Statistics and Probability Letters 127 (2017) 85–96

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Fluctuations of Marchenko–Pastur limit of random matrices with dependent entries

Ayako Hasegawa [a], Noriyoshi Sakuma [b,*], Hiroaki Yoshida [a]

[a] Department of Information Sciences, Ochanomizu University, 2-1-1, Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan

[b] Department of Mathematics, Aichi University of Education, 1 Hirosawa, Igaya-cho, Kariya-shi, Aichi 448-8542, Japan

ARTICLE INFO

Article history:
Received 26 June 2016
Received in revised form 21 March 2017
Accepted 21 March 2017
Available online 4 April 2017

Keywords:
Free probability
Non-commutative probability
Random matrix
Moving average
Compound Poisson

ABSTRACT

Pfaffel and Schlemm have investigated the sample covariance matrix with dependent entries motivated from moving average. Our former paper gave free probabilistic interpretation of it. In this paper, we suggest a new statistical hypothesis testing to coefficients of the moving average model using real second order freeness, which describes fluctuations of its almost sure limiting distribution.

1. Models of random matrices with dependent entries

Random Matrix Theory(RMT) is one of the most important areas in statistics. After an epoch making discovery by Marçenko and Pastur (1967), studying the almost sure limiting empirical distribution of eigenvalues has attracted many researchers from theoretical researches to applications for much here fifty years. For details, see Bai and Silverstein (2010); Paul and Aue (2014). Recently, random matrices whose entries are the moving average model have been interested (Pfaffel and Schlemm, 2011, 2012). We call them moving average random matrices. In this case, entries of random matrices are not independent. They proved convergence of their spectral measures.

We treated moving average random matrices from free probability theory (FPT) in our former paper (Hasegawa et al., 2013). Here, we shall give brief explanation of FPT and why we used it. Voiculescu established FPT and found an important connection between FPT and RMT. This is called asymptotic free independence. In FPT, we treat random matrices as random variables from non-commutative probabilistic way and give analytic and combinatorial tools to study their limiting empirical distributions using free analogue of cumulant. We call it free cumulant. Since the sample covariance matrix with the moving average entries can be viewed as the compound Wishart matrix and we have the moment-cumulant formula for free cumulant, we can easily compute theoretical trace moments for moving average random matrices. It allowed us to establish new statistics. This is why we can investigate limiting distributions of them using free probabilistic tools. We gave free probabilistic proof and new interpretation for moving average random matrices. For details, see Voiculescu et al. (1992) and Hasegawa et al. (2013).

In this paper we will give new tests from a deeper analysis in FPT. It is called second or higher order freeness. Second order freeness is to see fluctuations of tracial moments $\mathrm{tr}(\boldsymbol{X}^k)$ of random matrix $\boldsymbol{X}$. Especially, for suitable random matrices their

* Corresponding author.
E-mail address: nori01mar@gmail.com (N. Sakuma).

http://dx.doi.org/10.1016/j.spl.2017.03.022
0167-7152/

fluctuations become Gaussian and we can compute their covariance using so-called second order free cumulants. Especially, for compound Wishart random matrix, we have explicit formulas of mean and variance of fluctuations via its second order free cumulants so that we can apply second order freeness to time series models. Therefore we introduce a new test for moving average random matrix.

Our proposal test has the following benefits: in a test for the goodness of the fitting to given model, our proposal method is simpler than usual method for an estimation of the coefficients of time series model via sample autocorrelation function(ACF). The amount of calculation in our method is much less than the sample ACF method. In usual method, the test is required for each sample ACF of the model but our method does not require to estimate any coefficient of the model. Moreover, our method to calculate a tracial moment of the compound Wishart matrix is applicable to the classification of a set of time series without any estimation of the coefficients of time series model. Namely, the tracial moment can be used as the characteristics of given time series.

In Section 2, a brief review of second order freeness is given. In Section 3, our new method are suggested. Hereafter, we recall the model and former results (Hasegawa et al., 2013).

Suppose that $\{c_\ell\}_{\ell=0}^\infty$ is the sequences in $\ell^2(\mathbb{R}) := \left\{\{\xi_i\}_{i=0}^\infty\,;\ \sum_{i=0}^\infty \xi_i^2 < \infty\right\}$, and let $\left\{Z_{i,j}\right\}_{i,j\geq 1}$ be an independent family of standard Gaussian random variables.

We set

$$X_{i,j} = \sum_{\ell=0}^{\infty} c_\ell\, Z_{i,j+\ell}, \tag{1}$$

which is the MA (moving average) modeled Gaussian process, and define the $N \times M$ random matrix $\boldsymbol{X}_{N,M}$ by

$$\boldsymbol{X}_{N,M} = \left(X_{i,j}\right)_{i,j}.$$

Let $\boldsymbol{Z}_{N,M}$ be an $N \times M$ matrix the entries of which are independent *real* standard Gaussian random variables. Then we consider the $N \times N$ symmetric matrix $\boldsymbol{W}_N$ as

$$\boldsymbol{W}_N = \frac{1}{N}\boldsymbol{X}_{N,M}{}^t\boldsymbol{X}_{N,M} = \frac{1}{N}\boldsymbol{Z}_{N,M}\boldsymbol{\Gamma}_c{}^t\boldsymbol{Z}_{N,M},$$

where $\boldsymbol{\Gamma}_c = \left(\gamma(i-j)\right)_{i,j}$ It is corresponding to the sample covariance matrix.

In Pfaffel and Schlemm (2011) and Hasegawa et al. (2013), we determine the limit of the empirical spectral measure of the random matrix $\boldsymbol{W}_N$ as $N \to \infty$. Here in the limit $N \to \infty$, we assume that M has the same order as N, that is, $M = M_N \to \infty$ such that $\lim_{N\to\infty} \frac{M_N}{N} = \lambda$ for some $\lambda > 0$. We call such a limit *the Marchenko–Pastur limit with the asymptotic ratio* λ.

Let ρ_c be the weak limit as $M \to \infty$ of the spectral measure of the $M \times M$ Toeplitz matrix $\boldsymbol{\Gamma}_c = \left(\gamma_c(i-j)\right)_{i,j}^M$, where $\gamma_c(h)$ is given by

$$\gamma_c(h) = \sum_{j=0}^{\infty} c_j c_{j+|h|}.$$

The functions $\gamma_c(h)$ is the autocovariance functions of the MA-model in the rows direction.

Remark 1.1. The existence of the above weak limit measures ρ_c is due to the ℓ^2 conditions on the sequence $\{c_\ell\}$. Since the Toeplitz matrices $\boldsymbol{\Gamma}$ is constituted of the autocovariance functions and positive definite, thus the measure ρ_c becomes the compactly supported probability measures with the support being in $\mathbb{R}_{\geq 0}$.

Theorem 1.2. *In the Marchenko–Pastur limit $N \to \infty$ with the asymptotic ratio $\frac{M}{N} \to \lambda$, the empirical spectral measure of $\frac{1}{N}\boldsymbol{X}_{N,M}{}^t\boldsymbol{X}_{N,M}$ converges almost surely to the compound free Poisson distribution $\pi(\lambda, \rho_c)$, where $\boxtimes$ denotes the free multiplicative convolution.*

About the compound free Poisson distribution, see Hasegawa et al. (2013). In this paper, we consider second order fluctuation of this model to estimate coefficients of MA model.

Remark 1.3. It is known by Szegö's theorem (see, for instance, Grenander and Szego, 1958) that the moments of the weak limit measure ρ of the spectral measure of symmetric Toeplitz matrix of the form $\boldsymbol{\Gamma} = \left(\gamma(i-j)\right)_{i,j}$ are calculable.

Namely, we denote by $f(\omega)$ the Fourier transform of the function $\gamma(h)$, that is,

$$f(\omega) = \sum_{h\in\mathbb{Z}} \gamma(h) e^{-\sqrt{-1}h\omega}.$$

Then for any analytic function F, it holds that

$$\int_{\mathbb{R}} F(x)\,\rho(dx) = \frac{1}{2\pi}\int_0^{2\pi} F\big(f(\omega)\big)d\omega.$$

Especially, the kth moment $m_k(\rho)$ of measure ρ is given by

$$m_k(\rho) = \frac{1}{2\pi} \int_0^{2\pi} \big(f(\omega)\big)^k d\omega.$$

Thus the moments of the limiting spectral measure of the compound free Poisson type in the above Theorem 1.2 can be calculated from the coefficient $\{c_\ell\}$ by using the free moment-cumulant formula.

2. The fluctuation formulas of the compound Wishart matrices

In this section, we shall review the fluctuation formulas of Wishart (including the case of compound Wishart) random matrices investigated in Mingo and Speicher (2006).

2.1. Target quantity

Consider a sequence of hermitian random matrices $(\boldsymbol{X}_N)_{N\in\mathbb{N}}$ and look at the random variable $\mathrm{tr}(\boldsymbol{X}_N^k)$. We assume that the first order limits, namely the moments

$$\alpha_k = \lim_{N\to\infty} \mathrm{E}\big[\mathrm{tr}(\boldsymbol{X}_N^k)\big]$$

exist for $k \geq 1$, where tr denotes the normalized trace.

Then one can find that, in many cases including Wishart random matrices, the fluctuation $\mathrm{tr}(\boldsymbol{X}_N^k) - \alpha_k$ is asymptotically Gaussian of order $1/N$, that is, the random variable

$$N\big(\mathrm{tr}(\boldsymbol{X}_N^k) - \alpha_k\big) = \mathrm{Tr}\big(\boldsymbol{X}_N^k - \alpha_k \mathbf{1}\big)$$

converges in distribution to a centered Gaussian random variable as $N \to \infty$, where Tr denotes the non-normalized trace, the simple sum of the diagonal entries.

The main information about the fluctuations of the family of the asymptotically Gaussian random variables $\{\mathrm{Tr}(\boldsymbol{X}_N^k - \alpha_k\mathbf{1})\}_{k\geq 1}$ is given by the covariances

$$\begin{aligned}\alpha_{p,q} &= \lim_{N\to\infty} \mathrm{Cov}\big(\mathrm{Tr}(\boldsymbol{X}_N^p), \mathrm{Tr}(\boldsymbol{X}_N^q)\big)\\ &= \lim_{N\to\infty} \mathrm{E}\big[\mathrm{Tr}(\boldsymbol{X}_N^p - \alpha_p\mathbf{1}) \cdot \mathrm{Tr}(\boldsymbol{X}_N^q - \alpha_q\mathbf{1})\big].\end{aligned}$$

Such the second order limits are called *the fluctuation moments*.

Let $\boldsymbol{Z}_{N,M}$ be an $N \times M$ matrix the entries of which are independent *complex* standard Gaussian random variables, and let $\boldsymbol{W}_N = \frac{1}{N}\boldsymbol{Z}_{N,M}\boldsymbol{Z}_{N,M}^*$. The $N \times N$ random matrix $\boldsymbol{W}_N$ is called *complex* Wishart matrix.

We will consider a sequence $(\boldsymbol{W}_N)$ of $N \times N$ *complex* Wishart matrices and take the Marchenko–Pastur limit with the asymptotic ratio $M/N \to \lambda > 0$. Then it is known that the first order and the second order limits exist, which are given by the following combinatorial formulas:

The first order limits (the moments):

$$\alpha_p = \lim_{N\to\infty} \mathrm{E}\big[\mathrm{tr}(\boldsymbol{W}_N^p)\big] = \sum_{\pi\in NC(p)} \lambda^{b(\pi)},$$

where $NC(p)$ is the set of the non-crossing partitions of $[1, p]$ and $b(\pi)$ denotes the number of blocks in the partition π.

The second order limits (the fluctuation moments):

$$\alpha_{p,q} = \lim_{N\to\infty} \mathrm{Cov}\big(\mathrm{Tr}(\boldsymbol{W}_N^p), \mathrm{Tr}(\boldsymbol{W}_N^q)\big) = \sum_{\pi\in S_{annNC}(p,q)} \lambda^{c(\pi)},$$

where $S_{annNC}(p, q)$ is the set of the non-crossing (p, q)-annular permutations and $c(\pi)$ denotes the number of cycles in the permutation π.

The combinatorial object, the non-crossing partitions, used in the formula for the first order limits is well-known (for instance, Nica and Speicher, 2006), but the annular non-crossing permutations used in the second order limits is lesser-known than the non-crossing partitions. Thus for our convenience, we shall briefly recall the definition of the annular non-crossing permutations which is the combinatorial object introduced by Mingo and Nica in Mingo and Nica (2004) in order to describe the nature of the expectation of the trace for the products of independent Gaussian random matrices.

We fix two positive integers p and q. A (p, q)-*annulus* is an annulus in which the numbers $1, 2, \ldots, p$ are arranged in clockwise order on the external circle, and the numbers $p + 1, p + 2, \ldots, p + q$ are arranged in anti-clockwise order on the internal circle. Let $C = (c_1, c_2, \ldots, c_k)$ be a cycle of order k with $c_i \in [1, p + q]$. We will represent this cycle inside the (p, q)-annulus by drawing an arrow from c_i to $c_{i+1} (i = 1, 2, \ldots, k)$ where $c_{k+1} = c_1$ with an orientation in clockwise.

A cycle C is called *internal* (resp. *external*) if all of whose elements are in the internal (resp. external) circle. A cycle C is called *connected* if it is a cycle which contains both elements in the external and the internal circles.

Let π be the permutation on $[1, p + q]$. Then $[1, p + q]$ is partitioned into the orbits of π. The orbit of π is called *cycle*, and we will usually write a permutation in the cycle notation. Although it is customary to omit the orbits with one element from the cycle notation, we will not remove any of them in our case. We denote

$c(\pi)$ = the number of cycles (orbits) of π,

which is one of the fundamental permutation statistics. For a cycle (an orbit) C in π, we denote by $|C|$ the order of the cycle (the length of the orbit) C.

Definition 1. *A non-crossing* (p, q)*-annular permutation* is a permutation π on $[1, p+q]$ for which we can draw the cycles in the cycle notation of π as oriented cycles inside (p, q)-annulus in non-crossing way and it has *at least one* connected cycle. We denote by $S_{annNC}(p, q)$ the set of all non-crossing (p, q)-annular permutations.

Remark 2.1. (1) If there is no connected cycle inside (p, q)-annulus on drawing the cycles in the permutation π on $[1, p+q]$, then it is essentially a disjoint union of two non-crossing permutations on $[1, p]$ and $[p + 1, p + q]$. Hence we will only consider connected annular non-crossing permutations.

(2) It is obvious that the map changing each cycle to a block does not always give a one-to-one correspondence between annular non-crossing permutations and annular non-crossing partitions. For instance, two (2, 1)-annular non-crossing permutations (1, 2, 3) and (2, 1, 3) are sent to the partition with only one block {1, 2, 3}. However, as shown in Mingo and Nica (2004, Proposition 4.4), if there are at least two connected cycles, then this map becomes a bijection. For more about non-crossing (p, q)-annular permutations, see the papers (Mingo and Nica, 2004; Mingo and Speicher, 2006; Collins et al., 2007; Mingo and Popa, 2013).

2.2. Compound case

Here we shall recall the formulas for compound case. We consider the sequence of complex *compound* Wishart matrices $(\boldsymbol{W}_N)$. Namely, $\boldsymbol{W}_N$ is given as $\boldsymbol{W}_N = \frac{1}{N}\boldsymbol{Z}_{N,M}\boldsymbol{D}_M\boldsymbol{Z}^*_{N,M}$, where $\boldsymbol{Z}_{N,M}$ is an $N \times M$ matrix the entries of which are independent complex standard Gaussian random variables, and $(\boldsymbol{D}_M)_{M\in\mathbb{N}}$ is a sequence of $M \times M$ non-random symmetric matrices for which the weak limit distribution ρ exists as $M \to \infty$.

Taking the Marchenko–Pastur limit with the asymptotic ratio $M/N \to \lambda$, the limit distribution of $(\boldsymbol{W}_N)$ is known as $\pi(\rho, \lambda)$, the compound free Poisson law.

We denote by $m_k(\rho)$ the kth moment of the limit distribution ρ of the sequence of matrices $(\boldsymbol{D}_M)_{M\in\mathbb{N}}$. Then the first order limit, the moment of $\pi(\rho, \lambda)$, is given by the combinatorial formula

$$\alpha_p = \lim_{N\to\infty} \mathrm{E}\big[\mathrm{tr}\big(\boldsymbol{W}_N^p\big)\big] = \sum_{\pi\in NC(p)} \prod_{B\in\pi} \lambda m_{|B|}(\rho),$$

where $NC(p)$ is the set of the non-crossing partitions on $[1, p]$ and $|B|$ stands for the size of a block B in the partition π.

In the case of compound, the formula of the moments depends not only on the number of blocks but also on the size of each block, which is well-known from the free moment-cumulant formula.

It has been derived by Mingo and Speicher in Mingo and Speicher (2006) that the formula of the second order limit, the fluctuation moments $\alpha_{p,q}$, for the sequence of complex compound Wishart matrices can be written in terms of non-crossing annular permutations as follows: in the Marchenko–Pastur limit with the asymptotic ratio $M/N \to \lambda$, the fluctuation moments are given by the combinatorial formula

$$\alpha_{p,q} = \lim_{N\to\infty} \mathrm{Cov}\big(\mathrm{Tr}\big(\boldsymbol{W}_N^p\big), \mathrm{Tr}\big(\boldsymbol{W}_N^q\big)\big) = \sum_{\pi\in S_{annNC}(p,q)} \prod_{C\in\pi} \lambda m_{|C|}(\rho),$$

where $S_{annNC}(p, q)$ is the set of the non-crossing (p, q)-annular permutations and $|C|$ stands for the order of the cycle C in the permutation π.

Similar to the first order case, the formula of the second order limit depends not only on the number of cycles but also on the order of each cycle.

Example 2.2. The non-crossing (2, 2)-annular permutations $S_{annNC}(2, 2)$ are listed below:

Type:	$[1^2\,2^1]$	$[2^2]$	$[1^1\,3^1]$	$[4^1]$
	(1)(3)(2, 4)	(1, 3)(2, 4)	(1)(2, 3, 4)	(1, 2, 3, 4)
	(1)(4)(2, 3)	(1, 4)(2, 3)	(1)(2, 4, 3)	(1, 2, 4, 3)
	(2)(3)(1, 4)		(2)(1, 3, 4)	(1, 3, 4, 2)
	(2)(4)(1, 3)		(2)(1, 4, 3)	(1, 4, 3, 2)
			(3)(1, 2, 4)	
			(3)(1, 4, 2)	
			(4)(1, 2, 3)	
			(4)(1, 3, 2)	
Weight:	$\lambda^3 m_1^2 m_2$	$\lambda^2 m_2^2$	$\lambda^2 m_1 m_3$	λm_4

The weight for each permutation of the type $[k_1^{\ell_1} k_2^{\ell_2} \cdots k_m^{\ell_m}]$ is given by

$$(\lambda\, m_{k_1})^{\ell_1} (\lambda\, m_{k_2})^{\ell_2} \cdots (\lambda\, m_{k_m})^{\ell_m}.$$

Thus the fluctuation moment $\alpha_{2,2}$ is given by

$$\begin{aligned}\alpha_{2,2} &= \lim_{N\to\infty} \mathrm{Cov}\big(\mathrm{Tr}(\boldsymbol{W}_N^2), \mathrm{Tr}(\boldsymbol{W}_N^2)\big) = \lim_{N\to\infty} \mathrm{Var}\big(\mathrm{Tr}(\boldsymbol{W}_N^2)\big) \\ &= 4\lambda^3 m_1^2 m_2 + 2\lambda^2 m_2^2 + 8\lambda^2 m_1 m_3 + 4\lambda m_4.\end{aligned}$$

Similarly, by enumerating the $(1, 3)$-annular permutations, we have the following formula:

$$\begin{aligned}\alpha_{3,1} &= \lim_{N\to\infty} \mathrm{Cov}\big(\mathrm{Tr}(\boldsymbol{W}_N^3), \mathrm{Tr}(\boldsymbol{W}_N)\big) \\ &= 3\lambda^3 m_1^2 m_2 + 3\lambda^2 m_2^2 + 6\lambda^2 m_1 m_3 + 3\lambda m_4,\end{aligned}$$

where we have abbreviated $m_k(\rho)$ by m_k.

For the case of Wishart matrices, the analytical formula for the first and the second order moments has been derived by the study of the higher order free cumulants in Collins et al. (2007).

Let $\{\alpha_n\}_{n\geq 1}$ be the moments (the first order limit) of the ensemble of random matrices. Then we write the generating power series $M(x)$ as

$$M(x) = 1 + \sum_{n\geq 1} \alpha_n x^n,$$

and consider the Cauchy transform

$$G(x) = \frac{1}{x} M\Big(\frac{1}{x}\Big)$$

as usual. Let $\{\alpha_{m,n}\}_{m,n\geq 1}$ be the fluctuation moments (second order limit) of the ensemble of random matrices. Similar to the first order limit, we define the bivariate generating power series $M(x, y)$ by

$$M(x, y) = \sum_{m,n\geq 1} \alpha_{m,n} x^m y^n,$$

and consider a kind of the second order Cauchy transform defined by

$$G(x, y) = \frac{M\Big(\frac{1}{x}, \frac{1}{y}\Big)}{xy}.$$

In case of the Wishart matrices, it has been found in Collins et al. (2007) that all of the higher order free cumulants will vanish and that the second order Cauchy transform is totally determined in terms of the first order one via the functional equality

$$G(x, y) = \frac{G'(x)\, G'(y)}{\big(G(x) - G(y)\big)^2} - \frac{1}{(x-y)^2},$$

from which, for instance, we can obtain the relations:

$$\begin{aligned}\alpha_{1,1} &= \alpha_2 - \alpha_1^2, \\ \alpha_{1,2} &= 2\alpha_3 - 4\alpha_2\alpha_1 + 2\alpha_1^3, \\ \alpha_{2,2} &= 4\alpha_4 - 8\alpha_3\alpha_1 - 6\alpha_2^2 + 16\alpha_2\alpha_1^2 - 6\alpha_1^4, \\ \alpha_{1,3} &= 3\alpha_4 - 6\alpha_3\alpha_1 - 2\alpha_2^2 + 6\alpha_2\alpha_1^2 - 2\alpha_1^4.\end{aligned}$$

This formula on the fluctuations of Wishart matrices has been also derived by Bai and Silverstein in Bai and Silverstein (2004, Lemma 1.1).

In case of compound free Poisson law, since the kth free cumulant r_k of the limit distribution is given by

$$r_k = \lambda\, m_k(\rho), \quad k \geq 1.$$

With the help of the free moment-cumulant formula, we can recover the fluctuation formulas above, that is, it is easy to find that the fluctuation moments can be expressed by the moments $\{m_k(\rho)\}$ and the asymptotic ratio λ.

The fluctuation formula reviewed above in this section is the one for *complex* Wishart matrices. But in our application, we need the formula for *real* Wishart matrices. Fortunately, the asymptotics and fluctuations for the real case have been deeply studied by Redelmeier in Redelmeier (2011) with the genus expansion. Then she has derived that the covariances (the second cumulants) for the case of *real* Wishart matrices are given as twice their value in the complex case. Namely, we have the following formulas:

Theorem 2.3. *Let $\boldsymbol{Z}_{N,M}$ be an $N \times M$ matrix the entries of which are independent real standard Gaussian random variables, and let $\left(\boldsymbol{W}_N\right)$ be a sequence of real compound Wishart matrices with $\boldsymbol{W}_N = \frac{1}{N}\boldsymbol{Z}_{N,M}\boldsymbol{D}_M\, {}^t\boldsymbol{Z}_{N,M}$, where the sequence of deterministic diagonal matrices $\left(\boldsymbol{D}_M\right)_{M\in\mathbb{N}}$ has the weak limiting spectral measure ρ.*

Then in the Marchenko–Pastur limit with the asymptotic ratio $M/N \to \lambda$, the fluctuation moment $\alpha_{p,q}$ is given by the formula,

$$\alpha_{p,q} = \lim_{N\to\infty} \mathrm{Cov}\left(Tr\left(\boldsymbol{W}_N^p\right), Tr\left(\boldsymbol{W}_N^q\right)\right) = 2\left(\sum_{\pi\in S_{annNC}(p,q)} \prod_{B\in\pi} \lambda m_{|C|}(\rho)\right),$$

where $m_k(\rho)$ stands for the kth moment of ρ, $S_{annNC}(p,q)$ is the set of the non-crossing (p,q)-annular permutations, and $|C|$ stands for the order of cycle C in the permutation π.

3. An application to time series analysis

We shall apply the formulas of the moments and the fluctuations for the real Wishart matrices to the statistical hypothesis testing of the time series models.

Let $\{c_\ell\}_{\ell=0}^{\infty}$ be an infinite real sequence such that the absolute values of which decay to 0 at least polynomially, and let $\left\{X_j\right\}_{j\in\mathbb{Z}}$ be an MA modeled Gaussian process described with the coefficients c_ℓ, that is,

$$X_j = \sum_{\ell=0}^{\infty} c_\ell Z_{j+\ell}.$$

Using this process, Pfaffel and Schlemm in Pfaffel and Schlemm (2012) considered the $N\times M$ random matrix of the form

$$\boldsymbol{X}_{N,M} = \left(X_{i,j}\right) = \left(X_{(i-1)M+j}\right)_{i,j} = \begin{pmatrix} X_1 & \dots & X_M \\ X_{M+1} & \dots & X_{2M} \\ \vdots & & \vdots \\ X_{(N-1)M+1} & \dots & X_{NM} \end{pmatrix} \in \mathbb{R}^{N\times M},$$

which has an interesting dependence across both rows and columns, that is, not all entries far away from each other are asymptotically independent, for instance, the correlation between the entries $X_{i,M}$ and $X_{i+1,1}$ $(i = 1,\dots,N-1)$ does not depend on M.

Then they have investigated the limiting spectral measure of the empirical spectral measure of the sample covariance matrix $\frac{1}{N}\boldsymbol{X}_{N,M}\, {}^t\boldsymbol{X}_{N,M}$ in the Marchenko–Pastur limit with the asymptotic ratio $M/N \to \lambda$. Although the rows of the above matrix $\boldsymbol{X}_{N,M}$ are not exactly independent of each other, the limiting spectral measure is the same one for the random matrix whose rows are independent copies of the MA process given by (1).

Hence the limiting spectral measure is given as the compound free Poisson law $\pi\left(\lambda,\rho\right)$, where the measure ρ is the weak limit of the spectral measure of the Toeplitz matrix $\boldsymbol{\Gamma}_c = \left(\gamma_c(i-j)\right)$ constituted of the autocovariance function $\gamma_c(h) = \sum_{j=0}^{\infty} c_j c_{j+|h|}$ of the MA process in (1).

Remark 3.1. As it claimed in Pfaffel and Schlemm (2012), the assumption that the sequence of the coefficients $\{c_\ell\}$ decays at least polynomially is not very restrictive. The MA model of finite degree is, of course, allowed, even more it also allows an ARMA model. Moreover, actual data for the time series analysis are not so strongly correlated in long-terms.

Now we shall give some illustrative examples of the statistical hypothesis testing on an MA models of finite degree.

For a given time series data $\{x_j\}_{j=1}^{L}$, we consider the MA model of degree q,

$$X_j = c_0 Z_j + c_1 Z_{j+1} + \cdots + c_q Z_{j+q}, \quad q\in\mathbb{N}.$$

Then we should like to give a statistical testing for whether the time series data $\{x_j\}$ can be regarded as a sample path of the above MA model or not. For this, we will use the moment and the fluctuation for the real Wishart matrices.

For some positive integers M and N with $NM \le L$, we construct the $N\times M$ matrix $\boldsymbol{x}$ from given data $\{x_j\}_{j=1}^{L}$ as follows:

$$\boldsymbol{x} = \left(\xi_{i,j}\right) \quad \text{with } \xi_{i,j} = x_{(i-1)M+j}.$$

Then we compute, for instance, the second moment μ_2 of the covariance matrix $\frac{1}{N}\boldsymbol{x}^t\boldsymbol{x}$, that is, $\mu_2 = \frac{1}{N}\mathrm{Tr}\left((\frac{1}{N}\boldsymbol{x}^t\boldsymbol{x})^2\right)$, which will be thought as an observed value of the testing statistic, the second moment of the real compound Wishart matrix $\frac{1}{N}\left(\boldsymbol{X}_{N,M}{}^t\boldsymbol{X}_{N,M}\right)$.

As we have mentioned, for large M and N, the second moment of the real compound Wishart matrix $\frac{1}{N}\left(\boldsymbol{X}_{N,M}{}^t\boldsymbol{X}_{N,M}\right)$ can be regarded as almost Gaussian distributed random variable, the mean and the variance of which depend only on the MA parameters $\{c_\ell\}$, the ratio $\lambda = M/N$, and N. In theory, the Gaussianity will be valid, of course, for large M and N asymptotically, but it will be found from the numerical computations below that, in the size 50×50 or even 30×30, we can regard this statistic (the second moment) is already well approximated by the Gaussian distribution.

As claimed in Remark 1.3, we know that the kth moment $m_k(\rho)$ of the weak limit measure ρ is given by

$$m_k(\rho) = \frac{1}{2\pi}\int_0^{2\pi} \big(f(\omega)\big)^k\, d\omega,$$

where $f(\omega)$ is the Fourier transform of the autocovariance function $\gamma_c(h)$. Thus $m_k(\rho)$ is determined only by the model parameters $\{c_\ell\}_{\ell=0}^{q}$.

Applying the fluctuation formula on the moments of real compound Wishart matrices reviewed in the previous section, we can regard that the testing statistic, the second moment, is Gaussian distributed by

$$\mathcal{N}\big(\alpha_2,\ \alpha_{2.2}/N^2\big) = \mathcal{N}\Big(\lambda^2 m_1^2 + \lambda m_2,\ 2\big(4\lambda^3 m_1^2 m_2 + 2\lambda^2 m_2^2 + 8\lambda^2 m_1 m_3 + 4\lambda m_4\big)/N^2\Big),$$

where we denote $m_k(\rho)$ simply by m_k.

3.1. Numerical computations

We consider the MA model of degree 2,

$$X_j = 1.0\,Z_j + 1.0\,Z_{j+1} + 1.0\,Z_{j+2},$$

and generate its sample paths for our numerical computations. Since the MA parameters are given as

$$c_0 = 1, \qquad c_1 = 1, \quad c_2 = 1, \quad \text{and} \quad c_k = 0 \quad (k \geq 3),$$

the autocovariance function $\gamma_c(h)$ becomes

$$\gamma_c(0) = 3, \qquad \gamma_c(1) = 2, \qquad \gamma_c(2) = 1, \quad \text{and} \quad \gamma_c(k) = 0 \quad (k \geq 3).$$

Thus its Fourier transform is obtained as

$$f(\omega) = 3 + 2\big(2\cos\omega + \cos 2\omega\big),$$

which gives the first four moments of the limit measure ρ as

$$m_1(\rho) = 3, \qquad m_2(\rho) = 19, \qquad m_3(\rho) = 141, \quad \text{and} \quad m_4(\rho) = 1107.$$

The numerical computations below have been performed by using *Mathematica* 7.

(A) *Case of $M = 50$ and $N = 50$.*

As the ratio is $\lambda = 1.0$, the first and second asymptotics, the moment and fluctuation for the second moment of the real compound Wishart matrix are given by $\alpha_2 = 28$ and $\alpha_{2,2} = 18436$, respectively. This is the case of $N = 50$, hence, the second moment μ_2 could be observed according to the Gaussian distribution

$$\mathcal{N}\big(\alpha_2,\ \alpha_{2.2}/N^2\big) = \mathcal{N}\big(28,\ 7.3852\big).$$

Here we shall give the results of the numerical computation on the second moment μ_2 for 50 000 samples of 50×50 data matrices. The mean and the variance of the samples are 28.1163 and 7.3327, respectively.

The histogram of 50 000 samples of μ_2 is in Fig. 5.1 with the dashed line for the scaled normal curve of $\mathcal{N}\big(28, 7.3852\big)$. The cumulative distribution function of 50 000 samples in Fig. 5.2, where the dashed line indicates the cumulative distribution function of the Gaussian distribution $\mathcal{N}\big(28,\ 7.3852\big)$.

It can be seen from Figs. 5.1 and 5.2 that 50 000 of the observed data for μ_2 are well approximated by the Gaussian distribution.

(B) *Case of $M = 30$ and $N = 30$.*

In this case, the ratio remains $\lambda = 1.0$ and the moment and fluctuation are unchanged as $\alpha_2 = 28$ and $\alpha_{2,2} = 18436$. Since this is the case of $N = 30$, the second moment μ_2 could be observed according to the Gaussian distribution

$$\mathcal{N}\big(\alpha_2,\ \alpha_{2.2}/N^2\big) = \mathcal{N}\big(28,\ 20.48444\big).$$

The mean and the variance of 50 000 samples of 30×30 matrices are 28.2428 and 20.7449, respectively. The histogram and the cumulative distribution function of 50 000 samples are found in Figs. 5.3 and 5.4, respectively, where the dashed lines stand for the Gaussian distribution $\mathcal{N}\big(28,\ 20.48444\big)$ similarly to the numerical computation (A).

We can see from Figs. 5.3 and 5.4 that even in the case of 30×30, the observed data for μ_2 are still well approximated by the Gaussian distribution.

(C) *Case of $M = 60$ and $N = 40$*

The ratio is given as $\lambda = 3/2$, thus the moment and fluctuation become $\alpha_2 = 195/4 = 48.75$ and $\alpha_{2,2} = 36378$, respectively. Since this is the case of $N = 40$, the second moment μ_2 could be observed according to the Gaussian distribution

$$\mathcal{N}\big(\mu_2, \alpha_{2.2}/N^2\big) = \mathcal{N}\big(48.75,\ 22.7363\big).$$

The mean and the variance of 50 000 samples of 40×60 matrices are 49.1418 and 23.3413, respectively. The histogram and the cumulative distribution function for them are in Figs. 5.5 and 5.6, respectively, where the dashed lines stand for the Gaussian distribution $\mathcal{N}\big(28,\ 20.48444\big)$ similarly to the above numerical computations.

We can also see from Figs. 5.5 and 5.6 that even in case of the different ratio $\lambda = 3/2$, the observed data for μ_2 are still well approximated by the Gaussian distribution.

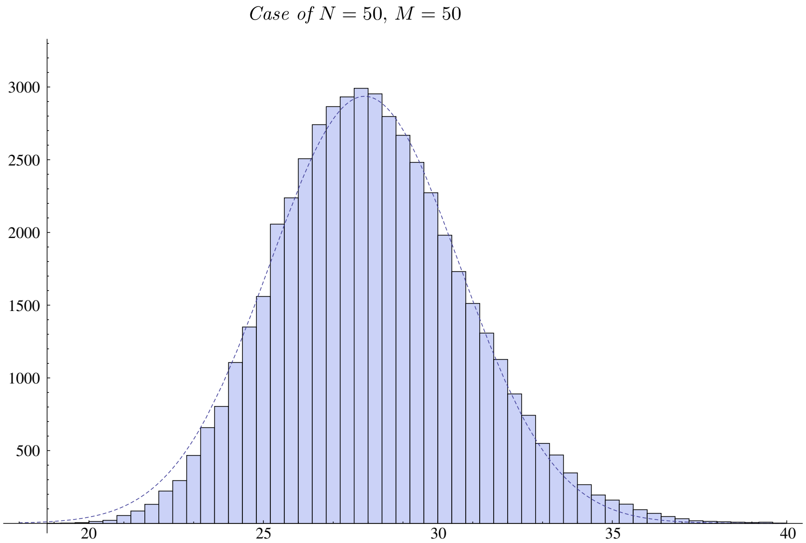

Fig. 5.1. Histogram of the second moments of 50 000 samples of 50×50.

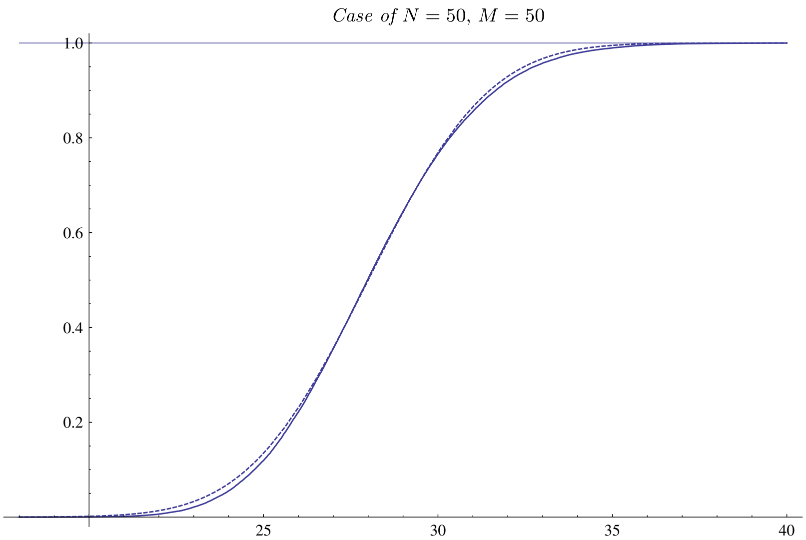

Fig. 5.2. Cumulative distribution function of 50 000 samples of 50×50.

3.2. Statistical hypothesis testing

From the above numerical observations, we should like to propose the statistical hypothesis testing on time series models.

(1) Let the time series data $\{x_j\}_{j=1}^{L}$ be given, and consider a certain time series model, for instance, an MA model of finite degree.
(2) Here the null hypothesis H_0 is that *given time series data could be generated by the model that we have assumed*.
(3) We construct the $N \times M$ matrix $\boldsymbol{x}$ from the given data as in (1), and compute the second moment of the sample covariance matrix $\mu_2 = \frac{1}{N}\mathrm{Tr}\big((\frac{1}{N}\mathbf{x}^t\mathbf{x})^2\big)$.
(4) Using the parameters of our assumed model, we estimate the mean and the variance of the Gaussian distribution to which the second moment should be distributed under the above null hypothesis.
(5) Then we shall apply the Z-test of two-tailed to see the statistical significance.

Remark 3.2. In above observation, we have used the second moment as the testing statistic, we can use, however, the moments of another orders alternately. Furthermore, we can choose the various ratio λ if the given time series data is long enough.

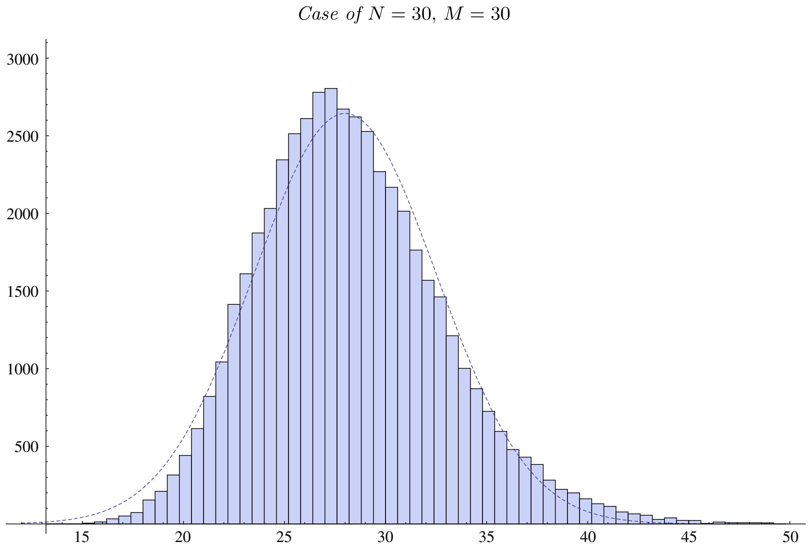

Fig. 5.3. Histogram of the second moments of 50 000 samples of 30 × 30.

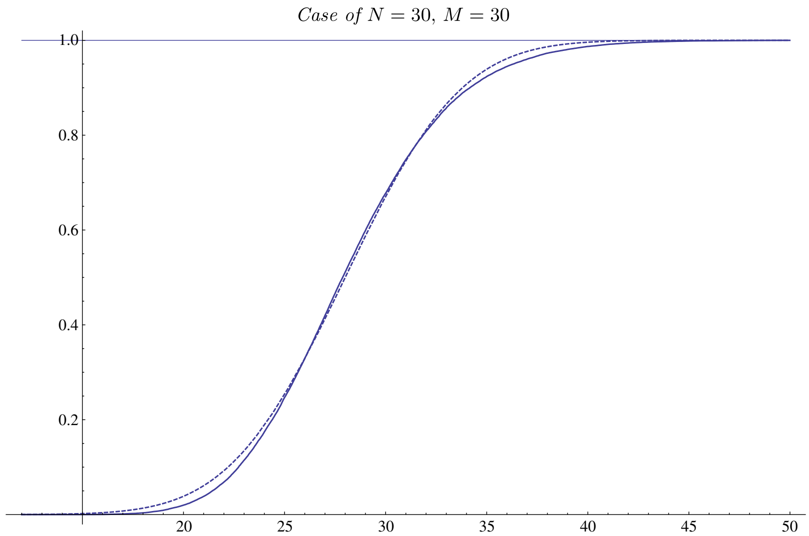

Fig. 5.4. Cumulative distribution function of 50 000 samples of 30 × 30.

Remark 3.3. The moments m_k of the weak limit ρ of the Toeplitz matrix $\mathbf{\Gamma}\big(\gamma(i-j)\big)$ with the autocovariance function γ of the time series model are needed to estimate the mean and the variance of the normal distribution for the Z-test. The explicit form of the autocovariance function, however, is not exactly needed but its Fourier transform is essential as we have seen above.

The stationary solution to the stochastic difference equation

$$b_0X_j + b_1X_{j+1} + \cdots + b_pX_{j+p} = c_0Z_j + c_1Z_{j+1} + \cdots + c_qZ_{j+q}$$

is called an ARMA(p, q) process. Although the autocovariance function of a general ARMA process does not have a simple closed form generally, its Fourier transform is simply given by

$$f(\omega) = \left| \frac{Q\left(e^{\sqrt{-1}\omega}\right)}{P\left(e^{\sqrt{-1}\omega}\right)} \right|^2, \quad \omega \in [0, 2\pi),$$

where P and Q are polynomials respectively given as

$$P(z) = b_0 + e_1z + b_2z^2 + \cdots + b_pz^p \quad \text{and} \quad Q(z) = c_0 + c_1z + c_2z^2 + \cdots + c_qz^q.$$

Thus the statistical hypothesis testing that we have proposed can be extended to an ARMA model without any change.

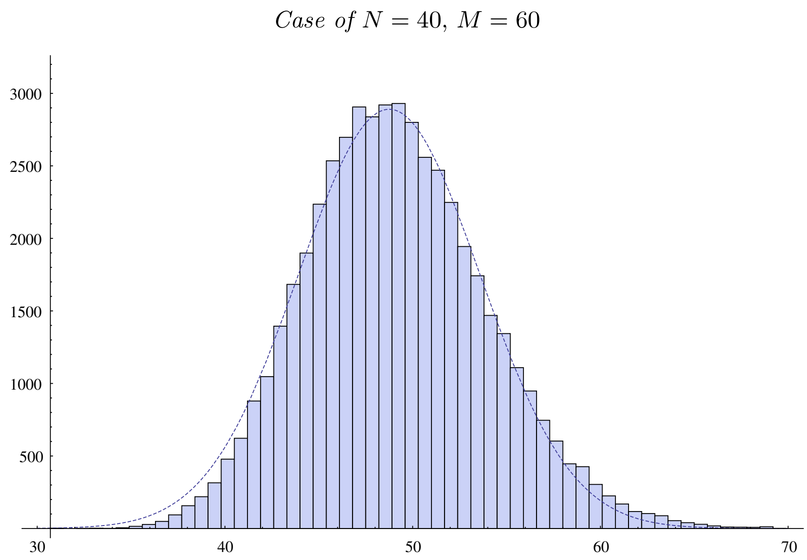

Fig. 5.5. Histogram of the second moments of 50 000 samples of 40 × 60.

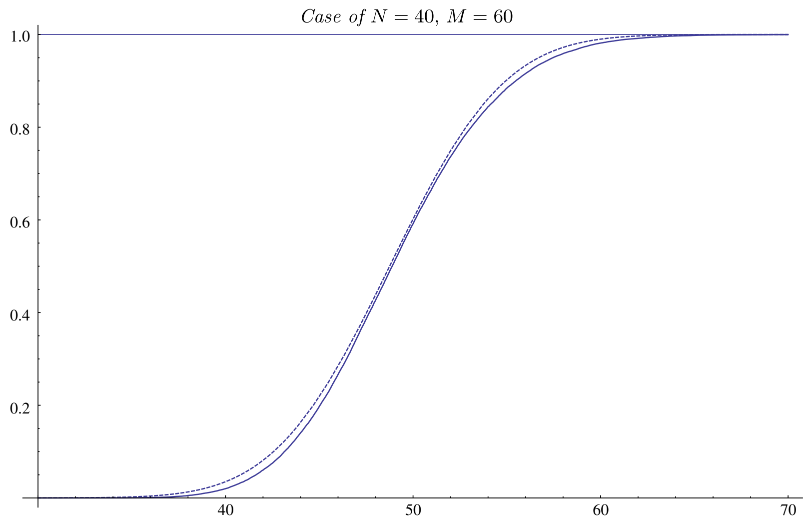

Fig. 5.6. Cumulative distribution function of 50 000 samples of 40 × 60.

3.3. Separation of alternative hypotheses

In a test of validity for a statistical model or goodness of fit, it is usual that an alternative hypothesis would not be set up explicitly. However we will illustrate below by numerical simulations how our testing statistic the z-value behaves under various alternative hypotheses and also see how the alternative hypothesis can be separated from the null hypothesis.

As the data for our simulations, we have generated 10 sample paths of length 2500 according to the MA model of degree 2 with the parameters $c_0 = 1.0,\ c_1 = 1.0,\ c_2 = 1.0$, and $c_k = 0\ (k \geq 3)$. For each generated sample path, we have made 50×50 square matrix $\boldsymbol{x}$ $(\lambda = 1)$ and computed the second moment of the covariance matrix, $\mu_2 = \frac{1}{N}\mathrm{Tr}\big((\frac{1}{N}\boldsymbol{x}^t\boldsymbol{x})^2\big)$. Then we have calculated the statistics the z-values under the alternative hypotheses given by varying the parameter c_1 or c_2.

In Table 1, the absolute values of the z-values are listed for 10 sample paths in the case where the parameter c_1 varies from 0.1 to 2.0 by 0.1 increment ($c_0 = 1.0$ and $c_2 = 1.0$ are unchanged). It is natural to regard that the most expected value of c_1 for each sample path should be given as one for the least absolute value of the z-values. Since we have generated sample paths for our simulation with $c_1 = 1.0$, thus $c_1 = 1.0$ is, of course, our desired value. Indeed it can be seen from Table 1 that in many cases we obtain $c_1 = 1.0$ as the most expected value. Sometimes $c_1 = 0.9$ or $c_1 = 1.1$ are selected but these are due to the random numbers in the simulated data.

In Table 2, the parameter c_2 varies from 0.1 to 2.0 by 0.1 increment ($c_0 = 1.0$ and $c_1 = 1.0$ are unchanged), and in many cases we obtain $c_2 = 1.0$ as the most expected value which is our desire.

Table 1
Case of $c_1 = 0.1 \times j$ ($j = 1, 2, \ldots, 20$), $c_2 = 1.0$.

c_1	1	2	3	4	5	6	7	8	9	10
0.1	23.8649	18.7611	23.7436	24.9143	23.2810	23.8328	19.1118	24.3576	16.5755	16.1874
0.2	21.5146	16.7949	21.4024	22.4850	20.9747	21.4849	17.1192	21.9702	14.7738	14.4150
0.3	18.3635	14.1526	18.2635	19.2293	17.8818	18.3371	14.4420	18.7700	12.3494	12.0292
0.4	15.0067	11.3288	14.9193	15.7629	14.5859	14.9836	11.5815	15.3617	9.7538	9.4741
0.5	11.7986	8.6204	11.7231	12.4521	11.4350	11.7787	8.8388	12.1054	7.2594	7.0177
0.6	8.8943	6.1593	8.8293	9.4566	8.5814	8.8771	6.3472	9.1583	4.9881	4.7801
0.7	6.3346	3.9822	6.2787	6.8183	6.0655	6.3198	4.1439	6.5617	2.9748	2.7960
0.8	4.1076	2.0814	4.0595	4.5242	3.8758	4.0949	2.2206	4.3032	1.2136	1.0596
0.9	2.1807	0.4310	2.1391	2.5404	1.9805	2.1697	0.5512	2.3496	0.3182	0.4513
1.0	0.5158	0.9995	0.4798	0.8274	0.3425	0.5063	0.8953	0.6621	1.6483	1.7636
1.1	0.9238	2.2403	0.9551	0.6531	1.0744	0.9321	2.1498	0.7967	2.8041	2.9042
1.2	2.1715	3.3190	2.1988	1.9355	2.3028	2.1787	3.2402	2.0607	3.8105	3.8977
1.3	3.2563	4.2599	3.2801	3.0499	3.3711	3.2626	4.1909	3.1594	4.6896	4.7659
1.4	4.2031	5.0836	4.2240	4.0221	4.3038	4.2086	5.0231	4.1181	5.4606	5.5276
1.5	5.0330	5.8080	5.0514	4.8737	5.1217	5.0379	5.7547	4.9582	6.1398	6.1987
1.6	5.7638	6.4480	5.7801	5.6232	5.8421	5.7681	6.4010	5.6978	6.7409	6.7930
1.7	6.4105	7.0163	6.4249	6.2859	6.4798	6.4143	6.9747	6.3520	7.2757	7.3218
1.8	6.9856	7.5235	6.9984	6.8750	7.0471	6.9890	7.4866	6.9337	7.7539	7.7948
1.9	7.4996	7.9786	7.5110	7.4011	7.5544	7.5026	7.9457	7.4534	8.1838	8.2202
2.0	7.9614	8.3891	7.9716	7.8735	8.0103	7.9641	8.3597	7.9201	8.5722	8.6048

Table 2
Case of $c_1 = 1.0$, $c_2 = 0.1 \times j$ ($j = 1, 2, \ldots, 20$).

c_2	1	2	3	4	5	6	7	8	9	10
0.1	21.9000	17.1271	21.7866	22.8813	21.3540	21.8700	17.4551	22.3607	15.0832	14.7203
0.2	18.8151	14.5419	18.7136	19.6938	18.3263	18.7883	14.8355	19.2276	12.7119	12.3870
0.3	15.7431	11.9558	15.6531	16.5218	15.3098	15.7193	12.2160	16.1087	10.3339	10.0459
0.4	12.8315	9.4960	12.7522	13.5173	12.4499	12.8106	9.7252	13.1535	8.0676	7.8140
0.5	10.1549	7.2279	10.0854	10.7568	9.8201	10.1365	7.4290	10.4375	5.9744	5.7518
0.6	7.7406	5.1764	7.6797	8.2678	7.4473	7.7245	5.3526	7.9881	4.0783	3.8833
0.7	5.5880	3.3427	5.5347	6.0497	5.3311	5.5739	3.4969	5.8048	2.3811	2.2104
0.8	3.6817	1.7146	3.6350	4.0862	3.4567	3.6694	1.8498	3.8716	0.8723	0.7227
0.9	1.9992	0.2741	1.9582	2.3539	1.8019	1.9884	0.3926	2.1658	0.4647	0.5959
1.0	0.5158	0.9995	0.4798	0.8274	0.3425	0.5063	0.8953	0.6621	1.6483	1.7636
1.1	0.7930	2.1263	0.8247	0.5189	0.9456	0.8014	2.0347	0.6643	2.6973	2.7986
1.2	1.9501	3.1256	1.9780	1.7084	2.0846	1.9575	3.0448	1.8366	3.6289	3.7183
1.3	2.9760	4.0145	3.0007	2.7624	3.0948	2.9825	3.9431	2.8757	4.4592	4.5382
1.4	3.8888	4.8083	3.9106	3.6997	3.9939	3.8945	4.7451	3.8000	5.2020	5.2719
1.5	4.7042	5.5201	4.7236	4.5364	4.7975	4.7093	5.4640	4.6254	5.8695	5.9315
1.6	5.4358	6.1614	5.4530	5.2866	5.5188	5.4404	6.1115	5.3658	6.4721	6.5273
1.7	6.0951	6.7418	6.1105	5.9622	6.1691	6.0992	6.6974	6.0327	7.0187	7.0679
1.8	6.6920	7.2696	6.7058	6.5733	6.7581	6.6957	7.2299	6.6363	7.5170	7.5609
1.9	7.2348	7.7518	7.2471	7.1285	7.2939	7.2380	7.7163	7.1849	7.9732	8.0125
2.0	7.7305	8.1942	7.7415	7.6351	7.7835	7.7334	8.1623	7.6857	8.3928	8.4280

It can be found from Tables 1 and 2 that undoubtedly wrong alternative hypotheses can be separated to be eliminated.

Acknowledgments

The authors are grateful to the anonymous referee for correcting inaccuracies and useful suggestions. Especially, the numerical simulations in Section 3.3 are suggested by the referee's comments.

N. Sakuma is supported by Scientific Research(C) 15K04923, JSPS. H. Yoshida is supported by Scientific Research(C) 26400112, JSPS.

References

Bai, Z.D., Silverstein, J.W., 2004. CLT for linear spectral statistics of large-dimensional sample covariance matrices. Ann. Probab. 32 (1A), 553–605.
Bai, Z., Silverstein, J.W., 2010. Spectral Analysis of Large Dimensional Random Matrices, second ed.. Springer, New York.
Collins, B., Mingo, J.A., Śniady, P., Speicher, R., 2007. Second order freeness and fluctuations of random matrices. III. Higher order freeness and free cumulants. Doc. Math. 12, 1–70.
Grenander, U., Szego, G., 1958. Toeplitz Forms and Their Applications. University of California Press, Berkeley and Los Angeles.
Hasegawa, A., Sakuma, N., Yoshida, H., 2013. Random matrices by MA models and compound free Poisson laws. Probab. Math. Statist. 33 (2), 243–254.
Marçenko, V.A., Pastur, L.A., 1967. Distribution of eigenvalues in certain sets of random matrices. Mat. Sb. (N.S.) 72 (114), 507–536.
Mingo, J.A., Nica, A., 2004. Annular noncrossing permutations and partitions, and second-order asymptotics for random matrices. Int. Math. Res. Not. IMRN 1413–1460.

Mingo, J.A., Popa, M., 2013. Real second order freeness and Haar orthogonal matrices. J. Math. Phys. 54, 051701.
Mingo, J.A., Speicher, R., 2006. Second order freeness and fluctuations of random matrices. I. Gaussian and Wishart matrices and cyclic Fock spaces. J. Funct. Anal. 235, 226–270.
Nica, A., Speicher, R., 2006. Lectures on the Combinatorics of Free Probability. In: London Mathematical Society Lecture Note Series, vol. 335. Cambridge University Press, Cambridge, p. xvi+417.
Paul, D., Aue, A., 2014. Random matrix theory in statistics: a review. J. Statist. Plann. Inference 150, 1–29.
Pfaffel, O., Schlemm, E., 2011. Eigenvalue distribution of large sample covariance matrices of linear processes. Probab. Math. Statist. 31 (2), 313–329.
Pfaffel, O., Schlemm, E., 2012. Limiting spectral distribution of a new random matrix model with dependence across rows and columns. Linear Algebra Appl. 436 (9), 2966–2979.
Redelmeier, C.E.I., 2011. Genus expansion for real Wishart matrices. J. Theoret. Probab. 24, 1044–1062.
Voiculescu, D.V., Dykema, K.J., Nica, A., 1992. Free Random Variables. A Noncommutative Probability Approach to Free Products with Applications to Random Matrices, Operator Algebras and Harmonic Analysis on Free Groups. In: CRM Monograph Series, vol. 1. American Mathematical Society, Providence, RI, p. vi+70.