

УДК 519.688.336.76.066

Запропонована модель, що дозволяє обчислювати значення ймовірності розвороту тренду на фінансовому ринку. Модель заснована на припущенні, що час появи розладнання статистичних характеристик ряду котирувань є випадковою величиною, функція розподілу якої повільно змінюється у часі. Наведені методи знаходження точок розладнання, відтворення щільності розподілу форми ділянок між розладнаннями, обчислення функції шуканої ймовірності

Ключові слова: ймовірність, В-сплайн, відтворення щільності розподілу, методи пошуку розладнання, фінансовий ринок

Предложена модель, позволяющая вычислять значение вероятности события разворота тренда на финансовом рынке. Модель основана на допущении, что время появления разладки статистических характеристик ряда котировок является случайной величиной, функция распределения которой медленно изменяется во времени. Приведены методы нахождения точек разладки, восстановления плотности распределения формы участков между разладками, вычисления функции искомой вероятности

Ключевые слова: вероятность, В-сплайн, восстановление плотности распределения, методы поиска разладки, финансовый рынок

MODEL OF PROBABILISTIC ASSESSMENT OF TREND STABILITY AT FINANCIAL MARKET

O. Lutsenko

Post-graduate student*

E-mail: lutsenkoop@yandex.ua

O. Baybuz

Professor*

*Department of Mathematical Software
Dnipropetrovsk National University

Oles Honchar

K. Marks avenue, 35, Dnipropetrovsk,
Ukraine, 49010

E-mail: obaybuz@inbox.ru

1. Introduction

Systems of decision-making support in the financial sphere are one of the most topical directions in the development of information technologies. Such systems can provide significant competitive advantage as compared to other market participants, and, as a result, attract considerable interest in the sphere of financial trade. The development of new technologies for the market condition forecasting and risk evaluation, as well as the study of the possibility of using modern methods of analysis for forecasting the currency market are especially actual problems, considering the increase in popularity of Forex in recent years.

2. Analysis of literature data and problem formulation

The basis of the technical analysis is the statement that the market prices move directionally, depending on a certain tendency (trend). The ascending, descending and lateral tendencies can be distinguished. The task of the technical analysis is the definition of the nature of current tendency and identification of the moments of its change by the analysis of the history of price changes at the market.

The methods, which are the most frequently used in the technical analysis include auxiliary constructions on the diagrams of currency exchange rates (lines of resistance and support, trend lines, etc.), indicators and oscillators. There are many variations and modifications of these methods, but most of them are characterized by the following common disadvantages [1, 2]:

1) the subjectivity of auxiliary constructions, which excludes its automation;

2) significant time delay, caused by the use of indicators and oscillators of the moving averages in calculations;

3) the large number of false signals of the trend reversal when using oscillators.

Recently, due to the fact that computers achieved the necessary level of computing opportunities, more integrated approaches to the solution of technical analysis problems, namely approaches with the use of neural networks [3], genetic algorithms [4, 5], visual image recognition technologies [6] started gaining popularity. The commercial software products, using the technologies of neural network construction and complex indicators optimization on the basis of existing ones (Neuroshell Trader, BioComp Profit, Brain Maker Professional) and technology of market strategy optimization using genetic algorithms (Neuroshell Trader, Wave59) are widely presented in the software market.

3. Objective and research tasks

The objective of the paper is the development of the model, which favorably differs from the existing approaches of technical analysis by the presence of physical sense: the value at the output of the model is not abstract, it is a probability of a quite certain event - reversal of the market trend during the time interval in future.

The probabilistic nature of the model provides the maximum transparency for the user, making decisions in the course of trade that is especially topical in conditions of abundance of the factors, influencing the changes in the exchange rate and complicating the deterministic forecasting.

Furthermore, the distinctive feature of the risk evaluation method, suggested by the authors of this paper, is the

consideration of the time series of currency quotations not as a set of separate points, but as a set of segments, into which it is split by the change points. Finding of the change points of the series allows to divide the series into segments with similar statistical properties that corresponds to one of the principle paradigms of the technical analysis – division of the time series into segments with a constant tendency.

4. Description of suggested model of probabilistic assessment of trend stability

The model is based on mathematical techniques of detecting the change-points of statistical characteristics of the series and the method of probability density estimation using splines. At the first stage of data processing, the input numerical series is split into segments, on which the trend direction is invariable. After that, there is the estimation of two-dimensional probability density function of distribution of duration of such segments and the difference between final and initial series values on the segment. Based on the obtained distribution function, the probability function of the required event – trend reversal, is calculated.

4. 1. Splitting of the series by the change points

For the solution of the problem of the series splitting, two methods are used: the graphic method of searching the inflection points of the diagram, aimed at finding the reversal points a posteriori, and the CUSUM test, which belongs to the methods of sequential change-point detection.

The graphic method is highly accurate and is assigned for the construction of the main sample of the stationarity segments, on the basis of which the initial value of the probability density function of the change-point distribution is estimated. This method is based on the following principle.

Let's determine the initial and the final moment in the process, in which we search the change point: m – the initial time moment, n – final time moment. For this segment, let's construct the array:

$$\{d_i, i = \overline{n, m}\},$$

$$d_i = \frac{|kx_i - y_i + b|}{\sqrt{k^2 + 1}}, \tag{1}$$

where k, b - found coefficients of the equation of the straight line ($y=kx+b$), passing through the points (m, x_m) and (n, x_n) .

The change point corresponds to the time moment r , such that:

$$x_r = \max_{i=m, n} d_i, r \in (m, n), x_r > \text{bar}, \tag{2}$$

where bar is the sensitivity barrier of the method.

Once the change point is found, the time moment r breaks the interval $[m, n]$ into two intervals, which then we consider in the similar way, initiating the function recursively. The process is finished when none of the values of d_i on the interval exceeds the established barrier.

The CUSUM algorithm is assigned for the change-point detection in the latest obtained data. It lies in the analysis of the behavior of the value [7]:

$$S_t = S_{t-1} + \ln(\omega(x_t / \theta_2) / \omega(x_t / \theta_1)), \tag{3}$$

where $\omega(x_t / \theta)$ is the probability density function of x_t , θ is the scalar parameter of density.

At each step the sum is compared to the set threshold. If at step t $S_t > h$, then the signal of change-point is set, the accumulation of the sum starts again from zero. A reflecting barrier of the function is set at the point 0.

$$S_t = \max(0, S_{t-1} + \Delta S),$$

$$\Delta S = \ln(\omega(x_t / \theta_2) / \omega(x_t / \theta_1)). \tag{4}$$

As we can see, such form of CUSUM requires the information about the value of the parameter θ after the change-point. Nevertheless, the use of this method in this case is complicated, as the exact value after the change-point is unknown. There is a way to weaken this requirement. The algorithm is adjusted so that to respond to any changes of the characteristics θ . Let m_2 be an unknown value, but the direction of its change is known, for example $m_2 > m_1$. Then (4) can be written as:

$$\Delta g_t = x - m_1 - k \quad \text{при } m_2 > m_1,$$

$$\Delta g_t = -x + m_1 + k \quad \text{при } m_2 < m_1, \tag{5}$$

where x is the value of observation at time t ;

m_1 is the expected value before the change-point;

$k \geq 0$ is a sensitivity threshold of the method.

Note that there is the procedure, equivalent to the simultaneous application of the formulas (5) [8], but for this task, the separation of conditions of setting the signal of change-point to “up” and “down” is important. The changes of characteristics towards the strengthening the current trend do not lead to the trend reversal and should not be considered.

Since the stable trend itself is characterized by changing of the mean values of observations, in this case it is expedient to apply CUSUM not to the initial series of observations, but to its first time derivative. The physical sense of such derivative is the rate of change of the function values. The sharp change of the first derivative will mean the change of the histogram direction that is the change-point of the initial process. In case when it is necessary to determine the change of the current trend, the reversal point of the diagram will be characterized by the fact that the value of the first derivative crosses the zero point. Eliminating the change-points, in which the observation change is directed to the opposite side from the zero point (trend strengthening), and also the change-points, lying far from the zero point (trend slowdown, which, nevertheless, is still far from the reversal), we obtain the points, dividing the action segments of three major market trends.

4. 2. Construction of the probability density function of change-points

Suppose that after the diagram splitting we have n segments of stationarity. Each of the segments is characterized by the vector of parameters, including its duration in time (the length of the stationary segment along the axis $x - x_s$) and the difference between the prices at the beginning and the end of the segment (difference on the axis $y - y_s$). Thus, the process described by the vector of random variables (x_s, y_s) .

The model, proposed by the authors, is based on the assumption that the time of the change-point appearance in the

statistical characteristics of the number of quotations is the random variable, the distribution function of which slowly changes in time. Having estimated the probability density of the distribution, we get the opportunity to judge on the probability of changing the direction of the market trend at every step of the observations.

To solve the problem of the density estimation, it is necessary to find a class of functions, which will allow to approximate the empirical data on the distribution with such degree of accuracy that the obtained result of approximation was the closest to the estimated theoretical dependence. However, it is necessary to consider that during solving the problem of the density estimation, the estimation of the quality measure of approximation with the use of the mean square error is not always acceptable. More important is the compliance of the resultant function to the distinctive properties of the input data, such as: the number of modes, asymmetry, tail behavior [9]. This is especially important in this case, since the distribution has a pronounced bimodal shape (Fig. 1) and varies over time. In view of the fact that such features are not retained in the approximation neither by any of the typical two-dimensional distributions, nor a mixture of distributions (using the EM-algorithm) for the distributions, obtained in this task, the transition to non-parametric methods for density estimation is expedient.

The problem of nonparametric distribution density estimation on the histogram lies in the construction of the approximating function on a rectangular grid of values with the restoration of the intermediate values. The first results in this direction were obtained in the papers [10, 11], where the histogram was considered as the estimation of the unknown probability density function. Almost certain convergence of the histogram to the continuous probability density function was established in [10], the limit law of distribution for the maximum of the absolute value of the normalized deviation of the histogram from the theoretical smooth density was obtained in [11].

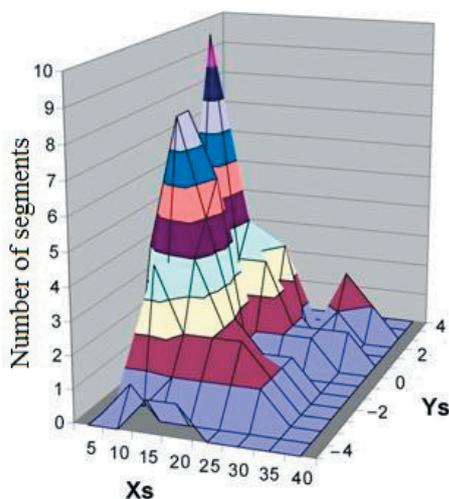


Fig. 1. Histogram of distribution of the vector (xs, ys)

B-splines were selected as the approximating function. At low computing complexity [12, 13], B-splines have good approximating properties. In addition, B-splines have the property of locality that allows to recalculate only a part of the function during the change of the source data. The

latter is especially actual, given the fact that in the course of obtaining the new market data it is necessary to update the density function.

The use of B-splines in the process of probability density estimation, in particular for the two-dimensional case was studied in [9, 14, 15].

Let's divide the distribution chart of (xs, ys) into rectangular segments with the size of dx to dy, count the number of hits on each segment and construct the histogram on the basis of the obtained values. The points of the histogram will serve as the spline nodes. We obtain the two-dimensional array of the node points with the dimension of m x n. The two-dimensional B-spline function can be written as the tensor product of the one-dimensional B-splines, constructed on each of the two axes of coordinates. In this case, the value of B-spline surface at any point will be equal [16]:

$$x(s, t) = \sum_{i=0}^n \sum_{j=0}^m P_{i,j} B_i(s) B_j(t),$$

$$B_{k,l} = \begin{cases} 1, & t_k < t < t_{k+1} \\ 0, & \text{otherwise} \end{cases}$$

$$B_{k,d} = \left(\frac{t - t_k}{t_{k+d-1} - t_k} \right) B_{k,d-1}(t) + \left(\frac{t_{k+d} - t}{t_{k+d} - t_{k+1}} \right) B_{k+1,d-1}(t), \tag{8}$$

where s, t are point coordinates;

P_{i,j} is the value of the approximated function at the point i, j;

B_{k,d}(t) is the value of the blending function of the order d at the point t, belonging to the interval (t_k; t_{k+1});

t_k; t_{k+1} are pairs of adjacent nodal points of the spline.

During the solution of the problem, the cubic blending functions of order 4 and open uniform knot vectors U = (-3, -2, -1, 0, 1, ..., n + 1), V = (-3, -2, -1, 0, 1, ..., m + 1) were used.

The example of the result of estimation of the probability density distribution of the vector (xs, ys) is shown in Fig. 2.

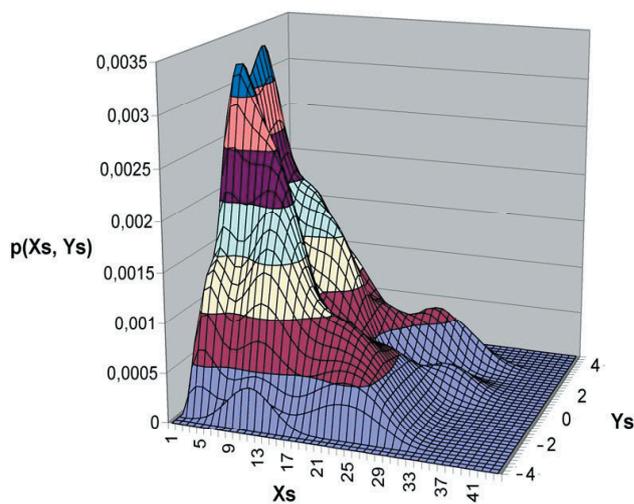


Fig. 2. Estimated density function (xs, ys)

4. 3. Calculation of the probability function values

The obtained density function is the density function of the trend reversal depending on the increment of the value of

the exchange rates from the moment of the last reversal and time. Knowing this density, as well as the current (since the last recorded change-point to the present moment) value (x_s, y_s), we can calculate the risk function, which expresses the probability of trend reversal at the next step of observations. To do this:

1) let's assume that the prices from disorder to disorder move directionally: either mainly up, or mainly down. The orientation of movement corresponds to one of the axioms of the technical analysis of the market and, in addition, the significant deviation of direction to the opposite side is the trend reversal itself. It's enough to define the current direction with the accuracy up to the sign of the price change from the moment of the last reversal. In other words, the assumption lies in the fact that the current price of the currency pair until the moment of the next trend reversal will change in the same direction, in which it changed from the moment of last reversal to the present moment, not the opposite;

2) define the expected value of the price change at the next step of observations. This value can be either a constant of the order of increment of the series, or the average value of the series increment in the selected direction for a number of l of previous time intervals. Possibility of insignificant shifts in the direction opposite to the trend does not contradict the first hypothesis, therefore, it is expedient to consider the possibility of changes in both directions (Δy_1 and Δy_2). Thus, we limit the area ($y_s - \Delta y_1; y_s + \Delta y_2$), to which the next value y_s the most likely belongs. As for x , we can definitely say that at the next step of observations, the value of the vector will increase by 1 as x is the time axis.

Knowing the distribution of intervals of stationarity, as well as the fact that the trend reversal is the event, which will happen sooner or later, we calculate the probability of the trend reversal by the analogy to the reliability function: as the relation of the probability P_1 of occurrence of reversal at the next observations to the probability P_2 of occurrence of this event at all other segments, to which the observed value may belong at future steps, according to the assumption 1 and taking into account the non-decreasing nature of x .

In the mathematical form:

$$P = \frac{P_1}{P_2},$$

$$P_1 = \int_x^{x+\Delta y_2} \int_{y-\Delta y_1}^y p(x,y),$$

$$P_2 = \begin{cases} \int_x^{x_{\max}} \int_{y-\Delta y_1}^{y_{\max}} p(x,y), & y \geq 0 \\ \int_x^{x_{\max}} \int_{y+\Delta y_2}^{y_{\min}} p(x,y), & y < 0 \end{cases}, \quad (9)$$

where $p(x, y)$ is the estimated probability density function;

P is the function, the physical meaning of which is the probability of the fact that the trend reversal will occur at

the next step of observation under the conditions, taken as assumptions above.

5. Approbation of the research results

Fig. 3 shows the example of the result of calculation of the function values on the historical values of the currency pair AUDJPY at the following input parameters: the size of the time interval of the past for the formation of the sample for density estimation is 45 days time interval is 10 minutes, time horizon for the calculation of the probability function is one step forward (10 minutes).

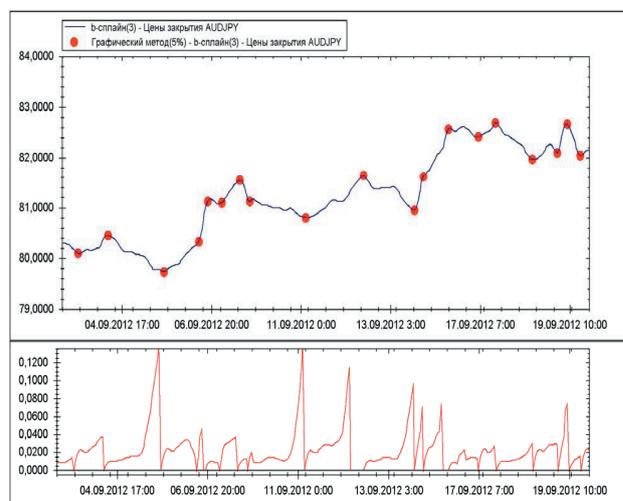


Fig. 3. Example of the graph of function P calculated on the currency pair AUDJPY with a 10-minute interval of time

The results of practical application of the function allow to conclude on its following characteristic features:

1. Function increases with the increase of distance from the starting point of the current trend, and the increase is quickly accelerated at the big distance from the previous change-point.

2. As the vector (x_s, y_s) reaches the anomalous value, the function turns into the uncertainty of the type (0/0) that is associated with the fact that the integrand in the system (9) takes the zero value. The feature was taken into account during the software realization of the model.

3. When calculating the function one step forward, the change-point occurs at relatively low values of probability. It is connected with the small value of the time horizon with respect to the length of the segment without the change-point (low probability of the event at a large number of experiments).

The obtained function can be used in the construction of trade strategies: either independently (rules of the type "if the probability of the termination of the ascending trend within the next 10 minutes is higher than 10 %, it is necessary to sell"), and as the specifying value ("to sell when the indicator gives the signal to sale, and the probability of the termination of the ascending trend in the next hour is higher than 30 %"). In addition, the function can be used as the function of confidence of the change-point detection under the condition of the bottom limit establishment of the values of P , on reaching of which the signals of change-point are considered as true.

6. Conclusions

The model of evaluation of the financial market state, based on the methods of change-point detection and probability density estimation using the spline functions, is proposed.

The model allows to calculate the probability function of the market trend reversal within the predetermined time interval in the future.

The knowledge of the exact value of the risk of the market trend reversal potentially allows to automate the choice of the take-profit value for the opened positions, based on

the expected profit that would be impossible when using the deterministic approaches; to formalize the decision-making about closing of the trade positions opened earlier by setting the threshold value of the acceptable risk. It is also possible to use the function value as the technical indicator, which has a high degree of visibility.

The prospects of further research in this direction: the determination of optimal sample size for the density estimation, selection of optimal density estimation parameters, depending on the time scale of the data, analysis of the effectiveness of the trade strategies based on the use of the obtained function.

References

1. Kahn, M. N. Technical Analysis Plain and Simple: Charting the Markets in Your Language [Text] / M. N. Kahn // FT Press – 2010 – 352 p.
2. Demark, T. R. The New Science of Technical Analysis [Text] / T.R. Demark // John Wiley – 1994 – 247 p.
3. Kuan, C. M. Forecasting exchange rates using feedforward and recurrent neural networks [Text] / C. M. Kuan, T. Liu // Journal of Applied Econometrics – 1995 – 10 – P. 347–364.
4. Mendes, L. A Forex trading system based on a genetic algorithm [Text] / L. Mendes, P. Godinho, J. Dias // Journal of Heuristics – 2012 – vol. 18, N 4 – 627–656.
5. Chan, L. Automated Trading with Genetic-Algorithm Neural-Network Risk Cybernetics: An Application on FX Markets [Text] / L. Chan, A. WK Wong // Finamatrix, Research and Markets – 2011 – Режим доступа : \WWW/ URL: <http://ssrn.com/abstract=1687763> – 01.11.2013.
6. Ciskowski, P. Neural Pattern Recognition with Self-organizing Maps for Efficient Processing of Forex Market Data Streams [Text] / P. Ciskowski, M. Zaton // Artificial Intelligence and Soft Computing – 2010 – 6113 – P. 307–314.
7. Page, E. S. Continuous Inspection Schemes [Текст] / E. S. Page // Biometrika – 1954 – vol. 41, N 1 – P.100–115.
8. Nadler, J. Some characteristics of Page's twosided procedure for detecting a change in a location parameter [Text] / J. Naddler, N. B. Robbins // Ann. Math. Statist. – 1971 – vol. 42, N 2 – P. 538–551.
9. Kooperberg, C. A study of logspline density estimation [Text] / C. Kooperang, C. Stone // Computational Statistics & Data Analysis – 1991. – vol. 12, N 3 – P. 327–347.
10. Гливенко, В. И. Курс теории вероятностей [Текст]: учебник для физико-математических факультетов государственных университетов / В. И. Гливенко. – М., Л.: ГОНТИ НКТП СССР, Редакция технико-теоретической литературы – 1939. – 220 с.
11. Смирнов, Н. В. Приближение законов распределения случайных величин по эмпирическим данным [Текст] / Н. В. Смирнов // Успехи Матем.Наук – 1944 – N 10 – С. 179-206.
12. Завьялов, Ю. С. Методы сплайн-функций [Текст] / Ю. С. Завьялов, Б. И. Квасов, В. Л. Мирошниченко – М.: "Наука" – 1980. – 352 с.
13. De Boor, C. Package for Calculating with B-Splines [Текст] / C. De Boor // SJAM J. On Numer. Anal. – 1977 – vol. 14, N3 – p. 441–472.
14. Koo, J.-Y., Tensor product splines in the estimation of regression, exponential response functions and multivariate densities [Text]: Ph.D. thesis / J. Y. Koo – Berkeley – 1988 – 132 p.
15. Stone, C. J. The use of polynomial splines and their tensor products in multivariate function estimation (with discussion) [Text] / C. J. Stone // Ann. Statist. – 1994 – N 22 – p. 118–184.
16. Rogers, D. F. An Introduction to NURBS with Historical Perspective [Text] / D. F. Rogers – San Francisco: Morgan Kaufmann Publishers – 2000 – 344p.