International Journal of Applied Mathematics

Volume 27 No. 3 2014, 237-243
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
doi: http://dx.doi.org/10.12732/ijam.v27i3.4

A SHORT NOTE ON THE PATTERN OF THE SINGULAR VALUES OF A SCALED RANDOM HANKEL MATRIX

Hossein Hassani[1 §], Nader Alharbi[2], Mansi Ghodsi[3]
[1,2,3]Statistical Research Centre
Business School
Bournemouth University, 89, Holdenhnrst Road
Bournemonth, BH8 8EB, UK

Abstract: This note considers some of the properties and studies the distribution of the eigenvalues of the matrix $\mathbf{XX}^T$ divided by its trace, where $\mathbf{X}$ is a Hankel random matrix. The results make a novel contribution in the area of signal processing and noise reduction.

AMS Subject Classification: 15A18, 15A99, 15B05, 15B52
Key Words: Hankel matrix, eigenvalue, singular value, random process, noise

1. Introduction

Consider a one-dimensional series $Y_N = (y_1, \dots, y_N)$ of length N. Transferring this series into the multi-dimensional series $X_1, \dots, X_K$ with vectors $X_i = (y_i, \dots, y_{i+L-1})^T \in \mathbf{R}^L$ provides the following trajectory matrix

$$\mathbf{X} = (x_{i,j})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_N \end{pmatrix}, \tag{1}$$

where L $(2 \leqslant L \leqslant N-1)$ is the window length and $K = N - L + 1$. The

Received: April 8, 2014

§Correspondence author

trajectory matrix $\mathbf{X}$ is a Hankel matrix, which means all the elements along the diagonal $i + j =$ const are equal. The square root of the eigenvalues of the L by L matrix $\mathbf{X}\mathbf{X}^T$, where $\mathbf{X}^T$ is the conjugate transpose, are called singular values of $\mathbf{X}$. The ratio of each eigenvalue $\lambda_i / \sum_{i=1}^{L} \lambda_i$ is the contribution of the matrix $\mathbf{X}_i$ to $\mathbf{X}$, since $||\mathbf{X}||_F^2 = tr(\mathbf{X}\mathbf{X}^T) = \sum_{i=1}^{L} \lambda_i$ and $||\mathbf{X}_i|| = \lambda_i$, where λ_i $(i = 1, \dots, L)$ are the eigenvalues of $\mathbf{X}\mathbf{X}^T$ and $|| \, ||_F$ denotes the Frobenius norm.

The Hankel matrix $\mathbf{X}$ and its corresponding singular values are important in many areas including time series analysis [8], [10], biomedical signal processing [17], mathematics [15], energy [11, 4, 16], econometrics [9] and physics [6]. The distribution of eigenvalues/singular values and their closed form are of great interest, but this issue has not been considered adequately [14].

Note also that if the series Y_N is a white noise process, then the trajectory matrix $\mathbf{X}$ will be called a random matrix where each column of $\mathbf{X}$ forms a L-variate normal distribution with zero mean [13], [7], [2]:

$$X_i = (y_i, \dots, y_{i+L-1})^T \sim N_L(\mathbf{0}, \mathbf{G}), \tag{2}$$

where, $\mathbf{G}$ is a $L \times L$ positive definite matrix, and $\mathbf{0}$ is a vector of zeros. Then, the Wishart distribution [18] is the probability distribution of the $L \times L$ random matrix $\mathbf{A} = \mathbf{X}\mathbf{X}^T$:

$$\mathbf{A} \sim W_L(\mathbf{G}, v), \tag{3}$$

where the positive integer v is the number of degrees of freedom, [12].

Theorem 1. *Let $\mathbf{G}$ be a positive-definite matrix with distinct eigenvalues, $\mathbf{X}\mathbf{X}^T \sim W_L(\mathbf{G}, v)$, and set $\mathbf{J} = v^{-1}\mathbf{X}\mathbf{X}^T$. Consider spectral decomposition $\mathbf{G} = \mathbf{Z}\boldsymbol{\Lambda}\mathbf{Z}^T$ and $\mathbf{J} = \mathbf{Q}\boldsymbol{\Gamma}\mathbf{Q}^T$, and let $\eta = (\eta_1, \dots, \eta_L)$ and $\lambda = (\lambda_1, \dots, \lambda_L)$ be the vectors of diagonal elements in $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$. Then, the following asymptotic distribution holds as $v \to \infty$:*

$$\lambda \sim N_L(\eta, 2\boldsymbol{\Lambda}^2/v), \tag{4}$$

where the eigenvalues of $\mathbf{J}$ are asymptotically normal, unbiased, and independent, with λ_i recording a variance of $2\eta_i^2/v$, see [2].

The above theorem works for the situation where the vectors X_i are distributed independently whilst for the Hankel matrix this is not applicable as the lagged vectors X_i and X_j are correlated. For example, X_i and X_{i+1}

$(i = 1, \ldots, K-1)$ have $L-1$ similar observations with the following covariance matrix:

$$Cov(X_i, X_{i+1}) = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ \sigma^2 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \sigma^2 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \sigma^2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \sigma^2 & 0 \end{pmatrix}, \tag{5}$$

where σ^2 is the variance of y_i. Moreover, it is always of interest to have bounded eigenvalues whilst in the above case, the magnitude of singular values change with the series length; increasing the sample size N leads to the increase of λ_i. To overcome this issue, we divide $\mathbf{XX}^T$ by its trace, $\mathbf{XX}^T/\sum\limits_{i=1}^{L}\lambda_i$. This in turn provides several important properties.

Proposition 1. *Let $\zeta_1, \ldots, \zeta_L$ denote eigenvalues of the matrix $\mathbf{XX}^T/\sum\limits_{i=1}^{L}\lambda_i$, where $\mathbf{X}$ is a Hankel trajectory matrix with L rows, and λ_i $(i = 1, \ldots, L)$ are the eigenvalues of $\mathbf{XX}^T$. Thus, we have the following properties:*

1. $0 < \zeta_L \leq \ldots \leq \zeta_1 < 1$,
2. $\sum_{i=1}^{L} \zeta_i = 1$,
3. $\zeta_1 \geq \frac{1}{L}$,
4. $\zeta_L \leq \frac{1}{L}$,
5. $\zeta_i \in (\frac{1}{L} - a, \frac{1}{L} + b) \quad (i = 2, \ldots, L-1)$, *where* $a, b \in [0, 1]$.

Proof. The first two properties are simply obtained from matrix algebra and thus not provided here. To prove the third property, the first two properties are used as follows. The second property confirms

$$\zeta_1 + \zeta_2 + ... + \zeta_L = 1.$$

Thus, using the first property, $\zeta_1 \geq \zeta_i$ $(i = 2, ..., L)$, we obtain

$$\underbrace{\zeta_1 + \zeta_1 + ... + \zeta_1}_{L \text{ elements}} = L\zeta_1 \geq 1 \Rightarrow \zeta_1 \geq 1/L.$$

Similarly, for the fourth property, it is straightforward to show that

$$\underbrace{\zeta_L + \zeta_L + ... + \zeta_L}_{L \text{ elements}} = L\zeta_L \leq 1 \Rightarrow \zeta_L \leq 1/L,$$

since $\zeta_L \leq \zeta_i$, $i = (1, 2, ..., L-1)$, and $\sum_{i=1}^{L} \zeta_i = 1$.

To prove part 5, let us first prove that there exists ζ_2 between real numbers ζ_1 and ζ_L. It is clear that $\zeta_L \leq \zeta_1$ for $L \geq 2$. Since $\zeta_1 - \zeta_L \geq 0$, we can then choose a natural number n, large enough to make $\frac{1}{n} < \zeta_1 - \zeta_L$. Now, from the numbers $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{k}{n}$ select the largest possible natural number k such that $\frac{k}{n} \leq \zeta_L$. Therefore, $\zeta_L < \frac{k+1}{n}$. Note that $\frac{k+1}{n} < \zeta_1$ since if we assume $\frac{k+1}{n} \leq \zeta_1$ then $\frac{1}{n} = \frac{k+1}{n} - \frac{k}{n} \geq \zeta_1 - \zeta_L$, which is false as n was picked such that $\frac{1}{n} < \zeta_1 - \zeta_L$. Thus, $\zeta_2 = \frac{k+1}{n}$ satisfies $\frac{1}{L} \leq \zeta_1 < \zeta_2 < \zeta_L \leq \frac{1}{L}$. This approach can be used for other ζ_i.

The above properties indicate that the distribution of ζ_i might not even be symmetric. Particularly the first and last eigenvalues tend to have a skewed distribution whilst the middle eigenvalue may have an asymptotically symmetric distribution. Furthermore, it indicates that ζ_i, particularly ζ_1 and ζ_L, converge asymptotically to $\frac{1}{L}$. Let us first evaluate the asymptotical behaviour of ζ_1 and ζ_L, for different values of N, generated from a white noise series (for simplicity, $L = 10$, ζ_1 and ζ_{10} are considered here). Fig. 1 displays the results for $m = 5 \times 10^3$ simulations, where $\overline{\zeta}_i = \left(\sum_{j=1}^{m} \zeta_{i,j}\right)/m$, $i = 1, 10$. As it appears from Fig. 1, the gap between ζ_1 and ζ_{10} becomes smaller as the sample size increases, and both converge to $\frac{1}{L}$. Thus, according to property 5, other eigenvalues tend to $\frac{1}{L}$.

Let us now consider the theoretical results for $L = 2$. Consider the random trajectory matrix $\mathbf{X}$ defined in Eq. (1). In this case $\mathbf{A} = \mathbf{X}\mathbf{X}^T$ is a square-symmetric matrix with the following eigenvalues:

$$\lambda_i = \frac{tr(\mathbf{A}) \pm \sqrt{tr^2(\mathbf{A}) - 4\,det(\mathbf{A})}}{2}, \qquad i = 1, 2.$$

Consequently, the eigenvalues of $\mathbf{A}/tr(\mathbf{A})$, ζ_1 and ζ_2 are as follows:

$$\zeta_i = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{4\,det(\mathbf{A})}{tr^2(\mathbf{A})}}, \qquad i = 1, 2.$$

In this case, we expect both ζ_1 and ζ_2 (or their averages after simulations, $\overline{\zeta}_1$, and $\overline{\zeta}_2$, respectively) would converge to 0.5 as there are only two eigenvalues.

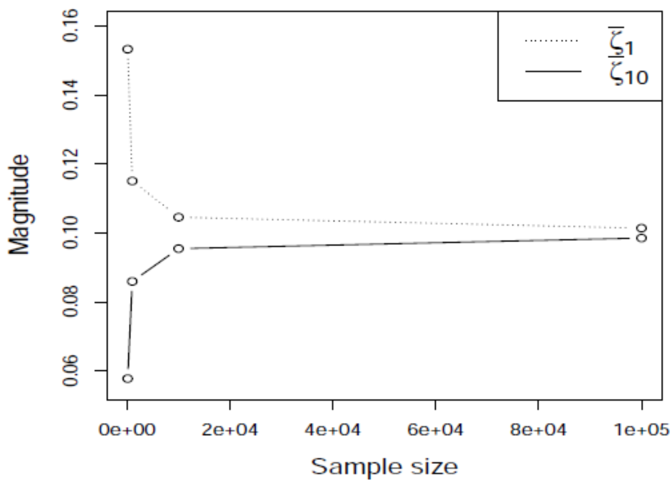

Figure 1: The plot of $\overline{\zeta}_i$, $(i = 1, 10)$ for different sample size N for a white noise series.

2. Conclusion

The distribution of the eigenvalues of the matrix $\mathbf{XX}^T / \sum_{i=1}^{L} \lambda_i$ was studied and several properties were introduced. As our future research, the theoretical distribution of the matrix $\mathbf{XX}^T / \sum_{i=1}^{L} \lambda_i$ is of interest to us. Furthermore, in an ongoing research we are evaluating the applicability of the results found here for noise reduction of the chaotic series. Additionally, we are applying the properties obtained here as extra criteria for filtering series with complex structure.

References

[1] R.B. D'Agostino, In: RB D'Agostino and MA Stephens (Eds.), *Goodness-Of-Fit Techniques*, Marcel Dekker (1986).

[2] G.W. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University Press, Cambridge (2010).

[3] T.W. Anderson, Asympototic theory for principal component analysis, *Ann. Math. Statist.* **34** (1963), 122–148.

[4] C. Beneki, and E. Silva, Analysing and forecasting European Union energy data, *International J. of Energy and Statistics*, **1** (2013), 127–141.

[5] M. Bulmer, *Principles of Statistics*, Dover, New York (1979).

[6] V.N. Chugunov, On the parametrization of classes of normal Hankel matrices, *Computational Math. and Math. Physics*, **51** (2011), 1823–1836.

[7] A. Edelman and N.R. Rao, Random matrix theory, *Acta Numer.*, **14** (2005), 233–297.

[8] H. Hassani, A. Soofi and A. Zhigljavsky, Predicting inflation dynamics with singular spectrum analysis, *J. of the Royal Statistical Society - Ser. A*, **176** (2013), 743–760.

[9] H. Hassani, D. Thomakos, A review on singular spectrum analysis for economic and financial time series, *Statistics and Its Interface*, **3** (2010), 377–397.

[10] H. Hassani and R. Mahmoudvand, Multivariate singular spectrum analysis: A general view and new vector forecasting approach, *International J. of Energy and Statistics*, **1** (2013), 55–83.

[11] H. Iranmanesh, M. Abdollahzadeh, A. Miranian, and H. Hassani, A developed wavelet-based local neuro fuzzy model for the forecasting of crude oil price, *International J. of Energy and Statistics*, **1** (2013), 171–193.

[12] K. Mardia, J. Kenet and J. Dibby, *Multivariate Analysis*, Academic Press, London (1995).

[13] M.L. Mehta, *Random Matrices*, Elsevier/Academic Press, Amsterdam (2004).

[14] L.A. Pastur, A simple approach to the global regime of Gaussian ensembles of random matrices, *Ukrainian Math. J.*, **57** (2005), 936–966.

[15] V. Peller, *Hankel Operators and Their Applications*, Springer, New York (2003).

[16] M. Qadrdan, M. Ghodsi, and J. Wu, Probabilistic wind power forecasting using a single forecast, *International J. of Energy and Statistics*, **1** (2013), 99–111.

[17] S. Sanei, M. Ghodsi and H. Hassani, An adaptive singular spectrum analysis approach to murmur detection from heart sounds, *Medical Engineering and Pphysics*, **33** (2011), 362–367.

[18] J. Wishart, Generalized product moment distribution in samples, *Biometrika*, **20** (1928), 32–52.