UDC 519.25:681.5

O.P. Sarychev, B.A. Perviy

# OPTIMAL REGRESSORS SEARCH SUBJECTED TO VECTOR AUTOREGRESSION OF UNEVENLY SPACED TLE SERIES

*An iterative procedure of the parametric identification of autoregressive models with unequally spaced observations has been developed. The task of the Sich-2 spacecraft dynamics modeling using its unequally spaced TLE elements is considered. For all elements, satisfactory quality models were obtained.*

## Introduction

The problem of improving the accuracy of forecasting the satellite position is relevant for the tasks of determining time of their existence, cataloging space debris, navigation, etc.

The only open source of orbital data for solving such problem are the two-line elements (TLE), which regularly and promptly are updated on the website of the American Space Monitoring System (SSS) [1]. The values of the orbit parameters contained in the TLE files are calculated by averaging over specific SGP4 or SDP4 models [2].

Prediction techniques based on statistical models of time series [3, 4], or on methods of machine learning [5, 6], are aimed at modeling and reproducing the missing dynamics of previously calculated approximations of SGP4 or SDP4 models. This combination improves the accuracy of conventional numerical, analytical, and semi-analytical methods for determining the position and velocity of any satellite or space debris object.

A distinctive feature of TLE-elements series is their time positioning not on a uniform temporal grid, but with irregular time intervals between observations, so-called "unequal observations". When solving problems of modeling time series with non-equal observations, they usually try to move to a uniform grid based on various local approximation procedures on a sliding interval [7]. In this way, the problem of choosing the type and optimal order of the smoothing polynomial arises, and so, in fact another replaces the previous task. When building statistical models of TLE elements series, this distinctive feature can be used to modify parameter estimation procedures.

The development of a method for constructing autoregressive models with unequally spaced time observations and its application in modeling the dynamics of large fragments of space debris in the problem of their removal from orbits is the goal of this work.

**Major part**

**1 A priori assumptions about a dynamic object represented by TLE elements series**

Assume the functioning of a dynamic object obeys the law in the form of an autoregressive equation

$$\overset{*}{x}_i = \left(\overset{*}{x}_{i-1}, \overset{*}{x}_{i-2}, ..., \overset{*}{x}_{i-p}\right)\left(\overset{o}{\theta}_{1i}, \overset{o}{\theta}_{2i}, ..., \overset{o}{\theta}_{pi}\right)^{\mathrm{T}} + \zeta_{i-1} = \overset{*}{\mathbf{Z}}_{i,\bullet}(p)\overset{o}{\mathbf{\Theta}}_{\bullet,i}(p) + \zeta_{i-1}, \quad (1)$$

where $\overset{*}{x}_i$ is unobserved value of the output variable of the object at discrete time $t = t_i$, $i = 1, 2, ..., n$; $n$ is total number of observations; $p$ is the number of previous values of the output variable, which affects its current value; $\zeta_{i-1}$ is unobserved random variable.

In the model of functioning process (1), the $(n \times p)$-matrix $\overset{*}{\mathbf{Z}}(p)$ is the matrix $p$ of the previous unobserved values of the variable:

$$\overset{*}{\mathbf{Z}}(p) = \begin{bmatrix} \overset{*}{x}_0 & \overset{*}{x}_{-1} & \cdots & \overset{*}{x}_{1-p} \\ \overset{*}{x}_1 & \overset{*}{x}_0 & \cdots & \overset{*}{x}_{2-p} \\ \vdots & \vdots & \ddots & \vdots \\ \overset{*}{x}_{i-1} & \overset{*}{x}_{i-2} & \cdots & \overset{*}{x}_{i-p} \\ \vdots & \vdots & \ddots & \vdots \\ \overset{*}{x}_{n-1} & \overset{*}{x}_{n-2} & \cdots & \overset{*}{x}_{n-p} \end{bmatrix} = \begin{bmatrix} \overset{*}{\mathbf{Z}}_{1,\bullet}(p) \\ \hline \overset{*}{\mathbf{Z}}_{2,\bullet}(p) \\ \hline \vdots \\ \hline \overset{*}{\mathbf{Z}}_{i,\bullet}(p) \\ \hline \vdots \\ \hline \overset{*}{\mathbf{Z}}_{n,\bullet}(p) \end{bmatrix}, \quad (2)$$

where, $\overset{*}{\mathbf{Z}}_{i,\bullet}(p)$ represents the row $\overset{*}{\mathbf{Z}}(p)$, the first element of which is $\overset{*}{x}_{i-1}$; in the designation of this matrix, " $p$ " means that the formation of a quantity $\overset{*}{x}_i$

involves $p$ quantities of $\overset{*}{x}_{i-1}, \overset{*}{x}_{i-2}, \ldots, \overset{*}{x}_{i-p}$, i.e., the $i$-th $(p \times 1)$-matrix row is multiplied by the $i$-th $(p \times 1)$-matrix column $\overset{\mathrm{o}}{\boldsymbol{\Theta}}(p)$.

Assume $\tau_i = t_i - t_{i-1}$, $i = 1, 2, \ldots, n$ are the values of the time intervals between adjacent pairs of observations.

Suppose $(p \times 1)$-vector of coeffients $\overset{\mathrm{o}}{\Theta}_{\bullet,i}(p)$ fulfill the equality:

$$\overset{\mathrm{o}}{\boldsymbol{\Theta}}_{\bullet,i}(p) = \left( \overset{\mathrm{o}}{\theta}_{1i}, \overset{\mathrm{o}}{\theta}_{2i}, \ldots, \overset{\mathrm{o}}{\theta}_{pi} \right)^{\mathrm{T}} = \left( \theta_1^{\mu_i}, \theta_2^{\mu_i}, \ldots, \theta_p^{\mu_i} \right)^{\mathrm{T}}, \quad i = 1, 2, \ldots, n \tag{3}$$

where

$$\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)^{\mathrm{T}} \tag{4}$$

is $(p \times 1)$-vector of unknown deterministic coefficients that must be determined by the results of the object; $\mu_i = \tau_i / \delta_t$ is the exponent in which the vector components $\boldsymbol{\theta}$ are raised; $\delta_t$ is some given value.

For example, if we take $\delta_t$ equal to the average value of the time intervals between adjacent pairs of observations:

$$\delta_t = \frac{1}{n} \sum_{i=1}^{n} \tau_i = \frac{1}{n}(t_n - t_0), \tag{5}$$

then this value will correspond to the time interval of observations in the case of equidistant observations, i.e. $\mu_i = 1$, $i = 1, 2, \ldots, n$

Taking into account (2)–(5) the law of functioning (1), according to which the output variable is formed, is written in the form

$$\overset{*}{x}_i = \overset{*}{\mathbf{Z}}_{i,\bullet}(p) \left( \theta_1^{\mu_i}, \theta_2^{\mu_i}, \ldots, \theta_p^{\mu_i} \right)^{\mathrm{T}} + \zeta_{i-1}, \quad i = 1, 2, \ldots, n, \tag{6}$$

where $\boldsymbol{\theta}$ is $(p \times 1)$-the vector of unknown deterministic coefficients.

Introduce the notation

$$\overset{=}{x}_i = \overset{*}{\mathbf{Z}}_{i,\bullet}(p) \left( \theta_1^{\mu_i}, \theta_2^{\mu_i}, \ldots, \theta_p^{\mu_i} \right)^{\mathrm{T}}; \tag{7}$$

given (7), then write (6)

$$\overset{*}{x}_i = \overset{=}{x}_i + \zeta_{i-1}, \; i = 1, 2, \ldots, n. \tag{8}$$

Add another notation

$$\overset{*}{\mathbf{x}} = (\overset{*}{x}_{i-1}, \overset{*}{x}_{i-2}, \ldots, \overset{*}{x}_{i-p})^{\mathrm{T}}, \quad \overset{=}{\mathbf{x}} = (\overset{=}{x}_{i-1}, \overset{=}{x}_{i-2}, \ldots, \overset{=}{x}_{i-p})^{\mathrm{T}}, \tag{9}$$

$$\zeta(-1) = \left(\zeta_0, \zeta_1, \ldots, \zeta_{n-1}\right)^{\mathrm{T}}, \tag{10}$$

where $\boldsymbol{\zeta}(-1)$ is the unobserved random $(n \times 1)$-vector; "$-1$" means that the quantity $\zeta(-1) = \zeta_{i-1}$ in (1) and (8) additively participates in the formation of the quantity $\overset{*}{x}_i$.

Considering (9)–(10), we write the model of functioning in the vector form

$$\overset{*}{\mathbf{x}} = \overset{=}{\mathbf{x}} + \boldsymbol{\zeta}(-1), \tag{11}$$

where $\overset{=}{\mathbf{x}}$ is the unobservable component of the $(n \times 1)$-vector of the variable values.

For observations of the output variable of the object, the next equation is introduced:

$$x_i = \overset{*}{x}_i + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{12}$$

where $x_i$ is the observed value of the variable, measured at the time $t = t_i$, $i = 1, 2, \ldots, n$; $\overset{*}{x}_i$ is unobserved value, which is formed according to (1), (6) and (7); $\varepsilon_i$ is random unobserved measurement error.

Given (12), we write the model of the observation of the object in a vector form

$$\mathbf{x} = \overset{*}{\mathbf{x}} + \boldsymbol{\varepsilon}. \tag{13}$$

We formulate assumptions about the statistical properties of random variables in the models of operation and observation. Let the following assumptions be fulfilled with $\zeta(-1)$ respect to (11):

$$E\{\boldsymbol{\zeta}(-1)\} = \mathbf{0}_n, \quad E\{\boldsymbol{\zeta}(-1)\boldsymbol{\zeta}^{\mathrm{T}}(-1)\} = \sigma_\zeta \cdot \mathbf{I}_n, \tag{14}$$

where $E\{\cdot\}$ is the sign of the expectation of possible implementations of the vector $\boldsymbol{\zeta}(-1)$; $\mathbf{0}_n$ is zero $(n \times 1)$-vector; $\sigma_\zeta$ is variance of the random variable $\zeta_i(-1)$, $i = 1,2,\ldots,n$, limited value; $\mathbf{I}_n$ is the unit $(n \times n)$-matrix.

Let the following assumptions be made with respect to $\varepsilon$ (13):

$$E\{\boldsymbol{\varepsilon}\} = \mathbf{0}_n, \quad E\{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\} = \sigma_\varepsilon \cdot \mathbf{I}_n, \tag{15}$$

 ISSN 1562-9945

where $E\{\cdot\}$ is the sign of the expectation of possible implementations of the vector $\varepsilon$; $\sigma_\varepsilon$ is variance of magnitude $\varepsilon_i$, $i = 1, 2, \dots, n$ limited value.

We will also assume that random vectors $\zeta(-1)$ and $\varepsilon$ are statistically independent:

$$E\{\zeta(-1)\boldsymbol{\varepsilon}^{\mathrm{T}}\} = \mathbf{O}_{(n\times n)}, \tag{16}$$

where $\mathbf{O}_{(n\times n)}$ is the zero $(n\times n)$-matrix.

Presume that at $t = t_i$, $i = 1-2p,\ 2-2p, \dots, 0, 1, 2, \dots, n$ time points, the $(n+2p)$-vector of observations of the output variable was obtained

$$\left(x_{1-2p}, x_{2-2p}, \dots, x_0, x_1, x_2, \dots, x_n\right)^{\mathrm{T}} = \begin{pmatrix} \mathbf{x}(0) \\ \hdashline \mathbf{x} \end{pmatrix}, \tag{17}$$

where the $(2p\times 1)$-vector $\mathbf{x}(0)$ will be used as the initial conditions.

To estimate the unknown coefficients $\boldsymbol{\theta}$ from the observations of the object (17), we use the results of [8]–[10], where an iterative parametric identification procedure was developed and investigated for models in the class of autoregressive equation systems.

**2 Estimation of coefficients in autoregressive equations subjected to unequally spaced observations**

From the model of functioning (7) and the generalized form (11) it follows

$$\overset{*}{\mathbf{Z}}(p) = \overset{=}{\mathbf{Z}}(p) + \boldsymbol{\Gamma}(-2;Z), \tag{18}$$

where $\overset{=}{\mathbf{Z}}(p)$ is the $(n\times p)$-matrix of unobserved values of the output variable of the object, its structure is similar to the matrix $\overset{*}{\mathbf{Z}}(p)$ in (1)–(2):

$$\overset{=}{\mathbf{Z}}(p) = \begin{bmatrix} \overset{=}{x}_0 & \overset{=}{x}_{-1} & \cdots & \overset{=}{x}_{1-p} \\ \overset{=}{x}_1 & \overset{=}{x}_0 & \cdots & \overset{=}{x}_{2-p} \\ \vdots & \vdots & \ddots & \vdots \\ \overset{=}{x}_{i-1} & \overset{=}{x}_{i-2} & \cdots & \overset{=}{x}_{i-p} \\ \vdots & \vdots & \ddots & \vdots \\ \overset{=}{x}_{n-1} & \overset{=}{x}_{n-2} & \cdots & \overset{=}{x}_{n-p} \end{bmatrix}; \tag{19}$$

$\boldsymbol{\Gamma}(-2;Z)$ is matrix of unobserved random variables

$$\mathbf{\Gamma}(-2;Z)=\begin{bmatrix}\zeta_{-1} & \zeta_{-2} & \cdots & \zeta_{-p}\\ \zeta_{0} & \zeta_{-1} & \cdots & \zeta_{1-p}\\ \vdots & \vdots & \ddots & \vdots\\ \zeta_{i-2} & \zeta_{i-3} & \cdots & \zeta_{i-1-p}\\ \vdots & \vdots & \ddots & \vdots\\ \zeta_{n-2} & \zeta_{n-3} & \cdots & \zeta_{n-1-p}\end{bmatrix}, \quad (20)$$

in the designation of which "−2" means that in (18) the value $\zeta_{i-2}$ is additively involved in the formation of the value $\overset{*}{x}_{i-1}$.

Substitute in (13) the vector $\overset{*}{\mathbf{x}}$ of (11) and use (18) for $\overset{*}{\mathbf{Z}}(p)$:

$$x_i=\overline{\overline{\mathbf{Z}}}_{i,\bullet}(p)\left(\theta_1^{\mu_i},\theta_2^{\mu_i},...,\theta_p^{\mu_i}\right)^{\mathrm{T}}+\left\{\varepsilon_i+\mathbf{\Gamma}_{i,\bullet}(-2;Z)\left(\theta_1^{\mu_i},\theta_2^{\mu_i},...,\theta_p^{\mu_i}\right)^{\mathrm{T}}+\zeta_i(-1)\right\} \quad (21)$$

or

$$x_i=\overline{\overline{\mathbf{Z}}}_{i,\bullet}(p)\left(\theta_1^{\mu_i},\theta_2^{\mu_i},...,\theta_p^{\mu_i}\right)^{\mathrm{T}}+\xi_i,\quad i=1,2,\dots,n, \quad (22)$$

where $\xi_i$ is a random variable enclosed in braces in (21).

Using (18)–(22) and given that the random vectors $\boldsymbol{\varepsilon}(k)$, $\boldsymbol{\zeta}(-1)$ and the random matrix $\mathbf{\Gamma}(-2;Z)$ have zero mathematical expectations, and all these values are statistically independent, for the mathematical expectation $\boldsymbol{\xi}=(\xi_i,\xi_2,...\xi_n)^{\mathrm{T}}$ we get

$$E\{\boldsymbol{\xi}\}=\mathbf{0}_n, \quad (23)$$

where $\mathbf{0}_n$ is zero $(n\times1)$ - vector.

Itroduce the notation

$$\mathbf{y}=\mathbf{x},\quad \mathbf{R}=\overline{\overline{\mathbf{Z}}}(p), \quad (24)$$

where $\mathbf{y}$ is the $(n\times1)$-vector of observations of the output variable; $\mathbf{R}$ is the regressor $(n\times p)$-matrix for the output variable.

Given (24), model (22) can be written


ISSN 1562-9945

$$y_i = \mathbf{R}_{i,\bullet} \begin{pmatrix} \theta_1^{\mu_i} \\ \theta_2^{\mu_i} \\ \vdots \\ \theta_p^{\mu_i} \end{pmatrix} + \xi_i = \mathbf{R}_{i,\bullet}\, \boldsymbol{\theta}^{(\mu_i)} + \xi_i = \mathbf{R}_{i,\bullet}\, \overset{\circ}{\boldsymbol{\theta}}_i + \xi_i = \overset{\circ}{y}_i + \xi_i ,$$

$i = 1, 2, \ldots, n$, (25)

or in vector form

$$\mathbf{y} = \overset{\circ}{\mathbf{y}} + \boldsymbol{\xi} , \tag{26}$$

where $\boldsymbol{\theta}^{(\mu_i)}$ is the designation for component-wise exponentiation $\mu_i$ of all components of the $(p \times 1)$-vector $\boldsymbol{\theta}$.

Finding the $(p \times 1)$-vector of unknown deterministic coefficients $\boldsymbol{\theta}$ is a difficult task due to the fact that in model (25) the components of this vector are included in the degree depending on the time intervals between adjacent pairs of observations, which are different for different observation numbers.

We write (25) in a form convenient for constructing an iterative procedure:

$$y_i = \mathbf{R}_{i,\bullet} \left(\theta_1^{\mu_i}, \theta_2^{\mu_i}, \ldots, \theta_p^{\mu_i}\right)^{\mathrm{T}} + \xi_i = \sum_{j=1}^{p} \mathbf{R}_{i,j} \theta_j^{\mu_i} + \xi_i = \sum_{j=1}^{p} \mathbf{R}_{i,j} \theta_j^{\mu_i - 1} \theta_j + \xi_i =$$

$$= \sum_{j=1}^{p} \underline{\underline{\mathbf{R}}}_{i,j} \theta_j + \xi_i = \underline{\underline{\mathbf{R}}}_{i,\bullet} \left(\theta_1, \theta_2, \ldots, \theta_p\right)^{\mathrm{T}} + \xi_i, \quad i = 1, 2, \ldots, n , \tag{27}$$

where $\underline{\underline{\mathbf{R}}}_{i,j} = \mathbf{R}_{i,j} \theta_j^{\mu_i - 1}$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, p$ is the matrix of new regressors.

We write (27) in vector form

$$\mathbf{y} = \underline{\underline{\mathbf{R}}}\, \boldsymbol{\theta} = \overset{\circ}{\mathbf{y}} + \boldsymbol{\xi} , \tag{28}$$

where $\boldsymbol{\theta}$ is the $(p \times 1)$-vector of unknown deterministic coefficients that must be determined.

According to [10], to estimate the coefficients $\boldsymbol{\theta}$ in (28), the following is fair:

$$\hat{\mathbf{d}} = \mathbf{C}\ \mathbf{y} , \tag{29}$$

where for $(p \times n)$-matrix $\mathbf{C}$, equality is fulfilled

$$\mathbf{C} = (\underline{\underline{\mathbf{R}}}^{\mathrm{T}} \boldsymbol{\Sigma}_{\xi}^{-1} \underline{\underline{\mathbf{R}}})^{-1} \underline{\underline{\mathbf{R}}}^{\mathrm{T}} \boldsymbol{\Sigma}_{\xi}^{-1} , \tag{30}$$

and $\mathbf{\Sigma}_{\xi}$ is covariance $(n \times n)$-matrix, mentioned in (22) $(n \times 1)$-vector of unobserved additive random components $\boldsymbol{\xi}$.

For the covariance matrix $\mathbf{\Sigma}_{\xi}$, the equality is fulfilled [10]:

$$\mathbf{\Sigma}_{\xi} = \sigma_{\varepsilon} \cdot \mathbf{I}_n + \mathbf{\Psi} + \sigma_{\zeta} \cdot \mathbf{I}_n , \tag{31}$$

where $\sigma_{\varepsilon}$ is the variance in the observation model introduced in (15); $\sigma_{\zeta}$ is variance in the functioning model introduced in (14); $\mathbf{\Psi}$ is $(n \times n)$-matrix has the form

$$\mathbf{\Psi} = \begin{bmatrix} \psi(0) & \psi(+1) & \cdots & \psi(p-1) & 0 & \cdots & 0 & 0 \\ \psi(-1) & \psi(0) & \cdots & \psi(p-2) & \psi(p-1) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \psi(1-p) & \psi(2-p) & \cdots & \psi(0) & \psi(+1) & \cdots & 0 & 0 \\ 0 & \psi(1-p) & \cdots & \psi(-1) & \psi(0) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \psi(0) & \psi(+1) \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \psi(-1) & \psi(0) \end{bmatrix}. \tag{32}$$

In (32), the quantities $\psi(\Delta)$, $\Delta = -p+1, -p+2, \ldots, p-2, p-1$ are determined by the formulas

$$\psi(\Delta) = \mathrm{Cov}\{\xi_{i_1} \xi_{i_2}\} = \sigma_{\zeta} \cdot \boldsymbol{\theta}^{\mathrm{T}} \mathbf{I}(i_1 - i_2) \boldsymbol{\theta} , \tag{33}$$

where $\mathbf{I}_p(i_1 - i_2)$ is a $(p \times p)$-matrix, in which all elements are equal to zero, except elements of a diagonal equal to one: if, $\Delta = i_1 - i_2 = 0$ then this is the main diagonal; if, $\Delta > 0$ then this is the diagonal located above the main diagonal on the $\Delta$ lines; if, $\Delta < 0$ then this is the diagonal located below the main diagonal the $|\Delta|$ rows apart.

Taking into account (30)–(33) for the estimates of the coefficients,

$$\hat{\mathbf{d}} = (\underline{\underline{\mathbf{R}}}^{\mathrm{T}} \mathbf{\Sigma}_{\xi}^{-1} \underline{\underline{\mathbf{R}}})^{-1} \underline{\underline{\mathbf{R}}}^{\mathrm{T}} \mathbf{\Sigma}_{\xi}^{-1} \mathbf{y} . \tag{34}$$

In (30) for the matrix $\mathbf{C}$, the regressor matrix $\underline{\underline{\mathbf{R}}}$ depends on the unobservable matrix $\mathbf{R}$ and on the unknown coefficient vector $\boldsymbol{\theta}$. In (31) for the matrix $\mathbf{\Sigma}_{\xi}$, the elements of the matrix $\mathbf{\Psi}$, as follows from (32) - (33), depend on $\boldsymbol{\theta}$. This was used in [10] to form an iterative procedure for calculating the unknown coefficients in the form (34).

 ISSN 1562-9945

**3 An iterative procedure for estimating coefficients in autoregressive equations subjected to unequally spaced observations**

Let $\hat{\mathbf{d}}(r)$ be the estimate $\boldsymbol{\theta}$ in the form of (34) obtained at the iteration $r$ of the coefficient vector $\boldsymbol{\theta}$ estimation procedure; matrix $\hat{\mathbf{R}}(r-1)$ is the estimate of the regressor matrix $\mathbf{R}$ obtained at the iteration $r-1$; $\hat{y}_i(r)$, $i=1,2,\dots,n$ is $(n\times 1)$-vector of the outputs of the regression model; $u_i(r)$, $i=1,2,\dots,n$ is $(n\times 1)$-vector of residuals of the regression model [11]. We write the regression model (27) in the form convenient for implementing an iterative estimation procedure.

$$y_i = \sum_{j=1}^{p} \hat{\mathbf{R}}_{i,j}(r-1)\,\hat{d}_j^{\,\mu_i-1}(r-1)\,\hat{d}_j(r) + u_i(r) = \sum_{j=1}^{p} \underline{\underline{\mathbf{R}}}_{i,j}(r-1)\,\hat{d}_j(r) + u_i(r) =$$
$$= \underline{\underline{\mathbf{R}}}_{i,\bullet}(r-1)\left(\hat{d}_1(r), \hat{d}_2(r), \dots, \hat{d}_p(r)\right)^{\mathrm{T}} + u_i(r) = \hat{y}_i(r) + u_i(r),\ i=1,2,\dots,n \tag{35}$$

where $\underline{\underline{\mathbf{R}}}_{i,j}(r-1) = \hat{\mathbf{R}}_{i,j}(r-1)\,\hat{d}_j^{\,\mu_i-1}(r-1),\ i=1,2,\dots,n,\ j=1,2,\dots,p$.

The iterative procedure for estimating the unknown coefficients of the regression model in the form (25)–(27) involves three steps.

Stage I. Initial approximation.

Step 1. Form the matrix of observed previous values of the variable (similar to (19))

$$\mathbf{Z}(p;0) = \begin{bmatrix} x_0 & x_{-1} & \cdots & x_{1-p} \\ x_1 & x_0 & \cdots & x_{2-p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i-1} & x_{i-2} & \cdots & x_{i-p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & \cdots & x_{n-p} \end{bmatrix}. \tag{36}$$

Step 2. Form a matrix of observed regressors (similar to (24))

$$\hat{\mathbf{R}}(0) = \mathbf{Z}(p;0)\,. \tag{37}$$

Step 3. We form the matrix of new auxiliary regressors in accordance with (35), assuming $\mu_i = 1,\ i=1,2,\dots,n$:

$$\underline{\underline{\mathbf{R}}}(0) = \hat{\mathbf{R}}(0). \tag{38}$$

Step 4. We assume $\hat{\mathbf{\Psi}}(0) = \mathbf{O}_{n\times n}$ is zero $(n \times n)$-matrix.

Step 5. Calculate the estimate of the coefficients:

$$\hat{\mathbf{d}}(0) = \left([\underline{\underline{\mathbf{R}}}(0)]^{\mathrm{T}}[\sigma_\varepsilon \cdot \mathbf{I}_n + \sigma_\zeta \cdot \mathbf{I}_n]^{-1}\underline{\underline{\mathbf{R}}}(0)\right)^{-1}[\underline{\underline{\mathbf{R}}}(0)]^{\mathrm{T}}[\sigma_\varepsilon \cdot \mathbf{I}_n + \sigma_\zeta \cdot \mathbf{I}_n]^{-1}\mathbf{y}. \tag{39}$$

Step 6. Calculate the model outputs:

$$\hat{\mathbf{y}}(0) = \underline{\underline{\mathbf{R}}}(0)\hat{\mathbf{d}}(0). \tag{40}$$

Step 7. Calculate model residuals:

$$\mathbf{u}(0) = \mathbf{y} - \hat{\mathbf{y}}(0). \tag{41}$$

Step 8. Calculate the target functional:

$$\Phi(0) = ((n-1)^{-1} \cdot \mathbf{u}^{\mathrm{T}}(0)\mathbf{u}(0))^{1/2}. \tag{42}$$

Stage II. The main stage. At iterations $r = 1,2,\ldots,r^*$ the operations are performed:

Step 1. Form the matrix of estimates of the previous values of the variable (similar to (19))

$$\hat{\mathbf{Z}}(p;r-1) = \begin{bmatrix} \hat{y}_0(r-1) & \hat{y}_{-1}(r-1) & \cdots & \hat{y}_{1-p}(r-1) \\ \hat{y}_1(r-1) & \hat{y}_0(r-1) & \cdots & \hat{y}_{2-p}(r-1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{i-1}(r-1) & \hat{y}_{i-2}(r-1) & \cdots & \hat{y}_{i-p}(r-1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_{n-1}(r-1) & \hat{y}_{n-2}(r-1) & \cdots & \hat{y}_{n-p}(r-1) \end{bmatrix}. \tag{43}$$

Step 2. Form the regressor matrix (similar to (24))

$$\hat{\mathbf{R}}(r-1) = \hat{\mathbf{Z}}(p;r-1). \tag{44}$$

Step 3. Form the matrix of new auxiliary regressors in accordance with (35)

$$\underline{\underline{\mathbf{R}}}_{i,j}(r-1) = \hat{\mathbf{R}}_{i,j}(r-1)\,\hat{d}_j^{\,\mu_i - 1}(r-1),\ i = 1,2,\ldots,n,\ j = 1,2,\ldots,p. \tag{45}$$

Step 4. Calculate the matrix $\hat{\mathbf{\Psi}}(r-1)$ - values $\hat{\psi}(\Delta;r-1)$, $\Delta = -p+1,-p+2,\ldots,-1,0,1,\ldots,p-2,p-1$ calculated by (32)–(33), using as an estimate $\boldsymbol{\theta}$ the approximation $\hat{\mathbf{d}}(r-1)$ obtained at the iteration $r-1$; at iteration $r=1$, we use the initial approximation estimate $\hat{\mathbf{d}}(0)$, obtained in (39).

Step 5. Calculate the estimate of the coefficients:

$$\hat{\mathbf{d}}(r) = \left( [\underline{\underline{\mathbf{R}}}(r-1)]^{\mathrm{T}} [\sigma_{\varepsilon} \cdot \mathbf{I}_n + \hat{\mathbf{\Psi}}(r-1) + \sigma_{\zeta} \cdot \mathbf{I}_n]^{-1} \underline{\underline{\mathbf{R}}}(r-1) \right)^{-1} \times$$

$$\times [\underline{\underline{\mathbf{R}}}(r-1)]^{\mathrm{T}} [\sigma_{\varepsilon} \cdot \mathbf{I}_n + \hat{\mathbf{\Psi}}(r-1) + \sigma_{\zeta} \cdot \mathbf{I}_n]^{-1} \mathbf{y} . \quad (46)$$

Step 6. Calculate the outputs of the models:

$$\hat{\mathbf{y}}(r) = \underline{\underline{\mathbf{R}}}(r-1)\hat{\mathbf{d}}(r) . \quad (47)$$

Step 7. Calculate the residuals of the models:

$$\mathbf{u}(r) = \mathbf{y} - \hat{\mathbf{y}}(r) . \quad (48)$$

Step 8. Calculate the target functional:

$$\Phi(r) = ((n-1)^{-1} \cdot \mathbf{u}^{\mathrm{T}}(r)\mathbf{u}(r))^{1/2} . \quad (49)$$

Stage III. Breakpoint. The iteration process ends at an iteration $r^*$ if the condition

$$\delta = \Phi(r^* - 1) - \Phi(r^*) \le \delta_0 , \quad (50)$$

where $\delta_0$ is the given value.

Note that in the case of first-order autoregression $(p=1)$, the matrix $\mathbf{\Sigma}_{\xi}$ in (31) has the form $\mathbf{\Sigma}_{\xi} = \sigma \cdot \mathbf{I}_n$, where, $\sigma = \sigma_{\varepsilon} + \sigma_{\zeta} \cdot \theta_1^2 + \sigma_{\zeta}$ that is, the estimation in the fifth steps of the first and second stages of the procedure is the usual least squares method.

If the dispersions $\sigma_{\zeta}$ and $\sigma_{\varepsilon}$ are a priori unknown, then the covariance matrix $\mathbf{\Sigma}_{\xi}$ can be estimated (taking into account its structure) iteratively using the residuals of the autoregressive model, as was done in the procedure [9]. In this case, the estimates of the usual MLS are taken as the initial approximation.

Note also that the considered problem of estimating autoregressive coefficients subjected to unequally spaced observations can be posed and solved under conditions of structural uncertainty, when the autoregressive order is a priori unknown. In this case, the evaluation of autoregression coefficients can be carried out based on the results of [12]–[13].

**4 TLE time series simulation for large space debris objects using the Ukrainian Sich-2 spacecraft**

The developed iterative procedure for estimating coefficients in autoregressive equations under unequal observation conditions was applied to simulate the TLE time series of the Sich-2 spacecraft [1]. Sich-2 is a Ukrainian small-sized remote sensing spacecraft (SC) operating from 2011–2012. It was intended to observe the surface of the Earth in the optical and mid-infrared ranges. In December 2012, communication with the Sich-2 spacecraft was lost. The TLE time series of the Sich-2 satellite are represented by seven main and three additional variables: (see Table 1).

Table 1

List of variables for Sich-2 TLE data

| Designation | Title | Unit of measurement |
|---|---|---|
| $x_1$ | Apogee | km |
| $x_2$ | Perigee | km |
| $x_3$ | Eccentricity | - |
| $x_4$ | Inclination | deg |
| $x_5$ | Right ascension of the ascending node | deg |
| $x_6$ | Argument of perigee | deg |
| $x_7$ | Mean anomaly | deg |
| $x_8$ | Revolution number at epoch | revs |
| $x_9$, $t_{nak}$ | Accumulated time | hrs |
| $x_{10}$, $\tau_i$ | The time interval between the current and the previous observation | hrs |

Simulations were performed for seven key variables. Additional variables $x_8$, $x_9$ were used to build the figures, and the variable $x_{10}$ was used for the calculation $\mu_i = \tau_i / \delta_t$, $i = 1,2,\ldots,n$, where $\mu_i$ is the exponent in which the vector $\boldsymbol{\theta}$ components are raised in (4). Attempts to model directly in the class of autoregressive models (1) were unsuccessful due to the strong correlation of autoregressors (correlations between the columns of the matrix (2)). The

 ISSN 1562-9945

way out of this situation is possible with the help of a priori data that contains information about the ratio between the coefficients of the autoregressors.

In [14] a method of structural-parametric identification was developed in the problem of modeling objects by observing their functioning in the class of beta-autoregressive equations, in which such weighting ratios for autoregressors in the autoregressive model are determined based on the density functions of beta-distributions. These results were used to solve this problem of simulating TLE-elements of the Sich-2 satellite.

By trial calculations based on the developed iterative procedure and the method of structural-parametric identification in the class of beta-autoregressive models [14], it was found that the maximum number of previous values of output variables $x_1, x_2, \dots, x_7$ affecting their current value is sufficient to be equal $p = 7$.

At the first preliminary stage of modeling based on the beta-autoregression method, the search for the optimal parameter value was carried out for each main variable $\beta$. This parameter indicates the degree of increase in the weight coefficients of the autoregression when approaching the current value. Having $\beta = 1$, all weights of autoregression are same: $a_j = 1$, $j = 1, 2, \dots, 7$.,. Having $\beta = 13$ the weight of the first "lag" in the model is approximately 0.90, and the sum of the weights of all other delays (their number is equal $p - 1 = 6$) is approximately 0.10. When $\beta = 25$ the weight of the first "lag" is approximately 0.99, and the sum of the weights of all other delays is approximately 0.01. Graphs of weight functions for $\beta = 1, 3, \dots, 25$ when $p = 7$ presented in Fig. 1.

Further, for each main variable, beta-autoregressive models were constructed for a set of parameter $\beta = 1, 2, \dots, 15$ values. Analysis of the dependence of the mean square error (MSE) of the models on parameter $\beta$ for seven variables showed that the optimal values of parameter $\beta$ is $\beta* = (7, 7, 3, 1, 10, 6, 13)$. Further $\beta$ increasing doesn’t cause significant MSE reducing.

The coefficients of the constructed beta-autoregressive models for seven main variables for the values $\boldsymbol{\beta}_*$ found are given in Table 2. Table 3 is listing the standard deviations of the model outputs from the observed values for the seven main variables. The resulting quality of the models can be considered satisfactory for all variables. It should be noted that the indicator $x_4$ (inclination) in the source data has a constant value (98.2 degrees). The root-mean-square error of the variable $x_7$ (mean anomaly) cannot be considered satisfactory.

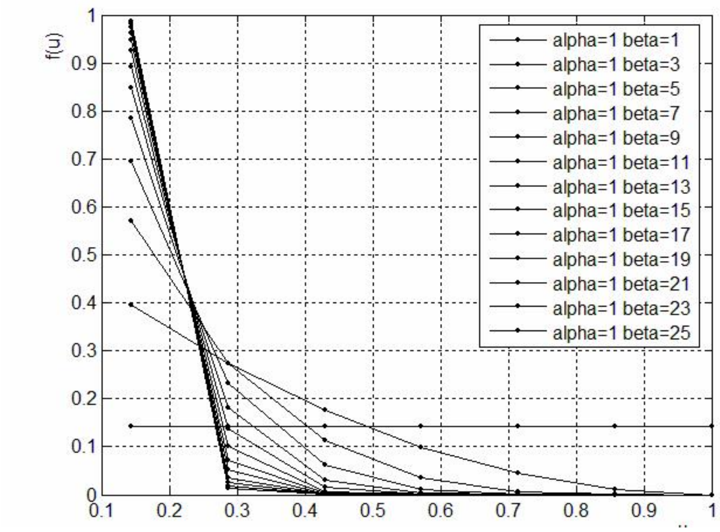


Figure 1 – Weight functions of coefficients for beta autoregression ($p = 7$) (weight functions are presented as values of probability density functions $f(u)$ for a random variable $u$ having beta distribution)

Table 2

Coefficients of autoregressive models for seven main TLE variables of the Sich-2 satellite data

| Variable | Coefficients of the autoregressive model | | | | | | |
|---|---|---|---|---|---|---|---|
| | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ |
| $x_1$ | 0,6946 | 0,2326 | 0,06098 | 0,01085 | 0,0009528 | 1,489e-005 | 0,0 |
| $x_2$ | 0,6946 | 0,2326 | 0,06098 | 0,01085 | 0,0009528 | 1,489e-005 | 0,0 |
| $x_3$ | 0,3955 | 0,2746 | 0,1758 | 0,09887 | 0,04394 | 0,01099 | 0,0 |
| $x_4$ | 0,1429 | 0,1429 | 0,14286 | 0,14290 | 0,1429 | 0,1429 | 0,1429 |
| $x_5$ | 0,81881 | 0,1587 | 0,02130 | 0,00160 | 4,16e-005 | 8,13e-008 | 0,0 |
| $x_6$ | 0,63785 | 0,2563 | 0,08400 | 0,01993 | 0,002625 | 8,203e-005 | 0,0 |
| $x_7$ | 0,89338 | 0,1002 | 0,006886 | 0,000218 | 1,68e-006 | 4,104e-010 | 0,0 |

 ISSN 1562-9945

Table 3

Standard deviations of model outputs for the seven main TLE variables of Sich-2 satellite

| $x_1$ (km.) | $x_2$ (km.) | $x_3$ | $x_4$ (deg.) | $x_5$ (deg.) | $x_6$ (deg.) | $x_7$ (deg.) |
|---|---|---|---|---|---|---|
| 0,001 | 0,001 | 1,30e-05 | 0,00 | 0,05 | 0,57 | 0,91 |

**Conclusion**

The method of constructing autoregressive models of the large fragments of space debris motion, represented by unequally spaced TLE time series has been developed. An iterative procedure has been developed for parametric identification of autoregressive equations subjected to unequally spaced observations, which has been studied by statistical testing. Based on the developed method for constructing autoregressive models, the problem of modeling the dynamics of the Sich-2 spacecraft using its TLE time series in the class of autoregressive models was considered. For all elements, satisfactory quality models were obtained.

**REFERENCES**

1. JFSCC TLE Source. – Access mode via https://www.space-track.org/#Landing

2. Ivanov N. M. Ballistics and navigation of spacecraft. / N. M. Ivanov, L. N. Lysenko // M.: Drofa, 2004. - 544 p.

3. San-Juan J. F. An economic hybrid J2 analytical orbit propagator program based on SARIMA models / J. F. San-Juan, M. San-Martin, I. P´erez // Math. Probl. Eng, 2012. – 15 p.

4. San-Martin M., M´etodos de propagaci´on h´ıbridos aplicados al problema delsat´elite artificial / M. San-Martin // T´ecnicas de suavizado exponencial, University of La Rioja, 2014.

5. Ampatzis C. Machine learning techniques for approximation of objective functions in trajectory optimisation / C. Ampatzis, D. Izzo // Proceedings of the International Joint Conference on Artificial Intelligence, 2009.

6. Peng, H., Bai, X. Improving orbit prediction accuracy through supervised machine learning. / H. Peng, X. Bai // Advanced Space Research. – 2018. – № 61 (10). – 2628–2646 p.

7. Lyubushin A. A. Analysis of data from geophysical and environmental monitoring systems / A. A. Lyubushin. - M.: Science, 2007. - 228 p.

8. Sarychev A. P. Identification of parameters of systems of autoregression equations with known covariance matrices / A. P. Sarychev // International Scientific and Technical Journal "Control and Informatics Problems". - 2012. - № 3. - 14–30 p.

9. Sarychev A. P. A study by the method of statistical tests of an iterative procedure for identifying the parameters of a system of autoregression equations / A. P. Sarychev // System Technologies. - 2014. - №3 (92). - 77–89 p.

10. Sarychev A. P. Linear autoregression based on the method of group accounting of arguments in the conditions of quasi-repeated observations / A. P. Sarychev // Artificial Intelligence. - 2015. - № 3-4 (69-70). - 105-123 p.

11. Seber J. Linear regression analysis: Per. from English / J. Seber. - M.: Mir, 1980. - 456 p.

12. Sarychev A. P. Modeling in the class of systems of autoregression equations under conditions of structural uncertainty / A. P. Sarychev // International Scientific and Technical Journal "Control and Informatics Problems". - 2015. - № 4. - 79–103 p.

13. Sarychev A. P. Modeling of complex systems under conditions of structural uncertainty: regression and autoregression models / A. P. Sarychev. - LAP LAMBERT Academic Publishing RU, Saarbrücken, Deutschland, 2016. - 274 p.

14. Analysis of global trends in space activities, development of information-analytical and methodological support for the creation of the latest models of rocket and space technology (intermediate) in 4 volumes, volume 1 / ITM of NASU and SSAU; Head of Research A. Alpatov. - Dnepr, 2016. - 245 p. - state registration number 0116U004129. - Inv. 22 - 9/2016.