MOTROL. Commission of Motorization and Energetics in Agriculture – 2015. Vol.17. No.2. 3-7

METHOD OF OPTIMAL LINEAR EXTRAPOLATION OF VECTOR RANDOM SEQUENCES WITH FULL CONSIDERATION OF CORRELATION CONNECTIONS FOR EACH COMPONENT

Igor Atamanyuk[1], Yuriy Kondratenko[2]
[1]Nikolaev National Agrarian University
St. Paris Commune, 9, Nikolaev, Ukraine. E-mail: Atamanyuk_igor@mail.ru
[2]Petro Mohyla Black Sea State University
St. 68 Desantnikiv, 10, Nikolaev, Ukraine. E-mail: Kondratenko@mail.ru

Игорь Атаманюк[1], Юрий Кондратенко[2]
[1]Николаевский национальный аграрный университет
Ул. Парижской коммуны, 9, г. Николаев, Украина. E-mail: Atamanyuk_igor@mail.ru
[2]Черноморский государственный университет им. Петра Могилы.
Ул. 68 Десантников, 10, г. Николаев, Украина. E-mail: Kondratenko@mail.ru

Summary. The work is devoted to solving important scientific and technical problems of formation of the optimal method of mean-square linear extrapolation implementations of vector random sequences of any number of known values used for the forecast. The resulting method, in contrast to existing solutions prediction problem, take full account of a priori information about the target sequence for each component.

Forecast model is synthesized based on the linear vector of the canonical decomposition of the random sequences. The formula for determining the mean square error of extrapolation, which allows us to estimate the accuracy of solving the problem of the prediction by the proposed method for some fixed number of known values and components of the vector sequence. The paper also shows a block-diagram of an algorithm for determining the parameters of the proposed method.

The method of extrapolation, as well as vector canonical decomposition, put in its basis, does not impose any significant restrictions on the class of projected random sequences (Markov, stationary, scalar, monotony, etc.). Given the recurrent nature of the calculation of estimates of future values of the target sequence method is quite simple computationally.

Since most of the physical, technical, economic or other real processes are stochastic, the proposed method has the wide range of applications in solving problems of governance in various fields of science and technology: predictive control of the reliability of technical devices, medical diagnostics, radar, control of technological objects and so on.

Key words: vector random sequences, linear canonical decomposition algorithm extrapolation.

STATEMENT OF THE PROBLEM

The solution of many actual scientific and technical problems associated with the use of algorithms and extrapolating devices, which are known ie observable part of the process make it possible to estimate the unknown inaccessible part of it. In particular extrapolating algorithms used in automatic control systems and objects inertial systems with delay. Exceptionally widely spread algorithm linear prediction vocoders used in modern digital communication systems, in the compressed audio and video signal [1]. It is also widely used predictive algorithms based on neural networks, Kalman-Bucy filter, group method of data and some others [2-16]. However, despite this diversity, the need for high-speed, robust and highly accurate algorithms and devices of the forecast continues to be relevant in the present and in the future.

ANALYSIS OF RECENT RESEARCH AND PUBLICATIONS

Assume that the random vector sequence $\{\overline{X}\}=\{X_1(i)...X_h(i)...X_H(i)\}$ in the study a number of points $t_i, i=\overline{1,I}$ given full matrix functions $M\left[X_h(i)\right]$, $h=\overline{1,H}$, $i=\overline{1,I}$; $M\left[X_\lambda(\nu)X_h(i)\right]$, $\nu,i=\overline{1,I}$, $\lambda,h=\overline{1,H}$. It is necessary to synthesize a method of predicting the future values of a random sequence $\{\overline{X}\}$ of known values $x_h(i)$, $i=\overline{1,k}$, $k<I$, $h=\overline{1,H}$, which are obtained by measuring the target sequence on the observation interval $[t_1...t_k]$.

OBJECTIVE OF RESEARCH

Method, in contrast to existing solutions prediction problem, take full account of a priori information about the target sequence for each component.

MAIN MATERIAL

The most universal method in terms of the restrictions that are imposed on the predicted sequences is the algorithm of extrapolation [17]:

$$m_h^{(r_1,\dots,r_k)}(i)=\begin{cases}M\left[X_h(i)\right],\ k=0,\ i=\overline{1,I},\\ m_h^{(r_1,\dots,r_k-1)}(i)+\left[x_{r_k}(k)-m_{r_k}^{(r_1,\dots,r_k-1)}(\mu)\right]\times\\ \times\varphi_{hk}^{(r_k)}(i),\ h=\overline{1,H},\ i=\overline{k+1,I}.\end{cases}\quad(1)$$

In the expression (1) $r_\mu,\ \mu=\overline{1,k}$ - number of components in the section t_μ. The parameters of the algorithm (1) are elements of the canonical decomposition:

$$X_h(i)=M[X_h(i)]+\sum_{\lambda=1}^{h}\sum_{\nu=1}^{i}V_\nu^{(\lambda)}\varphi_{h\nu}^{(\lambda)}(i),\quad h=\overline{1,H}, i=\overline{1,I}. \tag{2}$$

Relations for their definitions are of the form:

$$V_1^{(1)}=X_1(1)-M[X_1(i)],$$
$$V_i^{(h)}=X_h(i)-M[X_h(i)]-\sum_{\lambda=1}^{h}\sum_{\nu=1}^{i-1}V_\nu^{(\lambda)}\varphi_{h\nu}^{(\lambda)}(i),\quad h=\overline{1,H}, i=\overline{1,I}; \tag{3}$$
$$D_1^{(1)}=D_1(1),$$
$$D_i^{(h)}=D_h(i)-\sum_{\lambda=1}^{h}\sum_{\nu=1}^{i-1}D_\nu^{(\lambda)}\left[\varphi_{h\nu}^{(\lambda)}(i)\right]^2,\quad h=\overline{1,H}, i=\overline{1,I}; \tag{4}$$
$$\varphi_{h\nu}^{(\lambda)}(i)=\frac{1}{D_\nu^{(\lambda)}}M\left[V_\nu^{(\lambda)}\left(X_h(i)-M[X_h(i)]\right)\right],\quad h=\overline{1,H},\ \lambda=\overline{1,h},\ \nu,i=\overline{1,I}. \tag{5}$$

Expression (1) in the framework described in the canonical decomposition (2) the probability linear relations of sequence $\{\overline{X}\}$ allows to get the best result in the mean-square values of extrapolation $x_h(i),\ i=\overline{k+1,I},\ h=\overline{1,H}$. However, the full properties of the target sequence $\{\overline{X}\}$ in (2) takes into account only for component $\{X_H\}$ (for $\{X_h\},\ h<H$ it is not used in the formula (2) of interrelation communication $\{X_h\}$ with $\{X_{h+j}\},\ j=\overline{1,H-h}$) and, thus, only this component results extrapolation algorithm (1) can be considered strictly optimal for the available capacity of a priori information about the investigated vector random sequence. For other component characteristics estimation accuracy (1) can be improved by increasing the amount of a priori information, which is used for the forecast.

In order to eliminate this disadvantage use to generate a extrapolation algorithm a canonical decomposition [18-20] sequences $\{\overline{X}\}$ with full consideration of interrelation connections for each component:

$$X_h(i)=M[X_h(i)]+\sum_{\nu=1}^{i-1}\sum_{\lambda=1}^{H}V_\nu^{(\lambda)}\varphi_{h\nu}^{(\lambda)}(i)+\sum_{\lambda=1}^{h}V_i^{(\lambda)}\varphi_{hi}^{(\lambda)}(i),\ i=\overline{1,I}. \tag{6}$$

Elements of the expansion (6) are given by:

$$V_\nu^{(\lambda)}=X_\lambda(\nu)-M[X_\lambda(\nu)]-\sum_{\mu=1}^{\nu-1}\sum_{j=1}^{H}V_\mu^{(j)}\varphi_{\lambda\mu}^{(j)}(\nu)-\sum_{j=1}^{\lambda-1}V_\nu^{(j)}\varphi_{\lambda\nu}^{(j)}(\nu),\nu=\overline{1,I}; \tag{7}$$

$$D_\lambda(\nu)=M\left[\left\{V_\nu^{(\lambda)}\right\}^2\right]=M\left[\left\{X_\lambda(\nu)\right\}^2\right]-M^2[X_\lambda(\nu)]-\sum_{\mu=1}^{\nu-1}\sum_{j=1}^{H}D_j(\mu)\left\{\varphi_{\lambda\mu}^{(j)}(\nu)\right\}^2-\sum_{j=1}^{\lambda-1}D_j(\nu)\left\{\varphi_{\lambda\nu}^{(j)}(\nu)\right\}^2,\ \nu=\overline{1,I}; \tag{8}$$

$$\varphi_{h\nu}^{(\lambda)}(i)=\frac{M\left[V_\nu^{(\lambda)}\left(X_h(i)-M[X_h(i)]\right)\right]}{M\left[\left\{V_\nu^{(\lambda)}\right\}^2\right]}=\frac{1}{D_\lambda(\nu)}(M[X_\lambda(\nu)X_h(i)]-M[X_\lambda(\nu)]\times M[X_h(i)]-\sum_{\mu=1}^{\nu-1}\sum_{j=1}^{H}D_j(\mu)\varphi_{\lambda\mu}^{(j)}(\nu)\varphi_{h\mu}^{(j)}(i)-\sum_{j=1}^{\lambda-1}D_j(\nu)\varphi_{\lambda\nu}^{(j)}(\nu)\varphi_{h\nu}^{(j)}(i)),\ \lambda=\overline{1,h},\ \nu=\overline{1,i}. \tag{9}$$

The coordinate functions $\varphi_{h\nu}^{(\lambda)}(i),\quad h,\lambda=\overline{1,H},\ \nu,i=\overline{1,I}$ are characterized by the following properties:

$$\varphi_{h\nu}^{(\lambda)}(i)=\begin{cases}1,\ h=\lambda\ \&\ \nu=i;\\ 0,\ i<\nu.\end{cases} \tag{10}$$

Block diagram of the algorithm for calculating the parameters of the canonical decomposition (6) is shown in Fig. 1.

Suppose that at time $\mu=1$ known value $X_1(1)=x_1(1)$ of the first component $\{X_1\}$ sequence $\{\overline{X}\}$ and thus knows the value of the random coefficient $V_1^{(1)}=v_1^{(1)}:\ v_1^{(1)}=x_1(1)-M[X_1(1)]$.

Substituting $v_1^{(1)}$ into (6) gives:

$$X_h^{(1,1)}(i)=M[X_h(i)]+\left(x_1(1)-M[X_1(1)]\right)\varphi_{h1}^{(1)}(i)+\sum_{\lambda=2}^{H}V_1^{(\lambda)}\varphi_{h1}^{(\lambda)}(i)+\sum_{\nu=2}^{i-1}\sum_{\lambda=1}^{H}V_\nu^{(\lambda)}\varphi_{h\nu}^{(\lambda)}(i)+\sum_{\lambda=1}^{h}V_i^{(\lambda)}\varphi_{hi}^{(\lambda)}(i),\ i=\overline{1,I}. \tag{11}$$

$X_h^{(1,1)}(i)=X_h(i/x_1(1))$ - a posteriori random sequence in which the component $\{X_1\}$ passes through the coordinate $x_1(1)$ at time $\mu=1$.

Application to the operation of the expectation (11) provides an estimate of the future value:

$$m_h^{(1,1)}(i)=M[X_h(i)]+\left(x_1(1)-M[X_1(1)]\right)\varphi_{h1}^{(1)}(i). \tag{12}$$

Let us consider the value $x_2(1)$ of the same implementation. For it is the expansion (11) that allows you to specify the value of the coefficient $V_2^{(1)}=v_2^{(1)}$. In view of (12) the expression for the coefficient $v_2^{(1)}$ can be written as:

$$v_2^{(1)}=x_2(1)-m_2^{(1,1)}(1). \tag{13}$$

that allows to record:

$$m_h^{(1,2)}(i)=m_h^{(1,1)}(i)+\left[x_2(1)-m_2^{(1,1)}(i)\right]\varphi_{h1}^{(2)}(i). \tag{14}$$

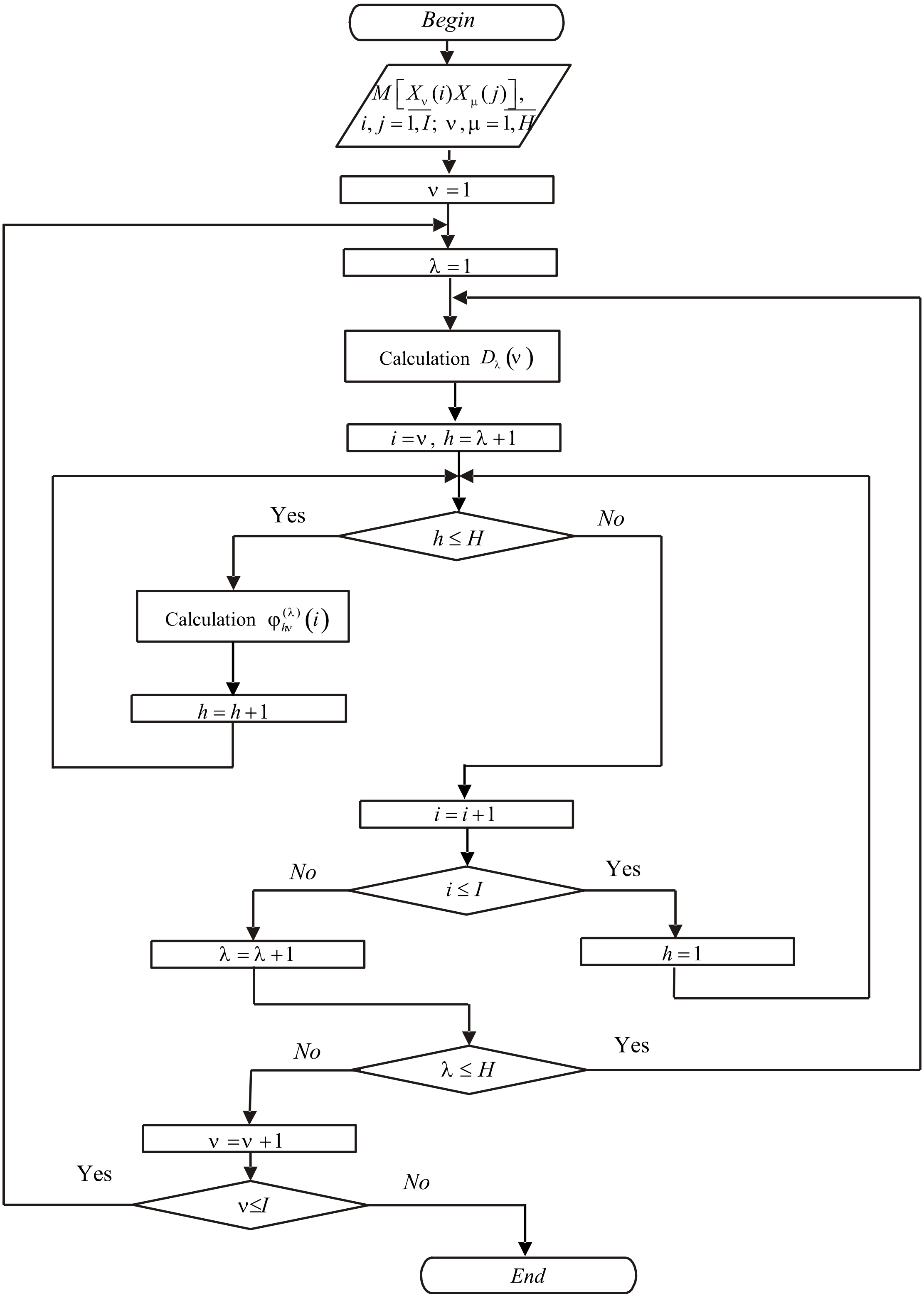

Fig. 1. Block diagram of the algorithm for calculating the parameters of the canonical decomposition (6)

With a further increase in the a posteriori information is used to forecast the resulting pattern gives a generalization of extrapolation algorithm for an arbitrary number of measurement points

$$m_h^{(\mu,l)}(i)=\begin{cases} M\left[X_h(i)\right],\ \mu=0; \\ m_h^{(\mu,l-1)}(i)+\left[x_l(\mu)-m_l^{(\mu,l-1)}(\mu)\right]\varphi_{h\mu}^{(l)}(i), \\ l\neq 1,\ \mu=\overline{1,k},\ i=\overline{k+1,I}; \\ m_h^{(\mu,H)}(i)+\left[x_1(\mu)-m_1^{(\mu-1,H)}(\mu)\right]\varphi_{h\mu}^{(1)}(i), \\ l=1,\ \mu=\overline{1,k},\ i=\overline{k+1,I}. \end{cases} \quad (14)$$

$m_h^{(\mu,l)}(i)=M\left[X_h(i)/x_\lambda(\nu),\lambda=\overline{1,H},\nu=\overline{1,\mu-1};x_j(\mu),j=\overline{1,l}\right]$, $h=\overline{1,H},\ i=\overline{k,I}$ - optimal for the criterion of the minimum mean square error of the forecast estimates of future values of the target sequence, provided that the values of $x_\lambda(\nu),\ \lambda=\overline{1,H},\ \nu=\overline{1,\mu-1};\ x_j(\mu),\ j=\overline{1,l}$

The first expression of the algorithm (15) corresponds to the case where a posteriori information is not provided, in the second ratio consistently recurrently known value is accounted for vector random sequences for the fixed time, and the third expression moves to the next point in time for the further accumulation of information, which is used to forecast .

The mean square error of extrapolation algorithm (15) is given by

$$E_h^{(\mu,l)}(i)=D_{X_h}(i)-\sum_{\nu=1}^{\mu-1}\sum_{\lambda=1}^{H}V_\nu^{(\lambda)}\varphi_{h\nu}^{(\lambda)}(i)-\sum_{j=1}^{l}V_\mu^{(j)}\varphi_{h\mu}^{(j)}(i),\ i=\overline{k+1,I}.$$

CONCLUSIONS

In the paper is formed prediction algorithm vector random sequences. Method well as the canonical decomposition, put in its basis, fully take into account for each component of all known information about the target sequence. This ensures the absolute minimum mean square error linear prediction for an arbitrary component. It is also an expression for the mean square error of extrapolation, which allows to evaluate the quality of solving the problem of forecasting for any number of dimensions and the number of components of the study of vector random sequences.

REFERENCES

1. **Oppenheim E. 1980.** Digital processing of speech signals. Moscow: Mir, 323.
2. **Kolmogorov A. 1941.** Interpolation and extrapolation of stationary random sequences. Izv. USSR Academy of Sciences. Ser. Math, 3–14.
3. **Wiener Y. 1961.** The interpolation, extrapolation and smoothing of stationary time series. N.Y.: J. Wiley,. 341.
4. **Kalman R. 1961.** New Results in Linear Filtering and Prediction. Journal of Basic Engineering. № 83, 95–108.
5. **Ivakhnenko A. 1991.** Begin inductive theory of the uneven recognition and forecasting of the casual processes and event. Kiev: Institute of the cybernetics AN Ukraine, 348.
6. **Ivakhnenko A.** 1975. Long–term forecasting and management complex system. Kiev: Tehnika, 389.
7. **Box J. 1974.** Time series analysis, forecasting and management. Moscow: Mir, 230.
8. **Kalman R.E. 1960.** A new approach to linear filtering and prediction problems. Basic Engg. v. 82 (Series D). 35– 45.
9. **Ivakhnenko A., Lapa V. 1971.** Prediction of random processes. Kiev: Naukova Dumka, 421.
10. **Pugachev V. 1962.** The theory of random functions and its prominence. Moscow: Fizmatgiz, 751.
11. **Kudritsky V. 1982.** Predicting the reliability of radio devices. Kiev: Tehnika, 217.
12. **Atamanyuk I.P. 2005.** Algorithm of Extrapolation of a Nonlinear Random Process on the Basis of Its Canonical Decomposition. Cybernetics and Systems Analysis. Volume 41, Issue 2, USA, 267-273.
13. **Atamanyuk I., Kondratenko V., Kozlov O., Kondratenko Y. 2012.** The algorithm of optimal polynomial extrapolation of random processes. Lecture Notes in Business Information Processing. 115 LNBIP, 78–87.
14. **Atamanyuk I., Kondratenko Y. 2012.** Polynomial algorithm for predicting the reliability of stochastic technical objects on the basis of the apparatus of canonical expansion. MOTROL. No2, 77–83.
15. **Atamanyuk I., Kondratenko Y. 2011.** Algorithm of extrapolation of nonlinear casual process on the base of its canonical decomposition: Proceedings of the First International Workshop Critical Infrastructure Safety and Security (CrISS-DESSERT'11). Volume 2, 308–314.
16. **Atamanyuk I.P. Kondratenko Y.P. 2013.** Information Technology of Determination of Descriptions of Optimum Polinomial Prognosis Algorithm of the State of Technical Systems. MOTROL. Motorization and Energetics in Agriculture, Lublin, Poland. Vol. 15 (2), 47-50.
17. **Kudritsky V.D. 2001.** Filtering, extrapolation and recognition implementations of random functions. K: Fada, Ltd. 176.
18. **Atamanyuk I.P. 1998.** Canonical decomposition of the vector random process taking fully into account the correlations for each component. Engineering. Bulletin ZITI. №8, 119-120.
19. **Atamanyuk I. 2001.** Implementation algorithm of nonlinear random sequence based on its canonical decomposition. Electronic simulation. №5, 38–46.
20. **Atamanyuk I.P. 2010.** Simulation algorithm implementations nonlinear vector random sequences based on the apparatus of canonical expansions. Cybernetics and Computer Science.№ 162, 82-91.

МЕТОД ОПТИМАЛЬНОЙ ЛИНЕЙНОЙ ЭКСТРАПОЛЯЦИЯ ВЕКТОРНЫХ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С ПОЛНЫМ УЧЕТОМ КОРРЕЛЯЦИОННЫХ СВЯЗЕЙ ДЛЯ КАЖДОГО КОМПОНЕНТА

Аннотация. Работа посвящена решению важной научной и технической проблемы формирования оптимального метода среднеквадратичной линейной экстраполяция реализаций векторных случайных последовательностей любого числа известных значений, используемых для прогноза. Разработанный метод в отличие от существующей проблемы решения прогнозирования, в полной мере учитывает априорную информацию о последовательности для каждого компонента.

Прогноз модели синтезировали на основе линейного вектора канонического разложения случайных последовательностей. Формула для определения среднего квадрата ошибки экстраполяции, что позволяет оценить точность решения задачи предсказания с помощью предложенного метода для некоторого фиксированного числа известных значений и компонентов векторной последовательности. В работе также показано блок-схему алгоритма для определения параметров предлагаемого способа.

Метод экстраполяции, а также вектор канонического распределения разложения, поставить на его основе, не накладывает каких-либо существенных ограничений на класс прогнозируемых последовательностей (Марков, стационарные, скаляр, монотонность, и т.д.). Учитывая рецидивирующий характер расчета оценочных значений будущих значений методом целевой последовательности достаточно просто в вычислительном.

Так как большинство физических, технических, экономических или других реальных процессов стохастический, предложенный метод имеет широкий спектр применения в решении задач управления в различных областях науки и техники: интеллектуального контроля надежности технических устройств, медицинской диагностики, радаров, управление технологическими объектами и так далее.

Ключевые слова: векторные случайные последовательности, линейные канонические алгоритмы разложения экстраполяции.