UDK 519.216

I.P. Atamanyuk, I.M. Hvozdeva, V.P. Prokopyshyn

METHOD OF OPTIMAL EXTRAPOLATION OF VECTOR RANDOM SEQUENCES REALIZATIONS ON THE BASIS OF NONLINEAR CANONICAL DECOMPOSITION

Annotation. *The given work is devoted to the solving of important scientific and technical problem of the method development of the optimal (in mean square sense) extrapolation of the realizations of vector random sequences for any quantity of the known values, used for forecasting, and for any order of nonlinear stochastic relations. Forecasting model is synthesized on the basis of polynomial canonical decomposition of vector random sequence. There is obtained the formula for determination of the mean square error of extrapolation that allows to estimate the solution accuracy of the forecasting problem with using the proposed method. Taking into account the recurrent character of the estimationprocesses of the future values of investigated sequence the method is quite simple in calculating respect. The developed method of extrapolation as well as the vector canonical expansion assumed as its basis don't put any essential limitations on the class of prognosticated random sequences (linearity, Markovian property, stationarity, scalarity, monotony etc.).*

Keywords: *optimal nonlinear extrapolation, vector random sequences, polynomial canonical decomposition.*

Introduction.

The peculiarityof a wide range of applied problems in different science and engineering fields is the probabilistic nature of an investigated phenomenon or presence of random factors influence on the investigated object, therefore the process of its state changing also has a probabilistic nature. The objects of such class are the objects with randomly changing conditions of operation and they are investigated in such applications as: the technical diagnostics, radiolocation, plane-to-plane navigation, predictive control of reliability, information security, synthesis of the models of chemical kinetics, technological objects control etc.

The specific peculiarity of mentioned problems is the presence of preliminary stage of gathering the information about the investigated object. Random character of external influences and coordinates (input and output) of the objects with randomly changing conditions of operation at sufficient volume of statistic data determines the necessity and reasonability of applying the deductive methods [1] of random sequences forecasting.

Analysis of the latest investigations and publications. It is known thatthe most general extrapolation form for the forecasting problems solving is the mathematical model in the form of Kolmogorov-Gabor polynomial [1, 2]. Such model permits to take into account any number of measurements of a random sequence and the order of exponential nonlinearity. But its practical application is limited by significant difficulties, associated with the formation of a large number of equations for the extrapolator parameters determination. The existing optimal methods, that are used for the applied problems solving, are obtained for certain classes of random sequences, in particular, Kolmogorov's [3] and Wiener's [4] methods – for stationary processes, Kalman's filter-extrapolator [5] – for markovian random sequences, Pugachev's [6] and Kudritskyi's [7-8] methods – for non-stationary Gaussian sequences etc. It should be mentioned that their application permits to obtain the optimal results only for the sequences with certain a priori known characteristics.

Thus there are exist the theoretically grounded solutions of random sequences forecasting problem but the known methods and models are based on the use of appropriate restrictions which don't permit to obtain the maximal extrapolation accuracy.

The purpose of the paper. The creation of the method of vector random sequences extrapolation at the most general assumptions concerning the stochastic qualities of investigated sequence (with use for forecasting any degree of nonlinear stochastic relations and measurements number).

Mathematical problem statement. Vector random sequence $\left\{\vec{X}\right\} = X_h(i),\ h = \overline{1,H}$, describing the time variation of H interdependent parameters of some object with a randomly varying operating conditions, fully specified in a discrete number of points $t_i,\ i = \overline{1,I}$ by moment functions $M\left[X_l^{\nu}(i)\right],\ M\left[X_l^{\nu}(i)X_h^{\mu}(j)\right],\ i,j = \overline{1,I};\ l,h = \overline{1,H}\ \nu,\mu = \overline{1,N}.$ It is necessary to obtain the best estimations $\overset{*}{x}_h(i), i = \overline{k+1,I}, h = \overline{1,H}$ of future values of the investigated random sequence for each of its components $X_h(i)$ under the stipulation that the values $x_h^{\mu}(j),\ j = \overline{1,k},\ \mu = \overline{1,N},\ h = \overline{1,H}$ of the first k observation points are known.

Main material statement. The most universal approach to the solution of the formulated problem from the viewpoint of the restrictions, imposed on the random process, consists of using the canonical decompositions [6-8]. For the vector case such decomposition with taking into account stochastic relations $M\left[X_l^{\nu}(i)\right]$, $M\left[X_l^{\nu}(i)X_h^{\mu}(j)\right]$, $i,j=\overline{1,I}$ has the form [9]:

$$X_h(i)=M\left[X_h(i)\right]+\sum_{\nu=1}^{i-1}\sum_{l=1}^{H}\sum_{\lambda=1}^{N}W_{\nu l}^{(\lambda)}\beta_{l\lambda}^{(h,1)}(\nu,i)+$$

$$\sum_{l=1}^{h-1}\sum_{\lambda=1}^{N}W_{il}^{(\lambda)}\beta_{l\lambda}^{(h,1)}(i,i)+W_{ih}^{(1)},\ i=\overline{1,I}, \qquad (1)$$

$$D_{l,\lambda}(\nu)=M\left[\left\{W_{\nu l}^{(\lambda)}\right\}^2\right]=M\left[X_l^{2\lambda}(\nu)\right]-M^2\left[X_l^{\lambda}(\nu)\right]-$$

$$-\sum_{\mu=1}^{\nu-1}\sum_{m=1}^{H}\sum_{j=1}^{N}D_{mj}(\mu)\left\{\beta_{mj}^{(l,\lambda)}(\mu,\nu)\right\}^2-\sum_{m=1}^{l-1}\sum_{j=1}^{N}D_{mj}(\nu)\left\{\beta_{mj}^{(l,\lambda)}(\nu,\nu)\right\}^2- \qquad (2)$$

$$-\sum_{j=1}^{\lambda-1}D_{lj}(\nu)\left\{\beta_{lj}^{(l,\lambda)}(\nu,\nu)\right\}^2,\ \nu=\overline{1,I};$$

$$\beta_{l\lambda}^{(h,s)}(\nu,i)=\frac{M\left[W_{\nu l}^{(\lambda)}\left(X_h^s(i)-M[X_h^s(i)]\right)\right]}{M\left[\left\{W_{\nu l}^{(\lambda)}\right\}^2\right]}=$$

$$=\frac{1}{D_{l\lambda}(\nu)}(M\left[X_l^{\lambda}(\nu)X_h^s(i)\right]-M\left[X_l^{\lambda}(\nu)\right]M\left[X_h^s(i)\right]- \qquad (3)$$

$$-\sum_{\mu=1}^{\nu-1}\sum_{m=1}^{H}\sum_{j=1}^{N}D_{mj}(\mu)\beta_{mj}^{(l,\lambda)}(\mu,\nu)\beta_{mj}^{(h,s)}(\mu,i)-$$

$$-\sum_{m=1}^{l-1}\sum_{j=1}^{N}D_{mj}(\nu)\beta_{mj}^{(l,\lambda)}(\nu,\nu)\beta_{mj}^{(h,s)}(\nu,i)-\sum_{j=1}^{\lambda-1}D_{lj}(\nu)\beta_{lj}^{(l,\lambda)}(\nu,\nu)\beta_{lj}^{(h,s)}(\nu,i),$$

$$\lambda=\overline{1,h},\ \nu=\overline{1,i}.$$

The random sequence $X_h(i), i=\overline{1,I}; h=\overline{1,H}$ is represented with $H\times N$ arrays $\{W_l^{(\lambda)}\}$, $\lambda=\overline{1,N}$; $l=\overline{1,H}$ of centered uncorrelated random coefficients $W_{\nu l}^{(\lambda)}$, $\nu=\overline{1,I}$; $\lambda=\overline{1,N}$; $l=\overline{1,H}$. Each of these coefficients contains information about corresponding value $X_l^{\lambda}(\nu)$, and coordinate functions $\beta_{l\lambda}^{(h,s)}(\nu,i)$ describe the probability connections of $\lambda+s$ order between the components $X_l(i)$ and $X_h(i)$ in sections t_ν and t_i. On Fig. 1 the algorithm block-diagram of the canonical model (1) parameters determination is represented.

The forecasting model, based on the canonical decomposition (1) of random sequence, has the form [10]:

$$m_{x;j,h}^{(\mu,l)}(s,i)=\begin{cases} M\left[X_h(i)\right], \text{ at } \mu=0, \\ m_{x;j,h}^{(\mu,l-1)}(s,i)+\left(x_j^l(\mu)-m_{x;j,j}^{(\mu,l-1)}(l,\mu)\right)\beta_{j,l}^{(h,s)}(\mu,i), \text{ at } l>1, j<H, \\ m_{x;j,h}^{(\mu,N)}(s,i)+\left(x_{j+1}(\mu)-m_{x;j,j+1}^{(\mu,1)}(N,\mu)\right)\beta_{j+1,1}^{(h,s)}(\mu,i), \text{ at } l=1, j<H, \\ m_{x;H,h}^{(\mu,N)}(s,i)+\left(x_1(\mu+1)-m_{x;H,1}^{(\mu,N)}(N,\mu+1)\right)\beta_{1,1}^{(h,s)}(\mu+1,i), \text{ for} \\ l=1, j=H, \end{cases} \quad (4)$$

were

$m_{x;j,h}^{(\mu,l)}(1,i)=M\left[X_h(i)/x_\lambda^n(\nu),\ \lambda=\overline{1,H}, n=\overline{1,N}, \nu=\overline{1,\mu-1};\ x_\lambda^n(\mu),\ \lambda=\overline{1,j}, n=\overline{1,l}\right]$

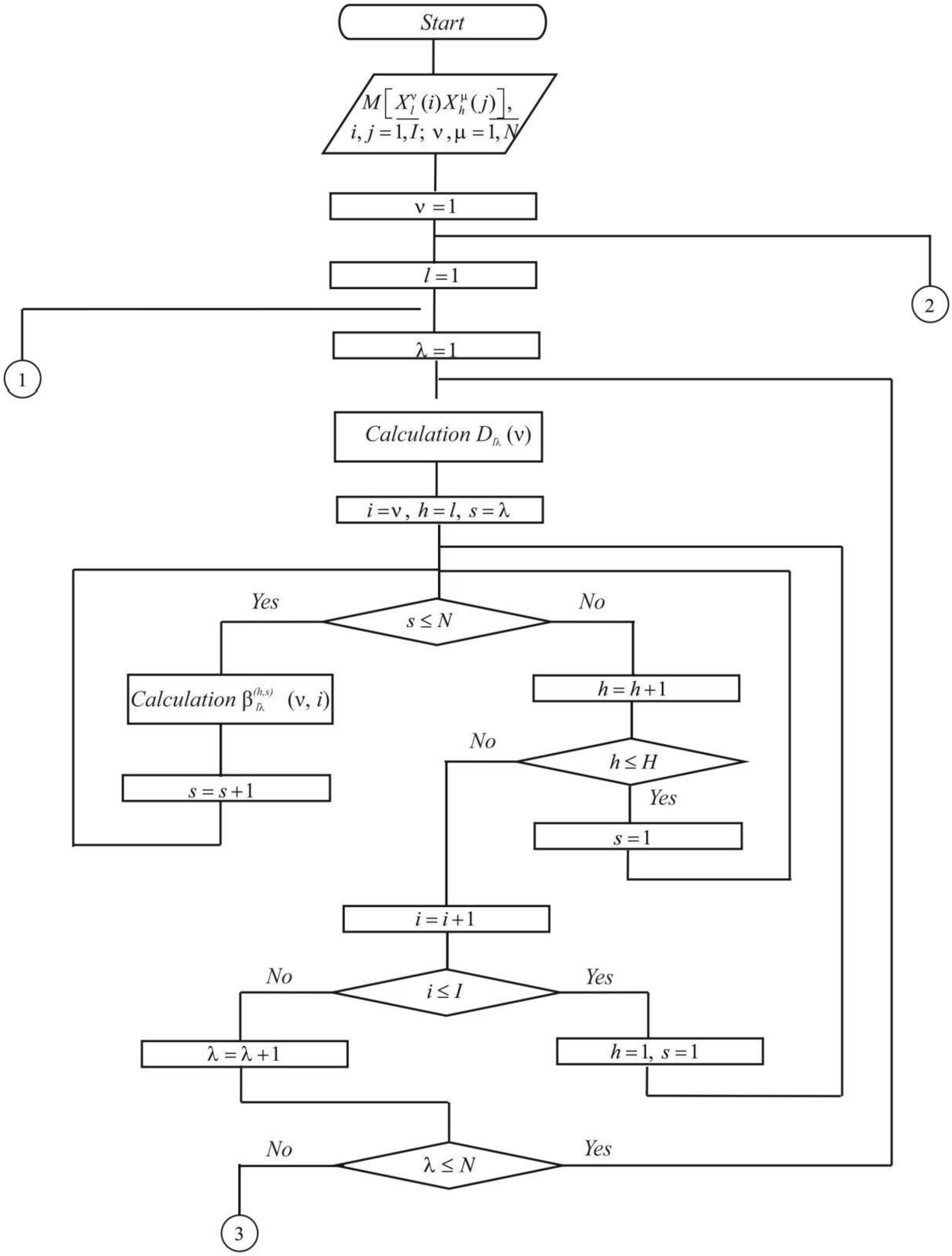
– optimal (in the root mean square sense) estimation of future values of the investigated random sequence at the assumption that for forecasting it is applied a posteriori information $x_\lambda^n(\nu), \lambda=\overline{1,H}, n=\overline{1,N}$, $\nu=\overline{1,\mu-1};\ x_\lambda^n(\mu),\ \lambda=\overline{1,j},\ n=\overline{1,l}$.

The diagram in Fig. 2 represents the distinctions of the calculation process of determination of a random sequence future values at using the extrapolation algorithm (4). The equation for the root mean square error of the extrapolation with using the forecasting model (4) according to known values $x_j^n(\mu),\ \mu=\overline{1,k};\ j=\overline{1,H};\ n=\overline{1,N}$ has the form

$$E_h^{(k,N)}(i)=M\left[X_h^2(i)\right]-M^2\left[X_h(i)\right]-\sum_{\mu=1}^{k}\sum_{j=1}^{H}\sum_{n=1}^{N}D_{jn}(\mu)\left\{\beta_{jn}^{(h,1)}(\mu,i)\right\}^2, \quad (5)$$

$$i=\overline{k+1,I}.$$

 ISSN 1562-9945


Start
$M\left[X_l^{\nu}(i)X_h^{\mu}(j)\right]$,
$i, j = \overline{1,I};\ \nu, \mu = \overline{1,N}$
$\nu = 1$
$l = 1$
2
1
$\lambda = 1$
Calculation $D_{D_\lambda}(\nu)$
$i = \nu,\ h = l,\ s = \lambda$
Yes
$s \le N$
No
Calculation $\beta_{D_\lambda}^{(h,s)}(\nu, i)$
$h = h + 1$
No
$h \le H$
Yes
$s = s + 1$
$s = 1$
$i = i + 1$
No
$i \le I$
Yes
$\lambda = \lambda + 1$
$h = 1,\ s = 1$
No
$\lambda \le N$
Yes
3

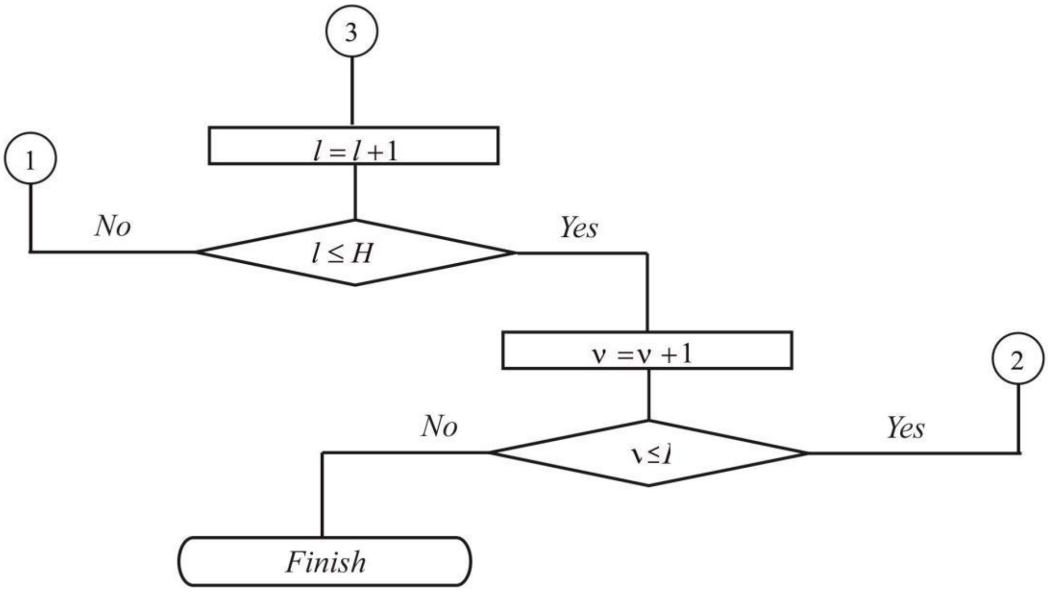


Fig.1. The algorithm block-diagram ofcalculatingthe parameters $D_{l,\lambda}(\nu)$, $\beta_{l\lambda}^{(h,s)}(\nu,i)$ of vectorcanonicaldecomposition (1)

The root mean square of extrapolation $E_h^{(k,N)}(i)$ error is equal to the dispersion of a posteriori random sequence

$$X_h^{(k,N)}(i) = X\left(i / x_l^{\nu}(j),\ \nu = \overline{1,N}, j = \overline{1,k}, l = \overline{1,H}\right) = m_{H,h}^{(k,N)}(1,i) +$$

$$+ \sum_{\nu=k+1}^{i-1} \sum_{l=1}^{H} \sum_{\lambda=1}^{N} W_{\nu l}^{(\lambda)} \beta_{l\lambda}^{(h,1)}(\nu,i) + \sum_{l=1}^{h-1} \sum_{\lambda=1}^{N} W_{il}^{(\lambda)} \beta_{l\lambda}^{(h,1)}(i,i) + W_{ih}^{(1)},\ i = \overline{k+1,I}.$$

The use of developed method of random sequences extrapolation on the basis of forecasting model (4) requires the implementation of the following stages:

Step 1: Collection of the statistics data about the investigated random sequence;

Step 2. Estimation of the moment functions $M\left[X_l^{\nu}(i)\right],\ M\left[X_l^{\nu}(i) X_h^{\mu}(j)\right],\ i,j = \overline{1,I}; l,h = \overline{1,H};\ \nu,\mu = \overline{1,N}$ on the basis of the accumulated realizations of the random sequence;

Step 3. Formation of canonical decomposition (1) for the investigated vector random sequence in accordance with the algorithm, represented on Fig. 1;

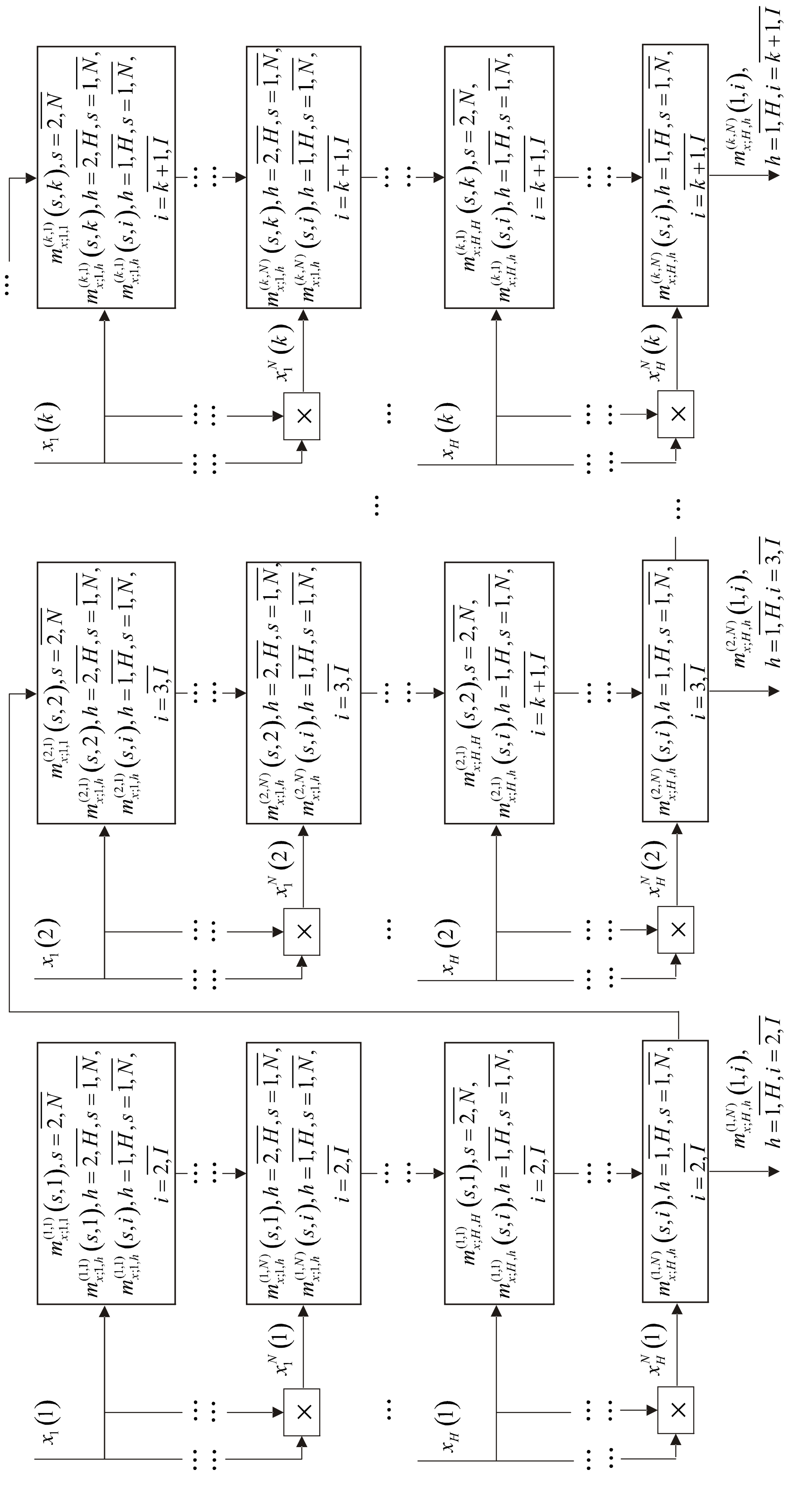


Fig.2Schematicprocedureof formation ofthe future values ofa random sequenceusing the algorithm(4)

Step 4: Calculating the estimates of future values of extrapolated realization on the basis of forecasting models (4) in accordance with the scheme on Fig. 2;

Step 5: Assessment of problem solution quality of investigated sequence forecasting with using equation (5).

At absence the stochastic relationships between the components the forecasting model (4) can be simplified to H equations [11–12] for scalar sequences extrapolation.

The method was tested for forecasting the random sequences, describing the average monthly air temperature changing in Odessa and Kiev. As statistics data have been used [13] the values of average (from January to December) temperature during hundred years (1910–2009 y.y.).

The Fig. 3 shows the diagrams of mathematical expectations of the investigated random sequences ($X_1(i),\ i=\overline{1,12}$ – random sequence of average monthly temperature changing in Odessa, $X_2(i),\ i=\overline{1,12}$ – a random sequence of average monthly temperature changing in Kyiv).

Numerical experiment was organized in the following way. On the basis of 99 realizations of random sequences $X_1(i), X_2(i)\ i=\overline{1,12}$ were determined the model's (4) parameters for $N=5$ and linear extrapolator (in model (4) used only correlations). For the remaining realization from a hundred existing statistical database the estimation of the future values was calculated and the forecasting error was determined. The procedure was performed a hundred iterations and forecasted realization was removed from the training sample, and in its place was placed the investigated in the previous experiment realization.

As a result of numerical experiments the standard deviations (SD) and the mean square errors of extrapolation of random sequences realizations $X_1(i), X_2(i)\ i=\overline{1,12}$ were obtained thought the linear algorithm and developed method based on the model (4) (Fig. 4, Fig. 5).

 ISSN 1562-9945

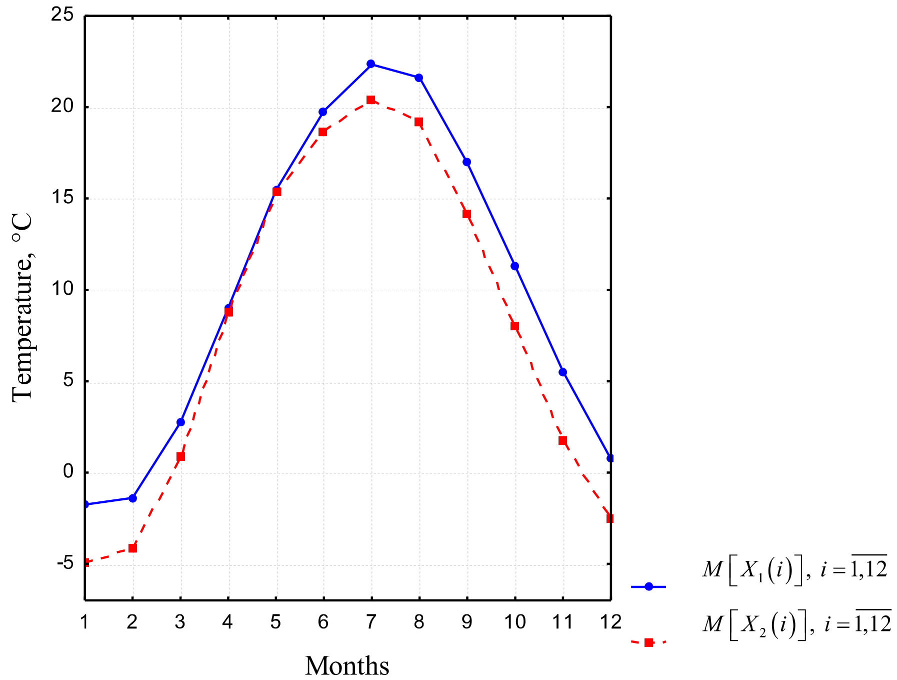


Fig. 3. The diagrams of mathematical expectations of random sequences of the average temperature in Odessa ($M[X_1(i)],\ i=\overline{1,12}$) and Kyiv ($M[X_2(i)],\ i=\overline{1,12}$)

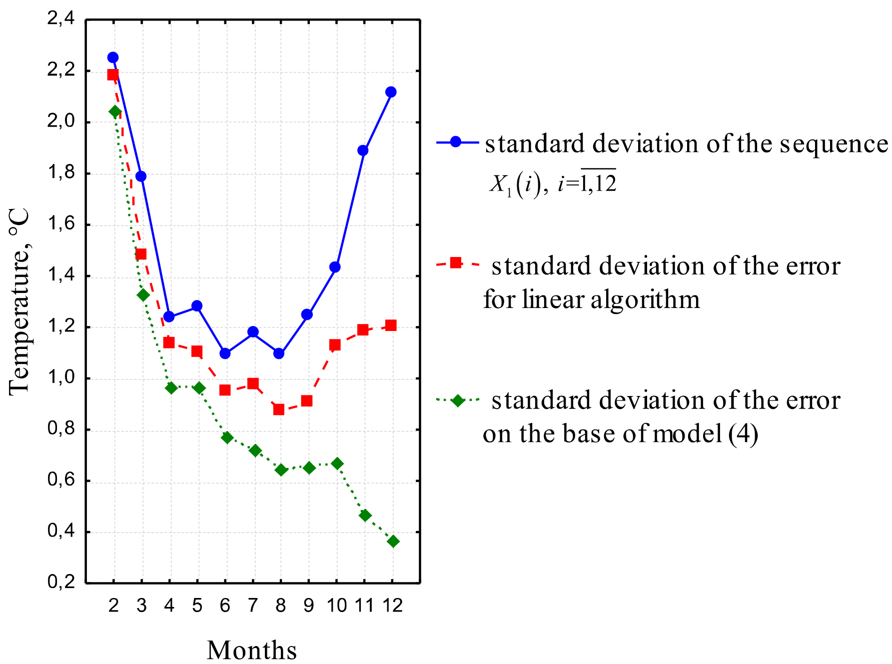


Fig. 4. The root mean square error of sequence realizations extrapolations using the linear algorithm on the basis of model (4)

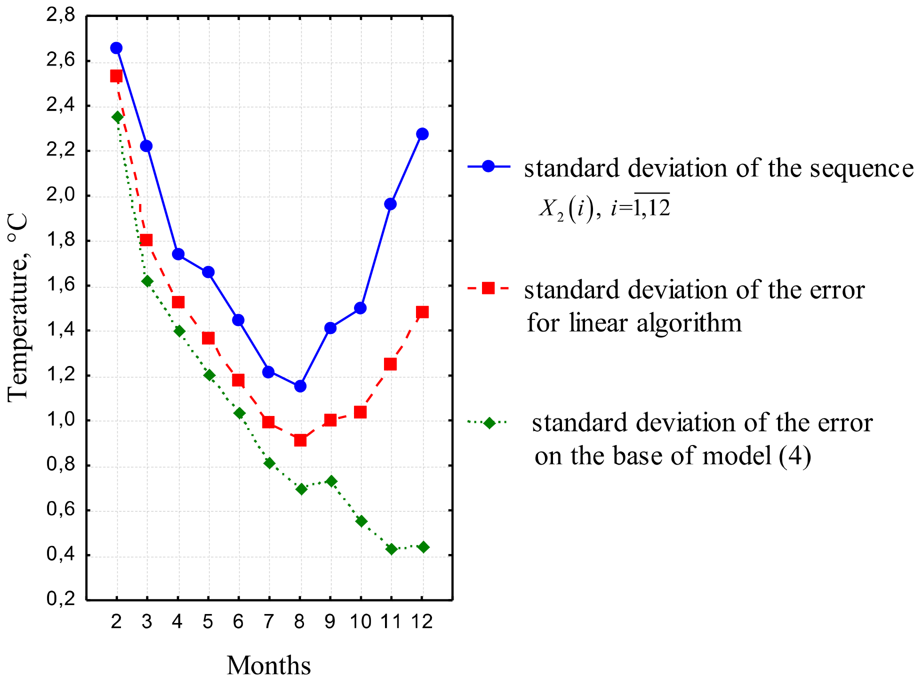


Fig. 5. The extrapolation root mean square error of sequence realizations $X_1(i),\ i=\overline{1,12}$ with using the linear algorithm and on the basis of the model (4)

Forecasting results show on the significant advantage in an extrapolation quality through using the nonlinear stochastic relations in forecasting model.

Conclusions.

There was developed the method of nonlinear extrapolation of vector random sequences, that do not impose any significant restrictions on the class of investigated sequences: stationarity, Markovian property, linearity, monotonicity, etc. The universality of the obtained solution is determined by the fact that there is a canonical decomposition and accurately describes in the discreteness points any random process with finite dispersion. The method is optimal in the sense of mean-square criterion and it allows the use of stochastic connections of any non-linearity order and any measurements number. Taking into account the recurrent nature of the extrapolator parameters calculations, its implementation on a computer is enough simple. The numerical experiments results shown high efficiency of the proposed method.

Since the most of investigated physical, technical, economic or other real processes are stochastic, the proposed method has the wide range of applications at solving the control problems in various fields of science

 ISSN 1562-9945

and engineering: forecasting control of technical devices reliability, medical diagnostics, radiolocation, technological objects control etc.

REFERENCES

1. Ivahnenko, A.G. Principals of Inductive Theory of Fuzzy Recognition and Prognosis of Random Processes and Events / A.G. Ivahnenko. Kyiv, 1991. – 47 p.
2. Box, G.E.P., G.M. (1970). Time–series Analysis: Forecasting and Control / G.E.P. Box, G.M. Jenkins.V.1. – M. World, 1974. – 406 p.
3. Kolmogorov, A. N. Interpolation and Extrapolation of Stationary Random Sequences / A.N. Kolmogorov // J. Proceedings of Academy of Sciences of the USSR. Mathematical Series, 1941. – V. 5, P. 3–14.
4. Wiener, N. Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications / N. Wiener. – N.Y.: MIT Press, 1949. – 250 p.
5. Kalman R.E. A New Approach to Linear Filtering and Prediction Problems / R.E. Kalman // Trans. ASME Series D, J. Basic Engg. – 1960. – V. 82 (Series D). – P. 35– 45.
6. Pugachev, V.S. The Theory of Random Functions and its Application / V.S. Pugachev. – M.: Fitmatgiz, 1962. – 720 p.
7. Kudritsky, V.D. Predictive control of radioelectronic devices / V.D. Kudritsky. – Kyiv: Technics, 1982. – 168 p.
8. Kudritsky, V.D. Filtering, Extrapolation and Recognition Realizations of Random Functions / V.D. Kudritsky. – Kyiv: FADA Ltd., 2001. – 176 p.
9. Atamanyuk, I.P. Vector Polynomial Canonical Decomposition of a Random Process / I.P. Atamanyuk. – J. Technical Diagnostics and Nondestructive Control. – 2003. – No. 1 – P. 36–40.
10. Atamanyuk, I.P. Vector Polinomial Algorithm of Optimal Extrapolation of a Random Process / I.P. Atamanyuk // the Vostochno-evropeyskiy Journal of advanced technologies. – 2010. – 2/4 (44). – P. 17–19.
11. Atamanyuk, I.P. The Algorythm of Optimal Nonlinear Extrapolation of the Realizations of Random Process with the Filtration of Errors Changes / I.P.Atamanyuk, Y.P.Kondratenko. – J. Electronic Modelling, 2012. – No. 4, P. 23–40.
12. Atamanyuk, I.P. The Algorithm of Optimal Polynomial Extrapolation of Random Processes / I.P. Atamanyuk,

V.Y. Kondratenko, O.V. Kozlov, Y.P. Kondratenko // Modeling and Simulation in Engineering, Economics and Management, K.J.Engemann, A.M.Gil-Lafuente, J.L.Merigo (Eds), International Conference MS 2012, New Rochelle, NY, USA (May 30 – June 1, 2012), Proceedings. Lecture Notes in Business Information Processing, Volume 115, Springer, 78-87.

13. datamarket.com/data/set/1loo/average-monthly-temperatures-across-the-world#q=

 ISSN 1562-9945