Volatility Analysis Strategy of Securitization Market Based on Markov Analysis Model

Junyi Du
junyi_du@ucsb.edu

Computer Science College of Engineering, University of California, Santa Barbara, Santa Barbara, California 93106

Abstract: Financial data analysis as well as its research on market fluctuation is an important research content in the financial field. A large number of analysts explore the operation law behind the data through the analysis of data change and fluctuation, to establish a reliable data financial analysis model and reduce the market risk. With the increasing complexity of data, the existing analysis models have been difficult to meet the market demand. Based on the summary of the discrete Markov analysis model system of volatility proposed by predecessors, the data envelopment and static tree analysis system are introduced to complete the data extraction and evaluation processing, and then the final data volatility results are obtained through the calculation of Brillouin monitoring data volatility. The analysis shows that the accuracy of data fluctuation evaluation of the method is significantly improved that can provide effective data theoretical guidance for investors' trading in finance, which has obvious practical advantages.

Key Words: finance analysis; data; static trees;

1 INTRODUCTION

In the field of modern financial development, data volatility analysis can be mainly divided into machine learning and traditional statistical analysis, but there are some problems for these two methods that have been difficult to meet the needs of financial analysis under the current social development. For example, the statistical model is so simplistic that in the process of data analysis, it only considers the trend of internal variables, and its core data must adopt stable and linearly related data [1]. Although machine learning can effectively deal with nonlinear data and high-dimensional data, its model is quite complex. In the process of calculation and analysis, it requires a large number of parameter optimization. There will be various computational logic problems once it is trapped into the deadlock of local extreme data extraction. Therefore, after summarizing the previous data analysis technology, the paper constructs a new financial data fluctuation analysis method based on Hidden Markov model (HMM). Markov model can effectively describe the data transfer between data hiding states and the observed measurement fluctuation under state data. These data analysis fluctuations are based on time series and changes over time. Therefore, the data fluctuation analysis strategy based on Markov model has stronger theoretical analysis and explanation. From the first application of Markov model in data analysis in the stock field by Hessian, a financial scientist, the analysis method of Markov model has formed a complete set of basic logic system, including data analysis and extraction. However, standardized operation is carried out

ICEMME 2022, November 18-20, Nanjing, People's Republic of China

DOI 10.4108/eai.18-11-2022.2326869

through parameter evaluation and algorithm training, and finally the data value is extracted, the predicted value is calculated to obtain the condition of the data fluctuation. A new logic system is added to the basic Markov model in this study. Firstly, data envelopment and static tree analysis are introduced, and data envelopment is used to make centralized decisions on various data indicators and influencing factors. Then, the data distribution structure is determined according to the production function, so as to establish the static tree analysis system, and use the linear specification to obtain the production frontier boundary of each economic data analysis door, and directly calculate the completion efficiency, input and output of the decision-making unit under each door; Markov model is introduced to complete parameter evaluation and preprocessing. Finally, the final data fluctuation results are obtained through Brillouin monitoring data volatility calculation. The analysis shows that the accuracy of data fluctuation evaluation of this method is significantly improved that has obvious practical advantages.

2 Data extraction

Firstly, we need to establish the central data flow to build a data envelopment analysis system. The flow data includes the actual analysis of various data analysis and evaluation indicators and relevant influencing factors. For different financial data, different data calculation methods can be established to complete the actual calculation. Specific parameters need to be considered in many aspects. After the data flow summary is completed, the data set can be extracted, and then the sampling extraction of the data center can be completed based on the development information of the financial region. In order to ensure the accuracy and scientificity of the data, it is designed to further extract the data from the data set through the data envelopment and static tree analysis system to complete the quantification of data logic analysis [2].

Data envelopment analysis (DEA) is an important derivative logical data algorithm in modern financial field operations research and mathematical economic analysis, which belongs to the cross derivative of mathematics and economics. Its core is to establish clear data logic evaluation indicators and data flow information and convert the fixed data in the data set into data output units or data output gates, being called data decision unit (DMU), and then the study can determine whether each data unit is valid data according to the data dissemination and data effectiveness among each decision unit, so as to complete the data extraction. In a word, the core of data analysis and decision-making unit is data efficiency production, and the effective data produced is extractable data [3].

In the field of data analysis, statistics and finance, the production of efficiency data is the premise of the construction of production function. This study uses the data envelopment analysis structure to determine the distribution of data production surface, so the data envelopment analysis method used can also be regarded as an unconventional comprehensive parameter evaluation template in the field of mathematics. Because the complexity of financial data leads to various decision-making units, the design further introduces the static tree analysis structure on the basis of basic data envelopment, and constructs a comprehensive analysis system of data envelopment and static tree, as shown in Figure 1.

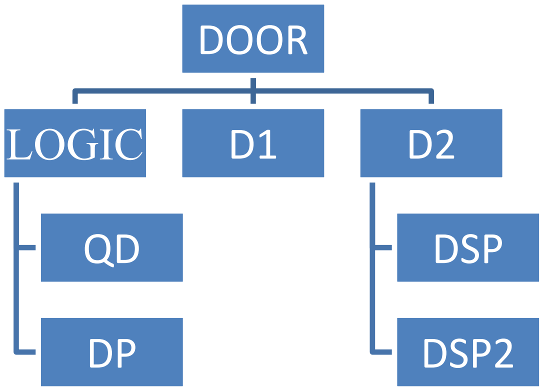

Figure1 Schematic Diagram of Static Tree Analysis Method

The static tree analysis method mainly uses the data linearization rules to directly calculate the data input and output of each decision-making unit under the data envelope by constructing the frontier data limit of economic data analysts with different structures. In the actual calculation, it is necessary to select different decision data first. Let the total amount of current financial data be N, and use M and S to represent the efficiency input and output of different decision-making units. Let the i-th input variable of the k-th decision-making unit in all decision-making units be represented by the symbol Xik (i = 1, 2, ... m), correspondingly, the j-th output of this series of decision-making units is represented by Yjk (i = 1, 2,... s), and the relevant calculation model is as follows:

$$\text{S,t}\begin{cases} \sum_{k=1}^{n} X_k H_k + s^- = \theta X_t \\ \sum_{k=1}^{n} X_k H_k - s^+ = Y_t \\ s^- \geq 0, s^+ \geq 0, H_k \geq 0, k = 1,2\ldots n \end{cases} \tag{1}$$

In the above formula, the s^- and s^+ in formula 1 respectively represent the economic coefficients in various data decision-making units; and H and θ unit data variables under different decision data; the value θ is the data influence variable under different financial information [4].

Formula (1) is mainly used for statistics and data analysis and calculation of the core decision-making efficiency of financial data, that is, the comprehensive efficiency is evaluated only at the technical level and scale level. When the value θ is 1, the value of s^- and s^+ is 0.

The efficiency obtained by formula (1) can be expressed by λ. In order to effectively distinguish the core values of the current data, different data areas are designed and divided:

λ12，λ13，λ24 andλ35. In the data evaluation, if the efficiency value of the current data is set as a fixed value, the above value range variables and the corresponding output are respectively

$$\lambda 12=\lambda 13=\lambda 24=\lambda 35=3.5752*10^{-5}/h \quad (2)$$

$$u12=u13=u24=u35=(1/8)/h \quad (3)$$

Where **h** is the comprehensive data coefficient of Solomon, which needs to be taken according to the actual situation of current financial data. Before the actual calculation, the static tree analysis structure is further changed into static data analysis chain, which can effectively reduce the operation law and improve the comprehensive operation results. As shown in Figure 2

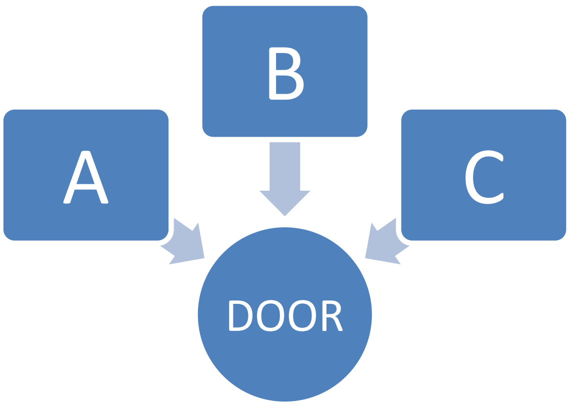

Figure2 Static Tree Analysis Chain

Based on Figure 2 and formulas (2) and (3), the input and output analysis matrix of the decision-making unit is established to calculate the boundary performance value A.

$$A=\begin{bmatrix} -(\lambda 12+\lambda 13) & \lambda 12 & \lambda 13 & 0 & 0 \\ u12 & -(u12+\lambda 24) & 0 & \lambda 24 & 0 \\ u13 & 0 & -(u13+\lambda 35) & 0 & \lambda 35 \\ 0 & u24 & 0 & -u24 & 0 \\ 0 & 0 & u35 & 0 & -u35 \end{bmatrix} \quad (4)$$

The justification effect value obtained by formula (4) can represent the data perfection degree in the current work. The above DEA analysis method is used to convert the input values of various decision-making units into boundary effect values, which are further projected into the geometric space, and the lowest input and highest output values are obtained as boundary values. Let the boundary performance of the decision-making unit be the product of spatial boundary. At this time, the decision-making unit (DMU) can be used as the effective value, and its relative efficiency refers to 1. At this time, it can also be considered that the decision-making unit has not effectively reduced and increased the data output. In this way, the output data is extracted. At this time, the DMU is the data within the boundary. If the value of given data is between 0 and 1, the input data remains unchanged, the output can be further increased, and the data can be extracted [5].

3 Parameter Evaluation and Preprocessing

After data processing, the parameters can be evaluated and preprocessed by Markov model. There are two kinds of preprocessed input data and target data, namely fractional value class and fractional ratio class. The latter has higher data dispersion and poor distribution, so standardized preprocessing is required in data analysis.

Design and introduce the standardized processing operation of Z-score with Markov model, specifically, firstly obtain the average value of the output data, divide it by its standard deviation, and construct its normal distribution. For the classified data, after standardization, it is necessary to understand the performance characteristics of relevant variable data and current data. In other words, users need to analyze the input variables under the current financial data variable model, which can also be used as a direct part of optimization. Firstly, the design starts to measure the correlation Pearson chi-square data. Generally, the set coefficient is Clem coefficient and the chi-square is LAMBDA CHI square coefficient. For different correlation data, the input variable data is different from the target variable data. Generally, the cross data needs to be transformed into continuous variable data. At this time, the larger the set Clem coefficient is, the greater the correlation between the numerical input data and the target linear data, the Pearson formula used is as follows:

$$r_{xy}=\frac{\sum_{i=1}^{n}(x_i-x^{-})(y_i-y^{-})}{\sqrt{\sum_{i=0}^{n}(x_i-x^{-})2\sum_{i}^{n}(y_i-y^{-})2}} \tag{5}$$

In the above formula, x is the input variable in the financial data set; y represents the current financial target variable (the variable needs to be divided according to the financial market); $\overline{x}$ represents the mean value of financial fluctuation variables; $\overline{y}$ represents the average value of financial data variables; x_i represents the observed value of variable x.

When the input variables that have been analyzed and calculated are in the classified state, the financial fluctuation data at this time also has the classification attribute. The design uses the likelihood ratio book tool for development and sick leave for data processing. Likelihood ratio is similar to Pearson chi-square, and its calculation results are more distinctive and consistent. The formula of likelihood ratio square is defined as follows:

$$T=2\sum_{i-1}^{r}\sum_{j=1}^{c} f_{ij}\ \ln\frac{f_{ij}^{0}}{f_{ij}^{c}} \tag{6}$$

In the formula, i represents the variable of input financial data; J represents financial target variable; f_{ij}^{o} represents the frequency of the observation of the fluctuation variable; f_{ij}^{c} indicates the prediction frequency of the expected variable;

When the input variable of financial data is classified data, the target variable financial data will also become classified data, and the proportion of Cramer coefficient begins to increase. At this time, Pearson formula can be used to adjust the variable of coefficient. Because the V coefficient can modify the Pearson coefficient, if the value of V coefficient starts to approach 1, it indicates that the correlation of the target variables of the data is enhanced. The V coefficient formula is as follows:

$$V=\sqrt{\frac{x^2}{n.min[(R-1),(C-1)]}} \tag{7}$$

In the formula, R represents the number of rows in the data contingency table; C indicates the number of columns of data concatenation;

4 CALCULATION OF BRILLOUIN MONITORING DATA VOLATILITY

Brillouin monitoring technology mainly relies on distributed financial linear data. Its principle is to accurately extract the fluctuation data of any point on the Markov data chain through the data feedback sensing sequence composed of frequency points, and determine the real-time fluctuation of the current monitoring target data through its own unique data calculation template and data input.

The core linear relationship formula is:

$$\Delta Va=Cv_T \Delta T+Cvc \Delta \varepsilon \tag{8}$$

$$\Delta Pa=Cp_T \Delta T+Cps \Delta \varepsilon \tag{9}$$

Solve the formula (4) and formula (5)

$$\begin{pmatrix}\Delta T\\ \Delta \varepsilon\end{pmatrix}\begin{pmatrix}C_{vT} & C_{vc}\\ C_{pT} & C_{ps}\end{pmatrix}\begin{pmatrix}\Delta V_a\\ \Delta P_a\end{pmatrix} \tag{10}$$

In the above formula, ΔT and $\Delta\varepsilon$ represents the financial change, C_{vT}, C_{vc}, C_{pT}, and C_{ps} respectively represent the current data hiding state probability, data displacement, risk strain coefficient and the observed value of system data. The following formula can be determined according to formula (11) (12):

$$\Delta T = \frac{1}{C_{vT}C_{ps} - C_{vs}C_{PT}} \left\{ -C_{ps} \Delta Va + C_{vc} \Delta P \right\} \tag{11}$$

$$\Delta \varepsilon = \frac{1}{lC_{vT}C_{ps} - C_{vs}C_{PTl}} \left\{ -C_{ps} \Delta Va + C_{vc} \Delta P \right\} \tag{12}$$

Therefore, when the current financial data and the current system strain system are known, the current pipeline strain information can be obtained according to the risk frequency offset to realize data generation.

According to the above established principle model, the volatility equations of system data can be established:

$$\nabla^2 E - \left(\frac{n}{c}\right)^2 \frac{\partial^2}{\partial t^2} - \frac{an}{c} E = \frac{\gamma_e}{p_0 c^2} \frac{\partial^2 (pE)}{\partial t^2} \tag{13}$$

Among which, α refers to the volatility coefficient; ρ indicates financial data density; the formula represents the observation sequence of Markov model as a whole.

According to the input items of the observation sequence, the equation is constructed:

$$\nabla^2 p - \frac{1}{V_A^2} \frac{p\partial^2}{\partial t} = \frac{\gamma}{V^2} \tag{14}$$

The above two sets of equations together constitute the coupled wave equations for data fluctuation monitoring. According to the scattering information, the adaptive boundary value of the wave data is obtained, and the discretized boundary conditions are as follows:

$$E_{s0}^{m} = 0 \tag{15}$$

The m is introduced as the time layer vector to obtain the actual monitoring equation, advance from each layer according to the pipeline level, and solve it by relying on the linear equation to obtain the final volatility.

5 FLUCTUATION DISTANCE RELATIONSHIP

According to the above analysis results of volatility, the distance relationship simulation analysis method is designed to complete the final fluctuation analysis of market data. The relationship formula of each financial entry point is:

$$P_{LS} = P_{IN} exp\ (-aL) . a_s \tag{16}$$

In the formula, $P_{IN\cdot}$ represents the current financial feedback power; α_S represents the real-time monitoring coefficient under fluctuation analysis. The relationship obtained according to the volatility is as follows:

$$T = \frac{2L}{v} \tag{17}$$

$$L = \frac{c}{n} \tag{18}$$

In the formula, L represents the discrete data of the implicit sequence, and V represents the data input point. According to the real-time analysis of the market data at this time, the linear physical change value of the data can be determined. According to the above scattering distance relationship, the optimal sequence of current fluctuation can be determined by relying on the data theorem. When the data fluctuation distance exceeds the rated distance, it will be displayed as abnormal fluctuation, which shows that there are some hidden dangers in the current financial data and further analysis is needed.

6 Conclusion

By establishing Markov model, the paper completes the analysis method innovation of financial data volatility. Aiming at the disadvantages of traditional prediction methods, a new prediction method is proposed. According to the data envelopment static tree analysis system, massive financial data are extracted; the Markov model is established to evaluate and preprocess the phasor parameters, then the Brillouin data volatility calculation formula is used to determine the fluctuation distance, and finally the fluctuation relationship distance is predicted and analyzed. Facts have proved that the proposed volatility calculation method has higher accuracy because it optimizes the data analysis process

References:

[1] Xu Ye. Analysis of Volatility Characteristics of Asset Securitization Market Based on Markov Regime Transfer Model [J]. Banker, 2019, No. 220(12):7+113-117.

[2] Wu Sirong. Analysis of Stock Market Volatility Based on Hidden Markov Model [J]. Graduate Journal of Zhongnan University of Economics and Law, 2018(3):47-55.

[3] Lu Hongqiang, Huang Dan, Zhang Zuoyang. Short Term Cross-Border Capital Fluctuations and the Robustness of Financial and Foreign Exchange Markets- Empirical Analysis Based on Markov Regime Transfer Model [J]. Western Finance, 2020(4):33-41.

[4] Wang Tao, Dong Meisheng. Research on the Regional Transformation of Income Fluctuation State of Shenzhen Hong Kong Stock Market under the “Shenzhen Hong Kong Stock Connect”--Empirical Analysis based on Markov State Transition Autoregressive Model[J]. Investment Research, 2018, v.37; No.418 (12):69-81.

[5] Zhang Fanlei. Analysis of Commodity Futures Price Behavior Based on Markov Transformation MSVAR Model [J]. Statistics and Decision-making, 2015(13):160-163.

[6] Wei Chuanqi. Stock Price Forecast of A-share Market Based on Markov Analysis [J]. Business Accountancy, 2012, 000(024):105-107.