2D3-OS-19a-1

Long Memory and Predictability in Financial Markets

Implications from Empirical Studies and Applications of the LSTM model

Takahashi, Shuntaro   Chen, Yu

Department of Human and Engineered Environmental Studies,
Graduate School of Frontier Sciences, The University of Tokyo

This paper examines the empirical properties of trading volume and the predictability of trading volume and absolute return with the Long Short Term Memory(LSTM). With the analysis on the S&P500 index/firms daily data, two properties are discovered: i) Trading volume has a long memory ii) Trading volume is essentially related to the price return distribution. The latter part aims to exploit long memory of financial markets for the prediction. The LSTM model, an architecture designed for modelling long memory is constructed for the prediction of trading volume and absolute return with their past values. The prediction with the model mainly achieved three results: iii) Trading volume is highly predictable with its past values. iv) The LSTM model overwhelms the performance of the GARCH(1,1) with input of temporally distant past values and without addition of variables. v) The contribution of trading volume to the prediction of absolute return is insignificant.

1. Introduction

Econophysics emerged in 1990s, and has provided a new perspective on financial markets. The researches suggest that price series are not simple random processes, which however contain complex spatio-temporal structures. Though the importance has been pointed out [Gallant 92], the property of trading volume has been less investigated in econophysics compared with its counterpart, i.e. volatility.

Deep learning (DL) is a framework, with which the layers of neural network learn the complex structures of data [LeCun 15]. DL improved the state of the art of various areas. Financial markets prediction with DL emerged in 1990s and is continuously studied up to the present.

This paper aims at bridging empirical researches inspired by econophysics and the financial markets prediction. This combination should be effective because the empirical researches lead to better understanding on the dataset, which contributes to the framework of DL. Specifically, this paper responds to the following questions. For the empirical researches, i) Does trading volume a have long memory? ii) To what extent does trading volume influence price return distribution? For the prediction task, iii) To what extent is trading volume predictable? iv) Does the temporally distant past values affect the accuracy of absolute return prediction? v) Does trading volume contribute to prediction of absolute return?

This paper uses the daily data of the S&P500 index and all firms retrieved from Yahoo! Finance. The period of time is from Jan 3rd 1950 to Oct 10th 2016 for the index and from Jan 3rd 1962 to Oct 10th 2016 for the firms. This difference comes from the data availability.

2. Empirical Researches

2.1 Long Memory of Trading Volume

A random process $X_{1..n}$ with mean μ and variance σ^2 is said to have long memory if it has an non-integrable autocorrelation function(ACF) [Lillo 04].

$$ACF(\tau) = \frac{E[(X_i - \mu)(X_{i+\tau} - \mu)]}{\sigma^2} \tag{1}$$

A power function $\tau^{-\alpha}$ where $\alpha < 1$ is an example of non-integrable function. A famous stylized fact, volatility clustering could be rephrased that the ACF of a series of absolute return follows such a power function.

This section tests whether trading volume also has long memory by its ACF. Due to the positive trend of trading volume in decades, the normalization procedure is necessary to detrend the data. The normalized trading volume is defined as follows.

$$V'_{t,k} = \frac{V_t - \mu_{t,k}}{\sigma_{t,k}} \tag{2}$$

V_t is the trading volume at time-step t and $\mu_{t,k}$ and $\sigma_{t,k}$ are mean and standard deviation of the trading volume between time-step $t-k$ and t.

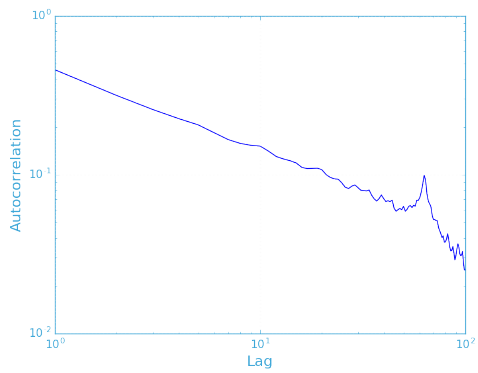


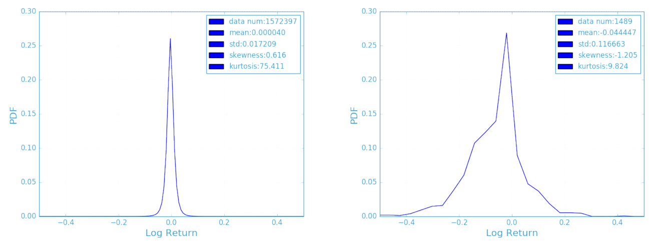
Figure 1: The ACF of the normalized trading volume of the S&P500 firms in log-log scale. $k = 300$ adopted.

The ACF in Figure 1 follows a power function with an exponent of $\alpha \approx 0.51$. Therefore the normalized trading volume satisfies the precondition for long memory.

Contact: Takahashi, Shuntaro, stakahashi@g.ecc.u-tokyo.ac.jp

2.2 The Difference in the Price Return Distribution Conditioning on Trading Volume

The functional form of price return distribution has been argued for decades [Mandelbrot 63]. However, except for the linear relationship between trading volume and standard deviation of price return distribution [Gallant 92], researches have not been conducted thoroughly on the relationship between trading volume and the form of price return distribution in daily scale. This paper examined the price return distributions conditioning in different ranges of the normalized trading volume.



(a) $-1\sigma \le V'_{t,100} < 0\sigma$ (b) $8\sigma \le V'_{t,100} < 9\sigma$

Figure 2: The PDFs of the log return conditioning on the normalized trading volume of the S&P500 firms. $k = 100$ adopted.

Figure 2a and Figure 2b are the price return distribution conditioning on small and large trading volume respectively. It is confirmed that larger trading volume not only leads to larger standard deviation, but also the differences in symmetry/asymmetry and sharpness of the peaks.

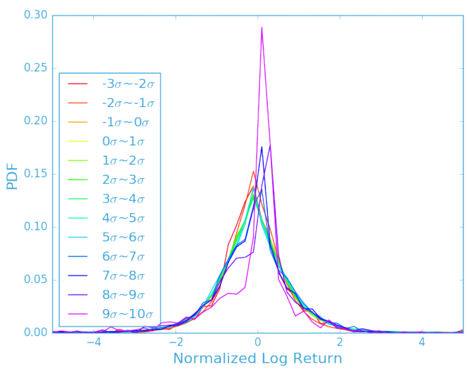


Figure 3: The PDFs of the normalized log return conditioning on the normalized trading volume. The normalized log return r' is defined as $r' = \frac{r-\mu_r}{\sigma_r}$ where r is log return and μ_r and σ_r are mean and standard deviation of the conditional price return distribution. Different colors represent the regions of the normalized trading volume as presented at the left below legend.

Figure 3 compares the conditional price return distributions by removing the effect of standard deviation and mean. It is clear that the distribution of large trading volume(blue to pink) is different from that of small trading volume(red to cyan).

3. Prediction with Deep Learning

3.1 Model and Experiment

The application of machine learning for financial markets prediction should take the property of the dataset, long memory into account. The long short term memory(LSTM) [Hochreiter 97] is a structure of units in neural networks, which is purposefully designed to reproduce long memory. The proposed model consists of three layers: input layer, the hidden layer with 50 LSTM blocks and output layer. The optimizer is Adam with the proposed hyper-parameters. It is implemented with Keras and coding is straightforward. The daily data of the S&P500 index between 1950 and 2016 is used in this work to be consistent with the empirical researches. The first 80% is used for training and the rest 20% for testing.

3.2 The Prediction of the Normalized Trading Volume

The predictability of trading volume has gained little attention. However, as the prediction of trading volume leads to the prediction of price return distribution which is dependent on trading volume, it is obviously meaningful for application usage.

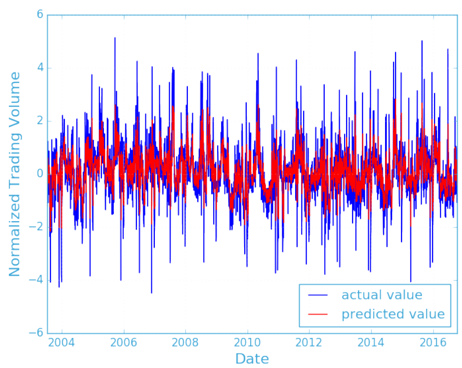


Figure 4: The prediction of the normalized trading volume of the S&P500 index with the LSTM model. The time-series of the past normalized trading volume is fed to the model, and input length is 15.

Figure 4 demonstrates that the prediction of the normalized trading volume is done effectively with the LSTM model from its past values. This result indicates the high predictability of trading volume, which is a practically and theoretically curious aspect of trading volume.

3.3 The Prediction of Absolute Return

3.3.1 The Prediction of Absolute Return with Its Past Time Series

Because a long memory has been found in the time series of absolute return, the distantly past values should contribute to the accurate prediction. To test it, the effect of input length is investigated. Input length refers to the number of past time-step fed to the model. For example, if input length is 5, then the values $|r_{t-5}|...|r_{t-1}|$ is fed to the model to predict $|r_t|$.

Table 1 is the accuracy of the prediction measured by root mean squared error(RMSE) for different input length. It is clear that longer input length leads to better prediction. The LSTM model with sufficient input length, 10 and 15 overwhelmed the

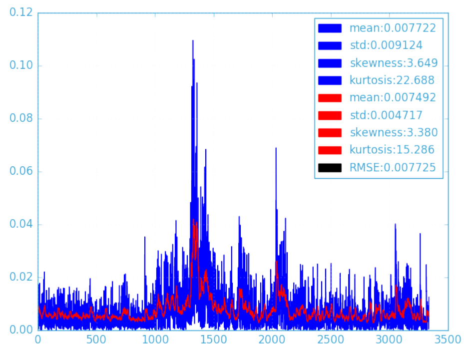


Figure 5: The absolute return prediction of the S&P500 index with the LSTM model with input length 15. The blue and red lines are the actual and the predicted values respectively.

GARCH(1,1) model, the dominant price fluctuation model without addition of variables.

3.3.2 The Prediction of Absolute Return with the Past Absolute Return and Normalized Trading Volume

The addition of the normalized trading volume is expected to improve the accuracy of the prediction of absolute return. However, as presented in Table 1, it did not significantly contribute to the improvement of accuracy in this framework. This result implies that the process of absolute return and trading volume is similar to each other.

Table 1: The list of input length and RMSE for the absolute return prediction of the S&P500 index. The GARCH model parameters are $p = 1$ and $q = 1$.

Input Length	1	5	10
RMSE	0.008812	0.008243	0.007820
Input Length	15	15(trading volume added)	GARCH
RMSE	0.007725	0.007696	0.007871

4. Conclusion

Throughout the paper, the empirical researches and the financial markets prediction with DL are conducted with respect to their connection. In the empirical researches, the property of trading volume, long memory and the influence on price return distribution were found. They were expectable to contribute to the better prediction of financial markets. In the prediction section, the LSTM model was applied because of long memory in absolute return and trading volume. With the model, high predictability of trading volume, good modelling of absolute return overwhelming the GARCH(1,1) and the insignificant contribution of trading volume to the prediction of absolute return were confirmed. The following is the list of the responses to the questions proposed in the introduction. i) Trading volume has long memory. The ACF of trading volume follows power function with an exponent $\alpha \approx 0.51$. ii) Trading volume essentially influences price return distribution in terms of symmetry/asymmetry and sharpness of the peaks. iii) Trading volume is highly predictable with the LSTM model. iv) Longer input length contributes to the better prediction of absolute return. The accuracy overwhelmed the GARCH(1,1) with input length 15 without addition of variables. v) Trading volume did not significantly contribute to the improvement of predictive power of absolute return with the LSTM model.

References

[Gallant 92] Gallant, A.R. and Rossi, P.E. and Tauchen, G.: Stock Price and Volume, The Review of Financial Studies, Vol.5, No.2, pp.199-242 (1992)

[Hochreiter 97] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Computation, Vol.9, No.8, pp.1735-1780 (1997)

[LeCun 15] LeCun, Y. and Bengio, Y. and Hinton, G.: Deep Learning, Nature Vol.521, No.7553, pp.436-444 (2015)

[Lillo 04] Lillo, F and Farmer, J.: The long memory of the efficient market, Studies in Nonlinear Dynamics & Econometrics, Vol.8, No.3, pp.- (2004)

[Mandelbrot 63] Mandelbrot, B.: The Variation of Certain Speculative Prices, The Journal of Business, Vol.36, No.4, pp.394-419 (1963)