Proof difficulties for spectral matching

Jiaxin Hu

Jun 2, 2022

Consider a super-symmetric random tensor $\mathcal{A} \in \mathbb{R}^{n^{\otimes m}}$ and the decomposition

$$\mathcal{A} = \mathcal{C} \times_1 \boldsymbol{U} \times_2 \cdots \times_m \boldsymbol{U}, \quad \boldsymbol{U} = \text{SVD}_n(\text{Mat}_1(\mathcal{A})), \tag{1}$$

where the columns in $\boldsymbol{U} \in \mathbb{R}^{n \times n}$ are the singular vectors of matricized $\mathcal{A}$, and $\mathcal{C} \in \mathbb{R}^{n^{\otimes m}}$ is the super-symmetric core tensor. Recall that we want to show the diagonal dominance of the similarity matrix

$$\boldsymbol{S}^* = \sum_{i,j \in [n]} \frac{1}{(\|\mathcal{C}_{i:}\|_F - \|\mathcal{C}_{j:}\|_F)^2 + \eta^2} \boldsymbol{U}_i \boldsymbol{U}_i^T \boldsymbol{J} \boldsymbol{U}_j \boldsymbol{U}_j^T,$$

where $\boldsymbol{J} \in \mathbb{R}^{n \times n}$ is the all-one matrix. We need to answer the following questions:

1. Under which setting (e.g. Gaussian Orthogonal Ensemble(GOE) for random matrix) we can obtain useful properties about the $\boldsymbol{U}$ and $\mathcal{C}$?
2. What is the distribution of $\boldsymbol{U}$?
3. What is the distribution of Frobenius norm of the slices of $\mathcal{C}$?

My current answers for these questions are following:

1. Basic random matrix theory (RMT) usually considers the setting that the distribution of random matrix is invariant to the orthogonal rotation. For example, if $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ follows GOE, the density of $\boldsymbol{A}$ is

   $$f(\boldsymbol{A}) = C \exp\left(-\frac{\|\boldsymbol{A}\|_F^2}{2}\right), \tag{2}$$

   where C is the normalization constant to make $f(\cdot)$ be a probability density, and the density $f(\boldsymbol{A})$ is invariant to the multiplication of orthonormal matrix to the matrix $\boldsymbol{A}$.

   Similarly, we can extend the GOE to the tensor case. Assume the tensor $\mathcal{A} \in \mathbb{R}^{n^{\otimes m}}$ follows tensor GOE. The density of $\mathcal{A}$ is

   $$f(\mathcal{A}) = C' \exp\left(-\frac{\|\mathcal{A}\|_F^2}{2}\right), \tag{3}$$

   where C' is the normalization constant to make $f(\cdot)$ be a probability density. Then, $f(\mathcal{A})$ is invariant to the multiplication of orthonormal matrix on any mode of the tensor $\mathcal{A}$.

2. Assume $\mathcal{A}$ follows the tensor GOE. The invariant law of $\mathcal{A}$ implies the independence between the factor matrix and the core tensor.

Lemma 1 (Factor matrix)**.** *Assume $\mathcal{A}$ follows the tensor GOE. The factor matrix $\boldsymbol{U}$ defined in (1) is independent to the core tensor $\mathcal{C}$, and $\boldsymbol{U}$ is a uniform random orthonormal matrix.*

Proof of Lemma 1. Follow the same proof of Corollary 2.5.4 in Anderson et al. (2010). □

3. I have not found a direct reference for this question.

Note that
$$\mathrm{Mat}_1(\mathcal{A}) = \boldsymbol{U}\mathrm{Mat}_1(\mathcal{C})(\boldsymbol{U}^{\otimes m-1})^T = \boldsymbol{U}\Lambda\boldsymbol{W}^T,$$
where $\boldsymbol{U}^{\otimes m-1}$ denotes the Kronecker product of $m-1$ matrices $\boldsymbol{U}$, the first equation follows by the tensor algebra with decomposition (1), the second equation follows by the definition of $\boldsymbol{U}$ and the matrix SVD with diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and orthogonal matrix $\boldsymbol{W} \in \mathbb{R}^{n^{m-1}\times n}$. Note that $\|\mathrm{Mat}_1(\mathcal{C})_{i:}\|_F = \|\mathcal{C}_{i:}\|_F$, and $\boldsymbol{U}^{\otimes m-1}$ is still orthogonal matrix. We have
$$\|\mathcal{C}_{i:}\|_F = \lambda_i,$$
and thus the distribution of $\|\mathcal{C}_{i:}\|_F$ follows the distribution of λ_i's.

One possible way to figure out the distribution of λ_i's is to apply the RMT to rectangle matrix $\mathrm{Mat}_1(\mathcal{A})$. For rectangle matrix and SVD, we usually consider the Wishart ensemble. The empirical distribution of the eigenvalues (empirical level density) of Wishart ensemble converges to law of Marcenko-Pastus; see reference (Edelman and Rao, 2005; Fan, 2010) (Both are not the original reference for Wishart ensemble, but their organizations are the clearest to me). However, due to the symmetry of $\mathcal{A}$, we can not apply the results of Wishart ensemble to $\mathrm{Mat}_1(\mathcal{A})$, because Wishart ensemble requires the entries in the rectangle matrix distribute independently.

Another possible way is to work on the random tensor directly. There are some literature discuss the spectral property of random tensors (Richard and Montanari, 2014; Tomioka and Suzuki, 2014; de Morais Goulart et al., 2021); however, these works define the spectral norm using the contraction; i.e., $\|\mathcal{A}\|_{op} = \max_{\boldsymbol{v}\in\mathbb{R}^n, \|\boldsymbol{v}\|=1} \langle \mathcal{A}, \boldsymbol{v}^{\otimes m}\rangle$. Then, the spectral norm $\|\mathcal{A}\|_{op}$ is more related to the coefficients in the CP decomposition rather than the norm of core tensor slices $\|\mathcal{C}_{i:}\|$. Also, we need the distribution of all "eigenvalues" and the spectral norm only considers the property of the largest "eigenvalue".

If we want to derive the random tensor theory under the tucker decomposition and tensor GOE, we may extend the RMT for matrix GOE. The RMT (Mehta, 2004) firstly obtains the joint distribution of eigenvalues in GOE by "integrating" out the eigenvectors (Edelman and Rao, 2005) from the joint distribution (2) and then derives the level density (is equal to the semicircle law under GOE) from the joint pdf of eigenvalues. Since the joint element distributions (2) and (3) are similar in the formulation, we may extend the technique in RMT to the tensor case. But I think I am not able to finish the extension in a short time.

References

Anderson, G. W., Guionnet, A., and Zeitouni, O. (2010). *An introduction to random matrices*. Number 118. Cambridge university press.

de Morais Goulart, J. H., Couillet, R., and Comon, P. (2021). A random matrix perspective on random tensors. *stat*, 1050:2.

Edelman, A. and Rao, N. R. (2005). Random matrix theory. *Acta numerica*, 14:233–297.

Fan, Z. (2010). *Global and Local Limit Laws for Eigenvalues of the Gaussian Unitary Ensemble and the Wishart Ensemble*. PhD thesis, Harvard University.

Mehta, M. L. (2004). *Random matrices*. Elsevier.

Richard, E. and Montanari, A. (2014). A statistical model for tensor pca. *Advances in Neural Information Processing Systems*, 27.

Tomioka, R. and Suzuki, T. (2014). Spectral norm of random tensors. *arXiv preprint arXiv:1407.1870*.