them to the required format.

As a conclusion I want to note that the search for effective solutions to problems of processing of big-data is an important task and that development of systems handling such data should be considered as features of processed data and scenarios of their use.

REFERENCES

1. *Vendrov A. M.* The software Design of economic information systems. M. : "finances and statistics", 2006. 376 p.

2. *Argerich L.* Professional PHP programming, 2nd edition. SPb. : Symbol-Plus, 2003. 1048 p.

3. *Tindova M. G., Zhuravskaya K. G.* Development applications for maintenance loans // Evraziyskiy Union of scientists. 2015. № 8 (17). Part 1. P. 83-85.

4. Tindova M. G. the Use of modern information technologies in the development of instrumental means of assessment of natural resources // Information security of regions. 2013. № 1. P. 47.

5. *Orlov S. A.* Technology development programs. SPb.: Piter, 2004. 609 p.

MODELING MULTIVALUED DYNAMIC SERIES OF FINANCIAL INDEXES ON THE BASIS OF MINIMAX APPROXIMATION IN THE HAUSDORFF METRIC WITH THE CONSTRAINT

I. Yu. Vygodchikova

Saratov State University named after N. G. Chernyshevsky, Russia
E-mail: irinavigod@yandex.ru, vigodchikovaiy@info.sgu.ru

The problems of modeling a time series by various methods were studied by many authors, including, D.A. Dickey, W.A. Fuller [1], C.A. Sims, J.H. Stock, M.W. Watson [2], Schumaker L. L. [3]. However, the question of mathematical modeling of the multivalued dynamic series regardless of the distribution law, with additional restrictions, little studied in the literature. This paper examines the model of approximation of the multivalued time series, represented as a series of ranges of numerical values of the indicator of financial market, with constraints to approximating function.

For approximation of the dynamic number of applicable optimization problem of minimizing the maximum of the Hausdorff distances between the ranges of the dynamic series and the values of the approximating function.

The purpose of this paper is to develop a methodology of mathematical modeling of multivalued dynamic series, in the presence of restrictions on the values of the approximating function, the development of rational algorithm.

Introduction. Will call by the multivalued dynamic series a sequence ordered in time ranges of numerical values of some object. In the literature was not considered the criterion of uniform approximation to be used for multivalued mappings using the Hausdorff distance [4], including the restrictions on the approximating function. One of the effective methods for the analysis of unambiguous time series of different nature is the criterion of uniform approximation by Chebyshev [5]. However, when considering the ranks of the specified ranges, there is a problem with the rationale for the selection of the point inside the range, which will allow you to achieve the most adequate model approximation. It is especially difficult to resolve this issue, if there is no information about the law distribution of financial index within the range.

1. Model examples. We have the multivalued dynamic series of financial's market index (for example, price of share). For every point, t_k, $T = \{t_0 < ...t_k ... < t_N\}$, $y_{2,k}$ is top and $y_{1,k}$ is bottom values of price, so range is $[y_{1,k}; y_{2,k}]$, $y_{1,k} \le y_{2,k}$, k=0,…,N.

Model example 1. Suppose that for one of the periods the price is fixed accurately, which can be connected, for example, with the presence of forward contract on purchase of securities during this period, or deliberate action of an investor to acquire a controlling stake (Fig. 1).

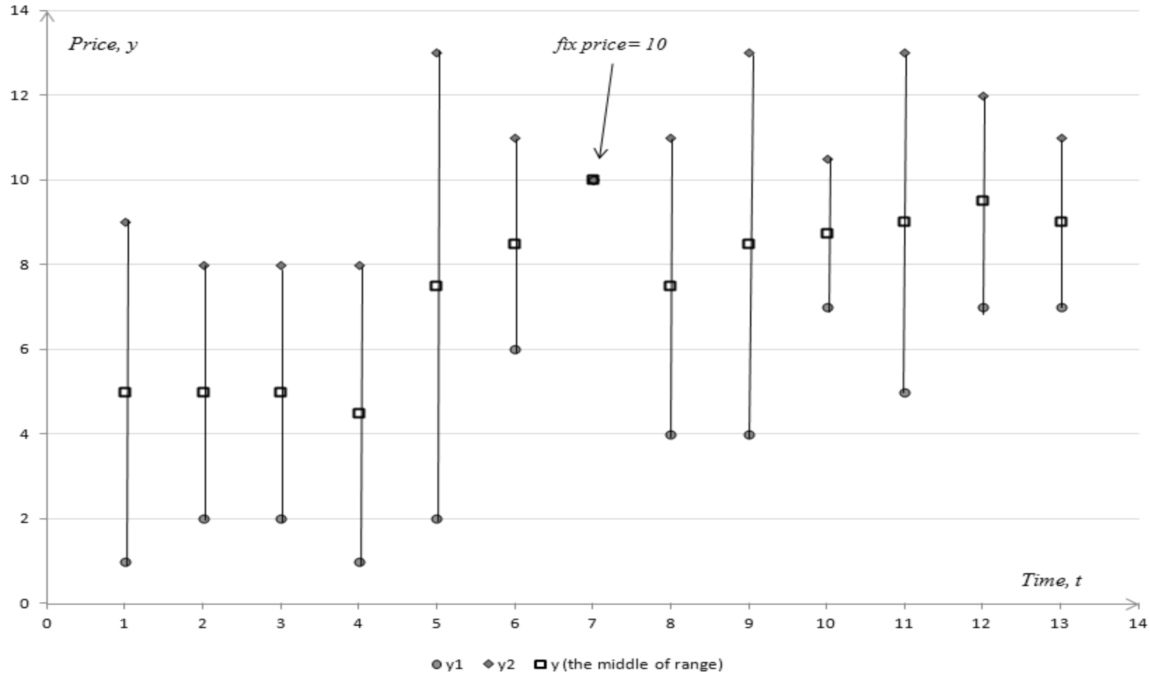

Fig. 1. Ranges and fix price for *t=7*

A mathematical model of multivalued number is represented as a polynomial $p_n(A, t_k) = a_0 + a_1 t_k + \ldots + a_n t_k{}^n$, we need to determine the coefficients of the polynomial $A = (a_0, a_1, \ldots, a_n) \in R^{n+1}$. If apply Ordinary Least Squares, OLS with constraint, for time series, composed of the midpoints of the ranges, we get the following result (fig. 2).

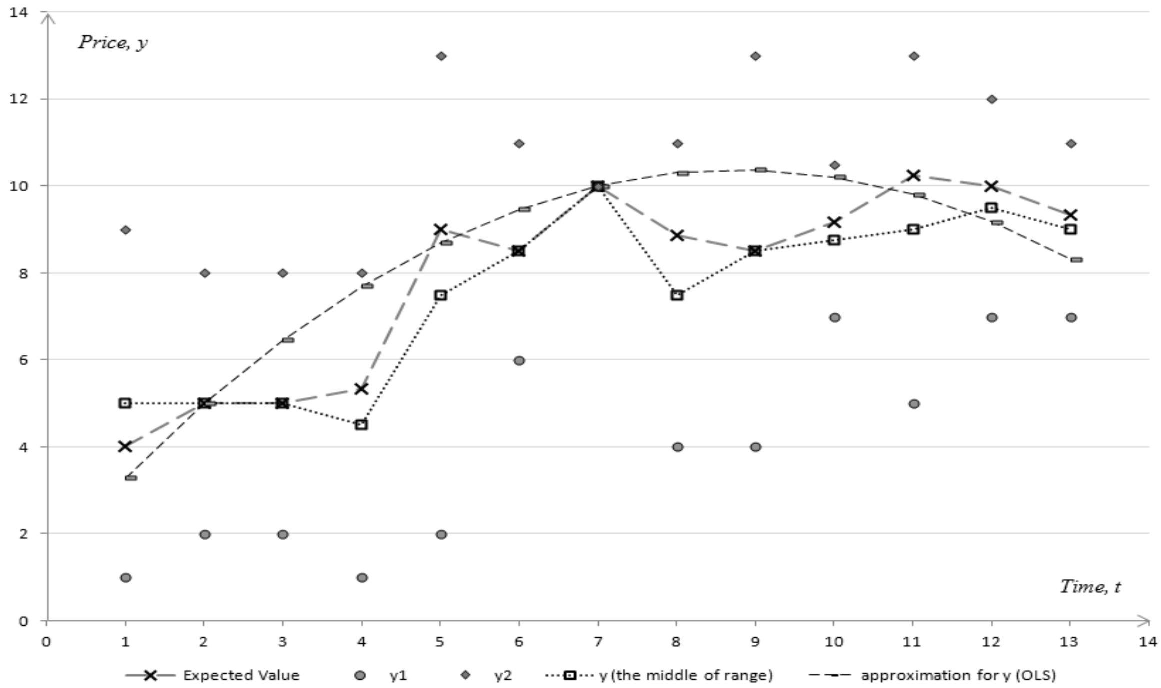

Fig. 2. The panorama distribution of prices for the model of example 1

Model example 1. Suppose that for two periods (*t=9, t=13*) prices fixed in ranges, correspondingly, from 8 to 9, and from 8,5 to 9. This may be due, for example, with the additional issue and primary placement of securities, option strategies, aimed at keeping prices in a fixed price corridor or deliberate actions of the owner of a controlling stake (Fig. 3).

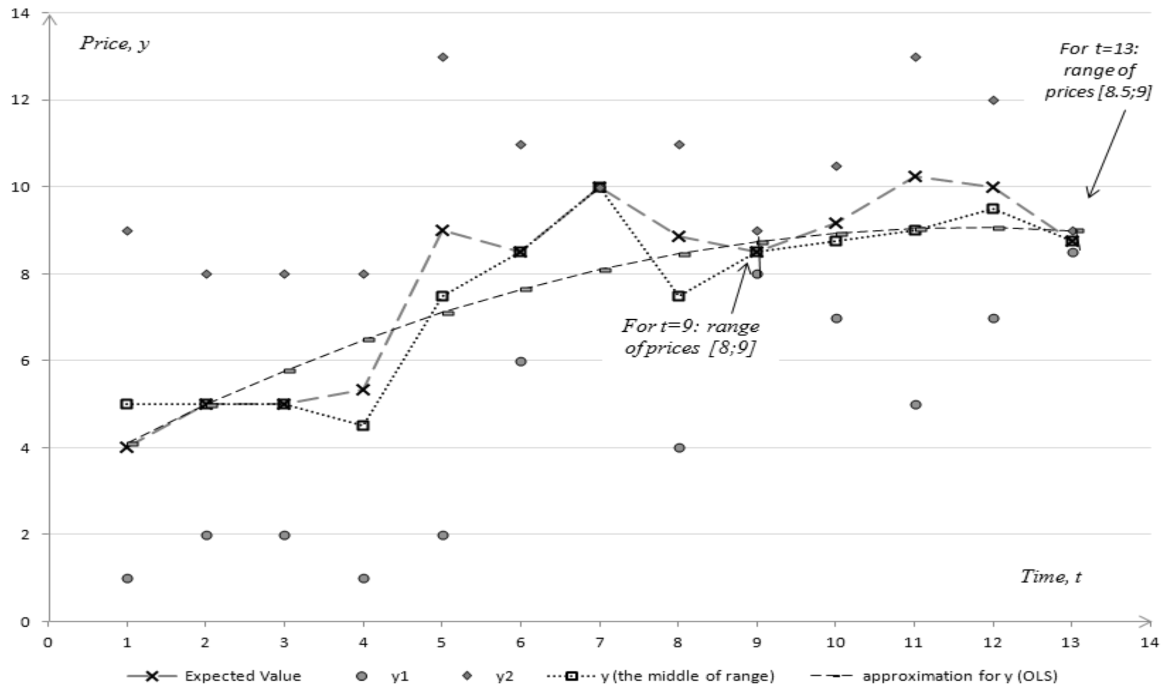

Fig. 3. The panorama distribution of prices for the model of example 2

2. Preliminary information on the Hausdorff metric. The purpose of the further analysis was to develop a method of approximation for ranges, more successful than that obtained using the method of least squares.

With this purpose, as the approximation criterion we use the distance (metric) Hausdorff [4]. Hausdorff distance for some discrete values t_k equals to maximum of the difference between the highest value of the index $y_{2,k}$ and the value of algebraic polynomial, and the difference between the value of algebraic polynomial and the lowest value of the index $y_{1,k}$.

This implies that the polynomial coefficients can be found based on the condition of minimizing the maximum of the Hausdorff distance for a given discrete time series [6], [7]:

$$\rho(A) = \max_{k \in \overline{0,N}} \max\{ y_{2,k} - p_n(A,t_k); p_n(A,t_k) - y_{1,k} \} \longrightarrow \min_{A \in R^{n+1}} . \qquad (1)$$

Note that before in the literature has not considered the application of the criterion of uniform approximation for multivalued mappings using the Hausdorff distance.

Further assume that the solution to the problem without restrictions will not satisfy the required constraints, that is constraints are significant.

3. Constraints of type equality. Consider the criterion (1), for example, to limits associated with condition a fixed location of a point of an algebraic polynomial at a grid node, $t_s \in T$: $p_n(A,t_s) = y_s$ [8]:

$$\rho(A) = \max_{k=\overline{0,N}} \{ y_{2,k} - p_n(A,t_k); p_n(A,t_k) - y_{1,k} \} \to \min_{A \in D} , \qquad (2)$$

$$D = \{ A \in R^{n+1} : p_n(A,t_s) = y_s \},\ t_s \in T .$$

Give algorithm for the situation when $y_{2,s} = y_{1,s} := y_s$. In [8] is proved, that decision will be the only. From [8] imply that to obtain the coefficients of the approximating polynomial in problem (2) it is necessary to carry out the iterative solution of systems of linear equations:

$$\varphi_{0,j_k} - p_n(A_0, t_{j_k}) = (-1)^k h_0,\ k = \overline{0,n+1} \setminus \{r\},\quad p_n(A, t_{j_r}) = y_s ,$$

or

$$\varphi_{1,j_k} - p_n(A_1, t_{j_k}) = (-1)^k h_1,\ k = \overline{0,n+1} \setminus \{r\},\quad p_n(A, t_{j_r}) = y_s ,$$

$$t_s \in \{ t_{j_0} < \ldots < t_{j_{n+1}} \} \subset T .$$

The solution will be polynomial coefficients A_i, for which turned out $\rho(A) = h_i$ for *i=0* or *i=1*.

That is from the original multivalued time series is selected arbitrarily (*n+2)* selector corresponding to a discrete ascending points in time t_{j_k}, $k = 0,\ldots,n+1$, and among these nodes necessarily must be the knot t_s.

Are considered two variants of choices of the tops and bottoms of the ranges.

Each time alternates – the upper, the lower bound of the range, as before, by considering the two choices of bound of the range directly left & right of the point constraints: the first option is taken the upper bounds of the ranges, and in the second embodiment, the lower bounds of the ranges. The procedure for changing of basis is optimized [8].

4. Constraints as ranges. Consider the criterion (1), for example, to constraints [9]:

$$\rho(A) = \max_{k=\overline{0,N}} \{ y_{2,k} - p_n(A,t_k); p_n(A,t_k) - y_{1,k} \} \to \min_{A \in D}, \qquad (3)$$

where $D = \{A \in R^{n+1} : u_{1,s_j} \le p_n(A,t_{s_j}) \le u_{2,s_j}\}$, $u_{1,s_j} < u_{2,s_j}$, $t_{s_j} \in T$, $s_j \in \{0,...,N\}$.

The polynomial in the all points t_{s_j} must be included in ranges $[u_{1,s_j}; u_{2,s_j}]$.

In [9] it is proved that to obtain the coefficients of the approximating polynomial is necessary to carry out the iterative solution of systems of linear equations, the iterative process is finite and leads to a decision that is proved.

That is from the original multivalued time series is selected arbitrarily *(n+2)* selector corresponding to a discrete ascending points in time t_{j_k}, $k = 0,...,n+1$, and among these nodes necessarily must be the knot all t_{s_j}.

The choice of the selector of a multivalued mapping from (N+2) points similar to the paragraph 3., When we consider the top bound of constraint "from left and from right" we use top bounds of the ranges, but when considering the lower bound of constraint "from left and from right" we use the lower ends of the ranges.

5. Computational experiments. Consider the results of applying the algorithm for example 1 (Fig. 4).

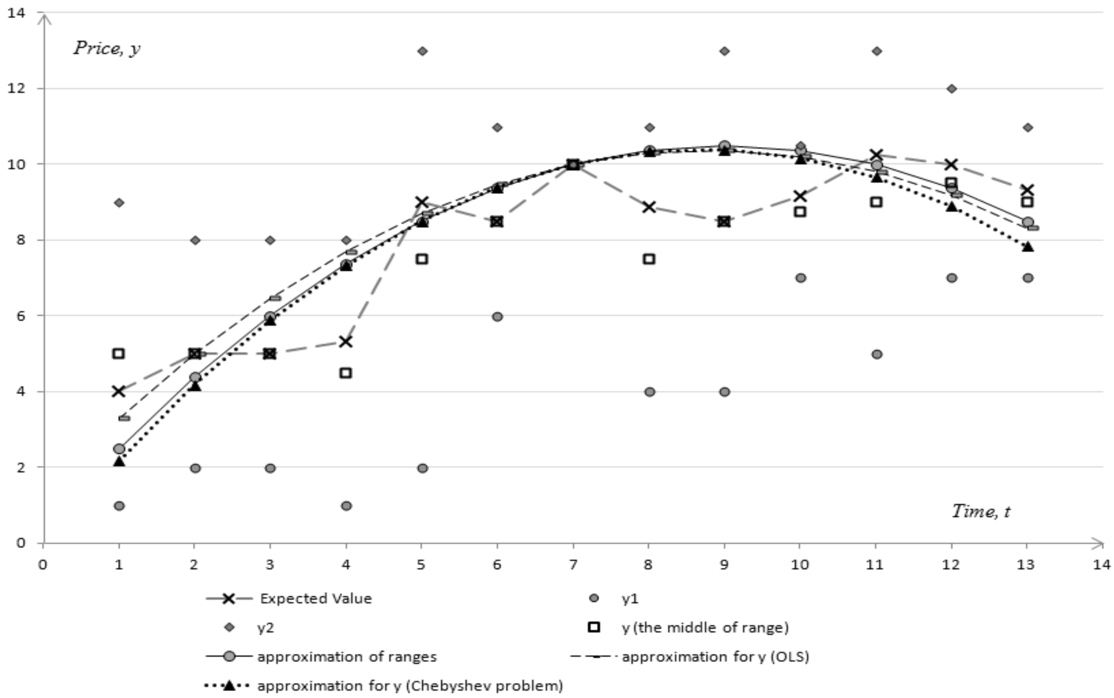

Fig. 4. Results of the experiment for the example 1

Assessing the adequacy of the approximation models obtained using the three methods, for example 1, are presented in table 1.

Table 1

Analysis of the significance of models of approximation for example 1

The significance of the correlation coefficient according to t-criterion			
	approximation of the ranges	approximation for the middle of ranges (OLS)	approximation for the middle of ranges (Chebyshev problem)
correlation coefficient	0,887695477	0,87244394	0,863366878
$t_{experim.}$	6,39434049	5,920779837	5,674982726
$t_{cr.}$	2,200985159	2,200985159	2,200985159
significance	significant	significant	significant

Consider the results of applying the algorithm for example 1 (Fig. 5).

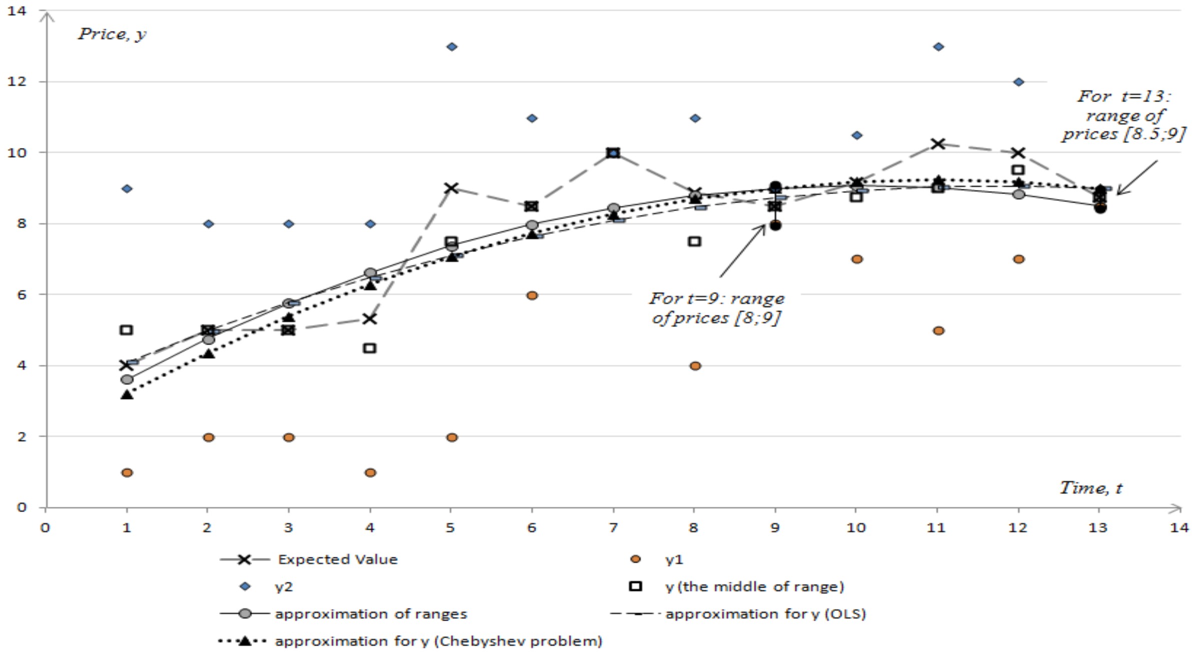

Fig. 5. Results of the experiment for the example 2

Assessing the adequacy of the approximation models obtained using the three methods, for example 2, are presented in table 2.

Table 2

Analysis of the significance of models of approximation for example 2

The significance of the correlation coefficient according to t-criterion			
	approximation of the ranges	approximation for the middle of ranges (OLS)	approximation for the middle of ranges (Chebyshev problem)
correlation coefficient	0,921398544	0,916088451	0,920735588
$t_{experim.}$	7,863571487	7,577310875	7,82633367
$t_{cr.}$	2,200985159	2,200985159	2,200985159
significance	significant	significant	significant

Conclusion. Was considered model of approximation of price ranges by algebraic polynomial. Model is based at the far-reaching generalization of the minimax criterion of Chebyshev. Was given account the possibility of applying for approximation of the data represented by ranges, as well as accounting additional restrictions on the approximating function. Developed a methodology to solve optimization problems, allowing to obtain high-quality approximation. The algorithm developed for solving discrete problems of the best uniform approximation of multivalued mapping with the specified ranges, algebraic polynomial with constraints on the value of the approximating polynomial on the top and bottom in several nodes of the grid. Computational experiments showed high-quality of the suggested method of approximation.

Research was supported by grant RFBR 16-06-00582.

REFERENCES

1. *Dickey D. A., Fuller W. A.* Distribution of the Estimators for Autoregressive Time Series with a Unit Root // Journal of the American Statistical Association. 1979. № 74. P. 427- 431.

2. *Sims C. A., Stock J. H., Watson M. W.* Inference in Linear Time Series Models with Some Unit Roots // Econometrica. 1990. № 58. P. 113-144.

3. *Schumaker L. L.* Spline Functions. Basic Theory // Cambridge University Press. 2007. P. 534-575.

4. *Sendov Blagovest Khristov* Hausdorff approximations. Sofia, 1979. 372 p.

5. *Dem'yanov V. F., Malozemov V. N.* Vvedenie v minimaks [Introduction to minimax]. M. : Nauka, 1972. 368 p.

6. *Vygodchikova I. Yu.* O metode approksimatsii mnogoznachnogo otobrazheniya algebraicheskim polinomom // Vestnik SGTU. Seriya : Matematika i mekhanika. 2013. № 2 (70). P. 7-12.

7. *Vygodchikova I. Yu.* O edinstvennosti resheniya zadachi nailuchshego priblizheniya mnogoznachnogo otobrazheniya algebraicheskim polinomom // Proceedings of Saratov University. The new series. Series: Mathematics. Mechanics. Informatics. 2006. T. 6. № 1-2. P. 11-19.

8. *Vygodchikova I. Yu.* O modifikatsii algoritma Valle-Pussena dlya approksimatsii mnogoznachnogo otobrazheniya algebraicheskim polinomom s ogranicheniem tipa ravenstva // Proceedings of Saratov University. The new series. Series : Mathematics. Mechanics. Informatics. 2014. T. 14. № 4-2. P. 526-532.

9. *Vygodchikova I. Yu.* O priblizhenii dvuznachnoi funktsii algebraicheskim polinomom [Approximation of double-valued function by an algebraic polynomial] // Izvestiya vysshikh uchebnykh zavedenii. Matematika. 2016. № 4. P. 8-13.