Empirical deviations of semicircle law in mixed-matrix ensembles

Mehmet Süzen

Empirical deviations of semicircle law in mixed-matrix ensembles

Mehmet Süzen*

(Dated: December 3, 2021)

An algorithm is introduced for sampling a set of matrices from mixed orders random matrix ensembles, i.e., Mixed Matrix Ensemble Sampling (MMES). The concept of *the degree of mixture* of the matrix ensemble provides a balanced sampling of the mixed matrix ensemble. As an application of MMES, we have shown how the semicircle law deviates from the conventional behaviour in mixed Gaussian Orthogonal Ensemble (mGOE) as a novel finding.

INTRODUCTION

Random matrix theory (RMT) [1–3] originates from mathematical statistics [4] and quantum statistical mechanics [5–8]. The applications of RMT appear in diverse fields, such as multivariate statistics [9], quantum chaos [10], neuroscience [11], and in deep learning [12]. The primary operation of RMT techniques is focused on sampling random matrices from a given conventional ensembles, i.e., random matrix generation procedure. However, it is established that conventional ensembles set matrix order to a single value.

Sampling from a mixed matrix ensemble that enables a balanced sampling over different matrix orders requires a formal procedure. This procedure should be applicable to any existing matrix ensemble as an additional step. Adjustment of how much mixture should there be in the sample, distribution of different order set of matrices, should also be available as a free parameter.

In this direction, we have developed a concept of *the degree of mixture*. This concept enables us to draw different order matrices from the base non-mixed ensemble to generate a sample of the mixed ensemble in a more representative and controlled manner, i.e., Mixed Matrix Ensemble Sampling (MMES).

Due to different order matrices in the mixed ensemble, periodic spectra is imposed for each matrix member in the spectral analysis. Hence, we investigated the deviation of mGOE from the semicircle law and also report the nearest-neighbour spacing density over different degree of mixture values, as an application of MMES.

MMES in Statistics and Machine Learning

Apart from the pure theoretical interest in RMT on how to construct a mixed matrix ensemble, there are many possible practical usages of MMES procedure that provides a fair sampling of processes that generates or requires different order matrices. However, an immeditate usage of MMES algorithm in machine learning appears to be initialisation and understanding of neural networks weight matrices [13, 14]. While, different size neural network layers generates different order weight matrices, MMES can provide a balanced sampling from a designated matrix ensemble for initialisation. On the otherhand, how learned weight matrices would look like structurally compare to a known random matrix ensemble's mixed versions is an active interest in Neural Architecture Search (NAS) [15]

As MMES can also be used in any random matrix ensemble's mixed version, sampling from the mixed Wishart ensemble in multivariate statistics [9] can be achived with MMES algorithm as well.

METHODS

The primary contribution of the work is two fold. We introduce an algoritm for a balanced sampling from a mixed order matrix ensemble. To demonstrate, the utility of the algorithm, an approach to spectral analysis for mixed matrix ensembles is placed in a comparable setting to non-mixed matrix ensemble's spectral analysis.

MMES Algorithm

A random matrix ensemble is defined as a mixed ensemble $\mathscr{M}$ if its M constituent matrices appear in varying orders s_i, the i-th matrix, $s_i \in [2, N]$. Whereby non-mixed ensemble is the conventional matrix ensembles of matrix orders of only N.

The degree of the mixture expresses the ratio of different matrix orders in M matrices, i.e., m_i is the number of matrices of order s_i. This definition brings a constraint on the mixing distribution, with distribution of m_i, $\rho^m(m_i; s_i)$, such that $\sum_i m_i = M$. This definition of degree of mixture appears as a Binomial process. This enables a sample of mixed matrix ensemble drawn in a balanced way given degree of mixture. Hence resulting set of matrices which are fair representation of the mixed matrix ensemble.

The Binomial distribution can be used to define the degree of the mixture $\mu \in [0, 1.0]$ naturally obeying the mentioned constrained. The success probability would play the role of degree of mixture. Given number of success would be the matrix order of the non-mixed ensemble M. Hence, Binomial distribution's density, ρ^{Bin} will give us the distribution of m_is, $\rho^m(m_i; s_i) = \rho^{Bin}(m_i^{'})$,

where $m_i^{'}$s are drawn from the Binomial distribution, $m_i^{'} \sim Bin(\mu, M)$,μ is the degree of the mixture on the ensemble size of M. This procedure explicity explained in Algoritm 1 in detail.

The spectral densities of each element from the mixed ensemble $\mathscr{M}$ contains different orders of spectra. We use so called periodic boundary conditions on the spectra of matrices in the mixed ensemble. It implies, given ith matrix with the order of $m_i < N$ and its eigenvalue spectrum $[\lambda_1, \lambda_2, ..., \lambda_{m_i}]$ will be transformed into vector of size N by repeating the eigenvalues in a cycle. This procedure would align all matrices spectrum sizes in the mixed ensemble and make the size of the collected of eigenvalues in the ensemble equivalent in sample size to the standard base ensemble, i.e., M matrices of order N making eigenvalue collected size of $M \times N$. This concept introduced earlier [15] in the context of deep learning.

Algorithm 1: Mixed Matrix Ensemble Sampling (MMES)

Input:
$\mu \in [0, 1]$: Degree of mixture.
N : Maximum matrix order in the mixed matrix ensemble.
M : Number of matrices to draw.
$\mathscr{G}(m)$: Non-mixed ensemble sampler, draws a matrix of order m from the base matrix ensemble
Output:
$\mathscr{S}$: A sample from a mixed matrix ensemble $\mathscr{M}(2; N)$, M matrices of variying orders, $s_i \in [2, N]$, and varying counts of m_i s.t. $\sum_i m_i = M$.

```
/* Mixed Matrix Ensemble Sampler          */
S ← []  // Initalise vector of mixed order matrices.
s, m = Bin_non-zero(μ, M)  // Vectors of size k , non-zero orders s and corresponding counts m drawn from Binomial distribution.
for i = 0 up to k do
    /* Draw m_i times an order s_i matrices.  */
    for j = 0 up to m_i do
        /* Draw order s_i matrix from non-mixed ensemble.  */
        S = G(s_i)
        /* Add order s_i matrix to mixed sample.  */
    S ← S
Return S
```

MIXED GAUSSIAN ORTHOGONAL ENSEMBLE

Gaussian Orthogonal Ensemble (GOE) is oe of the most common ensemble introduced in RMT studies in Physics [2]. In this section we study mixed GOE (mGOE), the deviation from the semicircle law and how nearest-spacing distribution given various values of the degree of mixtures. Conventional GOE would be generated with drawing number from Gaussian distribution for a matrix A of order N, $A \in \mathcal{N}(0, 1)$ then $GOE(N) = \frac{1}{2}(A + A^T)$, with further adjusting off-diagonal elements with $offdiag(GOE) \in \mathcal{N}(0, 1/2)$, resulting a symmetric real matrix. The usage of A in sampling the ensemble is crucial, as this guarantees real eigenvalues are outputted from numerical libraries [2]. We see deviation from the semicircle law in lower degree of mixture, in the tails, forming an M shape rather than perfect semicircle, in Figure 1. This behaviour vanishes where the degree of mixture approaches the vanishing value of 1.0. For the nearest-neighbour spacing distribution, we see that distribution shifts slightly, but this is not prominent, in Figure 2 for smaller values of the degree of mixture..

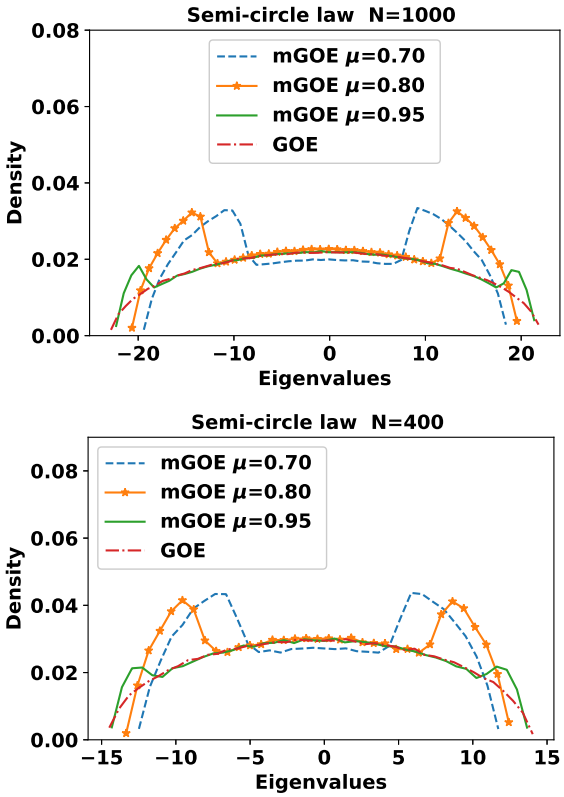

FIG. 1. Semicircle law of mixed GOE with different degree of mixing and non-mixing for N=400 and N=1000 over 100 draws, i.e., ensemble size.

CONCLUSION

We have developed a sampling algorithm to draw a set of matrices from mixed random matrix ensembles. Introduction of the concept of *the degree of the mixture*,

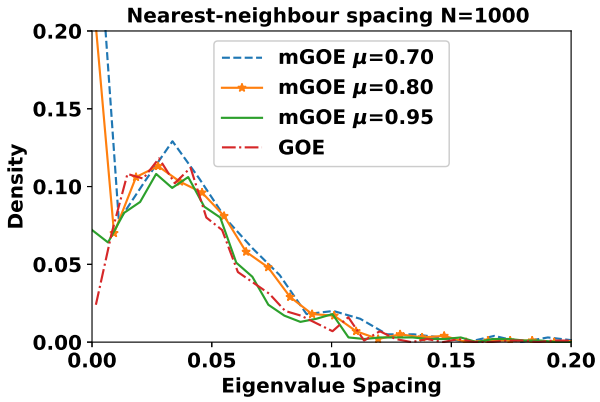

FIG. 2. Nearest-neighbour spacing for mixed GOE with different degree of mixing and non-mixing for N=1000.

how different orders in the ensemble distributed, provides a robust way of controlling mixture levels. In the spectral analysis example, the observed deviation from semicircle law in the mixed setting GOE (mGOE), appeared as an M-shape, is identified as a novel property of mGOE.

The author expresses gratitude for the tools provided by the scientific Python community [16, 17], Shanto Roy for LaTeX instruction for typesetting algorithm environment and Galin Jones for the suggestion of documenting the algorithm explicitly.

* suzen@acm.org

[1] M. L. Mehta, *Random Matrices* (Elsevier, 2004).
[2] A. Edelman and N. R. Rao, Acta Numerica **14**, 233 (2005).
[3] T. Tao, *Topics in Random Matrix Theory*, Vol. 132 (American Mathematical Society, 2012).
[4] J. Wishart, Biometrika **20**, 32 (1928).
[5] E. P. Wigner, Annals of Mathematics , 325 (1958).
[6] E. P. Wigner, SIAM Review **9**, 1 (1967).
[7] F. J. Dyson, Journal of Mathematical Physics **3**, 1199 (1962).
[8] F. J. Dyson, Communications in Mathematical Physics **19**, 235 (1970).
[9] R. J. Muirhead, *Aspects of multivariate statistical theory*, Vol. 197 (John Wiley & Sons, 2009).
[10] H.-J. Stöckmann, *Quantum Chaos: An Introduction* (Cambridge University Press, 2006).
[11] K. Rajan and L. F. Abbott, Physical Review Letters **97**, 188104 (2006).
[12] J. Pennington and P. Worah, Journal of Statistical Mechanics: Theory and Experiment **2019**, 124005 (2019).
[13] J. Y. Yam and T. W. Chow, Neurocomputing **30**, 219 (2000).
[14] J. Frankle and M. Carbin, in *International Conference on Learning Representations* (2018).
[15] M. Süzen, J. J. Cerdà, and C. Weber, arXiv preprint arXiv:1911.07831 (2019).
[16] C. R. Harris and et. al., Nature **585**, 357 (2020).
[17] J. D. Hunter, Computing in Science & Engineering **9**, 90 (2007).