Exploration of Multivariable Financial Time Series based on Algorithm

Yixin Zhang
School of International Education, Wuhan University of Technology, Wuhan, China

Abstract. To explore the multivariable financial time series, first of all, the concepts related to financial time series are introduced. The phase space reconstruction theory is described firstly and then the methods for testing the nonlinear characteristics of financial time series are expounded, namely Hurst index and BDS testing. Secondly, the phase space reconstruction and nonlinear test of multivariable financial time series are discussed. Several sets of composite index and industry index of Shanghai security market in China are selected as the objects for the analysis of multivariable financial time series. The multivariable reconstruction methods are used to study the system reconstruction issue. Finally, the computation of nonlinear invariants in multivariate financial time series is explored by using the maximum Lyapunov exponent. The research results show that the maximum Lyapunov exponent and correlation dimension calculated by multivariate reconstruction are significantly higher than those calculated by univariate environment, showing stronger nonlinear characteristics. It can be seen that it is effective and feasible to reconstruct phase space using multivariate financial data and predict it by using nonlinear time series method.

Keywords: multivariable financial time series; phase space reconstruction; nonlinear test.

1. Introduction

The efficiency of financial investment is one of the goals pursued by governments and investment institutions. Economic and financial time series are the most important data types in the economic and financial fields. It is an important work to analyze, forecast and control these data for the economic and financial activities. Financial time series include bonds, interest rates, exchange rates, stock prices and financial futures prices, etc. From a macro perspective, it can also include investment, income and consumption and other macroeconomic data [1]. A variety of financial time series derived from the same time period constitute multivariate (dimensional) financial time series. The stock market has always been the subject of the financial market and plays an important role. Especially in today's era when China's economy and the world's economy are becoming more and more integrated, the financial industry is facing greater and new opportunities and challenges for development. Moreover, under the international background of the rapid development of global financial innovation activities and the increasing linkage effect of financial markets in various countries, the understanding and grasp of the essential laws of financial markets is more directly related to the stability, efficiency and security of financial markets [2]. After more than ten years of rapid development of China's financial industry, a considerable scale of the market has been formed. The financial market is huge system affected by a variety of factors, with very complex rules of motion, and financial time series data is its comprehensive external manifestation. "Essence determines phenomena, phenomena reflect the essence", so the financial time series is bound to contain a lot of objective law information of the financial system. It is undoubtedly of great significance for financial investment and financing forecasting, decision-making and risk management activities to find out all kinds of information and better understand, grasp and use its laws.

Financial market is a complex and interactive nonlinear dynamic system. In practice, the univariate financial time series are often limited in length and accompanied by noise, and the information contained in the univariate financial time series is incomplete and uncertain, which cannot accurately reflect the dynamic characteristics of the financial system [3]. Compared with the univariate financial time series, the multivariate financial time series contains more relevant information, which can overcome the influence of some noise and provide more complete dynamic information. Therefore, it is necessary to study multivariable financial time series. Firstly, the major concepts and methods related to financial time series are described, namely phase space reconstruction theory and nonlinear

characteristics testing method for financial time series, including Hurst exponent and BDS testing. Secondly, the analysis objects of multivariable financial time series are introduced, and the system reconstruction is discussed. At last, the nonlinear invariant calculation is illustrated.

2. Methodology

2.1 Phase Space Reconstruction Theory

Reconstructing phase space is a method of reconstructing attractors to study the behaviour of observational dynamic systems in accordance with finite time series data. It was proposed by Packard, Ruell and Takens in the 1980s. There are three main methods for reconstructing phase space of time series: coordinate delay reconstruction method, derivative reconstruction method and basic component coordinate method. Among them, the coordinate delay reconstruction method is the most commonly used method to reconstruct the phase space of chaotic time series.

According to the coordinate delay reconstruction method proposed by Packad et al., for the one-dimensional time series $\{x(t)\}$, t=1, 2, …N, a m-dimensional vector can be constructed [4]:

$$X_n = \{x(n), x(n+\tau), ..., x(n+(m-1)\tau)\}, \quad n = 1, 2, ..., N-(m-1)\tau \tag{1}$$

In the above formula, m indicates the embedding dimension and τ refers to the delay time. The key to the coordinate delay reconstruction method lies in the determination of the embedding dimension and the delay time. In accordance with Takens theorem, the embedding dimension and the time delay can be arbitrary for an ideal one-dimensional time series with infinite length and no noise [5]. In the presence of noise, when the phase space is reconstructed, it should meet the following condition: $n \geq 2d+1$, in which d suggest the fractal dimension of attractors. Because the time series in practical applications are of finite length and noisy, the embedding dimension and time delay cannot be arbitrarily determined, but must be carefully determined. Takens theorem guarantees that we can reconstruct a phase space equal to its original dynamical system in topological sense from one-dimensional chaotic time series, thus grasping the properties and laws of chaotic time series. Characteristic judgment, analysis and prediction of chaotic time series are based on reconstructing phase space. Therefore, phase space reconstruction is the key to the study of chaotic time series.

2.2 Testing Method for Nonlinear Characteristics of Financial Time Series

Hurst exponent: it was proposed by hydroscientist Hurst in 1951, which is a widely used quantitative index used to describe the fractal characteristics of time series and can be determined by rescaled range statistics (R/S method). In 1995, Peter improved the calculation process of R/S method and its basic idea is: the time series Xi with length of M is divided into A equal length sub-intervals with length of N, recorded as $D_a, a = 1,2,...\text{A}$. The corresponding element of the sub-interval D_a is $x_{i,a}$,i=1,2,...,N, and the mean value is $\bar{x}_a$. The cumulative deviation of the sub-interval $D_a, a = 1,2,...\text{A}$ is calculated as:

$$X_{k,a} = \sum_{i=1}^{k}\left(x_{i,a} - \bar{x}_a\right), \text{k=1,2,...,N} \tag{2}$$

For each sub-interval, N cumulative deviations can be obtained. Then the range R_a of the sub-interval $D_a, a = 1,2,...\text{A}$ can be expressed as:

$$R_a = \max\left(X_{k,a}\right) - min\left(X_{k,a}\right) \tag{3}$$

The rescaled range can be expressed as $(R/S)_N = \frac{1}{A}\sum_{a=1}^{A} R_a / S_a$, where R_a refers to the range of data in the sub-interval a, S_a represents the standard deviation of data in the sub-interval a. For the different interval divisions N, there is the following relationship:

$$\left(R/S\right)_N = kN^H \tag{4}$$

In Formula (4), H refers to the Hurst exponent, and k is a consonant. When H=0.5, the time series is random, and the values of different time are irrelevant. When $0.5 < H \le 1$, time series have long-term persistence, and it is often called fractal time series. In other words, the previous time series is upward (down), then the next period is likely to continue upward (down). The closer H is to 1, the longer the persistence is. When $0 \le H \le 0.5$, time series are counter-persistent and are often referred to as mean recovery, i.e. if a sequence moves up (down) in the previous period, it is likely to turn down (up) in the next period. The closer H is to 0, the longer the mutagenicity is.

BDS testing: it was proposed by Brock et al. in 1987 on the basis of correlation integral. It is an effective statistical method to test whether the ordered data have nonlinear characteristics. The time series $x_t, t = 1,2,...,n$ is embedded into m-dimensional phase space, and there is $X_t = \left(x_t, x_{t+\tau}, \ldots, x_{t+(m-1)\tau}\right)$, $X_t \in R^m$.

3. Results and Discussion

Financial time series is the discrete output produced by the complex financial market system in the evolution process. According to different measurement methods, financial time series can be divided into high-frequency data and low-frequency data; according to the number of observation points, financial time series can be divided into univariate time series and multivariate time series. In multi-dimensional time series, some dimension time series have close relationship, even linear transformation relationship; while some dimension time series have large evolution differences. If the time series contributing greatly to the quality of reconstructed phase space can be selected from a large number of time series and reconstruction is completed, better reconstruction and prediction results can be achieved. Therefore, it is meaningful to reconstruct the complex financial market system by using multivariate financial time series and make relevant research.

According to the reconstructing technique of multivariate complex system, it can be known that reconstructing financial market by using multivariate financial time series may get better results. In addition, using multivariate financial time series can effectively reduce the impact of noise on the reconstructing process. The reconstructing technique technology is combined with multivariate financial time series for empirical research, the object of study is low-frequency (daily sampling unit) univariate and multivariate time series. The data source is the multidimensional time series produced by Shanghai security market and the sampling time is from January 2, 1997 to December 31, 2002.

3.1 Analysis Objects of Multivariate Financial Time Series

There are a variety of industry indices in the stock market, which can be used to describe the changing trend of the industry and the direct relationship between the industry and the overall economic situation. In Morgan Stanley's global industry index, the energy and raw materials sector has been far ahead of the rest this year, with the energy sector increased by 20%. From the perspective of return on investment, investors in the stock market can share the high premium of energy. This indicates that the change of energy industry index will deviate from the change of stock market composite index to a certain extent.

It is considered that both industry index and composite index are time series produced by the complex system of securities market. These time series interact with each other, and the same trend will evolve in a period of time while the other time will produce mutual deviation evolution. This evolution deviation indicates that the evolution information of complex systems contained in multiple time series is different. Therefore, the reconstruction of complex systems using multiple industry index time series should be closer to the real system than simply using a single observation. For this reason, several groups of composite index and industry index (daily data from January 2, 1997 to December 31, 2002) of Shanghai security market are selected as the empirical analysis objects. The specific research object is as follows:

Shanghai composite index: The sample stocks of the Shanghai composite index are all listed stocks, including A shares and B shares. They reflect the changes in the prices of listed stocks on the Shanghai Stock Exchange in general and have been issued since July 15, 1991.

Shanghai stock A-share index: The sample stocks of the Shanghai A-share index are all listed A-shares, reflecting the overall fluctuation of the A-share price, which has been officially issued since February 21, 1992.

Shanghai stock B-share index: The sample stocks of the Shanghai B-share index are all listed B-shares, reflecting the overall fluctuation of the stock price of B-shares, which has been officially issued since February 21, 1992.

Shanghai stock industry classification index: The Shanghai Stock Exchange divides listed companies into five categories according to their industries: industry, commerce, real estate, public utilities and comprehensive industries. The sample stocks of the industry classification index are all listed stocks in the industry, including A-shares and B-shares, reflecting the prosperity of different industries and the overall changes of their stock prices, which has been officially released since June 1, 1993.

The index code and name comparison table of these specific research objects are shown in Table 1.

Table 1. The index code and name comparison table of specific research objects

Shanghai Stock Index	1A001	Commerce index	1B002
A-share index	1A002	Real estate index	1B004
B-share index	1A003	Public utilities index	1B005
Industry index	1B001	Comprehensive industries index	1B006

3.2 System Reconstruction of Multivariate Financial Time Series

After testing the nonlinearity and determinacy of each subject, the multivariate reconstruction method will be used to study the reconstruction of multivariate financial time series.

Firstly, the univariate reconstruction of 1A0001, 1A0002, 1A0003 is investigated, and the bivariate reconstruction of 1A0001 and 1A0002, 1A0001 and 1A0003, 1A0002 and 1A0003 is further investigated. The results are shown in Figure 1 (A0001, 1A0002, 1A0003 are abbreviated to A1, A2, A3, the same below).

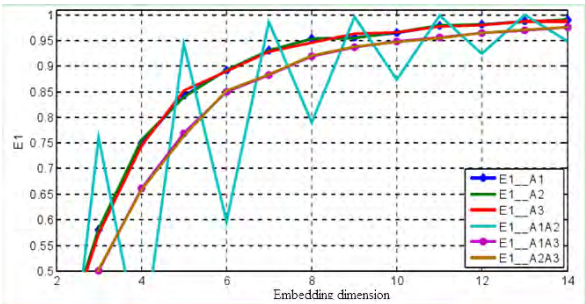

Figure. 1 Comparison of multivariate reconstruction parameters among Shanghai composite, A-share and B-share index returns

As can be seen from Figure 1, the performance of 1A0001, 1A0002 and 1A0003 is almost the same when reconstructing phase space with univariate, and there is no obvious difference. The parameters of global reconstructing are basically the same. But when using bivariate reconstruction, it can be seen that there is obvious difference. The wave curves of the global reconstruction dimensions of 1A0001 and 1A0003, 1A0002 and 1A0003 are basically the same, but the combination of 1A0001 and 1A0002 produces serious oscillation and cannot give stable convergence. In this case, it can be considered that the reconstruction time delay of 2 time series is considered independently, and the correlation between the 2 time series is not considered, so the information provided by the 2 time series interferes with each other and the satisfactory results cannot be obtained.

Figure 2 shows the comparison between the univariate reconstruction of Shanghai composite index and the bivariate reconstruction dimension of the return series of A-share and B-share indices. It can be seen from the graph that the univariate reconstruction of 1A0001 tends to be more stable than the bivariate reconstruction of 1A0002 and A0003, that is, the speed of stabilization is faster. However, from the smoothness of the curve, it can be seen that when bivariate reconstruction is used and the search embedding dimension reaches 8 dimension, the curve changes very smoothly and exceeds 0.95; although the search curve reconstructed by univariate is more stable than that of bivariate, there are obvious fluctuations when the search dimension reaches 9-10 dimension. From the graph, it can be concluded that bivariate reconstruction can converge faster than univariate reconstruction in searching global reconstruction dimension, that is, lower global reconstruction dimension can be used when multivariate phase space reconstruction is used.

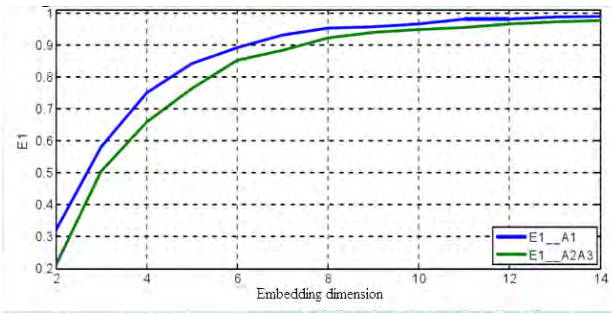

Figure. 2 Comparison of multivariate reconstruction parameters between Shanghai composite and A-share and B-share index returns

Figure 3 gives a comparison of the calculation process of the univariate reconstruction parameters of Shanghai's industry index return. From the figure, it can be seen that the dimension of phase space reconstructed by the time series of industry index return is basically the same. Among them, 1B0004 shows a unique fluctuation in the process of searching the embedding dimension. After the embedding dimension exceeds 11 dimensions, all industry index return time series reach a stable convergence.

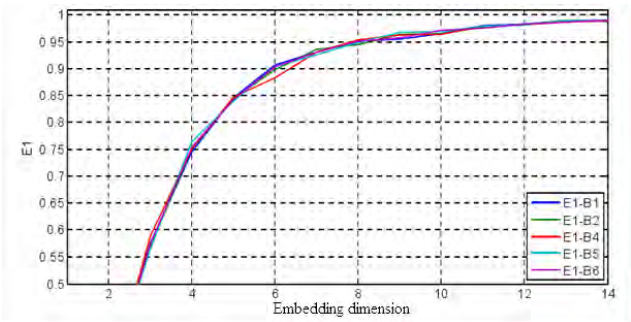

Figure. 3 Calculation and comparison of univariate reconstruction parameters of Shanghai stock index return

The specific calculation results of global dimension and local dimension of multivariable reconstruction phase space are shown in Table 2.

Table 2. The specific calculation results of global dimension and local dimension of multivariable reconstruction phase space

Reconstruction variable	Embedding dimension/time delay
Shanghai Stock Index	10/2
A-share index and B-share index	1/2, 10/2
Industry index	10/2

3.3 Computation of Nonlinear Invariants in Multivariate Financial Time Series

In order to quantitatively test the quality of the reconstructed multivariable phase space, the maximum Lyapunov exponent, correlation dimension and one-step forward prediction error are used for comparative analysis. Firstly, the phase space is reconstructed by using the results of Table 2, and the three discriminants are calculated and the results are shown in Table 3.

Table 3. Calculation of nonlinear invariants in multivariable reconstruction

Reconstruction variable	The maximum Lyapunov exponent	Correlation dimension	Prediction error
1A001	0.016	7.301	1.218
1A002, 1A003	0.052	8.353	1.216
1B001	0.028	6.701	1.246

Observing the variation of the maximum Lyapunov exponent in Table 3, it can be found that there is a special case in the process of multivariate reconstruction of the industry classification index. When four variables are used for reconstruction, the maximum Lyapunov exponent decreases. This phenomenon is mainly due to the fact that the nonlinear characteristics of the 1B0004 index return time series are not strong enough (the linear characteristics are obvious on the contrary), which largely leads to the decline of the maximum Lyapunov exponent (limiting the divergence of the orbits exponential velocity). Similarly, despite the bivariate reconstruction results of 1A0002 and 1A0003 can generate the maximum Lyapunov exponent, the nonlinear characteristics of 1A0003 are not strong enough to make the evolution orbit diverge exponentially more effectively. Because the correlation dimension itself does not reflect this nonlinear relationship, the results of correlation dimension calculation are not disturbed by the time series which are not strong enough for the nonlinear characteristics of 1A0003 and 1B0004. In terms of prediction error, there is no obvious difference between different situations, and the improvement effect of multivariate is not obvious. This is mainly due to the strong volatility of financial time series (lack of certainty).

As can be seen from Table 3, the maximum Lyapunov exponent and correlation dimension calculated by multivariate reconstruction are significantly higher than those calculated by univariate environment, showing stronger nonlinear characteristics.

4. Conclusion

Multivariable reconstruction parameter selection algorithm searches the reconstructed parameters of multivariate financial time series in Shanghai stock market. In the process of searching, it is found that some time series have close correlation. Due to the improper choice of time delay, the convergence of multivariate reconstruction is worse than that of univariate reconstruction. But on the whole, it can be concluded that the process of searching the optimal global embedding dimension with multivariate time series is more stable than that with univariate. The quality of reconstructed phase space is judged based on the calculation of nonlinear invariants. The simulation results show that the quality of multivariate reconstructed phase space is better than that of univariate reconstructed phase space. The experimental results show that it is meaningful to reconstruct the system of financial market by using multivariable financial time series, and the phase space reconstruction method is effective in multivariable environment.

References

[1]. Islam F, Shahbaz M, Ahmed A U, et al. Financial development and energy consumption nexus in Malaysia: a multivariate time series analysis. Economic Modelling, 2013, 30, pp. 435-441.

[2]. Dimitriou D, Kenourgios D, Simos T. Global financial crisis and emerging stock market contagion: A multivariate FIAPARCH–DCC approach. International Review of Financial Analysis, 2013, 30, pp. 46-56.

[3]. Brechmann E C, Czado C. COPAR—Multivariate time series modeling using the copula autoregressive model. Applied Stochastic Models in Business and Industry, 2015, 31(4), pp. 495-514.

[4]. Gao Z K, Yang Y X, Fang P C, et al. Multiscale complex network for analyzing experimental multivariate time series. EPL (Europhysics Letters), 2015, 109(3), pp. 30005.

[5]. Laih Y W. Measuring rank correlation coefficients between financial time series: A GARCH-copula based sequence alignment algorithm. European Journal of Operational Research, 2014, 232(2), pp. 375-382.