})x(t_{k+1}) + v(t_{k+1}), \quad k = 0, 1, \ldots, N-1, \]

where \( x(t_k) \) is the state \( N \)-vector; \( u(t_k) \) is the system noise \( p \)-vector; \( y(t_{k+1}) \) is the measurement \( m \)-vector (output); \( v(t_{k+1}) \) is the measurement noise \( m \)-vector.

Let us suppose that:
- the random vectors \( w(t_k) \) and \( v(t_{k+1}) \) is white Gaussian sequences with:
  \[ E[w(t_k)]=0, \quad E[w(t_k)w^T(t_i)]=Q(t_k)\delta_{ii}, \]
  \[ E[v(t_{k+1})]=0, \quad E[v(t_{k+1})v^T(t_{k+1})]=R(t_{k+1})\delta_{ii}, \]
  \[ E[w(t_k)v^T(t_i)]=0, \quad k, i = 0, \ldots, N-1 \]
  (here and below \( E[\cdot] \) denotes the mathematical expectation, \( \delta_{ii} \) – the Kronecker symbol);
- the initial state \( x(t_0) \) is normal distribution with parameters:
  \[ E[x(t_0)] = \bar{x}(t_0), \]
  \[ E[(x(t_0) - \bar{x}(t_0))(x(t_0) - \bar{x}(t_0))^T] = P(t_0) \]
  and is uncorrelated with \( w(t_k) \) and \( v(t_{k+1}) \) for all values of \( k \);
- output data may contain outliers;
- the unknown parameters are summarized in the \( s \)-vector \( \theta \) and can be contained in elements of the matrices \( \Phi(t_k), \Psi(t_k), \Gamma(t_k), H(t_{k+1}), Q(t_k), P(t_0) \) and initial state vector \( \bar{x}(t_0) \) in various combinations.

For the mathematical model (1), (2), taking into account the stated a priori assumptions, we need to develop a procedure for active parametric identification of stochastic dynamic systems based on the robust Särkkä–Nummenmaa filter and conduct a numerical study of its application effectiveness.

3. Procedure of active identification

The general scheme of the procedure of active identification of systems with a pre-selected model structure suggests the following steps:

**step 1.** Calculation of parameter estimates based on measurement data corresponding to a test input signal.

**step 2.** Synthesis based on the obtained estimates of the optimal input signal (experiment design).

**step 3.** Recalculation of estimates of unknown parameters according to the measured data corresponding to the synthesized signal.

**Step 1**

Unknown parameters estimations of the mathematical model (1), (2) are carried out according to observational data \( \Xi \) by using some criterion of identification \( \chi \). The collection of numerical data occurs during identification experiments which are carried out under some design \( \xi_v \).

Suppose that experimenter can make \( v \) system starts. He gives signal \( U_1 \) as the input of the system \( d_1 \) times, signal \( U_2 - d_2 \) times, etc., at last, signal \( U_q - d_q \) times. In this case, the discrete (exact) normalized experimental design \( \xi_v \) is:

\[ \xi_v = \begin{bmatrix} U_1 & U_2 & \cdots & U_q \\ d_1 & d_2 & \cdots & d_q \end{bmatrix}, \quad U_i \in \Omega_u, \quad i = 1, 2, \ldots, q, \]
where $\Omega_d \subset \mathbb{R}^{N_d}$ is area of allowable values of input signals.

Let us denote through $Y^T_j = \begin{bmatrix} y^j(t_0) \cdots y^j(t_N) \end{bmatrix}^T$ realization of the output signal with number $j$ ($j = 1, 2, \ldots, d_j$), corresponding to input signal $U^T_i = \begin{bmatrix} u^i(t_0) \cdots u^i(t_{N-i+1}) \end{bmatrix}^T$ with number $i$ ($i = 1, 2, \ldots, q$). Then, as a result of carrying out the identification experiment for design $\xi$, the following set will be generated:

$$\Xi = \left\{(U_j, Y_j) : j = 1, 2, \ldots, d_j, i = 1, 2, \ldots, q\right\}. \sum d_j = N_d.$$

Maximum likelihood method is chosen for unknown parameter estimation. In view of the fact that measurement data contain outliers (the distribution of the measurement noise may differ from the normal distribution), we will calculate quasi-likelihood estimates, solving the optimization problem:

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \left[ \chi(\theta; \Xi) \right] = \arg \min_{\theta \in \Theta} \left[ -\ln L(\theta; \Xi) \right].$$

Here

$$\chi(\theta; \Xi) = \frac{N_m}{2} \ln 2\pi + \frac{1}{2} \sum_{i=1}^q \sum_{j=1}^{d_j} \sum_{k=1}^{N_{i+j}} \ln \det B^j(t_{i+j}) + \frac{1}{2} \sum_{i=1}^q \sum_{j=1}^{d_j} \sum_{k=1}^{N_{i+j}} \left[ e^j(t_{i+j}) \right]^T \left[ B^j(t_{i+j}) \right]^{-1} e^j(t_{i+j}),$$

where $e^j(t_{i+j})$ and $B^j(t_{i+j})$ are determined by recursive equations of the robust filter. As such filter we choose Särkkä–Nummenmaa filter obtained on the basis of variational Bayesian estimation under the assumption that the covariance noise matrix of measurements $R(t_{i+j})$ of order $m$ is considered diagonal and unknown.

**Filter Särkkä–Nummenmaa [8]**

**Initialization:**

$$\hat{x}(t_0|t_0) = \bar{x}(t_0), \ P(t_0|t_0) = P(t_0) ; \ \alpha(t_0) = (\alpha_{00}, \alpha_{01}, \ldots, \alpha_{0m}), \ \beta(t_0) = (\beta_{01}, \beta_{02}, \ldots, \beta_{0m}), \ L = L_0.$$

**To run in a loop on $k = 0, \ldots, N-1$**

**Prediction:**

$$\hat{x}(t_{k+1}|t_k) = \Phi(t_k) \hat{x}(t_k|t_k) + \Sigma(t_k) u(t_k) ; \ P(t_{k+1}|t_k) = \Phi(t_k) P(t_k|t_k) \Phi^T(t_k) + \Gamma(t_k) \Sigma R^{-1}(t_k) \Gamma^T(t_k).$$

**Update:**

$$e(t_{k+1}) = y(t_{k+1}) - H(t_{k+1}) \hat{x}(t_{k+1}|t_k) ; \ \alpha(t_{k+1}) = \begin{bmatrix} 1 + \alpha_{1}(t_k) & \ldots & \frac{1}{2} + \alpha_{m}(t_k) \end{bmatrix} ; \ \beta(t_{k+1}) = \beta(t_k).$$

**To run in a loop on $\mu = 1, \ldots, L$**

$$\hat{R}^\mu(t_{k+1}) = \text{diag} \left( \frac{\beta_{1}^\mu(t_{k+1})}{\alpha_{1}(t_{k+1})}, \ldots, \frac{\beta_{m}^\mu(t_{k+1})}{\alpha_{m}(t_{k+1})} \right);$$

$$B^\mu(t_{k+1}) = H(t_{k+1}) P(t_{k+1}|t_k) H^T(t_{k+1}) + \hat{R}^\mu(t_{k+1});$$

$$K^\mu(t_{k+1}) = P(t_{k+1}|t_k) H^T(t_{k+1}) \left[ B^\mu(t_{k+1}) \right]^{-1};$$

$$\hat{\epsilon}^\mu(t_{k+1}|t_k) = \hat{x}(t_{k+1}|t_k) + K^\mu(t_{k+1}) e(t_{k+1});$$

$$P^\mu(t_{k+1}|t_k) = \left[ I - K^\mu(t_{k+1}) H(t_{k+1}) \right] P(t_{k+1}|t_k).$$

**To run in a loop on $\eta = 1, \ldots, m$**
\[
\beta^\mu_\eta(t_{k+1}) = \beta^\mu_\eta(t_i) + \frac{1}{2}\left[y\left(t_{k+1}\right) - H\left(t_{k+1}\right)\hat{x}^\mu\left(t_{k+1} \mid t_{k+1}\right)\right]^2 + \frac{1}{2}\left[H\left(t_{k+1}\right)P^\mu\left(t_{k+1} \mid t_{k+1}\right)H^T\left(t_{k+1}\right)\right]_{\eta \eta}
\]

End of loop \( \eta \)

End of loop \( \mu \)

\[
\hat{x}\left(t_{k+1} \mid t_{k+1}\right) = \hat{x}^L\left(t_{k+1} \mid t_{k+1}\right);
\]

\[
P\left(t_{k+1} \mid t_{k+1}\right) = P^L\left(t_{k+1} \mid t_{k+1}\right);
\]

\[
\beta\left(t_{k+1}\right) = \beta^L\left(t_{k+1}\right).
\]

End of loop \( k \)

We use Sequential Quadratic Programming (SQP) method [9,10] for solving a nonlinear programming problem (3). This method refers to first-order methods implemented within the Optimization Toolbox package of the MATLAB and it is one of the most advanced and effective at the moment.

**Step 2**

The application of optimal control theory in estimating unknown parameters makes it possible to improve the accuracy of estimating parameters by collecting the most informative data.

Under continuous normalized design \( \xi \) we will understand a set of values:

\[
\xi = \left\{ U_1, U_2, \ldots, U_q \right\}, \quad p_i \geq 0, \sum_{i=1}^{q} p_i = 1, U_i \in \Omega_U, i = 1,2,\ldots,q .
\]

Unlike discrete design \( \xi_v \), weights \( p_i \) in continuous design \( \xi \) can take any values, including irrational number.

Information matrix \( M(\xi) \) for design (4) is determined by the relation:

\[
M(\xi) = \sum_{i=1}^{q} p_i M\left(U_i, \hat{\theta}\right),
\]

where Fisher information matrices for one-point designs \( M\left(U_i, \hat{\theta}\right) \) depend on the vector of unknown parameter estimates \( \hat{\theta} \) obtained in the first step of the active parametric identification procedure (this fact allows us to speak further only about locally optimal design).

The optimal experimental design is found for some convex functional \( X \) of information matrix \( M(\xi) \) by solving the following extremal problem:

\[
\xi^* = \arg\min_{\xi \in \Omega_\xi} X\left[M(\xi)\right].
\]

Solving the problem of experiment design, we influence the lower bound of the Rao-Kramer inequality in a certain way: for D-optimal design we minimize the volume, for A-optimal design we minimize the sum of squares of the ellipsoid axis lengths of the scattering parameter estimates.

The first of these involves minimum search of the functional \( X\left[M(\xi)\right] \) directly (it is recommended to choose \( q = \frac{s(s+1)}{2} + 1 \)) and the second is the solution of the dual problem and based on the generalized equivalence theorem. The application of the so-called combined approach is useful when solving practical problems. In this approach, direct procedure is used at the beginning to improve the initial approximation of the design, and then dual procedure is used to obtain the final result. The combined procedure for constructing continuous A-optimal designs with pre-computed parameter estimates of \( \hat{\theta} \) is presented in [14]. Construction of D-optimal design requires minor changes.
Practical application of the synthesized optimal design is difficult, because weights are arbitrary real numbers, enclosed in the interval from zero to one. In the case of a given number $v$ of possible system starts, it is necessary to "round" the continuous design to discrete \cite{12}. As a result, we obtain a discrete design:

$$\xi_v = \left\{ \frac{U_1^*}{v}, \frac{U_2^*}{v}, \ldots, \frac{U_q^*}{v} \right\}$$

Step 3

Recalculation of unknown parameters estimates from experimental data corresponding to the synthesized signal.

4. Simulation

The presented procedure of active parametric identification is tested using the example of a position control system model consisting of an antenna and a direct current (DC) motor \cite{15}.

Let the first component of the state vector be responsible for the angular position of the antenna, the second for its angular velocity. The input signal is the voltage at the input of DC amplifier driving motor. The angular position is measured with the help of a potentiometer and all the a priori assumptions stated in problem definition section. Then state and observation models at discrete moments of time are determined by the relations:

$$x(t_{k+1}) = \begin{bmatrix} 1 & \frac{1}{\theta_1}(1 - e^{-\theta_1 T}) \\ 0 & e^{-\theta_1 T} \end{bmatrix} x(t_k) + \begin{bmatrix} \frac{\theta_1}{\theta_1} (T - \frac{1}{\theta_1} + \frac{1}{\theta_1} e^{-\theta_1 T}) \\ \frac{\theta_2}{\theta_1} (1 - e^{-\theta_1 T}) \end{bmatrix} u(t_k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t_k);$$

$$y(t_{k+1}) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t_{k+1}) + v(t_{k+1}), \quad k = 0, 1, \ldots, N-1,$$

where $\theta_1, \theta_2$ is unknown parameters and $\Omega_0 = \{1 < \theta_1 < 6, 0 < \theta_2 < 1\}$.

Let us set $N = 30$, $Q(t_k) = Q = 0.01$, $R(t_{k+1}) = R = 0.1$, $x(t_0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $P(t_0) = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}$ and set the values parameters of Särkkä–Nummenmaa filter $\alpha_0 = \beta_0 = 1$, $L_0 = 4$. We chose the area of allowable values of input signals $\Omega_U = \{ U \in \mathbb{R}^N | 2 \leq u(t_k) \leq 30, k = 1, 2, \ldots, N-1 \}$.

To reduce the dependence of the estimation results on the experimental data, we perform five independent starts of the system and average the obtained estimates of the unknown parameters. The implementation of the output signals will be obtained by computer simulation with true values of the parameters $\theta_1^* = 4.600, \theta_2^* = 0.787$ by setting the pollution coefficient of the sample $\lambda = 0.1$ and the noise dispersion of anomalous observations $R_A = 1000 R$. We assume the processed data from the random nature of the anomalous dimensions.

The quality of parametric identification will be judged by the value of the relative estimation error $\delta_\theta$ calculated by formula:

$$\delta_\theta = \sqrt{\frac{(\hat{\theta}_1 - \hat{\theta}_1)^2 + (\hat{\theta}_2 - \hat{\theta}_2)^2}{\sigma_1^2 + \sigma_2^2}}$$

where $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)$ is vector of unknown parameters estimates.

The results of the robust procedure of active parametric identification are presented in Table 1 (the optimal design is one-point).
Table 1. Results of robust procedure of active parametrical identification

| Experiment design and values of the A-optimality criterion | System start number | Estimates and estimation errors | \( \hat{\theta}_1 \) | \( \hat{\theta}_2 \) | \( \delta_0 \) |
|--------------------------------------------------------|---------------------|--------------------------------|----------------|----------------|----------------|
| \( \xi = \{U\} \)                                      | 1                   |                                | 3.802          | 0.636          | 0.173          |
|                                                        | 2                   |                                | 3.065          | 0.572          | 0.332          |
|                                                        | 3                   |                                | 3.629          | 0.618          | 0.211          |
| \( U \) corresponds to the signal shown in Figure 1.   | 4                   |                                | 4.306          | 0.761          | 0.063          |
|                                                        | 5                   |                                | 2.823          | 0.491          | 0.386          |
| \( X[M(\xi)] = 1.008 \)                               |                     | Average value for startups     | 3.525          | 0.616          | 0.233          |
| \( \xi^* = \{U^*\} \)                                  | 1                   |                                | 4.647          | 0.806          | 0.010          |
|                                                        | 2                   |                                | 5.781          | 1.001          | 0.257          |
|                                                        | 3                   |                                | 4.650          | 0.802          | 0.011          |
| \( U^* \) corresponds to the signal shown in Figure 2. | 4                   |                                | 4.816          | 0.820          | 0.047          |
|                                                        | 5                   |                                | 4.617          | 0.801          | 0.004          |
| \( X[M(\xi^*)] = 0.241 \)                             |                     | Average value for startups     | 4.902          | 0.846          | 0.066          |

Figure 1. Initial input signal.

Figure 2. Optimum input signal.

Analysis of contents in Table 2 shows that design of A-optimal input signal using the combined procedure makes it possible to improve the quality of estimation by 16.7% percent.

Thus, the authors consider that the applying of the active parametric identification procedures based on robust estimation and optimal design of input signals is helpful and advisable in the presence of outliers in the measurement data.

5. Conclusion
The robust procedure of active parametric identification for models of stochastic linear discrete systems including robust estimation of parameters based on the Särkkä–Nummenmaa filter and optimal design
of input signal is developed. The case of the entry of unknown parameters into the state and observations equations, the initial condition and the covariance noise matrices of the system and measurements is considered. The efficiency of the developed robust procedure of active parametric identification is demonstrated by the example of a DC motor control system.

6. Acknowledgments
The work was financially supported by the Ministry of Education and Science of the Russian Federation on the state task (project № 2.7996.2017/8.9).

References
[1] Ljung L 1991 System identification: Theory for the user (Moscow: Science) p 432 (in Russian)
[2] Walter E and Pronzato L 1997 Identification of parametric models from experimental data (Berlin: Springer-Verlag) p 413
[3] Jauberthie C, Bourronville F, Coton P and Rendell F 2006 Optimal input design for aircraft parameter estimation Aerospace science and technology 10 331–337
[4] Brasil N, Hemery E and Goes L 2009 Aircraft parameter estimation experiment design considering measurement colored residuals Journal of aircraft 46(6) 1857–1865
[5] Bernaerts K, Servaes R D, Kooyman S, Versyck K J and Van Impe J 2002 Optimal temperature input design for estimation of the Square Root model parameters: parameter accuracy and model validity restrictions International Journal of Food Microbiology 73 145–157
[6] Galvanin F, Marchesini R, Barolo M, Bezzo F and Fidaleo M 2016 Optimal design of experiments or parameter identification in electrodialysis models Chemical engineering research and design 105 107–119
[7] Chubich V M and Prokofieva A E 2017 Comparative analysis of some robust filters for non-stationary linear discrete systems Bulletin of Irkutsk state technical University 21(12) 123-137 (in Russian)
[8] Särkkä S and Nummenmaa A 2009 Recursive noise adaptive Kalman filtering by variational Bayesian approximations IEEE Transactions on Automatic control 54 596–600
[9] Antoniou A and Lu W-S 2007 Practical optimization: algorithms and engineering applications (New York: Springer) p 669
[10] Izmailov A F and Solodov M V 2008 Numerical optimization methods (Moscow: Phystmathlit Publ) p 320 (in Russian)
[11] Fedorov V V 1971 Theory of optimal experiment (Moscow: Science) p 312 (in Russian)
[12] Yermakov S M and Zhiglyavskiy A A 1987 Mathematical theory of optimal experiment (Moscow: Science) p 320 (in Russian)
[13] Pronzato L and Pazman A 2013 Design of experiments in nonlinear models. Asymptotic normality, optimality criteria and small-sample properties (New York: Springer) p 399
[14] Chubich V M and Prokofieva A E 2018 The application of robust estimation to active parametric identification of stochastic linear discrete systems Actual problems of electronic instrument engineering (APEIE–2018) 1(4) 53-57
[15] Kwakernaak H and Sivan R 1972 Linear optimal control systems (New York: John Wiley Sons) p 595