

Research Article

A Bayes Estimator of Parameters of Nonlinear Dynamic Systems

I. A. Boguslavsky

State Institute of Aviation Systems, Physical-Technical Institute, Russia

Correspondence should be addressed to I. A. Boguslavsky, bogus@gosniias.ru

Received 28 November 2008; Revised 28 February 2009; Accepted 18 May 2009

Recommended by David Chelidze

A new multipolynomial approximations algorithm (the MPA algorithm) is proposed for estimating the state vector θ of virtually any dynamical (evolutionary) system. The input of the algorithm consists of discrete-time observations Y . An adjustment of the algorithm is required to the generation of arrays of random sequences of state vectors and observations scalars corresponding to a given sequence of time instants. The distributions of the random factors (vectors of the initial states and random perturbations of the system, scalars of random observational errors) can be arbitrary but have to be prescribed beforehand. The output of the algorithm is a vector polynomial series with respect to products of nonnegative integer powers of the results of real observations or some functions of these results. The sum of the powers does not exceed some given integer d . The series is a vector polynomial approximation of the vector $E(\theta | Y)$, which is the conditional expectation of the vector under evaluation (or given functions of the components of that vector). The vector coefficients of the polynomial series are constructed in such a way that the approximation errors uniformly tend to zero as the integer d increases. These coefficients are found by the Monte-Carlo method and a process of recurrent calculations that do not require matrix inversion.

Copyright © 2009 I. A. Boguslavsky. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Consider a dynamical system whose mathematical model has the form

$$x(k+1) = f(x(k), \eta_k, k). \quad (1.1)$$

Here, the integers $0, 1, \dots, k, \dots$ refer to time instants $t_0, t_1, \dots, t_k, \dots$; $x(k) \in R^n$ is the state vector of the dynamical system at the instant t_k ; $\eta_k \in R^n$ is the vector of random perturbations; $f(x(k), \eta_k, k)$ is a given function (nonlinearly depending on its arguments, in general); $x(0), \eta_k$ are random vectors with given distributions.

Let the results of observations at the instants $t_1, \dots, t_k, \dots, t_N$ be described by a scalar sequence $y_1, \dots, y_k, \dots, y_N$ generated by the mathematical model

$$y_k = H_k(x(k), \xi_k), \quad (1.2)$$

where $H(k, x(k))$ is a given function of its arguments, and the distribution is given for the sequence of random variables $\xi_1, \dots, \xi_k, \dots, \xi_N$.

In what follows, it is assumed that y_1, \dots, y_k are components of the vector Y_k .

A *parameter vector* θ is defined as a vector whose components coincide with some components (or given functions of these) of state vectors of the dynamical system at given time instants. In the present paper, we describe a new algorithm for estimating the vector θ .

The problem of estimating the vector θ is called [1, 2] the *problem of smoothing*, if $\theta = x(0)$; the *problem of filtration*, if $\theta = x(N)$; the *problem of extrapolation*, if $\theta = x(N^*)$, $N^* > N$.

It is well known that the mean-square optimal estimate of the random vector θ on the basis of observations of the vector Y_N is the vector of conditional expectation $E(\theta | Y_N)$. Therefore, we consider the problem of creating the MPA algorithm for the construction of approximations converging to the vector $E(\theta | Y_N)$.

There is only one requirement with regard to computer power and the expressions (1.1) and (1.2): for given distributions of random variables $x(0)$, η_k , ξ_i , it is possible to generate a set that consists of a considerable number (several thousands) of sequences of random state vectors and observation results satisfying (1.1) and (1.2).

Elements of the sequence $y_1, \dots, y_k, \dots, y_N$ are sent to the input of the algorithm at the instants $t_1, \dots, t_k, \dots, t_N$.

Let us briefly consider some known algorithms that give an approximate solution to the above estimation problems.

(1) The algorithm of the nonlinear least-squares method should be regarded as the most general method of solution. This algorithm is mentioned in a huge number of publications such as, for instance, [3]. Thus, in the smoothing problem for $\eta(k) \equiv 0$, the algorithm determines the estimate vector $\hat{\theta} = z_{x(0)}(Y_N)$ that approximately minimizes the heuristic quality functional $S(Y_N, z_{x(0)}(Y_N))$ with the constraints (1.1). The estimation quality functional is the sum of squared residuals

$$S(Y_N, z_{x(0)}(Y_N)) = \sum_{i=1}^{i=N} (y_i - H(i, x_i(z_{x(0)}(Y_N))))^2. \quad (1.3)$$

Here, the vectors $x_i(z_{x(0)}(Y_N))$ are obtained by consecutive calculations of state vectors on the basis of (1.1) with the initial data $z_{x(0)}(Y_N)$. For the numerical solution of this problem one utilizes numerous versions of the gradient method or the Newton method. In order to apply the Newton method, one has to solve a system of algebraic equations, which is obtained if we equate to zero partial derivatives of the right-hand side of (1.3) with respect to the components of the vector $z_{x_0}(Y_N)$.

The basic drawbacks of the nonlinear least-squares method are the following:

- (i) the quality functional in the form of the sum of squared residuals is theoretically adequate for an estimate quality criterion only in the unrealistic case of the residuals being independent random variables with Gaussian distribution;

- (ii) the versions of the gradient method or the Newton method are applicable if the functions in the mathematical models (1.1) and (1.2) are differentiable; these methods require the existence of a good first approximation and the construction of a complicated interactive computation process, if the function $S(Y_N, z_{x_0}(Y_N))$ happens to have a set of local minima;
- (iii) to get an idea of the statistical characteristics of estimation errors is possible only by means of the Monte-Carlo method (there are no explicit formulas for the estimate error variance).

(2) The extended Kalman filter algorithm (EKF algorithm) yields an heuristic solution to the filtration problem (there are many publications dedicated to the exposition of the EKF algorithm; it suffices to mention [4]). The EKF algorithm is based on a sequence of linearizations of nonlinear functions involved in the mathematical model of a dynamical system, the linearizations being constructed in a neighborhood of a sequence of estimate vectors.

The recurrent scheme of the EKF algorithm has a two-step structure.

Step 1. Suppose that after the instant t_{k-1} , approximations have been found for the first and the second statistical moments of the state vector at the instant t_{k-1} (the estimation vector for this vector coincides with the first statistical moment). On the time interval $[t_{k-1}, t_k - 0]$ (prior to observation results at the instant t_k) prediction is made and its results are approximations for the first and the second statistical moments of the state vector at the instant $t_k - 0$. The vector and the matrix of the prediction are found by the calculation of a matrix formed by partial derivatives of the state vector (the Jacobian matrix). Therefore, realization of the prediction requires the differentiability of the right-hand side of (1.1) and the admissibility of its linearization with respect to the components of the state vector increments arising from t_{k-1} to t_k .

Step 2. After actual observations at the instant t_k and linearization of the right-hand side of (1.2), mean-square optimal linear correction of the first statistical moment of the state vector at the instant t_k is implemented, as well as the corresponding correction of the second statistical moment. This step results in approximations for the first and the second statistical moments of the state vector components at the instant t_k .

The two-step scheme of the recurrent EKF algorithm is quite credible and gives quite satisfactory results in many practically important cases.

The basic drawbacks of the EKF algorithm are similar to those of the nonlinear least-squares method:

- (i) the estimate quality functional is defined and coincides with the estimation error variance vector only if the functions in the right-hand sides of (1.1) and (1.2) linearly depend on their arguments (in this situation, the EKF turns into the standard discrete Kalman filter);
- (ii) the EKF algorithm is applicable only if the functions involved in the mathematical models (1.1) and (1.2) are differentiable;
- (iii) if the functions in (1.1) and (1.2) are essentially nonlinear, it is necessary to have a fairly good first approximation in order to ensure a convergent calculation process (the notion of a "fairly good approximation" is heuristic and is determined on the basis of numerical experiments);
- (iv) there are no explicit formulas for the estimate error variance.

(3) In last ten years the significant number of researches with representation basic solution of a problem by definition of conditional expectation of a vector of parameters for mathematical model nonlinear dynamic system [1, 5–10] was published. By means of multiple application of formula Bayes to vectors from (1.1) and (1.2) and by using numerical quadratures, the recurrent equations are easily discovered for the probability density function (pdf) of state vectors of dynamic system. The actual solutions of the previous recurrent equations are, however, unfeasible because of the unwieldy dimensions of the integrals involved [10]. Therefore several alternative strategies have been developed. One set of such alternatives consists of developing suboptimal filtering, such as those based on linearization or transformations, and the other consists of methods which employ Monte-Carlo simulation strategies (e.g., estimation by using the sequential importance sampling particle filter) to approximate evaluate the multidimensional integrals in a recursive manner.

The presented direction is perspective but is bulky and unsuitable for the solution of practical applied (instead of model!) problems of nonlinear identification.

2. Schematic Diagram of MPA Algorithm

Let us describe the basic structure of the MPA algorithm [11–14]. Denote by θ the vector of dimension $r \times 1$ to be estimated upon fixing the vector Y_N .

The MPA algorithm is a new recurrent algorithm, such as asymptotic accurately solve nonlinear problem to construction of the vectors conditional expectations of the vector of parameters for mathematical model nonlinear dynamic system by any random errors and perturbations with given distributions. Therefore MPA algorithm approximate solve problem the optimal mean-square estimation.

MPA algorithm in principle differs from the all above mentioned algorithms since

- (i) MPA algorithm does not make linearizations and do not attempt to find the posterior probability density function of state vectors of dynamic system and never uses Bayes's formula, although essentially MPA algorithm is Bayes—it uses a priori information about a priori distributions. For given Y_N , MPA algorithm once build approximations to vectors posterior expectation of the estimated parameters vector. The approximation determined in the form segments of the vectorial power series relatively products the integer non negative powers of components of the vector Y_N . For every term of series sum of powers bounded above given integer d . For given Y_N , the approximations uniformly bounded converge to the vectors posterior expectation of the estimated parameters vector by $d \rightarrow \infty$;
- (ii) the vectorial coefficients of power series are defined by means of an adjustment of MPA algorithm. The adjustment is realized be means of define the first and second statistical moments of basis random vector, which components are estimated parameters and the mentioned powers of components of the vector Y_N . This definition is realized by means of Monte-Carlo's method from (1.1) and (1.2). However, after these vectorial coefficients have been calculated and kept in computer memory, the estimation vectors are determined by simple calculations after any new observations of the vectors Y_N ;

- (iii) there is a sequence r vectors such as each consist from k leading to a component of vector θ , $k = 1, \dots, r$. After adjustment by given d, N , MPA algorithm determines the recurrent approximations to the conditional expectation of these vectors and matrices covariance of errors of these approximations. In this connection, formulas of evaluations are simple and do not require the inverse of matrices and evaluations of quadratures.

The new MPA algorithm proposed in this paper for estimating the vector of parameter (and in particular state vector) of a nonlinear dynamical system, for the most part, has no drawbacks mentioned above.

Step 1. Suppose that d is a given positive integer number and the set of integer numbers a_1, \dots, a_N consists of all nonnegative solutions of the integer inequality $a_1 + \dots + a_N \leq d$, the number of which we denote by $m(d, N)$. The value $m(d, N)$ is given by the recurrent formula proved by induction,

$$m(d, N) = m(d - 1, N) + (N + d - 1) \cdots \frac{k}{d!}, \quad m(1, N) = N. \quad (2.1)$$

We obtain the vector $W_N(d)$ of dimension $m(d, N) \times 1$, the components $w_1, \dots, w_{m(d, N)}$ of which are all possible values of the form that represent the powers of measurable values.

Next, we define a basic vector $V(d, N)$ of dimension $(r + m(d, N)) \times 1$, $V(d, N) = \|\theta \ W_N(d)\|^T$.

Step 2. We use a known statistical generator of random vectors $x(0), \eta_k, \xi_i$ to solve repeatedly the Cauchy problem for (1.1) for a given initial conditions $x(0)$, a control law u and various realizations of random vectors $x(0), \eta_k, \xi_i$ and k . As a result, the computer memory will contain a set of realizations of random basic vectors $V(d, N)$ sufficient for the calculation of the statistical characteristics of the basic vector $V(d, N)$. We apply the Monte-Carlo method to find the prior first and second statistical moments of the vector $V(d, N)$, that is, the mathematical expectation $\bar{V}(d, N) = E(V(d, N))$, and the covariance matrix

$$C_V(d, N) = E\left((V(d, N) - E(V(d, N)))(V(d, N) - E(V(d, N)))^T\right). \quad (2.2)$$

Implementation of Step 2 is a learning process for the algorithm, adjusting it to solve the particular problem described by (1.1) and (2.3).

Step 3. For given d and N and a fixed vector Y_N , we assign the vector $Z_{E(\theta|Y_N)}(W_N(d))$ to be the solution to the estimation problem. This vector gives an approximate estimate of the vector $E(\theta | Y_N)$ that is optimal in the root-mean-square sense on the set of vector linear combinations of components of the vector $W_N(d)$. Then the vector $Z_{E(\theta|Y_N)}(W_N(d))$ is an element of this set and is an approximate mean-square optimal estimate for the vector $E(\theta | Y_N)$ of the conditional expectation of the vector θ .

The vector $Z_{E(\theta|Y_N)}(W_N(d))$ can be presented in a following aspect:

$$Z_{E(\theta|Y_N)}(W_N(d)) = \sum_{a_1 + \dots + a_N = d} \lambda(a_1, \dots, a_N) y_1^{a_1} \dots y_N^{a_N}, \quad (2.3)$$

where vectors $\lambda(a_1, \dots, a_N)$ are some weight vectorial factors.

The vector $\bar{V}(d, N) = E(V(d, N))$ and the matrix $C_V(d, N)$ are the initial conditions for the process of recurrent calculations that realizes the principle of observation decomposition [6] and consists of $m(d, N)$ steps. Once the final step is performed, we obtain vector weight coefficients $\lambda(a_1, \dots, a_N)$ for (1.1). Moreover, we determine the matrix $C(d, N)$, which is the covariance matrix of the estimation errors for the vector of conditional mathematical expectation estimated by the vector θ . Calculating the elements of the matrix $C(d, N)$, we have the method of preliminary (prior to the actual flight) analysis of observability of identified parameters for the given control law, structure of measurements and their expected random errors. Recurrent calculations do not require matrix inversion and indicate the situations when the next component of the vector $W_N(d)$ is close to linear combination of its previous components. To implement the recursion, we process the components of the vector $W_N(d)$ one after another. However, the adjustment of the algorithm performed by applying the Monte-Carlo method to find the vector $\bar{V}(d, N) = E(V(d, N))$ and the matrix $C_V(d, N)$ takes into account a priori ideas on stochastic structure of components of the whole set of possible vectors $W_N(d)$ that can appear in any realizations of the random vectors $x(0)$, η_k , ξ_i and k allowed by a priori conditions. This adjustment is the price we have to pay if we want the MPA algorithm to solve nonlinear identification problems efficiently. This is what makes the MPA algorithm differs fundamentally from, for instance, the standard Kalman filter designed to solve linear identification problems only or from multiple variations of algorithms resulted from attempts to extend the Kalman filter to nonlinear filtration problems.

Using the multidimensional analog of the Weierstrass theorem (a corollary of the Stone theorem [15]), we prove that with the increase of the integer d the error estimate vector $Z_{E(\theta|Y_N)}(W_N(d)) - E(\theta | Y_N)$ tends to zero uniformly in some domain.

The basic structure of the MPA algorithm described above allows us to use a strict quality criterion for the estimates obtained, since on each step the algorithm tends to ensure the minimal mean-square error of the estimate for the conditional expectation vector with a given volume of observations. Its closeness to the minimum increases with the increase of the integer d and the increase of the number of realizations, if the Monte-Carlo method is used.

If the number $m(d, N)$ of the approximating polynomial series is not small, a large volume of calculations has to be performed (after the initial choice of the integers n, N, d) for the determination of the vector coefficients $\lambda(a_1, \dots, a_N)$.

3. Mean-Square Optimal Estimate for the Conditional Expectation Vector $E(\theta | Y_N)$

Consider the principal algorithm (PA) for solving the problem of the mean-square optimal estimate for the vector $E(\theta | Y_N)$. It is known that the vector $E(\theta | Y_N)$ is the mean-square optimal estimate for the vector θ after the vector Y_N has been fixed. Therefore, it makes sense that the PA should estimate the conditional expectation vector.

We are going to construct a PA that gives a polynomial approximation of the vector $E(\theta | Y_N)$. To that end, we obtain an estimate for the vector $E(\theta | Y_N)$ which is linear with respect to the components of the vector $W_N(d)$ and is mean-square optimal.

In what follows the vector of that estimate is denoted by $Z_{E(\theta|Y_N)}(W_N(d))$.

An explicit expression for the estimate vector is obtained after the calculation of the elements of the vector $\bar{V}(d, N) = E(V(d, N))$ and the covariance matrix $C_V(d, N)$, which coincide with the first and the second (centered) statistical moments for the vector $V(d, N)$. This vector and this matrix can be divided into blocks whose structure can be represented as follows:

$$\begin{aligned}
 E(E(\theta | Y_N)) &= E(\theta), \\
 E(E(W(d, N) | Y_N)) &= E(W(d, N)), \\
 C_\theta(0) &= E\left((E(\theta | Y_N) - E(\theta))(E(\theta | Y_N) - E(\theta))^T\right) \\
 &= E\left((\theta - E(\theta))(\theta - E(\theta))^T\right), \\
 L_N(d) &= E\left(\left(E(\theta Y_N) - E(E(\theta | Y_N(d)))\right)(W_N(d) - E(W_N(d)))^T\right) = E(\theta)W_N(d)^T \\
 &\quad - E(\theta)E(W_N(d))^T, \\
 Q_N(d) &= E\left((W_N(d) - E(W_N(d)))(W_N(d) - E(W_N(d)))^T\right).
 \end{aligned} \tag{3.1}$$

The right-hand sides of the above blocks coincide with the first and the second (centered) statistical moments determined by the Monte-Carlo method. On the other hand, the left-hand sides of the same blocks coincide with the first and the second (centered) statistical moments of the components of the conditional expectation vector. Therefore, the said statistical moments can be found experimentally on the basis of mathematical models (1.1) and (1.2) also for the conditional expectation vectors. This obvious statement is crucial for the practical numerical procedure of estimating the conditional expectation vector $E(\theta | Y_N)$.

Let

$$Z_{E(\theta|Y_N)}(W_N(d)) = E(\theta) + \Lambda_N(d)(W_N(d) - E(W_N(d))), \tag{3.2}$$

where $\Lambda_N(d)$ is an $n \times m(d, N)$ -matrix satisfying the equation

$$\Lambda_N(d)Q_N(d) = L_N(d). \tag{3.3}$$

Let

$$\tilde{Z}_\theta(W_N(d)) = z + \tilde{\Lambda}_N(d)(W_N(d) - E(W_N(d))), \tag{3.4}$$

where z is an arbitrary $n \times 1$ -vector and $\tilde{\Lambda}_N(d)$ is an arbitrary $n \times m(d, N)$ -matrix. Let $C_Z(d, N)$ and $\tilde{C}(d, N)$ be error estimate covariance matrices for the vector $E(\theta | Y_N)$ in terms of the estimate vectors $Z_{E(\theta|Y_N)}(W_N(d))$ and $\tilde{Z}_\theta(W_N(d))$.

Lemma 3.1. *A matrix $\tilde{C}(d, N) - C_Z(d, N)$ is the nonnegative definite matrix: $C_Z(d, N) \leq \tilde{C}(d, N)$.*

Lemma 3.1 follows from the identity

$$\begin{aligned} \tilde{C}(d, N) &= C_Z(d, N) + \left(\Lambda_N(d) - \tilde{\Lambda}_N(d) \right) \left(\Lambda_N(d) - \tilde{\Lambda}_N(d) \right)^T \\ &\quad + \left(\Lambda_N(d) Q_N(d) - L_N(d) \right) \left(\tilde{\Lambda}_N(d) - \Lambda_N(d) \right)^T \\ &\quad + \left(\left(\Lambda_N(d) Q_N(d) - L_N(d) \right) \left(\tilde{\Lambda}_N(d) - \Lambda_N(d) \right)^T \right)^T + \left(z - E(\theta)(z - E\theta)^T \right). \end{aligned} \quad (3.5)$$

Corollary from Lemma 3.1

The vector $Z_{E(\theta|Y_N)}(W_N(d))$ is a mean-square optimal estimate for the vector $E(\theta | Y_N)$ on the set of estimates linear with respect to the components of the vector $W_N(d)$.

If $Q_N(d) > 0$, then the estimate vector is unique and

$$Z_{E(\theta|Y_N)}(W_N(d)) = E(\theta) + L_N(d) Q_N(d)^{-1} (W_N(d) - E(W_N(d))). \quad (3.6)$$

The covariance matrix $C_Z(d, N)$ for the error of the estimate of vector $E(\theta | Y_N)$ is defined by

$$C_Z(d, N) = C_\theta(0) - \Lambda_N(d) L_N(d). \quad (3.7)$$

If $Q_N(d) \geq 0$, then the vectors giving the linear mean-square optimal estimate are nonunique, but the variance of the components of the difference of these vectors is equal to zero.

Formula (3.2) gives explicit expressions for the vector coefficients $\lambda(a_1, \dots, a_N)$ in (2.3). These expressions are obtained by equating the right-hand side of (3.2) to the right-hand side of (2.3) and writing out explicit expressions for the components of the vector $W_N(d)$.

Let us examine asymptotic errors of the estimates obtained when using formula (3.2).

For a given vector Y_N , suppose that the vector $E(\theta | Y_N)$ is defined as a function of the argument Y_N in some a priori given compact domain $\Omega_{Y_N} \subset R^N$ and is continuous in that domain. Then the following result holds.

Theorem 3.2.

$$\sup_{Y_N \in \Omega_{Y_N}} |Z_{E(\theta|Y_N)}(W_N(d)) - E(\theta | Y_N)| \Rightarrow 0, \quad d \Rightarrow \infty. \quad (3.8)$$

Proof. According to the multidimensional version of the Weierstrass theorem [7], for any $\varepsilon > 0$ there is a multidimensional polynomial $P(W_N(d_\varepsilon))$ such that

$$\sup_{Y_N \in \Omega_{Y_N}} |P(W_N(d_\varepsilon)) - E(\theta | Y_N)| < \varepsilon. \quad (3.9)$$

This relation can be rewritten as

$$\sup_{Y_N \in \Omega_{Y_N}} |P(W_N(d)) - E(\theta | Y_N)| \Rightarrow 0, \quad d \Rightarrow \infty. \quad (3.10)$$

Let C be the covariance matrix for the random vector $P(W_N(d)) - E(\theta | Y_N)$:

$$C = E(P(W_N(d)) - E(\theta | Y_N))(P(W_N(d)) - E(\theta | Y_N))^T. \quad (3.11)$$

From (3.10), it follows that

$$C \Rightarrow 0_n, \quad d \Rightarrow \infty. \quad (3.12)$$

By construction, the vector $Z_{E(\theta|Y_N)}(W_N(d))$ gives a mean-square optimal estimate for the vector θ and this estimate is linear with respect to the components of the vector $W_N(d)$. But Lemma 3.1 implies that for any other nonoptimal linear estimate, in particular that like the vector $P(W_N(d))$, we have $C \geq C_Z(d, N)$. Hence, taking into account (3.12), we get

$$C_Z(d, N) \Rightarrow 0_n, \quad d \Rightarrow \infty. \quad (3.13)$$

The statement (3.13) is equivalent to (3.8), if we take into account that

$$C_Z(d, N) = \int (Z_{E(\theta|Y_N)}(W_N(d)) - E(\theta | Y_N)) \times (Z_{E(\theta|Y_N)}(W_N(d)) - E(\theta | Y_N))^T p(\theta, Y_N) d\theta dY_N, \quad (3.14)$$

where $p(\theta, Y_N)$ is the joint distribution density for the random vectors θ, Y_N . Theorem 3.2 is proved. \square

Thus, the PA, by virtue of (2.3), determines a vector series which, with the increase of the number of its terms $m(d, N)$, approximates the conditional expectation for the vector θ of the estimated parameters with arbitrarily small uniform mean-square error.

4. Recurrently MPA Algorithm

Suppose that for a given observation vector Y_N and an integer d we have constructed the vector $W_N(d)$ with the components $w_1, w_2, \dots, w_{m(d,N)}$.

The implementation of the algorithm starts with approximate calculation (by the Monte-Carlo method) of $n + m(d, n)$ -dimensional integrals that determine the components of the vector $\bar{V}(d, N) = E(V(d, N))$ and the matrix $C_V(d, N) = E((V(d, N) - \bar{V}(d, N))(V(d, N) - \bar{V}(d, N))^T)$. Further, splitting this vector and this matrix into suitable blocks and using (2.3), we find the desired estimate vector $Z_{E(\theta|Y_N)}(W_N(d))$ and the estimate error covariation matrix $C_Z(d, N)$.

However, this approach is unreasonable, since it requires the inversion of the matrix $Q_N(d)$, which is a difficult task for $Q_N(d)$ of a large dimension $m(d, N) \times m(d, N)$ or $Q_N(d)$ being close to a singular matrix.

Below we describe a recurrent calculation process based on the principle of decomposition of observations expounded in [11].

We are going to construct a recurrently MPA algorithm that does not involve matrix inversion and consists of $m(d, N)$ steps of calculations of the first and the second statistical moments for a sequence of special vectors $V_1, \dots, V_i, \dots, V_{m(d,N)}$ after a priori moments $\bar{V}(d, N), C_V(d, N)$ for the basic vector $V(d, N)$ have been found.

Denote by V_1 the vector formed by the components of the basic vector $V(d, N)$ remaining after the elimination of the component $w_1; \dots$; by V_{i+1} the vector formed by the components of V_i remaining after the elimination of w_i ; and so forth.

The quantity $w_{m(d,N)}$ is the last component of the vector $W_N(d)$, and after its elimination, the last vector $V_{m(d,N)}$ turns out equal to the estimate vector $Z_{E(\theta|Y_N)}(W_N(d))$.

On Step 1, formulas (3.2) and (3.6) are used for the calculation of the following quantities:

- (i) the vector \bar{V}_1 that yields an estimate for the vector V_1 which is mean-square optimal and linear with respect to w_1 ;
- (ii) the estimate error covariance matrix $C(V_1)$.

The estimate vector consists of the conditional expectation estimate vector $E(\theta | w_1)$ and a vector of dimension $(m(d, N) - 1) \times 1$. For w_1 fixed, the latter coincides with the vector of statistical prediction of mean *future* values of the quantities $w_2, \dots, w_{m(d,N)}$.

Note that the calculations of Step 1 are based on a priori values $\bar{V}(d, N), C(V(d, N))$ found beforehand.

In a similar way, suppose that on steps $1, \dots, i$ of the calculation process, the vector \bar{V}_i and the matrix $C(V_i)$ have been found after fixing the quantities w_1, \dots, w_i .

On step $i + 1$, the following quantities are calculated by means of (3.2) and (3.6):

- (i) the vector $\bar{V}_{(i+1)}$ of the estimate of V_{i+1} , which is mean-square optimal and linear with respect to w_1, \dots, w_{i+1} ;
- (ii) the estimate error covariance matrix $C(V_{(i+1)})$.

The vector $\bar{V}_{(i+1)}$ is composed of the conditional expectation estimate $E(\theta | w_1, \dots, w_{i+1})$ and a vector which, after the quantities w_1, \dots, w_{i+1} have been fixed, is the vector of statistical prediction for the mean values of "future" quantities $w_{i+2}, \dots, w_{m(d,N)}$.

Note that the calculations of step $i + 1$ are based on the already determined quantities $\bar{V}_i, C(V_i)$, which it is natural to call *a priori first and second statistical moments* of the "future"

quantities $w_{i+1}, \dots, w_{m(d,N)}$. The vector \bar{V}_i and the matrix $C(V_i)$ represent a priori data regarding the statistical moments of the components of the vector V_{i+1} before the quantity w_{i+1} is sent to the input of the algorithm.

The formulas of the computation process described above are the following:

$$\begin{aligned}\bar{V}_{(i+1)} &= \bar{V}_i + q_i^{-1} b_i (w_{i+1} - z_{w_{i+1}}), \\ C(V_{i+1}) &= C(V_i)^1 - q_i^{-1} b_i b_i^T.\end{aligned}\quad (4.1)$$

Here, the scalar $z_{w_{i+1}}$ is the $(n+1)$ th component of the vector \bar{V}_i (this scalar is a linear mean-square optimal estimate for the component w_{i+1} after the algorithm has used the components w_1, \dots, w_i); the vector \bar{V}_i^1 is obtained from the vector \bar{V}_i by eliminating its component $z_{w_{i+1}}$; the scalar q_i is the $(n+1)$ th diagonal element of the matrix $C(V_i)$ (this scalar is the variance of the estimate error for the component w_{i+1} after the components w_1, \dots, w_i have been used); the matrix $C(V_i)^1$ is obtained from the matrix $C(V_i)$ by elimination of its $(n+1)$ th row and its $(n+1)$ th column; the vector b_i coincides with the $(n+1)$ th column of the matrix $C(V_i)$ with the $(n+1)$ th component excluded.

If the scalar q_i happens to be close to zero, then the component w_{i+1} becomes close to a linear combination of the components w_1, \dots, w_i . In this case, the component w_{i+1} gives no new information about θ and should be dropped from the computation process..

Note that the sequence of random variables $(w_{i+1} - z_{w_{i+1}})$ is a renewable sequence.

The top left block of the $(n \times n)$ -matrix $C(V_i)$ coincides with the estimate error covariance matrix $C_Z(d, i)$ for the vector $E(\theta \mid w_1, \dots, w_i)$ after the vector $W_i(d)$ has been used by the algorithm.

Denote by $l(i)$ the vector formed by the first n components of the vector b_i . The formula representing the evolution of the covariance matrix $C_Z(d, i)$ as a function of i (the number of the components of the vector $W_i(d)$) has the form

$$C_Z(d, i) = C_\theta(0) - q_1^{-1} l(1) l(1)^T - \dots - q_{i-1}^{-1} l(i-1) l(i-1)^T. \quad (4.2)$$

In what follows, the above MPA algorithm is tested in several numerical problems of estimating components of state vectors of essentially nonlinear dynamical systems. The components to be estimated are unknown random constant parameters $\theta_1, \dots, \theta_r$ of a dynamical system. The specific applied problems considered below are those of smoothing and filtration.

The Monte-Carlo method is used if the number of random realizations is from 5000 to 10000. This number has no strong effect on estimate errors arising in the MPA algorithm. It is assumed that the estimated random parameters are statistically independent and have uniform distribution a priori.

The quantity $\sigma(i, \text{cal})$ is determined theoretically by the calculation of variances, that is, the diagonal elements of the covariance matrix $C_Z(d, N)$.

The quantity $\sigma(i, \text{exp})$ is determined experimentally by the Monte-Carlo method with the number of realizations being 10000.

The fact that the experimental and the theoretical mean-square deviations are approximately the same proves the correctness of the MPA algorithm formulas specified above.

5. Estimates for the Principal Moments of Inertia and Orientation Angles of the Principal Axes of Inertia of a Solid

Consider a solid in space (e.g., a satellite) with a fixed reference trihedron, that is, orthogonal axes X, Y, Z and their origin O . The corresponding projections $\omega_x, \omega_y, \omega_z$ of the angular velocity of the body, ω , are measured by angular velocity sensors. The projections M_x, M_y, M_z of the moment M of external forces are assumed known.

In the general situation, a solid has 6 moments of inertia relative to a rectangular reference frame X, Y, Z , among which three are called centrifugal moments. It is known that for a point O there is a rectangular coordinate system X_o, Y_o, Z_o whose axes coincide with the principal axes of inertia. In this coordinate system, the centrifugal moments of inertia are equal to zero and the remaining ones, J_x, J_y, J_z , are called the principal moments of inertia relative to the point O .

Let α, β, γ be the Euler angles that determine the angular position of X_o, Y_o, Z_o with respect to X, Y, Z , let A be the orthogonal matrix of the direction cosines of these angles.

Generally, a solid in outer space is a complex nonsymmetric structure whose parts play different roles and are constructed of different materials. Therefore, it is difficult to calculate the quantities $\alpha, \beta, \gamma, J_x, J_y, J_z$ on the basis of design diagrams of these parts and their junctures. However, for the analysis of the dynamics of a rotating solid it is convenient to use the coordinate axes that coincide with the principal axes of inertia, since the equations of motion (dynamical Euler equations) have a simple form in these axes.

Thus, it is reasonable to state the problem as follows: using the MPA algorithm, find estimates for the quantities $\theta_1 = \alpha, \theta_2 = \beta, \theta_3 = \gamma, \theta_4 = J_x, \theta_5 = J_y, \theta_6 = J_z$, provided that the solid is rotating in space and in the reference frame the quantities $\omega_x, \omega_y, \omega_z, M_x, M_y, M_z$ are observed as discrete-time functions.

In the continuous-time model, the equations of motion of a nonlinear dynamical system have the form

$$\begin{aligned} \frac{J_x d\omega_x^o}{dt} + (J_z - J_y)\omega_y^o\omega_z^o &= M_x^o, \\ \frac{J_y d\omega_y^o}{dt} + (J_x - J_z)\omega_z\omega_x &= M_y^o, \\ \frac{J_z d\omega_z^o}{dt} + (J_y - J_x)\omega_x^o\omega_y^o &= M_z^o, \end{aligned} \quad (5.1)$$

$$\begin{aligned} \frac{dJ_x}{dt} = 0; \quad \frac{dJ_y}{dt} = 0; \quad \frac{dJ_z}{dt} = 0, \\ \frac{d\alpha}{dt} = 0; \quad \frac{d\beta}{dt} = 0; \quad \frac{d\gamma}{dt} = 0. \end{aligned} \quad (5.2)$$

Here, $\omega_x^o, \omega_y^o, \omega_z^o, M_x^o, M_y^o, M_z^o$ are projections of the vectors ω and M to the axes X_o, Y_o, Z_o :

$$\begin{aligned} \omega^o &= A\omega, \\ M^o &= AM. \end{aligned} \quad (5.3)$$

Further, assume that observations of the time-functions $\omega_x, \omega_y, \omega_z, M_x, M_y, M_z$ take place with intervals 0.1 second and have duration 10 seconds. In order to create a mathematical model (1.1) with discrete time, we use numerical integration by the fourth-order Runge-Kutta method with constant step 0.2 second between observation instants and the initial conditions $\omega_x(0) = 1/20, \omega_y(0) = -1/20, \omega_z(0) = 1/20$ and $M_x = 100 \text{ kg}, M_y = -100 \text{ kg}, M_z = 100 \text{ kg}$.

Setting $y_x(t_k) = \omega_x(t_k), y_y(t_k) = \omega_y(t_k), y_z(t_k) = \omega_z(t_k)$, we find that at the instant t_k the components of the initial observation vector $y(t_k)$ are related to nonobservable quantities α, β, γ and nonobservable functions $\omega_x^o, \omega_y^o, \omega_z^o$ by

$$y(t_k) = A^T \omega^o(t_k). \quad (5.4)$$

Using the sequence $y_x(t_k), y_y(t_k), y_z(t_k)$ of primary observation vectors and (5.1), it is required to estimate the parameters $J_x, J_y, J_z, \alpha, \beta, \gamma$.

The secondary observations $y_x(i), y_y(i), y_z(i), i = 1 \div 4$, which serve as the input data of the estimation algorithm, are found by summing up the components of each l vectors of consecutive primary observations:

$$\begin{aligned} y_x(i) &= \sum_{j=1}^{j=l} y_x(t_{l(i-1)+j}), \\ y_y(i) &= \sum_{j=1}^{j=l} y_y(t_{l(i-1)+j}), \\ y_z(i) &= \sum_{j=1}^{j=l} y_z(t_{l(i-1)+j}). \end{aligned} \quad (5.5)$$

In our model, it is assumed that $l = 25$ and the array of 300 primary observations has been converted to an array of $N = 12$ scalar secondary observations.

For the operation of the MPA algorithm it is necessary to find a priori first and second statistical moments of the basic 18-dimensional vector formed by the estimated quantities $\alpha, \beta, \gamma, J_x, J_y, J_z$ and $y_x(i), y_y(i), y_z(i), i = 1 \div 4$. These data were found by the Monte-Carlo method under the assumption that the random parameters to be estimated have uniform distribution on a priori given segments:

$$\begin{aligned} -\rho &\leq \alpha, \beta, \gamma \leq \rho, \\ J_x^o(1-\rho) &\leq J_x \leq J_x^o(1+\rho), \\ J_y^o(1-\rho) &\leq J_y \leq J_y^o(1+\rho), \\ J_z^o(1-\rho) &\leq J_z \leq J_z^o(1+\rho), \end{aligned} \quad (5.6)$$

where $J_x^o = 4000, J_y^o = 20000$, and $J_z^o = 15000$.

The quantity ρ describes the range in which a priori errors of the estimated parameters may vary. In our model, the values of ρ were the following: $\rho = 0.2$ (large a priori errors), $\rho = 0.1$ (medium a priori errors), and $\rho = 0.05$ (small a priori errors).

Let $\sigma_a(i)$ be a priori mean-square deviation (MSD) of the scattering of the unknown parameter θ_i , and let $\sigma(i)$ be an a posteriori MSD of the error of estimating this parameter found after the utilization of the MPA. In what follows, the quality of the estimate of the parameter θ_i is characterized by the ratios $\sigma(i)/\sigma_a(i)$, $i = 1, \dots, 6$, which show how many times a posteriori (after observation) dispersal of the estimated parameter is smaller than a priori dispersal.

Below, for different values of ρ we give the values of the said ratios for the MPA algorithm operating with $d = 2$ (the estimate of θ_i is realized by second-order polynomials with respect to 12 secondary scalar observations; the number of terms of the polynomials is equal to $m(2, 12) = 90$):

$$\rho = 0.2$$

$$\begin{array}{cccccc} \sigma(1)/\sigma_a(1) & \sigma(2)/\sigma_a(2) & \sigma(3)/\sigma_a(3) & \sigma(4)/\sigma_a(4) & \sigma(5)/\sigma_a(5) & \sigma(6)/\sigma_a(6) \\ 0.5568 & 0.0440 & 0.5614 & 0.1005 & 0.0862 & 0.0886 \end{array} \quad (5.7)$$

$$\rho = 0.1$$

$$\begin{array}{cccccc} \sigma(1)/\sigma_a(1) & \sigma(2)/\sigma_a(2) & \sigma(3)/\sigma_a(3) & \sigma(4)/\sigma_a(4) & \sigma(5)/\sigma_a(5) & \sigma(6)/\sigma_a(6) \\ 0.5069 & 0.0063 & 0.5079 & 0.0123 & 0.0164 & 0.0161 \end{array} \quad (5.8)$$

$$\rho = 0.05$$

$$\begin{array}{cccccc} \sigma(1)/\sigma_a(1) & \sigma(2)/\sigma_a(2) & \sigma(3)/\sigma_a(3) & \sigma(4)/\sigma_a(4) & \sigma(5)/\sigma_a(5) & \sigma(6)/\sigma_a(6) \\ 0.4983 & 0.0012 & 0.4985 & 0.0017 & 0.0028 & 0.0026 \end{array} \quad (5.9)$$

These data show that for $d = 2$ and the chosen characteristics for the vectors of primary and secondary observations, a posteriori dispersal of estimate errors is about 100 or more times less than a priori dispersal of the unknown parameters β, J_x, J_y, J_z , provided that a priori errors of the parameters belong to the classes of medium or small errors. A posteriori dispersal of the parameters α, γ is only about 2 times less than a priori dispersal. To increase the accuracy of the estimates of these parameters, it seems necessary to use larger values of d , as well as other characteristics of the vectors of primary and secondary observations. It is also possible to use iterations that on each step decrease the intervals of possible a priori errors.

For the general reduction of estimate errors it seems reasonable to split a priori error intervals into smaller intervals. Thus, suppose that we have large a priori errors. In order to obtain estimate errors corresponding to the class of medium a priori errors it suffices to divide every large a priori interval into two parts and apply the MPA algorithm to each of the 64 resulting families of a priori intervals from the class of medium error intervals. The MPA algorithm calculation of the MSDs of a posteriori dispersal allows us to find a family of a priori intervals with the minimal values of these MSDs.

6. Estimates of the State Vector of a Satellite by Means of Handling the Information on Angles of Vising

Many results of astronomical observations depend on measurements and subsequent processing of digital data with regard to a sequence of viewing angles of space objects obtained in the optical frequency band. The viewing angles are measured by an optical device or on the basis of images on a photographic plate.

These angles are often the only available data regarding high-orbit sun-exposed satellites with onboard radio communication system failure (the energy of terrestrial radio band viewing facilities is insufficient because of the long distances to a satellite). Therefore, it makes sense to develop an algorithm that would give fairly precise estimates for state vectors of objects in space on the basis of processing the data about viewing angles.

For high-orbit satellites such data is obtained by specially constructed optical-electronic systems (e.g., the optical-electronic system *Window* [16] which monitors near-the-earth-space in optical frequency band and gives viewing information about space bodies at distances up to 40000 km).

In an inertial geocentric coordinate system, the equations of motion of a body whose mass is not very large have the form

$$\begin{aligned}\frac{dX_i}{dt} &= V_i, \\ \frac{dV_i}{dt} &= -\frac{\mu X_i}{R}.\end{aligned}\quad (6.1)$$

Here, $R = (X_1^2 + X_2^2 + X_3^2)^{1/2}$, μ is the gravity constant, X_i , $i = 1 \div 3$ are rectangular coordinates of the body.

The viewing angles are the two angles that determine (with respect to the earth) the direction of the viewing vector, that is, the direction of the straight line from the observer to an object in space.

We assume that at the current time instant t the center of viewing is located at a point of the Earth surface with geographical coordinates $\lambda = \omega t$ (longitude), φ (latitude), h (altitude), where ω is the angular speed of the Earth rotation. In the rotating right geocentric coordinate system x_i , $i = 1 \div 3$, the coordinates of the center of viewing are $X'_1 = r_e \cos(\varphi) - h \sin(\varphi)$, $X'_2 = 0$, $X'_3 = r_e \sin(\varphi) + h \cos(\varphi)$, where r_e is the Earth radius.

The two angles of orientation of the viewing vector, in the local horizon plane of the viewing center, are represented by an angle y_1 and, in the local vertical plane, by an angle y_2 . These angles are defined by

$$y_1 = \arctan\left(\frac{b}{a}\right), \quad y_2 = \arctan\left(\frac{c}{(a^2 + b^2)^{1/2}}\right), \quad (6.2)$$

where

$$\begin{aligned}a &= \cos(\lambda) \cos(\varphi) X_1 + \sin(\lambda) \cos(\varphi) X_2 + \sin(\varphi) X_3 - r_e, \\ b &= -\sin(\lambda) X_1 + \cos(\lambda) X_2, \\ c &= -\sin(\varphi) \cos(\lambda) X_1 + \sin(\varphi) \sin(\lambda) X_2 + \sin(\varphi) X_3 - h.\end{aligned}\quad (6.3)$$

In what follows, we assume that the algorithm processing the data on viewing angles should solve a filtration problem: in an inertial geocentric coordinate system at certain current time instants, the algorithm should estimate three rectangular coordinates of an object in space and three rectangular components of its velocity vector.

The possibility of creating such an algorithm is based on the existence or correlation between the varying viewing angles and varying rectangular coordinates of a space object. This correlation is due to the law of gravitation responsible for the gravity acceleration vector that changes the trajectory of its motion and the viewing angles and is a known function of rectangular coordinates of the object.

In the absence of gravity acceleration, it would be impossible to estimate the parameters of motion by measuring the viewing angles. Indeed, the law of variation of the viewing angles is the same for all bodies in uniform rectilinear motion with parallel velocity vectors and the same constant ratio of the distance and the velocity modulus. Therefore, in this situation the parameters of motion are unobservable and cannot be estimated on the basis of information about the viewing angles.

The problem of the algorithm for estimating current components of the state vector satisfying the system of six nonlinear differential equations (6.1) is a problem of nonlinear filtration. The input of the algorithm consists of a sequence of observation results (6.2) nonlinearly depending on the current values X_i , $i = 1 \div 3$.

Let τ be the time interval between two consecutive observations of pairs of viewing angles of the space object. After the first k observations, we construct integral observation results $y_{\text{int},1}$, $y_{\text{int},2}$. Here, $y_{\text{int},1}$ is the sum of the first k fixed quantities y_1 ; similarly, $y_{\text{int},2}$ is the sum of the first k fixed quantities y_2 . In like manner, we construct integral pairs $y_{\text{int},3}$, $y_{\text{int},4}$, \dots , $y_{\text{int},2s-1}$, $y_{\text{int},2s}$, \dots , and so forth. The sequence of these random variables is a sequence of inputs to the recurrent algorithm at the instants $T, 2T, \dots, rT, \dots$, where $T = k\tau$. The summation operator reduces the effect of random tracking errors on the accuracy of the estimates.

Choosing an integer s , we introduce the following definitions. Let us divide the sequence of pairs of integral observation results into a sequence of segments, each containing s pairs.

The iteration vector 1 is the estimate vector obtained after the 1st pair segment has been sent to the input of the recurrent algorithm. The iteration vector 1 is determined at the instant sT after sk observations of viewing angle pairs.

The iteration vector 2 is the estimate vector obtained after sending to the input the second segment of the pairs of integral observation results. The iteration vector 2 is determined at the instant $2sT$ after $2sk$ observations.

Iteration vectors 3, 4, and so forth, are determined in a similar manner.

For the construction of iteration vector 1, it is necessary to find the vector and the matrix of a priori data: the first and the second statistical moments of the basic random vector 1 whose first six components coincide with the components of the actual (unknown) initial state vector of the space object; its other components have the form of products of integer powers $2s$ of integral observation results; the sum of these powers does not exceed the given integer d . For $d = 1$, the dimension of the basic vector 1 is equal to $6 + 2s$; for $d = 2$, the dimension of the basic vector 1 is equal to $6 + 2s + (2s + 1)s$, and so forth.

A priori scattering of the vectors $X(0)$, $V(0)$ determines a priori dispersal of the other components of the basic vector 1 by means of (6.1) and (6.2). Next, we set $X(0) = X(0)_N + \delta X(0)$, $V(0) = V(0)_N + \delta V(0)$, where $X(0)_N$ is the nominal initial radius-vector of the observed space object and $V(0)_N$ is the nominal initial vector of its velocity. It is assumed

that $V(0)_N$ is the velocity vector of an object moving on a circular orbit of radius $|X(0)_N|$. Uniform scattering of the components of random vectors $\delta X(0)$, $\delta V(0)$ is determined by the inequalities

$$\begin{aligned} -\rho|X(0)| &\leq \delta X_1, \delta X_2, \delta X_3 \leq \rho|X(0)|, \\ -\rho|V(0)| &\leq \delta V_1, \delta V_2, \delta V_3 \leq \rho|V(0)|, \end{aligned} \quad (6.4)$$

where $\rho = 0.2$.

As the modeling process goes on, the errors in the determination of the initial coordinates reach several thousands km, and the errors in the determination of the initial velocity vector reach several hundreds m/sec. In this situation, the extended Kalman filter based on linearization is inapplicable for estimation, since it diverges at a fast rate.

A priori first and second statistical moments of the components of the basic vector 1 are determined by integrals over domains Ω_X , Ω_V , which are calculated by the Monte-Carlo method.

After a priori statistical characteristics of the basic vector 1 have been defined, $2s$ random variables are sent to the input of the recurrent algorithm, namely, $2s$ integral observation results for the viewing angles on the time interval sT . After $2s$ steps of calculations, the recurrent algorithm determines the iteration vector 1, which is the estimate vector $Z_{X(sT)}$, $Z_{V(sT)}$ for the state vector of the space object at the instant sT .

The accuracy of the estimate is described by the estimate error covariance matrix calculated on each step of the recurrent algorithm.

The determination of iteration vector 2 is similar to the above. Set $X(0) = Z_{X(sT)} + \delta X(0)$, $V(0) = Z_{V(sT)} + \delta V(0)$ for

$$\begin{aligned} -\rho 3\sigma_{x,i} &\leq \delta X_i \leq \rho 3\sigma_{x,i}, \\ -\rho 3\sigma_{v,i} &\leq \delta V_i \leq \rho 3\sigma_{v,i}, \end{aligned} \quad (6.5)$$

where $i = 1 \div 3$.

The components of a priori initial state vector correspond to a satellite on a circular orbit:

$$\begin{aligned} X_1(0) &= 26 \times 10^6 \text{ m}, & X_2(0) &= X_3(0) = 0, \\ V_1(0) &= 0, & V_2(0) &= 3.685 \times 10^3 \text{ m/sec}, & V_3(0) &= 0. \end{aligned} \quad (6.6)$$

The viewing system is defined by the relations $t = 0 : \lambda(0) = 0$, $\phi = 0$, $h = 0$ ($y_1(0) = y_2(0) = 0$).

The components of the actual initial state vector are determined by the inequalities (6.4) for $\rho = 0.2$, $|X(0)| = 26 \times 10^6$, $|V(0)| = 3.685 \times 10^3$.

In the process of simulation, the recurrent algorithm was operating under the conditions $\tau = 10$ seconds, $k = 20$, $s = 6$. These conditions mean that each 20 pairs of viewing angles y_1 , y_2 observed during the sequence of time-intervals of length $T = 200$ seconds are replaced by pairs $y_{\text{int},1}$, $y_{\text{int},2}$ of integral observation results. Iteration vectors 1, 2, 3, ... form the output of the algorithm at the instants $6T, 12T, 18T, \dots$

In the process of simulation, the method of polynomial approximation for the components of the estimate vector was used for $d = 1$ and $d = 2$. For $d = 1$, on each step of the recurrent algorithm, the components of the estimate vector are linear combinations of the last 12 components of the basic vector and consist of 12 integral observation results. For $d = 2$, on each step of the recurrent algorithm, the components of the estimate vectors are linear combinations of 90 terms, which coincide with the last 90 components of the basic vector and consist of products of powers 0, 1, 2 of 12 integral observation results.

Below, we give the values of $\sigma_{x,i}$ in meters and $\sigma_{v,i}$ in meters per second attained in the process of several iterations.

A priori MSDs $\sigma_{x,i}^a, \sigma_{v,i}^a$ of the components of the state vector of the space object at the instant $6T$ have the form

$$\begin{array}{cccccc} \sigma_{x,1}^a & \sigma_{x,2}^a & \sigma_{x,3}^a & \sigma_{v,1}^a & \sigma_{v,2}^a & \sigma_{v,3}^a \\ 3.76 \times 10^6 & 3.73 \times 10^6 & 3.67 \times 10^6 & 5.45 \times 10^2 & 5.39 \times 10^2 & 5.23 \times 10^2 \end{array} \quad (6.7)$$

Iteration 1, $d = 2$:

$$\begin{array}{cccccc} \sigma_{x,1} & \sigma_{x,2} & \sigma_{x,3} & \sigma_{v,1} & \sigma_{v,2} & \sigma_{v,3} \\ 5.00 \times 10^5 & 1.60 \times 10^5 & 7.89 \times 10^4 & 3.44 \times 10 & 6.55 \times 10 & 1.62 \times 10 \end{array} \quad (6.8)$$

Iteration 2, $d = 2$:

$$\begin{array}{cccccc} \sigma_{x,1} & \sigma_{x,2} & \sigma_{x,3} & \sigma_{v,1} & \sigma_{v,2} & \sigma_{v,3} \\ 6.57 \times 10^4 & 3.18 \times 10^4 & 9.84 \times 10^1 & 6.26 \times 10^{-1} & 8.34 & 8.94 \times 10^{-3} \end{array} \quad (6.9)$$

Iteration 3, $d = 2$:

$$\begin{array}{cccccc} \sigma_{x,1} & \sigma_{x,2} & \sigma_{x,3} & \sigma_{v,1} & \sigma_{v,2} & \sigma_{v,3} \\ 2.03 \times 10^2 & 1.39 \times 10^2 & 1.65 \times 10^{-3} & 8.71 \times 10^{-3} & 3.25 \times 10^{-2} & 1.53 \times 10^{-7} \end{array} \quad (6.10)$$

Iteration 4, $d = 1$:

$$\begin{array}{cccccc} \sigma_{x,1} & \sigma_{x,2} & \sigma_{x,3} & \sigma_{v,1} & \sigma_{v,2} & \sigma_{v,3} \\ 5.80 & 4.22 & 8.78 \times 10^{-9} & 4.43 \times 10^{-4} & 9.40 \times 10^{-4} & 3.88 \times 10^{-12} \end{array} \quad (6.11)$$

The above simulation data show that each iteration decreases, about 100 times, the MSD of the dispersal of the components of the state vector of the space object. By the end of iteration 4 ($t = 24T$), its coordinates and the components of the velocity vector are practically precise.

7. Identification of Parameters of a Non Linear Servomechanism

We consider a task of identification of parameters of the nonlinear servo-mechanism, which consists sequentially of an inertial link from a constant of time T , an amplifier with factor K_1 , an integrating link 1, a nonlinear link *a gap 1* (magnitude of the gap is c_1), an amplifier with factor K_2 , an integrating link 2, a nonlinear link *a gap 2* (magnitude of the gap is c_2), a transmitter of a feedback, which exit a variable x_5 . A variable U is an input of the servo-mechanism which should be repeated by an exit of the servo-mechanism—a variable x_4 .

The block diagram of the servo-mechanism is similar to the hydraulic booster which is used for turn steering surfaces of an aircraft [17].

The equations of driving of the servo-mechanism look like

$$\begin{aligned}\frac{dx_1}{dt} &= \left(\frac{1}{T}\right)(-x_1 + \Delta), \\ \frac{dx_2}{dt} &= K_1 x_1, \\ x_3 &= F(x_1, x_2, c_1), \\ \frac{dx_4}{dt} &= K_2 x_3, \\ x_5 &= F(x_3, x_4, c_2), \\ \Delta &= U - x_1 - x_2 - \alpha x_5,\end{aligned}\tag{7.1}$$

where $F(a, b, c) = b - c$, if $a > 0$, $F(a, b, c) = b + c$, if $a < 0$, and $\alpha = 0$, if the system is open loop, $\alpha = 1$ if the system is closed loop.

We suppose that $U(t) = \sin(2\pi t)$ —to the servo-mechanism set harmonic motions with frequency 1 hertz and amplitude 1, a priori data about gaps: $0 < c_1, c_2 < 0.1$, remaining a priori data: $T = T(n)(1 + \theta_3)$, $K_1 = K_1(n)(1 + \theta_4)$, $K_2 = K_2(n)(1 + \theta_5)$, where $T(n)$, $K_1(n)$, $K_2(n)$ are nominal (a priori) magnitudes, $\theta_i = \rho v_i$, $0 < \rho < 1$, $-1 < v_i < 1$, $i = 3, \dots, 5$. Further we suppose $\rho = 0.2$ (we know parameters with errors of not surpassing 20 percents).

Outcomes of observations are variables x_5 measure with frequency 1000 hertz: $y_1 = x_5(t_1) + \varepsilon_1, \dots, y_k = x_5(t_k) + \varepsilon_k, \dots, y_N = x_5(t_N) + \varepsilon_N$. They are used for identification of parameters. Random errors of observations ε_k are uniformly distributed on a segment $[-\varepsilon_0, \varepsilon_0]$. The sequence of observations decomposes on segments in length L . The sums of outcomes of observations, belonging each segment, are sequence of inputs in MPA algorithm, which estimates 5 parameters. Parameters c_1, c_2, T, K_1, K_2 have numbers $1, \dots, 5$.

The exactitude of identification is presented by magnitudes $\sigma(i)/\sigma_a(i)$, $i = 1, \dots, 5$, and the rations are certain in Section 4.

At $T(n) = 0.01$ seconds, $K_1(n) = 60$, $K_2(n) = 30$, $d = 1$, $L = 20$, $\alpha = 0$, $\varepsilon_0 = 0.0001$ and various times of observations t_N the mentioned rations are presented as follows:

$\sigma(i)/\sigma_a(i)$	1	2	3	4	5
$t_N = 2 \text{ sec}$	0.11073	0.00030	0.14257	0.12981	0.00091
$t_N = 4 \text{ sec}$	0.05632	0.00019	0.07860	0.06738	0.00012

(7.2)

8. Conclusion

The algorithm described above gives a sequence of convergent estimates for the vector of conditional expectation and the corresponding estimate error covariance matrices practically in any situation, provided that one knows the mathematical models of the system and observations necessary for leaning of the algorithm on the basis of the Monte-Carlo method.

It is convenient to compress data by replacing the high-dimensional vector of primary observations with a vector of a substantially smaller dimension, whose components are functions of primary observations. Such a replacement is likely to cause an insignificant increase of theoretical estimate error variance.

Thus, in Sections 5 and 6 data compression was implemented by passing from a 200-dimensional vector of primary observations to a 12-dimensional vector whose components are formed by sums of 20 consecutive components of the primary vector.

The development of rational methods for the compression of primary data is necessary for practical utilization of the proposed general algorithm, which is a unique example of a situation in which the Mote-Carlo method provides synthesis of a new efficient method of nonlinear estimation.

Acknowledgment

This work has been supported by the Russian Foundation for Basic Research.

References

- [1] O. Cappé, E. Moulines, and T. Rydén, *Inference in Hidden Markov Models*, Springer Series in Statistics, Springer, New York, NY, USA, 2005.
- [2] V. Klein and A. G. Morelli, *Aircraft System Identification: Theory and Methods*, AIAA, 2006.
- [3] L. Lenart, *System Identification: Theory for the USER*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1987.
- [4] M. I. Ribeiro, *Kalman and Extended Kalman Filter: Concept, Derivation and Propertis*, Institute foe Systems and Robotics Instituto Superior, 2004.
- [5] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-gaussian Bayesian state estimation," *IEE Proceedings, Part F*, vol. 140, no. 2, pp. 107–113, 1993.
- [6] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.
- [7] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science, Springer, New York, NY, USA, 2001.
- [8] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, Boston, Mass, USA, 2004.
- [9] S. J. Ghosh, C. S. Manohar, and D. Roy, "A sequential importance sampling filter with a new proposal distribution for state and parameter estimation of nonlinear dynamical systems," *Proceedings of The Royal Society of London. Series A*, vol. 464, no. 2089, pp. 25–47, 2008.
- [10] V. Namdeo and C. S. Manohar, "Nonlinear structural dynamical system identification using adaptive particle filters," *Journal of Sound and Vibration*, vol. 306, no. 3-5, pp. 524–563, 2007.
- [11] J. A. Boguslavskiy, "A Bayes estimations of nonlinear regression and adjacent problems," *Journal of Computer and Systems Sciences International*, vol. 4, pp. 14–24, 1996.
- [12] J. A. Boguslavskiy, "Integrated method of the numerical decision of the algebraic equations," *Applied Mathematics and Computation*, vol. 166, no. 2, pp. 324–338, 2005.
- [13] J. A. Boguslavskiy, *Polynomial Approximations for Nonlinear Problems of Estimation and Control*, MAIK, Fizmat, Russia, 2006.
- [14] I. A. Boguslavskiy, "Method for the non-linear identification of aircraft parameters by testing maneuvers," in *Proceedings of the International Conference on Numerical Analysis and Applied Mathematics (AIP '08)*, vol. 1048, pp. 92–99, 2008.

- [15] M. H. Stone, "Applications of the theory of Boolean rings to general topology," *Transactions of the American Mathematical Society*, vol. 41, no. 3, pp. 375–481, 1937.
- [16] K. Lantratov, "Optiko-elektronnyj a complex of monitoring of a space," "WINDOW" News cosmic, 2002.
- [17] S. V. Konstantinov, P. G. Redko, and S. A. Ermakov, *Electric-Hydraulic Steering Drives of Control Systems of Flight*, Janus-K, Moscow, Russia, 2006.