Random Matrix Time Series

Peiyuan Teng · Min Xu

Accepted: 14 July 2023 / Published online: 3 August 2023
© Grace Scientific Publishing 2023

Abstract
In this paper, a time series model is proposed, where the coefficients take values from random matrix ensembles. Formal definitions, theoretical solutions, and statistical properties are derived. Estimation and forecast methodologies for random matrix time series are discussed, accompanied by examples. Additionally, the paper suggests random matrix differential equations and explores potential applications of the time series model at the end.

Keywords Random matrix theory · Time series · Maximum likelihood estimation

1 Introduction
Random matrix theories were first introduced in statistics by Wishart [1]. In 1951, they were introduced into nuclear physics by Wigner [2]. Since then, random matrix theories have become a rich research topic in both the Mathematics and Physics communities. In 1962, Dyson introduced the classification of random matrix ensembles based on symmetries [3], which are the Gaussian unitary ensemble (GUE), the Gaussian orthogonal ensemble (GOE), and the Gaussian symplectic ensemble (GSE). Various applications can be found in the research of statistics, integrable systems, operator algebras, and number theory, as well as in statistical physics, quantum chromodynamics, and quantum gravity [4].

The Gaussian orthogonal ensemble (GOE) is a widely used random matrix ensemble that can be defined by symmetric matrix with random elements. Specifically, the diagonal elements $h_{ii}$ of the random matrix $H$ are assumed to be normally distributed with mean 0 and variance 2, while the off-diagonal elements are independent and identically distributed with mean 0 and variance 1. This joint distribution of the independent matrix elements of the symmetric random matrix can be described by a density function of

$\Phi$ Peiyuan Teng
teng.73@osu.edu

1 Department of Physics, The Ohio State University, 191 West Woodruff Ave, Columbus 43210, OH, USA
\[
\frac{1}{Z_{GOE(n)}} e^{-\frac{n}{4} \text{tr} H^2},
\]

on the space of \( n \times n \) real symmetric matrices \( H \). Here \( Z_{GOE(n)} \) is a normalization constant.

The Gaussian unitary ensemble (GUE) is defined by a random Hermitian matrix whose diagonal elements \( H_{ii} \sim N(0, 1) \). Its off-diagonal elements satisfy \( H_{ij} \sim N(0, 0.5) + N(0, 0.5)i \) with \( i < j \). The joint distribution of the independent matrix elements is described by the density

\[
\frac{1}{Z_{GUE(n)}} e^{-\frac{n}{2} \text{tr} H^2},
\]

on the space of \( n \times n \) Hermitian matrices \( H \). \( Z_{GUE(n)} \) is a normalization constant.

The joint distribution of the independent matrix elements of the Gaussian symplectic ensemble (GSE) is described by the density

\[
\frac{1}{Z_{GSE(n)}} e^{-n \text{tr} H^2},
\]

on the space of \( n \times n \) Hermitian quaternionic matrices \( H \). \( Z_{GSE(n)} \) is a normalization constant.

In statistics, time series analysis refers to methods for analyzing time-series data. Statistics and other characteristics can be acquired by fitting a time series model with actual data. Those fitted characteristics or time series model parameters can also be used to generate forecasts. Typical time series models are autoregressive (AR) models, vector autoregression (VAR), autoregressive integrated moving average (ARIMA) models, and generalized autoregressive conditional heteroskedasticity (GARCH) models. Readers can refer to Ref. [5] for a detailed review of those models.

The vector autoregression (VAR) models can be viewed as another starting point for this paper. A pth-order VAR multivariate time series denoted as \( \text{VAR}(p) \), can be written as:

\[
X_T = c + A_1 X_{T-1} + A_2 X_{T-2} + \cdots + A_p X_{T-p} + e_T.
\]

The variable \( X_T \) represents a kth-order vector, which is a vector of multiple different variables all evaluated at the same time \( T \). The variable \( X_{T-i} \) represents the \( i \)th lag of \( X_T \). The variable \( c \) is the intercept of the model. \( A_p \) represents time-invariant \( k \times k \) matrices, and \( e_T \) represents the error term.

Since both random matrix theory and time series theory are very useful statistical methods, it is natural to consider combining these two theories. Several pioneering works have been published in this direction. For example, in Ref. [6], the authors use methods of random matrix theory to analyze the cross-correlation matrix of stock price changes and find that it exhibits the universal properties of the Gaussian orthogonal ensemble of random matrices. Subsequent research can be found in [7–9]. Time series have been employed to characterize random matrix eigenspectra. In Ref. [10], time series are used to characterize energy spectrum fluctuations of quantum systems. In
Ref. [11], time-series techniques are applied to separate or “detrend” the secular part of the spectrum from the local 1/f fluctuations. In Ref. [12], the same time series techniques are used to demonstrate that local fluctuations may exhibit crossover behavior for nonstandard ensembles. In Ref. [13], this crossover behavior appears to resolve the long-standing problem of nonergodicity in RMT. On the other hand, RMT has been employed to characterize the dynamics of time series. For example, Ref. [14] proposes a random matrix theory (RMT) approach to distinguish chaotic time series from noise. The authors of this paper initially set out to study generic time series models connected by a random matrix and eventually arrived at a form that they later realized could be seen as a generalization of VAR.

In this paper, these two theories are combined from a new perspective. Time series vectors are defined iteratively, with iteration coefficients taking values from random matrix ensembles. For simplicity, we refer to our time series models as nth-order time series with random matrix coefficients, or RMTS(n). A formal definition will be provided, and based on this definition, theoretical solutions and statistical properties such as expectation and variance will be derived. Furthermore, we extend these derivations to include complex random matrices. Estimation and forecast methodologies for random matrix time series will also be discussed, along with examples. To facilitate comparison, the first-order random matrix differential equation is defined in parallel to the random matrix time series. Finally, potential applications will be suggested at the end.

This paper is organized as follows: In Sect. 2, we research first-order random matrix time series models, RMTS(1). In Sect. 3, we discuss nth-order models, RMTS(n). In Sect. 4, we propose the first-order random matrix differential equation. Section 5 suggests some potential applications.

2 First-Order Time Series with Random Matrix Coefficients, RMTS(1)

The first-order time series with random matrix coefficients, or RMTS(1), can be defined similarly to VAR(1) (see Eq. 3). The difference is that the error term is no longer needed since it can be absorbed into \( b(T) \), as defined in Eq. (4), which is a random variable. Instead, the matrix coefficient, which is the random matrix, contains not only the estimation parameters but also absorbs randomness.

**Definition:** A first-order vector autoregressive time series with random matrix coefficients, denoted as RMTS(1), is defined by the iterative relation:

\[
X(T + 1) = A(T)X(T) + b(T). \tag{4}
\]

In this equation, \( X(T) \) represents the vector time series, \( A(T) \) is a random matrix at time \( T \), and \( b(T) \) is a random vector with independent Gaussian random variables as its elements. The random matrix \( A(T) \) can take values from a Gaussian ensemble, or more generally, a matrix with random number elements. The density distribution of \( A(T) \) is assumed to be time-invariant.

Readers may notice that if \( A(T) \) and \( b(T) \) are fixed numbers instead of random numbers, Eq. (4) essentially becomes a difference equation. Due to this similarity,
many methods in difference equations can be extended similarly to the study of this type of time series. For a comprehensive discussion of difference equations, readers can refer to Ref. [15].

2.1 Solutions of RMTS(1)

In this part, we will derive an analytical solution for Eq. (4). We can start with a homogeneous RMTS(1) with no constant term.

Consider the case when \( b(T) = 0 \), we have equation:

\[
X(T + 1) = A(T)X(T).
\]

We define the product of random matrices as:

\[
V(T) = \prod_{i=0}^{T} A(i). \tag{5}
\]

Therefore, using the iterative relation, we can easily obtain:

\[
X(T + 1) = V(T)X(0).
\]

Now let’s consider a generic case where \( b(T) \neq 0 \). Assuming \( V(T) \) has an inverse for every \( T \) (which is true in general due to the randomness of \( A(T) \)), we define \( C(T) \) by the equation (with \( C(0) = X(0) \)):

\[
X(T + 1) = V(T + 1)C(T + 1). \tag{6}
\]

Substituting this equation in Eq. (4), we have:

\[
V(T + 1)C(T + 1) = X(T + 1) = A(T)X(T) + b(T) = A(T)V(T)C(T) + b(T) = V(T + 1)C(T) + b(T),
\]

This can be written as:

\[
V(T + 1)C(T + 1) = V(T + 1)C(T) + b(T),
\]

Multiply both sides of this equation by \( V(T + 1)^{-1} \), we have

\[
C(T + 1) - C(T) = V(T + 1)^{-1}b(T).
\]

Therefore:

\[
C(T + 1) = C(0) + \sum_{i=0}^{T} V(i + 1)^{-1}b(i).
\]
The solutions for \( X(T + 1) \) can be written as:

\[
X(T + 1) = V(T + 1)(C(0) + \sum_{i=0}^{T} V(i + 1)^{-1} b(i)),
\]

(7)

where the matrix \( V(T) \) is defined as:

\[
V(T) = \prod_{i=0}^{T} A(i).
\]

**Products of random matrices:**

The variable \( V(T) \) represents a product of random matrices. The study of the product of random matrices is an ongoing research direction in random matrix theory, and it has yielded valuable theoretical results and practical applications. From a theoretical perspective, existing research has focused on topics such as the probability density of the products of random matrices [16, 17], the joint probability density of eigenvalues [18], and the singular values [19]. These results have found applications in various fields, including matrix-valued diffusions [20], quantum chromodynamics [21], and wireless telecommunication [22], among others. For a comprehensive discussion, readers can refer to Ref. [16], which provides an analytical representation of \( V(T) \) in the context of the product of random matrices.

In Sect. 4, the distributions of the product of random matrices will play a significant role in the discussion of solutions to differential equations.

### 2.2 Maximum Likelihood Estimation for RMST(1)

Generically, each element \( a_{ij} \) of the random matrix \( A \) is assumed to follow a distribution \( a_{ij} \sim N(r_{ij}, \sigma_{ij}^2) \), and these distributions are assumed to be time-invariant. When dealing with actual data, it is important to derive the methodology for parameter estimation.

From Eq.(4), utilizing the properties of the sum of Gaussian random variables, it is evident that the ith row of \( A(T)X(T) \) follows a distribution of 

\[
N \left( \sum_j (r_{ij} X_j(T)) + \sum_j \sigma_{ij}^2 X_j(T)^2, \sum_j \sigma_{ij}^2 X_j(T)^2 + \sigma_{bi}^2 \right).
\]

Assuming \( b_i(T) \sim N(b_i, \sigma_{bi}^2) \), the ith row of \( X(T + 1) \) then follows a distribution of

\[
N \left( \sum_j (r_{ij} X_j(T) + b_i), \sum_j \sigma_{ij}^2 X_j(T)^2 + \sigma_{bi}^2 \right).
\]

(8)

Given the observations from period 0 to \( T_f + 1 \), the likelihood function can be written as

\[
L(r_{ij}, \sigma_{ij}, b_i, \sigma_{bi}^2; X(0)|X(1), X(1)|X(2), ..., X(T_f)|X(T_f + 1))
\]
Table 1: Time series parameters

| Case | Matrix A dimensions | \( r_{ij} (i \neq j) \) | \( \sigma_{ij} (i \neq j) \) | \( r_{ij} (i = j) \) | \( \sigma_{ij} (i = j) \) | \( b_i \) | \( \sigma_{bi} \) | T |
|------|---------------------|-----------------|-----------------|-----------------|-----------------|---------|---------|-----|
| 1    | 5 \times 5          | 0               | 0.125           | 0               | 0.125           | 0       | 0.5     | 50,000 |
| 2    | 5 \times 5          | 0               | 0.125           | 0               | 0.125           | 0       | 0.5     | 500,000 |
| 3    | 10 \times 10        | 0               | 0.1             | 0.1             | 0.1             | 1       | 1       | 10,000 |

Table 2: Maximum likelihood estimation results

| Case | -log-likelihood | \( r_{ij} (i \neq j) \) | \( \sigma_{ij} (i \neq j) \) | \( r_{ij} (i = j) \) | \( \sigma_{ij} (i = j) \) | \( b_i \) | \( \sigma_{bi} \) | iterations |
|------|-----------------|-----------------|-----------------|-----------------|-----------------|---------|---------|---------------|
| 1    | 192,214.4       | 1.382e−4        | 1.239e−1        | 1.115e−3        | −1.360e−1       | −2.010e−3| 5.010e−1 | 393           |
| 2    | 1,916,179.2     | 4.219e−4        | −1.236e−1       | −1.932e−4       | −1.274e−1       | 5.142e−4 | 4.991e−1 | 495           |
| 3    | 765,322.9       | 1.336e−4        | 1.000e−1        | 1.025e−1        | −9.058e−2       | 0.9973  | 1.003   | 576           |

\[
= \prod_{T=0}^{T_f+1} \prod_i \frac{1}{\sqrt{2\pi (\sum_j \sigma_{ij}^2 X_j(T) + \sigma_{bi}^2)}} e^{-\frac{(X_i(T+1)-\sum_j r_{ij} X_j(T)-b_i)^2}{2(\sum_j \sigma_{ij}^2 X_j(T) + \sigma_{bi}^2)}}.
\] (9)

Given \( X(T) \), we can find the best-estimated parameters by maximizing this function. In practice, if we have more than one time series available (with the same parameter distribution), the likelihood function can be written accordingly by multiplying the likelihood of all the available time series.

For demonstration and verification purposes, we perform maximum likelihood estimations with synthetic data. Specifically, \( a_{ij} \sim N(r_{ij}, \sigma_{ij}^2) \) and \( b_i(T) \sim N(b_i, \sigma_{bi}^2) \) are sampled from predetermined distributions. The RMTS(1) can then be constructed by iteration using Eq. (4). With the time series constructed, the maximum likelihood estimation procedure will yield the estimated parameters in \( a_{ij} \sim N(r_{ij}, \sigma_{ij}^2) \) and \( b_i(T) \sim N(b_i, \sigma_{bi}^2) \). These estimated parameters can be compared with our initial setups to assess estimation accuracy.

Table 1 lists the parameters that are used to generate the time series. Three different cases are generated. For simplicity, we assume all off-diagonal elements have the same distribution \( N(r_{ij} (i \neq j), \sigma_{ij}^2 (i \neq j)) \), and all diagonal elements have the same distribution \( N(r_{ij} (i = j), \sigma_{ij}^2 (i = j)) \), as shown in Table 1. The initial condition of the time series is not shown here since it is not relevant in determining the parameter for the time series.

Using the generated synthetic time series, we present the estimation results in Table 2. All parameters are initialized with a value of one. The Nelder-Mead method is employed for the optimizations.

From this table, we can observe that the estimation achieves high accuracy, and the accuracy improves as the number of points (T) in the time series increases. It is worth noting that in practical scenarios, the estimations of standard deviations may have a negative sign, but this will not affect the variances.
2.3 Statistical Properties of X(T + 1)

In this section, we will derive statistical properties for X(T + 1), such as expectations, variance, and covariance.

We assume that each element $a_{ij}(T)$ in $A(T)$ is independent of $X(T)$, and follows a normal distribution with mean $r_{ij}$ and variance $\sigma^2_{ij}$. Similarly, $b_{i}(T)$ follows a normal distribution with mean $b_{i}$ and variance $\sigma^2_{b_{i}}$. In this paper, we consider the scenario where these random coefficients have the same distributions across different time points (time-invariant).

2.3.1 Expectation

To find the relations satisfied by the expectation value, we start with the iteration relation (Eq. 4) and apply the law of iterated expectations to this equation. According to the law of iterated expectations, the expectation values satisfy:

$$E(X(T + 1)) = E(E(X(T + 1)|X(T))) = E(RX(T) + b(T)) = RE(X(T)) + b.$$  \hspace{1cm} (10)

Here, $R$ represents the matrix of expectations corresponding to $r_{ij}$, and $b$ represents the vector of expectations corresponding to $b_{i}(T)$. It is important to note that the expectation value of the random matrix $A(T)$ is denoted as $R$.

If the iteration converges, the converged expectation $e$ can be obtained as follows:

$$e = (I - R)^{-1}b.$$  

By subtracting the equation $e = Re + b$ from the equation of iterated expectation, we obtain:

$$E(X(T + 1)) - (I - R)^{-1}b = R(E(X(T)) - (I - R)^{-1}b).$$

Thus, we have:

$$E(X(T)) - (I - R)^{-1}b = R^T(E(X(0)) - (I - R)^{-1}b).$$  \hspace{1cm} (11)

Here, $R^T$ represents the product of $T$ (an integer) $R$ matrices, not the transpose operation.

Assuming that $R$ has an eigendecomposition given by:

$$R = QAQ^{-1},$$  \hspace{1cm} (12)

where $\Lambda$ is the eigenvalue matrix with eigenvalues $\lambda_i$, we can conclude that a necessary condition for convergence is $|\lambda_i| < 1$. 

\[\square\]
2.3.2 Variance

Similar to expectation, we can derive properties of the variance. By the law of iterated variance, we have

\[ \text{Var}(X(T + 1)) = E(\text{Var}(X(T + 1)|X(T))) + \text{Var}(E(X(T + 1)|X(T))), \] (13)

specifically,

\[ \text{Var}(E(X(T + 1)|X(T))) = \text{Var}(RX(T) + b) = R^{\otimes 2}\text{Var}(X(T)), \]

and

\[ E(\text{Var}(X_i(T + 1)|X(T))) = E(\sum_j \sigma^2_{ij} X_j(T)^2 + \sigma^2_{bi}) = \sum_j \sigma^2_{ij} E(X_j(T)^2) + \sigma^2_{bi}. \]

In matrix notation, define \( \Sigma^{\otimes 2} \) (elementwise square) as the matrix representation of \( \sigma^2_{ij} \), and similarly for \( R^{\otimes 2} \). Plug into Eq. (13), we have

\[ \text{Var}(X(T + 1)) = \Sigma^{\otimes 2} E(X^2(T)) + \Sigma^2_b + R^{\otimes 2}\text{Var}(X(T)), \]

which can also be expressed as

\[ \text{Var}(X(T + 1)) = \Sigma^2_b + (\Sigma^{\otimes 2} + R^{\otimes 2})\text{Var}(X(T)) + \Sigma^{\otimes 2}(E(X(T)))^2. \]

When the expectation value converges, we have

\[ \text{Var}(X(T + 1)) = (\Sigma^{\otimes 2} + R^{\otimes 2})\text{Var}(X(T)) + \Sigma^2_b + \Sigma^{\otimes 2}((I - R)^{-1}b)^{\otimes 2}. \] (14)

Similar to the expectation, by requiring the iterated variance to be convergent, we have

\[ \text{Var}(X^*) = (I - \Sigma^{\otimes 2} - R^{\otimes 2})^{-1}(\Sigma^2_b + \Sigma^{\otimes 2}((I - R)^{-1}b)^{\otimes 2}). \] (15)

A necessary condition for convergence is: the eigenvalues of \( \Sigma^{\otimes 2} + R^{\otimes 2} \) are less than 1.

2.3.3 Covariance

By the law of iterated covariance, we have

\[ \text{Cov}(X_i(T + 1), X_j(T + 1)) = E(\text{Cov}(X_i(T + 1), X_j(T + 1)|X(T))) \]
\[ + \text{Cov}(E(X_i(T + 1)|X(T)), E(X_j(T + 1)|X(T))). \] (16)
When \( i \neq j \),

\[
E(Cov(X_i(T+1), X_j(T+1)|X(T))) = E \left( Cov \left( \sum_k a_{ik}(T)X_k(T), \sum_l a_{jl}(T)X_l(T)|X(T) \right) \right).
\]

With

\[
E(X_i(T+1)|X(T)) = \sum_k r_{ik}X_k(T) + b_i,
\]

we have

\[
Cov(E(X_i(T+1)|X(T)), E(X_j(T+1)|X(T))) = \sum_{kl} r_{ik}r_{jl}Cov(X_k(T), X_l(T)).
\]

So

\[
Cov(X_i(T+1), X_j(T+1)) = E \left( Cov \left( \sum_k a_{ik}(T)X_k(T), \sum_l a_{jl}(T)X_l(T)|X(T) \right) \right)
+ \sum_{kl} r_{ik}r_{jl}Cov(X_k(T), X_l(T)).
\]

(17)

Normally, the first term will not be zero unless \( a_{ik}(T) \) and \( a_{jl}(T) \) are independent. In the case of the Gaussian orthogonal ensemble (GOE) and Gaussian unitary ensemble (GUE), the symmetric constraints imposed on the random matrices will impact the iteration relation mentioned above. We will further discuss these cases in the examples below.

2.4 Examples

2.4.1 RMTS(1) from GOE

The Gaussian orthogonal ensemble (GOE) involves sampling from random symmetric matrices, where the upper diagonal entries are independent. In this ensemble, the diagonal elements satisfy \( A_{ii} \sim N(0, 2) \), while the off-diagonal elements follow \( A_{ij} \sim N(0, 1) \) for \( i < j \), with the additional constraint that \( A_{ij} = A_{ji} \).

In this case, we have \( R = 0 \), and the expectation \( E(X(T+1)) \) is equal to \( b \).

The variance is given by:

\[
Var(X(T+1)) = \Sigma^2_b + \Sigma^2 \circ b^2 + \Sigma^2 \circ Var(X(T)),
\]

(18)

The symmetric constraint does not affect the calculation of variance since the elements in the same row of a random matrix are independent.

For the GOE, if we define \( \text{Ones} \) as a matrix with all elements equal to one, then:

\[
\Sigma^2 \circ = \text{Ones} + I.
\]

It is important to note that under these parameters, the variance will not be stable.
However, if we introduce a scaling factor $f$ to $a_{ij}$, the deviation matrix becomes $f\Sigma$.

The variance may converge to:

$$Var(X_\infty) = (I - f^2\Sigma^{-2})^{-1}(\Sigma^2_b + f^2\Sigma^{o2}b^2).$$

(19)

For GOE, where $a_{ij} = a_{ji}$, Eq. (17) exhibits nonvanishing terms when $k = j$ and $l = i$.

$$Cov(X_i(T + 1), X_j(T + 1)) = E(Cov(a_i(T)X_j(T), a_{ji}(T)X_i(T)|X(T))) = Var(a_i(T))E(X_i(T)X_j(T)),$$

which equals

$$Cov(X_i(T + 1), X_j(T + 1)) = \sigma_{ij}^2E(X_i(T)X_j(T)) = \sigma_{ij}^2(E(X_i(T))E(X_j(T)) + Cov(X_i(T), X_j(T))).$$

So

$$Cov(X_i(T + 1), X_j(T + 1)) = \sigma_{ij}^2(b_ib_j + Cov(X_i(T), X_j(T))).$$

(20)

The covariance is stable or converges when $\sigma_{ij}^2 < 1$ (after multiplying a constant factor on the random matrix).

For the purpose of verification, we conducted simulations using synthetic time series. In these simulations, $a_{ij}$ and $b_i(T)$ were sampled from predetermined GOE distributions and normal distributions, respectively. The RMTS(1) was then constructed through iteration using Eq. (4). By comparing the simulated parameters with the theoretical results, we can assess the accuracy of the simulations.

Table 3 presents a comparison between the simulation results and the theoretical results.

The results above demonstrate the correctness of our derivations.

### 2.4.2 RMTS(1) from GUE

For complex random variables, the properties of variances and covariances are slightly different due to the modulus in the definition of variance and covariance.
For example, the variance of a complex random variable is defined as,

\[ \text{Var}(X(T + 1)) = E(|X(T) - E(X(T))|^2) = E(|X(T)|^2) - |E(X(T))|^2). \] (21)

From Eq. (13), we have

\[ \text{Var}(E(X(T + 1)|X(T))) = \text{Var}(RX(T) + b) = |R|^2 \text{Var}(X(T)). \]

With

\[ E(\text{Var}(X_i(T + 1)|X(T))) = E \left( \sum_j |\sigma_{ij}|^2 |X_j(T)|^2 + |\sigma_{bi}|^2 \right) \]

\[ = \sum_j |\sigma_{ij}|^2 E(|X_j(T)|^2) + |\sigma_{bi}|^2, \]

we have

\[ \text{Var}(X(T + 1)) = |\Sigma_b|^2 + (|\Sigma|)^2 + |R|^2 \text{Var}(X(T)) + |\Sigma|^2 (|E(X(T))|)^2. \]

which may be stable at

\[ \text{Var}(X^*) = (I - |\Sigma|^2 - |R|^2)^{-1}(|\Sigma_b|^2 + |\Sigma|^2 (|I - R|^{-1}b)|^2). \] (22)

The covariance for complex random variables is defined as

\[ \text{Cov}(X_i(T), X_j(T)) = E((X_i(T) - E(X_i(T)))(X_j(T) - E(X_j(T))). \] (23)

For GUE, \( a_{ij} = \bar{a}_{ji} \), Eq. (17) has nonvanishing terms when \( k = j \) and \( l = i \). So

\[ \text{Cov}(X_i(T + 1), X_j(T + 1)) = E(\text{Cov}(a_{ij}(T)X_j(T), a_{ji}(T)X_i(T)|X(T))) \]

\[ = \text{Cov}(a_{ij}(T), a_{ij}(T))E(X_j(T)X_i(T)), \]

which equals

\[ \text{Cov}(X_i(T + 1), X_j(T + 1)) \]

\[ = \text{Cov}(a_{ij}(T), a_{ij}(T))(E(X_j(T)E(X_i(T)) + \text{Cov}(X_i(T), X_j(T))). \]

\( \text{Cov}(a_{ij}(T), a_{ij}(T)) \) is actually the pseudo-variance of \( a_{ij} \), and in our special case of GUE, evaluate to \( \text{Re}(\sigma_{ij})^2 - \text{Im}(\sigma_{ij})^2 \)

So

\[ \text{Cov}(X_i(T + 1), X_j(T + 1)) = (\text{Re}(\sigma_{ij})^2 - \text{Im}(\sigma_{ij})^2)(\bar{b}_i b_j + \text{Cov}(X_i(T), X_j(T))) \] (24)
The Gaussian unitary ensemble (GUE) is represented by random Hermitian matrices. The diagonal elements $A_{ii} \sim N(0, 1)$, and the off-diagonal elements satisfy $A_{ij} \sim N(0, 0.5) + N(0, 0.5)i$ with $i < j$. In this case, $R = 0$ and $E(X(T+1)) = b$.

Similar to GOE, a comparison can be carried out to verify these relations.

Since the GSE (Gaussian symplectic ensemble) is represented using quaternion matrices, and the probability theory of quaternion numbers has received limited research attention due to their noncommutativity, we will not provide any examples of GSE in this context.

### 3 Nth-Order Time Series with Random Matrix Coefficients, RMTS(n)

Similar to RMTS(1), we can easily define RMTS(n) as follows:

**Definition:** An nth-order vector autoregressive time series with random matrix coefficients, denoted by RMTS(n), is defined by the iteration relation:

$$X(T+1) = \sum_{i=T-n+1}^{T} A(T_i)X(T_i)) + b(T). \quad (25)$$

This equation can be converted to an RMTS(1) by rewriting the iteration relation in matrix form:

$$\begin{bmatrix}
    X(T+1) \\
    X(T) \\
    \vdots \\
    X(T-n+2)
\end{bmatrix} =
\begin{bmatrix}
    A(T) & A(T-1) & \cdots & A(T-n+1) \\
    I & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & I
\end{bmatrix}
\begin{bmatrix}
    X(T) \\
    X(T-1) \\
    \vdots \\
    X(T-n+1)
\end{bmatrix} +
\begin{bmatrix}
    b(T) \\
    0 \\
    \vdots \\
    0
\end{bmatrix}. \quad (26)$$

As a result, the properties of RMTS(1) directly apply to the equation above.

### 4 First-Order Random Matrix Differential Equations

With the derived properties of random matrix time series, it is natural for us to consider the continuous case for comparison, which is the first-order random matrix differential equations.

A typical form of the first-order random matrix differential equations can be written as follows:

$$\frac{dX(T)}{dT} = A(T)X(T) + b(T). \quad (27)$$

Here, $A(T)$ and $b(T)$ have the same definition as in the time series. Notice that when this equation is discretized, it becomes an RMTS(1) with coefficients $\Delta T(A(T) + I)$ and $\Delta Tb(T)$.

We can find the solution to this equation by taking the limit of the corresponding discretized time series. For example, consider the trivial case where $b(T) = 0$. By separating $T$ into $n$ equally spaced intervals and iterating, it is easy to see that the solution is:
\[ X(T) = \lim_{n \to \infty} \prod_{n} \left( \frac{T A(t)}{n} + 1 \right)^{n} X(0). \]

We define the limit as \( \text{RMEXP}(A, \mathbf{T}) \) if it exists, and then our solution can be expressed as:

\[ X(T) = \text{RMEXP}(A, \mathbf{T}) X(0). \]

Using this definition in Eq. (7), the solution, when \( b(T) \neq 0 \), can be written as:

\[ X(T) = \text{RMEXP}(A, \mathbf{T}) \left( X(0) + \int_{t=0}^{T} \text{RMEXP}(A, t)^{-1} b(t) dt \right). \quad (28) \]

Here, the integral should be understood as the summation of a random process. Notice that the solution is similar to the corresponding one in ordinary differential equations.

In general, computing \( \text{RMEXP}(\mathbf{A}, \mathbf{T}) \) can be difficult, but a few known solutions exist. The simplest example is when \( A(t) \) is a 1x1 random scalar and \( b(T) = 0 \). Each scalar is drawn from a normal distribution with zero mean and unit variance.

In this case, exact solutions have been derived [23]. The limit, defined as the variable \( Y \),

\[ \text{RMEXP}(a, \mathbf{T}) = Y = \lim_{n \to \infty} \prod_{t} \left( \frac{T a(t)}{n} + 1 \right)^{n}, \]

has a probability density function of

\[ \mathcal{P}(Y) = \frac{1}{Y \sqrt{2\pi T}} \exp \left[ -\frac{(\log Y + T^2/2)^2}{2T^2} \right]. \]

which is lognormal.

When \( b(T) = 0 \) and \( A(t) \) is a 2x2 matrix, with each element of \( \mathbf{A}(\mathbf{T}) \) drawn from a normal distribution with zero mean and unit variance, the random matrix exponential \( \text{RMEXP}(\mathbf{A}, \mathbf{T}) \) is still exactly solvable. The limiting matrix probability distribution is very complicated, and its analytical expression is omitted here. Readers can refer to Ref. [23] for further details.

Normally, the statistical properties of \( X(T) \), such as expectation and variance, will diverge as \( T \) goes to infinity. This is due to the specific form of this differential equation 27 (since an identity matrix is added to \( \mathbf{A}(\mathbf{T}) \)).

With the ongoing fruitful research on the products of random matrices, we expect that more analytical solutions for such differential equations can be derived, and this topic can be systematically investigated.

5 Applications

Time series models are commonly used to uncover the underlying structure of observed data or to fit a model for making forecasts. These models have diverse applications, including economic forecasting and quality control, among others. In this article,
we propose a simple scenario where our random matrix time series models can be employed.

In this scenario, we interpret the vector $X(T)$ as the count of restaurants belonging to a specific category (e.g., traditional, fast food) in a city at time $T$. It is reasonable to assume that the number of restaurants in one category at the next timestamp is influenced by the number of restaurants in another category at the current timestamp, reflecting market competition. Additionally, the number of restaurants in the same category at the next timestamp depends on the number of restaurants in the same category at the current timestamp due to closures and internal competition. The coefficient matrix $A(T)$ is employed to capture these dependencies. The coefficient $b(T)$ within the same category can be interpreted as the count of newly opened restaurants in that category.

To illustrate this process, we generate synthetic data resembling the maximum likelihood estimation. Table 4 presents the parameters used to generate the time series. These values represent the dependency coefficients discussed in the scenario above.

Using this synthetic time series generated, we list the estimation results in Table 5. All parameters are initialized with one. Powell’s method is used in the optimizations.

From this table, we can see that the estimation achieves good accuracy and captures the dynamics between categories in the data.

The authors believe that our RMTS model can fit all data that can be fitted using VAR (Vector Autoregression) models. This is because our model is a well-defined generalization of VAR and only requires a time series vector as input. We believe that if the correlations of the actual time series components are random in nature, our RMTS model can provide valuable insights.
6 Conclusions

In this paper, we discussed random matrix time series with numerically verified properties. This approach is novel and raises interesting questions and promising directions for exploring more structures with random matrices and identifying potential applications.

Acknowledgements This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors would like to express their gratitude to the reviewers of this paper for their valuable suggestions.

Declarations

Conflict of interest No potential conflict of interest was reported by the author(s).

References

1. Wishart J (1928) The generalised product moment distribution in samples from a normal multivariate population. Biometrika 20(1/2):32–52
2. Wigner EP (1951) On the statistical distribution of the widths and spacings of nuclear resonance levels. In: Mathematical proceedings of the Cambridge philosophical society, vol 47, pp 790–798 (1951). Cambridge University Press
3. Dyson FJ (1962) Statistical theory of the energy levels of complex systems. I. J Math Phys 3(1):140–156
4. Forrester P, Snaith N, Verbaarschot J (2003) Developments in random matrix theory. J Phys A: Math Gen 36(12):1
5. Box GEP, Jenkins G (1990) Time series analysis. Forecasting and control. Holden-Day Inc, USA
6. Plerou V, Gopikrishnan P, Rosenow B, Amaral LAN, Stanley HE (1999) Universal and nonuniversal properties of cross correlations in financial time series. Phys Rev Lett 83(7):1471
7. Biely C, Thurner S (2008) Random matrix ensembles of time-lagged correlation matrices: derivation of eigenvalue spectra and analysis of financial time-series. Quant Finance 8(7):705–722
8. Laloux L, Cizeau P, Potters M, Bouchaud J-P (2000) Random matrix theory and financial correlations. Int J Theor Appl Finance 3(03):391–397
9. Ormerod P (2008) Random matrix theory and macro-economic time-series: an illustration using the evolution of business cycle synchronisation, 1886–2006. Economics 2(1)
10. Relano A, Gómez J, Molina R, Retamosa J, Faleiro E (2002) Quantum chaos and 1/f noise. Phys Rev Lett 89(24):244102
11. Fossion R, Vargas GT, Vieyra JL (2013) Random-matrix spectra as a time series. Phys Rev E 88(6):060902
12. Torres-Vargas G, Méndez-Bermúdez J, López-Vieyra J, Fossion R (2018) Crossover in nonstandard random-matrix spectral fluctuations in financial market systems. Phys Rev E 98(2):022110
13. Berkovits R (2020) Super-poissonian behavior of the Rosenzweig–Porter model in the nonergodic extended regime. Phys Rev B 102(16):165140
14. Ye B, Chen J, Ju C, Li H, Wang X (2017) Distinguishing chaotic time series from noise: a random matrix approach. Commun Nonlinear Sci Numer Simul 44:284–291
15. Agarwal RP (2000) Difference equations and inequalities: theory, methods, and applications. Chapman & Hall/CRC Pure and Applied Mathematics. CRC Press, USA
16. Ipsen JR (2015) Products of independent gaussian random matrices. arXiv preprint arXiv:1510.06128
17. Akemann G, Ipsen JR (2015) Recent exact and asymptotic results for products of independent random matrices. arXiv preprint arXiv:1502.01667
18. Ipsen JR, Kieburg M (2014) Weak commutation relations and eigenvalue statistics for products of rectangular random matrices. Phys Rev E 89(3):032106
19. Akemann G, Ipsen JR, Kieburg M (2013) Products of rectangular random matrices: singular values and progressive scattering. Phys Rev E 88(5):052118
20. Gudowska-Nowak E, Janik RA, Jurkiewicz J, Nowak MA (2003) Infinite products of large random matrices and matrix-valued diffusion. Nucl Phys B 670(3):479–507
21. Osborn JC (2004) Universal results from an alternate random-matrix model for qcd with a baryon chemical potential. Phys Rev Lett 93(22):222001
22. Tulino AM, Verdú S et al (2004) Random matrix theory and wireless communications. Found Trends Commun Inf Theory 1(1):1–182
23. Jackson AD, Lautrup B, Johansen P, Nielsen M (2002) Products of random matrices. Phys Rev E 66(6):066124

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.