Analysis of Dynamical Cross Correlations in Financial Time Series

Y. NAKAYAMA and H. IYETOMI[1]

Department of Physics, Niigata University, Nishi-ku, Niigata 950-2181, Japan
[2] *E-mail: hiyetomi@sc.niigata-u.ac.jp*
http://www.sc.niigata-u.ac.jp/sc/English/index.html

An analysis of dynamical cross correlations in financial time series are carried out. We first calculate 1-day return of 557 Japanese stocks for the 11-year period of 1996-2006, and take a discrete Fourier transform of the return. Then we construct a correlation matrix $\mathbf{C}(\omega)$ for each frequency ω, and calculate the eigenvalues λ of $\mathbf{C}(\omega)$. Also we repeat the same calculation using random data instead of real stock returns for reference. Comparison of the empirical eigenvalues with the reference results enables us to extract essential dynamical correlations involved in complicated behavior of the stock returns. Furthermore, we demonstrate the eigenvectors associated with the outliers of $\lambda' s$ significantly differ from the random case.

Keywords: Random Matrix Theory; Financial time series; Dynamical cross Correlation.

1. Introduction

Recently correlations between different stocks have been investigated very actively, and the results are applied to portfolio management. The successes [1-3] are remarkable in eliminating noise out of the equal-time correlations based on the random matrix theory (RMT). However, it is important to pay our attention to not only the static correlations but also the time-lagged correlations. Although some effort was made [4], the research to analyze such dynamical correlations has just been initiated. In this paper, we propose a new alternative method.

2. Static Cross-Correlations

We begin with defining return of the stock $i(=1, \cdots, N)$

$$r_i(t) \equiv \ln S_i(t+1) - \ln S_i(t), \tag{1}$$

where $S_i(t)$ is the stock price at time $t(= 1, \cdots, T)$. And then we normalize the return as

$$w_i(t) \equiv \frac{r_i(t) - \langle r_i \rangle}{\sigma_i}, \tag{2}$$

where $\sigma_i \equiv \sqrt{\langle (r_i - \langle r_i \rangle)^2 \rangle}$ is the standard deviation of r_i, and $\langle \cdots \rangle$ denotes a time-average over the period T.

The equal-time cross-correlation matrix $\mathbf{C}$ is defined as

$$\mathbf{C} \equiv \frac{1}{T}\mathbf{W}\mathbf{W}^t, \tag{3}$$

where $\mathbf{W}$ is a $N \times T$ rectangular matrix with elements $W_{it} \equiv w_i(t)$ and $\mathbf{W}^t$, its transpose. So the component of the cross-correlation matrix $\mathbf{C}$ is given as

$$C_{ij} = \langle w_i(t) w_j(t) \rangle. \tag{4}$$

According to the RMT, in the limit $N \to \infty$, $T \to \infty$, such that $Q = \frac{T}{N} \geq 1$ is fixed, the probability density $\rho(\lambda)$ for the eigenvalue λ of the corresponding random matrix is given [5-8] by

$$\rho(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}, \tag{5}$$

$$\lambda_\pm = \left(1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}\right). \tag{6}$$

To analyze a database from the Tokyo Stock Exchange, we use 1-day return for the 11-yr period of 1996-2006 ($T = 2707$ days). We extract from this database 557 stocks that were priced for all business days in that period. Figure 1(a) shows $\rho(\lambda)$ for the cross-correlation matrix $\mathbf{C}$ with the real data. There are 13 eigenvalues larger than the upper cutoff λ_+.

To construct the cross correlation matrix, we also use the Spearman's rank correlation coefficient. It is a non-parametric measure of correlations that may be more effective when the distribution is not normal. We first rank the return $r_i^{'}(t)(= 1 \sim T)$ and normalize the ranking as

$$w_i^{'}(t) \equiv \frac{r_i^{'}(t) - \langle r_i^{'} \rangle}{\sigma_i^{'}}, \tag{7}$$

where $\sigma_i^{'} \equiv \sqrt{\langle (r_i^{'} - \langle r_i^{'} \rangle)^2 \rangle}$ is the standard deviation of $r_i^{'}$. We then reconstruct the correlation matrix $\mathbf{C}^{'}$ using $w_i^{'}$ in place of w_i and caluclate the eignvalues of $\mathbf{C}^{'}$. The results as shown in Fig. 1(b) are almost indistinguishable from those in Fig. 1(a).

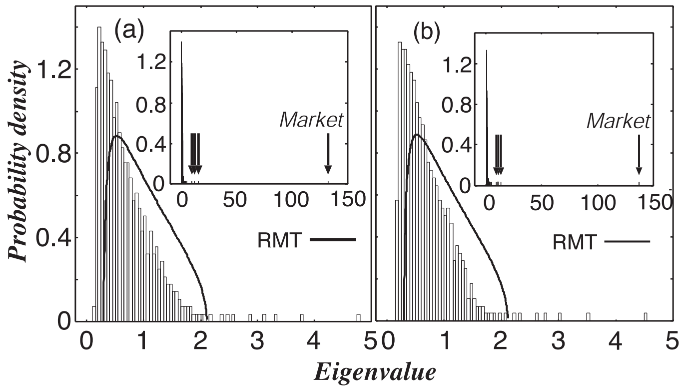

Fig. 1. (a) The eigenvalue distribution for the empirical **C**, where $Q = \frac{T}{N} = 4.86$. If the returns are completely random, the eigenvalues are distributed in the interval $0.30 \leq \lambda \leq 2.11$. However the largest eigenvalue, $\lambda \simeq 132$, is 60 times larger than the upper cutoff. (b) The eigenvalue distribution for the cross-correlation matrix constructed from the Spearman's rank correlation coefficients with the same data. The largest eigenvalue is $\lambda \simeq 139$.

The i-th component of the eigenvector corresponding to the eigenvalue λ_α will be denoted as $v_{\alpha,i}$. We normalize it such that $\sum_{i=1}^{N} {v_{\alpha,i}}^2 = N$. The RMT shows the components $v_{\alpha,i}$ of the random correlation matrix should distribute according to the Gaussian with mean zero and unit variance,

$$\rho(v) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right). \quad (8)$$

3. Dynamical Cross-Correlations

We define time-lagged correlation function as

$$C_{ij}(\tau) = \langle w_i(t+\tau) w_j(t) \rangle, \quad (9)$$

where τ is the time-lag. The discrete Fourier transform of $C_{ij}(\tau)$ is

$$\begin{aligned} \sum_\tau C_{ij}(\tau) \exp(-i\omega\tau) &= \frac{1}{T} \sum_\tau \sum_t w_i(t+\tau) w_j(t) \exp(-i\omega\tau) \\ &= \frac{1}{T} W_i(\omega) W_j^*(\omega) \\ &\equiv C_{ij}(\omega), \end{aligned} \quad (10)$$

where $W_i(\omega) \equiv \sum_t w_i(t) \exp(-i\omega t)$.

The correlation matrix $\mathbf{C}(\omega)$ in the Fourier space relates to the equal-time cross-correlation matrix **C** through the sum rule

$$\sum_{\omega} \mathbf{C}(\omega) = \mathbf{C}. \tag{11}$$

3.1. *Distribution of Eigenvalues*

The rank of the correlation matrix $\mathbf{C}(\omega)$ reduces to 1, so that $\mathbf{C}(\omega)$ has only a single eigenvalue $\lambda_\omega(\neq 0)$. The eigenvalue λ_ω is explicitly given as

$$\begin{aligned}\lambda_\omega &= \frac{1}{T}\sum_j |W_j(\omega)|^2 \\ &= \frac{1}{T}\sum_j\sum_t\sum_{t'} w_j(t)w_j(t')\exp\left\{i(t'-t)\omega\right\}.\end{aligned} \tag{12}$$

In the case where $w_j(t)$ are random numbers with zero mean and unit variance, the central limit theorem guarantees the distribution of the eigenvalues at any frequency takes the Gaussian form:

$$\rho(\lambda) = \frac{1}{\sqrt{2\pi N}}\exp\left(-\frac{(\lambda-N)^2}{2N}\right). \tag{13}$$

We calculated λ_ω's for the dynamical correlation matrix using the same real data as for the equal-time correlation matrix. There are a number of spikes in the frequency spectrum of the eigenvalues as seen in the panel (a) of Fig. 2. The corresponding panel of Fig. 3 shows the probability density $\rho(\lambda)$ of the eigenvalues for all the frequencies. The distribution remarkably deviates from Eq. (13). The panels (b) in Figs. 2 and 3 are the frequency spectrum of the eigenvalues and its distribution, respectively, but for the random data. The distribution is now in good agreement with Eq. (13). The largest eigenvalue (1769.7) is about 3.18 times larger than the average (557.0). Although the averages of the eigenvalues are almost identical for both data, but the standard deviation (149.0) of the eigenvalue distribution for the real stock data is much larger than that (23.5) for the random data.

3.2. *Distribution of Eigenvector Components*

The i-th component of the eigenvector corresponding to an eigenvalue λ_ω will be denoted as $u_{\omega,i}$. We normalize it such that $\sum_{i=1}^{N} |u_{\omega,i}|^2 = N$. For a random counterpart of the dynamical correlation matrix $\mathbf{C}(\omega)$, the distribution $\rho(|u|)$ of the eigenvector components should conform to

$$\rho(|u|) = 2|u|\exp\left(-|u|^2\right). \tag{14}$$

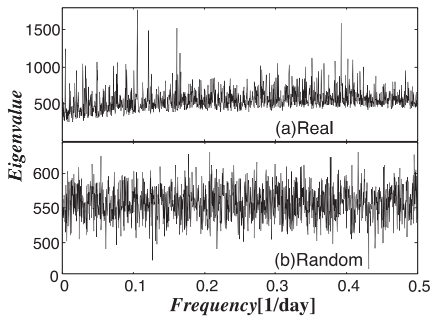

Fig. 2. (a) Frequency spectrum of the eigenvalue for the real market data. (b) Frequency spectrum of the eigenvalue for the random data.

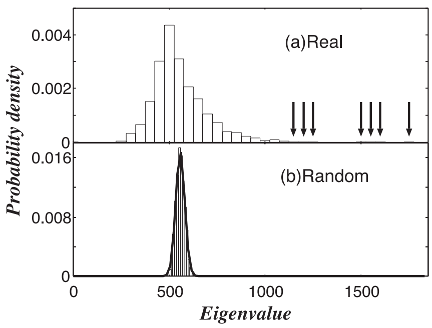

Fig. 3. The probability density of the eigenvalues for all frequencies; (a) using the real stock data, (b) using the random data.

This corresponds to Eq. (**??**) in the static RMT. Figure 4(a) shows $\rho(|u|)$ for the real correlation matrix appreciably deviates from Eq. (14). Figure 4(c) illustrates the eigenvector components u's themselves in the complex plane for the largest eigenvalue. We thus see there is a dynamical fluctuation in which most of the stocks behave in phase with a period of 9.47 days. In contrast, the results for the third-smallest eigenvalue agrees well with the random results as shown in the panels (b) and (d) of Fig. 4.

3.3. *Dynamical market trend*

All the components coherently participate in the eigenvector associated with the largest eigenvalue for the static correlation matrix. It thereby represents an influence that is common to all stocks. Also we observe such market trend is reflected in the eigenvectors with large eigenvalues for the dynamical correlation matrix as demonstrated in Figure 4(c); those outliers thus correspond to the market trend. Figure 5 plots the norm-ranking of the eigenvector components for the dynamical correlation matrix at ten characteristic frequencies. These periods arises from the first 10 largest eigenvalues for the dynamical correlation matrix.

We chose the top five dominant components of the market trend eigenvector for the static cross correlations to trace their behavior. As appreciated in the figure, the dominant components in the eigenvector of the largest eigenvalue for the static correlation matrix do not always lead in the dynamical correlations. Adopting the present analytical method, we were thus able to resolve the market trend into dynamically inter-correlated compo-

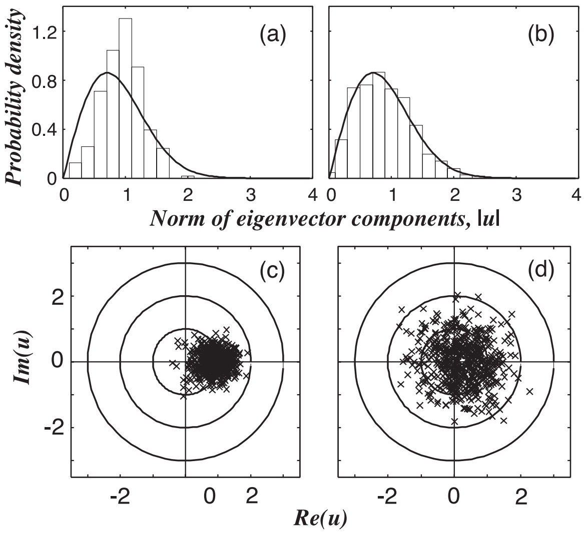

Fig. 4. (a) Distribution $\rho(|u|)$ of the norm of the eigenvector components for the largest eigenvalue $\lambda \approx 1769.7$ with period $T \approx 9.47 days$.(b)$\rho(|u|)$ for the third-smallest eigenvalue $\lambda \approx 264.0$ with $T \approx 142.5$ days. Note the good agreement with Eq. (4.3) (solid curve). (c) Eigenvector components of (a). (d) Eigenvector components of (b).

nents with characteristic frequencies.

4. Conclusion

We have confirmed that the market returns have definite dynamical correlations as well as static correlations through comparison of the eigenvalue distribution calculated from the dynamical correlation matrix based on the real market data with the random counterpart. Each component of the eigenvector reflects temporal behavior of the corresponding stock. We are thus successful in extracting essential dynamical correlations in the market by observing the eigenvectors associated with the eigenvalues far apart from the random reference. Stock prices data may be available at tick level. The present method is also applicable to the analysis of such high-frequency correlations.

We thank Ken Millennium Corporation for its financial support in carrying out the present work.

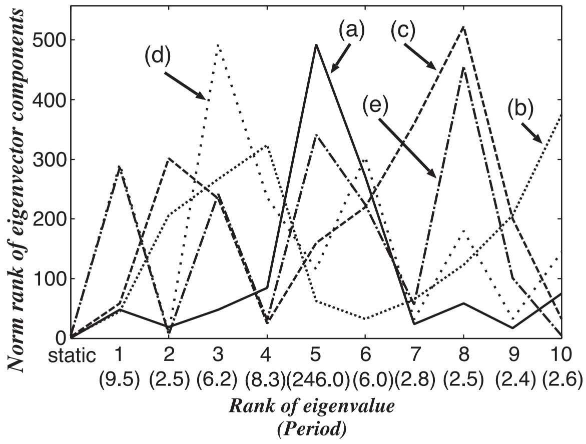

Fig. 5. The norm-ranking of the eigenvector components for the dynamical correlation matrix at ten characteristic frequencies (the corresponding periods are in units of day). The top five dominant components of the market trend eigenvector for the static cross correlations are traced. The types of industry for those chosen stocks are (a) real estate, (b) financial, (c) textile, (d) machinery, and (e)non-iron metal.

References

1. L. Laloux, P. Cizeau, J. P. Bouchaud, and M. Potters, *Phys. Rev. Lett.* **83**, 1467 (1999).
2. V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley, *Phys. Rev. Lett.* **83**, 1471 (1999).
3. V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, *Phys. Rev. E* **65**, 066126 (2002).
4. C. Biely and S. Thurner, *cond-mat* , 0609053 (2006).
5. A. Edelmann: Eigenvalues and condition numbers of random matrices. *SIMA Hournal of Matrix Analysis and Applications,* **9** (1988), 543-560.
6. A. M. Sengupta, and P. P. Mitra: Distributions of singular values for some random matrices. *Phys. Rev. E,* **60** (1999), 3389-3392.
7. M. L. Mehta: Random Matrices; 3rd ed. (New York Academic Press, New York, 2004).
8. J. P. Bouchaud and M. Potters: Theory of Financial Risk and Derivate Pricing: From Statistical Physics to Risk Management; 2nd ed. (Cambridge University Press, 2003).