ISSN: 2688-9323

Research on the Characteristics of Financial Time Series based on Denoising

Wanyi Ge

School of Mathematics and Statistics, Central South University, Hunan, 410083, China

Abstract

In this paper, the deep learning long short memory (LSTM) model is applied to VAR risk management by using financial high-frequency data trading volume information. Taking the Shanghai Composite Index return series RT as an example, the wavelet denoising method based on multi-resolution analysis is used to denoise the high-frequency noise containing signal RT. The processed signal is stable and the denoising effect is ideal, which provides a guarantee and support for the accurate prediction of the future return series RT trend.

Keywords

High Frequency; Filter; Wavelet Processing; LSTM.

1. Introduction

Financial market is an important part of national economic development. Predicting the trend of financial stock price is of great significance to the government, investors and investment institutions, and has attracted many scholars to study it[1]. However, the price trend of the stock market is affected by political, economic, legal, military and other factors. Wavelet analysis is a new subject with great application potential. It has rich mathematical content and widely applicable characteristics, and has attracted the attention of many academic researchers at home and abroad[2].

In the economic and financial field, information continuously affects the market development process, but the collection of economic and financial data is based on discretization, which is bound to lead to different degrees of information missing [3]. High frequency time series contain more market information than low frequency series and have long-term trends. In this paper, the ARIMA model is combined with the adaptive filtering method to adjust the pre-P period historical data and pre-Q period prediction error coefficient in the ARIMA model through the adaptive filtering method [4][5].

2. LSTM Model

Let σ represent sigmoid function between 0 and 1, which is used to control the opening and closing of the gate control unit, and tanh be the hyperbolic tangent function. LSTM cells can remember the long-term information of the past, and realize the screening and filtering of historical information through the ForgetGate control of self-cycling weight. Forgetgate is defined as:

$$g_t = \sigma\left(U_g^T x_t + W_g^T h_{t-1} + b_g\right) \tag{1}$$

Wherein, b_g, U_g, W_g are respectively the bias vector, the input weight matrix and the cyclic weight matrix of the forgetgate. x_t is the feature vector at time t of the sample, h_{t-1} is the

ISSN: 2688-9323

implicit state vector at time t-1, and g_t is the output vector of the forgetgate at time t. The cell state is also related to input, and the external input gate vector it is defined as

$$i_t = \sigma\left(U_i^T x_t + W_i^T h_{t-1} + b_i\right) \tag{2}$$

$$\tilde{c}_t = \tanh\left(U_c^T x_t + W_c^T h_{t-1} + b_c\right), c_t = g_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \tag{3}$$

The output of LSTM cells is controlled by the OutputGate. The OutputGate vector OT and the final hidden state output vector HT are defined as:

$$o_t = \sigma\left(U_o^T x_t + W_o^T h_{t-1} + b_o\right), h_t = \tanh\left(c_t\right) \cdot o_t \tag{4}$$

3. Wavelet Signal Analysis

The goal of wavelet denoising is to suppress the interference signal and restore the real signal. Wavelet denoising based on multi-resolution analysis is essentially the problem of filtering the noise signal, but it is different from the traditional low-pass filter. Let the finite energy signal $\psi(t) \in L^2(R)$, and define the Fourier transform of ψ(t) as follows:

$$\hat{\psi}(\omega) = \int_{-\infty}^{+\infty} e^{-it\omega} \psi(t) \mathrm{d}t \tag{5}$$

For $f(x) \in L^2(R)$, if ψ(t) satisfies Equation (3), then the following reconstruction formula can be obtained,

$$f(x) = \frac{2}{C_\psi} \int_0^\infty \left[\int_{-\infty}^{\infty} \left\{ \left(W_\psi f\right)(b,a) \right\} \left\{ \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a} \right) \right\} \mathrm{d}b \right] \frac{\mathrm{d}a}{a^2} \tag{6}$$

The goal of wavelet denoising is to suppress the interference signal and restore the real signal. Wavelet denoising based on multi-resolution analysis is essentially the problem of filtering the noise signal, but it is different from the traditional low-pass filter:

$$\mu = \sigma\sqrt{2\ln N} \tag{7}$$

4. Adaptive Filtering Process

Select the current period t=N, and take a set of initial weights $\omega_i \left(i = 1, 2, 3, \cdots, N\right)$, using Equation to calculate the predicted value $\hat{x}_{t+1}$ in T +1 period.

The error between the predicted value and the actual value was calculated. The calculation formula is shown in Equation (3).

$$e_{t+1} = \hat{x}_{t+1} - x_{t+1} \tag{8}$$

Adjust the weight and formula according to the error value e_{t+1}:

$$\omega_i^{'} = \omega_i + 2ke_{t+1}x_{t-i+1} \tag{9}$$

ISSN: 2688-9323

Adjust the weight and formula according to the error value ET +1.

5. Model Result

Taking RT of Shanghai Stock Exchange as an example, wavelet de-noising based on multi-resolution analysis is used to de-noising the index.

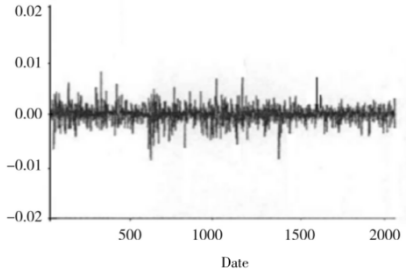

Fig 1. RT series chart of yields

In order to accurately extract the main characteristic components of the trend of return time series signals, it is necessary to de-noise RT of return time series signals. Due to the factors such as the instability of the yield sequence signal and the large fluctuation of data, and combined with the performance characteristics of various wavelet basis functions.

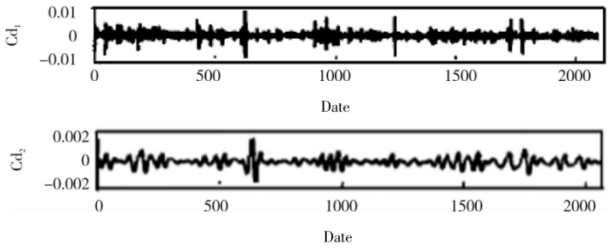

Fig 2. The result after denoising

6. Conclusion

With the advent of the era of big data, people's requirements for the accuracy of the collected data are constantly improving, and the collection of high-frequency data is the trend. Thanks to the development of science and computing technology, it is possible to collect economic and financial data on a minute-second basis. This paper proposes a research method of stock time series prediction based on WBED hybrid model under industry background differences. The traditional ARIMA model is optimized by using the adaptive filtering method, and the combined model is successfully applied in the 5-minute high-frequency financial time series. The research shows that the combined model can significantly improve the prediction accuracy of the traditional model compared with the traditional model.

References

[1] Box G.E.P, Jenkins G.M. Time series Analysis: Forecasting and Control[M]. San Francisco: Holden Day, 1976.

[2] Shadab Ahzam, Ahmad Shamshad, Said Saif. Spatial forecasting of solar radiation using ARIMA model [J]. Remote Sensing Applications: Society and Environment, 2020, 20.

ISSN: 2688-9323

[3] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. Neural Computation, 2006, 18(7): 1527-1554.

[4] Zhao J F,Mao X,Chen L J. Speech emotion recognition using deep 1D & 2D CNN LSTM networks. Biomedical Signal Processing and Control, 2019,47: 312-323.

[5] Candelon B, Colletaz G, Hurlin C, et al. Backtesting Value-at-Risk: A GMM Duration-based Test [J]. Journal of Financial Econometrics, 2011,9(2).