Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 18, No. 2, pp. 395–414, 2022

VON NEUMANN-TYPE INEQUALITY FOR COMPLETELY ORTHOGONALLY DECOMPOSABLE TENSORS*

Jie Wang

Abstract: This paper presents some generalizations of Von Neumann's trace inequality for matrices to the contents of completely orthogonally decomposable tensors. The angle between two completely orthogonally decomposable (symmetrical) tensors is defined and taken into account in the Von Neumann-type inequality. Moreover, the properties of spectral functions in the case of completely orthogonally decomposable asymmetrical and symmetrical tensors are studied, respectively.

Key words: *completely orthogonally decomposable tensor, Von Neumann's trace inequality*

Mathematics Subject Classification: *15A18, 15A69, 65F18*

---

1 Introduction

Von Neumann's trace inequality, a famous trace inequality of matrices related to singular values, was proposed by Von Neumann[16] in 1937. This inequality is of great significance not only in mathematical theory but also in practical applications, such as signal processing, communication engineering, systems engineering and so on. It plays a key role in many matrix approximation problems.

As we all know, matrix singular value decomposition (SVD) is a practical tool in matrix analysis, engineering, statistics, and many other fields. With the SVD method, an original matrix can yield a low-rank matrix that is closer to it. With the coming age of big data, tensor, as a tool for describing high-dimensional data, has attracted more and more attention and research[1, 4, 12]. Among these, tensor decomposition and approximation problems are some of the focuses[7, 9]. Therefore, it would be meaningful if these properties and decompositions could be directly generalized to higher-order tensors. Tensor decomposition and approximation problem is essentially a higher-order generalization of the matrix.

The tensor decompositions derived from different matrix decomposition forms are different. At present, the most commonly used extension from matrix SVD to higher-order tensors is the so-called Higher-Order Singular Value Decomposition (HOSVD)[15]. Chrétienand and Wei [3] extended Von Neumann's trace inequality to tensors using SVD based on Tucker decomposition of tensors.

The famous Von Neumann's trace inequality explores the relationship between matrices and their singular values. Inspired by it, we intend to study the intrinsic relationship between

*This work was supported by the National Natural Science Foundation of China (Grant No. 11901549 and 12171128) and the Natural Science Foundation of Zhejiang Province, China (Grant No. LY22A010022).

tensors and their SVDs. One of the difficulties is the appropriate description of orthogonal decompositions of tensors. Actually, Kolda[10] presented several orthogonal decompositions of tensors, where only the completely orthogonal decomposition of the tensor is a parallel generalization of the matrix SVD.

This paper is structured as follows. In section 2, we will review some definitions and notations for matrices and tensors. In section 3, we will propose Von Neumann-type inequality based on completely orthogonally decomposable tensors and Fan-type inequality based on completely orthogonally decomposable symmetrical tensors (SCODT), respectively. Subsequently, we will discuss the spectral function for CODT in section 4, and the spectral function for SCODT in section 5. Some conclusions will be presented in the last section to conclude this paper.

2 Preliminary

In this section, we will briefly review some concepts and notations for matrices and tensors. Given two (real) matrices $A = (a_{ij})$ and $B = (b_{ij})$ of the same size $m \times n$, their *Hadamard product* $A * B$ is a matrix with size $m \times n$, which elements are given by $(A * B)_{ij} = a_{ij}b_{ij}$. A matrix $C = (c_{ij}) \in \mathbb{R}^{n\times n}$ is said to be *absolutely doubly stochastic*, if its sum of absolute value of the elements in each row and column equals to 1, i.e., $|C|e = e$ and $e^{\mathsf{T}}|C| = e^{\mathsf{T}}$ where $|C| = (|c_{ij}|)$ and $e = \begin{bmatrix}1, \cdots, 1\end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^n$; a matrix $C \in \mathbb{R}^{n\times n}$ is called *absolutely doubly substochastic*, if the sum of absolute value of the elements in each row and column of C is at most 1. For every $\mathbf{x} \in \mathbb{R}^n$, we have the l_p-norm (for $p \geq 1$) $\|\mathbf{x}\|_p$, defined as $\|\mathbf{x}\|_p = \sqrt[p]{(\sum\limits_{i=1}^{n} |\mathbf{x}_i|^p)}$.

A (real) m-th order $I_1 \times \cdots \times I_m$-dimensional tensor (a.k.a. *hypermatrix* [13]) $\mathcal{A}$, whose element is specified as $\mathcal{A} = (a_{i_1\cdots i_m})$ where $i_j \in \{1, \cdots, I_j\}$ and $j \in \{1, \cdots, m\}$. The set of all tensors with size $I_1 \times \cdots \times I_m$ is denoted by $T(I_1, \cdots, I_m)$. Specially, when $I_1 = \cdots = I_m = n$, the set of all the m-th order n-dimensional tensors is denoted by $T_{m,n}$. Given a tensor $\mathcal{A} = (a_{i_1\cdots i_m}) \in T_{m,n}$, if the entries $a_{i_1\cdots i_m}$ are invariant under any permutation of their indices, then $\mathcal{A}$ is called a *symmetric tensor*. The set of all the m-th order n-dimensional symmetric tensors is denoted by $S_{m,n}$. For a tensor $\mathcal{A} = (a_{i_1\cdots i_m}) \in T_{m,n}$, $a_{i\cdots i}$ for $i \in \{1, \cdots, n\}$ are called diagonal entries of $\mathcal{A}$, and the other entries of $\mathcal{A}$ are called off-diagonal entries of $\mathcal{A}$. A tensor $\mathcal{A}$ is called *diagonal* if all of its off-diagonal entries are zero. For a tensor $\mathcal{A} \in T(m, n)$, we denote λ as the eigenvalue map $\lambda : \mathcal{A} \to \mathbb{R}^r$, where r is the rank of $\mathcal{A}$. Given a tensor $\mathcal{A} = (a_{i_1\cdots i_m}) \in T_{m,n}$ and a vector $\mathbf{x} \in \mathbb{R}^n$, then $\mathcal{A}\mathbf{x}^{m-1}$ is a vector in $\mathbb{R}^n$ with its ith component as

$$(\mathcal{A}\mathbf{x}^{m-1})_i := \sum_{i_2,\cdots,i_m=1}^{n} a_{ii_2\cdots i_m} x_{i_2} \cdots x_{i_m}$$

for $i \in \{1, \cdots, n\}$. $\mathcal{A}\mathbf{x}^m$ is a homogeneous polynomial of degree m, defined as

$$\mathcal{A}\mathbf{x}^m := \mathbf{x}^{\mathsf{T}}(\mathcal{A}\mathbf{x}^{m-1}) = \sum_{i_1,\cdots,i_m=1}^{n} a_{i_1\cdots i_m} x_{i_1} \cdots x_{i_m}.$$

Given tensors $\mathcal{A} = (a_{i_1\cdots i_m})$ and $\mathcal{B} = (b_{i_1\cdots i_m}) \in T(I_1, \cdots, I_m)$, their *inner product* is defined as

$$\langle \mathcal{A}, \mathcal{B}\rangle = \sum_{i_1=1}^{I_1} \cdots \sum_{i_m=1}^{I_m} a_{i_1\cdots i_m} b_{i_1\cdots i_m}.$$

$\boxed{2.1}$ Decomposable tensors

A tensor $\mathcal{U} \in T(I_1, \cdots, I_m)$ is called a *decomposable tensor*[10], if $\mathcal{U}$ can be written as

$$\mathcal{U} = \mathbf{u}^{(1)} \otimes \mathbf{u}^{(2)} \otimes \cdots \otimes \mathbf{u}^{(m)},$$

where " $\otimes$ " denotes outer product and each $\mathbf{u}^{(j)} \in \mathbb{R}^{I_j}$ for $j = 1, \cdots, m$. Specially, when $\mathbf{u}^{(1)} = \cdots = \mathbf{u}^{(m)} = \mathbf{u}$, $\mathcal{U}$ can be abbreviated as $\mathcal{U} = \mathbf{u}^{\otimes m}$. Each component of $\mathcal{U}$ is

$$\mathcal{U}_{i_1 \ldots i_m} = \mathbf{u}_{i_1}^{(1)} \mathbf{u}_{i_2}^{(2)} \cdots \mathbf{u}_{i_m}^{(m)}.$$

A decomposable tensor is a tensor of rank-1. A tensor is *rank-1* if it can be expressed as the outer product of a series of vectors, such as $\mathcal{U} = \mathbf{u}^{(1)} \otimes \cdots \otimes \mathbf{u}^{(m)}$ for $\mathbf{u}^{(j)} \in \mathbb{R}^{I_j}$. A tensor $\mathcal{A}$ is *rank-r* if r is the smallest integer such that $\mathcal{A}$ is the weighted sum of rank-1 tensors, i.e.,

$$\mathcal{A} = \sum_{i=1}^{r} \gamma_i \mathbf{u}_i^{(1)} \otimes \cdots \otimes \mathbf{u}_i^{(m)}.$$

A decomposable tensor has a higher-order singular value decomposition(HOSVD). The HOSVD is a kind of multidimensional generalization of the matrix singular value decomposition, which defined by De Lathauwer et al[15].

The symbol $St(m, n)$ denotes a *Stiefel manifold*, which is the set of all m-tuples orthonormal vectors in $\mathbb{R}^n$. Given column vectors $\mathbf{v}_1, \cdots, \mathbf{v}_m \in \mathbb{R}^n$, a matrix $[\mathbf{v}_1, \cdots, \mathbf{v}_m] \in St(m, n)$, if $\|\mathbf{v}_i\| = 1$ and $\mathbf{v}_i \perp \mathbf{v}_j$ for all $i \neq j \in \{1, \ldots, m\}$.

$\boxed{2.2}$ Completely orthogonally decomposable tensors

Given a real matrix $A \in \mathbb{R}^{m \times n}$, it can be expressed as the following formality,

$$A = U \Sigma V^{\mathsf{T}} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathsf{T}},$$

which is the SVD of A, $U = [\mathbf{u}_1, \cdots, \mathbf{u}_r] \in St(r, m)$, $V = [\mathbf{v}_1, \cdots, \mathbf{v}_r] \in St(r, n)$, and $\Sigma = diag(\sigma_1, \cdots, \sigma_r)$ is a diagonal matrix in $\mathbb{R}^{r \times r}$, where r is the rank of A. Therefore, a matrix can be decomposed into a sum of rank one matrices. Analogously, defined in the same way, a tensor that can be expressed as the sum of rank one tensors is called completely orthogonally decomposable tensor, refer to [10].

Definition 2.1 (CODT)**.** Let $\mathcal{A} \in T(I_1, \cdots, I_m)$ be a tensor. $\mathcal{A}$ is called a completely orthogonally decomposable tensor, if $\mathcal{A}$ can be represented as

$$\mathcal{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i^{(1)} \otimes \cdots \otimes \mathbf{u}_i^{(m)}, \tag{2.1}$$

where $[\mathbf{u}_1^{(j)}, \cdots, \mathbf{u}_r^{(j)}] \in St(r, I_j)$ for $j \in \{1, \cdots, m\}$ and $\sigma_i > 0$ for $i \in \{1, \cdots, r\}$, the corresponding r is the completely orthogonal rank of $\mathcal{A}$. In this case, σ_i are called a singular value of $\mathcal{A}$, $\mathbf{u}_i^{(1)}, \cdots, \mathbf{u}_i^{(m)}$ are called singular vectors corresponding to σ_i for all $i \in \{1, \cdots, r\}$, and the decomposition (2.1) is called the higher-order singular value decomposition of $\mathcal{A}$.

Definition 2.2 (SCODT)**.** Let $\mathcal{A} \in S_{m,n}$ be a tensor. $\mathcal{A}$ is called a completely orthogonally decomposable symmetrical tensor, if $\mathcal{A}$ can be represented as

$$\mathcal{A} = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i \otimes \cdots \otimes \mathbf{u}_i, \tag{2.2}$$

or abbreviated as $\sum\limits_{i=1}^{n} \lambda_i \mathbf{u}_i^{\otimes m}$. It also can be written as

$$\mathcal{A} = \Lambda \cdot (U, \cdots, U) \tag{2.3}$$

where $U = [\mathbf{u}_1, \cdots, \mathbf{u}_r] \in St(r, n)$, $\Lambda \in S_{m,r}$ is a diagonal tensor and its diagonal elements are order singular values of tensor $\mathcal{A}$, and $\lambda_i > 0$ with $i \in \{1, \cdots, r\}$, the corresponding r is the completely orthogonal symmetrical rank of $\mathcal{A}$.

Usually, the diagonal elements of tensor Λ are arranged in order by size, called *ordered singular value*. Let $\|\cdot\|$ denote a unitarily invariant norm, then $\|\mathcal{A}\| = \|\Lambda \cdot (U_1, \cdots, U_m)\| = \|\Lambda\|$ for the above tensor $\mathcal{A}$.

Throughout the rest of this paper, the set of all the completely orthogonally decomposable (symmetrical) tensors is denoted by CODT (SCODT) . If a tensor $\mathcal{A}$ is completely orthogonally decomposable (symmetrical), then it is denoted by $\mathcal{A} \in$ CODT (SCODT) .

3 Von Neumann-type Inequality for CODT

In this section, we intend to study the relationship between tensors and their SVDs based on the Von Neumann's trace inequality. At the beginning of this section, some lemmas are given for the requirement. After that, the angles between two CODTs are defined and the Von Neumann-type inequality for CODT is established. Meanwhile, through the similar analysis, the Fan's inequality for matrices is extended to the content of SCODT. For Lemma 3.1 and Lemma 3.2, see [6].

Lemma 3.1 (Von Neumann's trace inequality)**.** *Let the non-increasingly ordered singular values of* $A, B \in \mathbb{R}^{m\times n}$ *be* $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$ *and* $\sigma_1(B) \geq \cdots \geq \sigma_r(B)$*, where* $r = \min\{m, n\}$*. Then, the following inequality holds*

$$tr(AB^{\mathsf{T}}) \leq \sum_{i=1}^{r} \sigma_i(A)\sigma_i(B).$$

Lemma 3.2 (Fan's inequality)**.** *Any real symmetric matrices* X *and* Y *satisfy the inequality*

$$tr(XY) \leq \lambda(X)^{\mathsf{T}}\lambda(Y),$$

equality holds if and only if X *and* Y *have a simultaneous ordered spectral decomposition.*

Lemma 3.3. *Let* $U_1 = (u_{ij}^1)$, $\cdots$, $U_m = (u_{ij}^m) \in \mathbb{R}^{n\times n}$ *be orthogonal matrices, and denote* $C = U_1 * \cdots * U_m$*, then* C *is absolutely doubly substochastic.*

Proof. First, we can prove that the Hadamard product of two orthogonal matrices is doubly substochastic. Suppose that U_1 and U_2 be orthogonal matrices, it is easy to see that for each $i \in \{1, \cdots, n\}$,

$$\sum_{j=1}^{n} |[U_1 * U_2]_{ij}| \leq \sqrt{(u_{i1}^1)^2 + \cdots + (u_{im}^1)^2}\sqrt{(u_{i1}^2)^2 + \cdots + (u_{im}^2)^2} \leq 1,$$

and for each $j \in \{1, \cdots, n\}$ we similarly obtain $\sum_{i=1}^{n} |(U_1 * U_2)_{ij}| \leq 1$. Then $U_1 * U_2$ is absolutely doubly substochastic.

Furthermore, we can prove that the Hadamard product of an orthogonal matrix and an absolutely doubly substochastic matrix is absolutely doubly substochastic. Suppose that U_α is orthogonal and U_β is absolutely doubly substochastic, then it is easy to verify that for each $i \in \{1, \cdots, n\}$

$$\sum_{j=1}^{n} |(U_\alpha * U_\beta)_{ij}| \leq \sum_{j=1}^{n} |(U_\beta)_{ij}| \leq 1,$$

and for each $j \in \{1, \cdots, n\}$ we can similarly obtain $\sum_{i=1}^{n} |(U_\alpha * U_\beta)_{ij}| \leq 1$.

According to the above deduction, it is naturally deduced that C is absolutely doubly substochastic. □

Lemma 3.4. *Let $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ be absolutely doubly substochastic, then there exists an absolutely doubly stochastic matrix $W = (w_{ij}) \in \mathbb{R}^{n\times n}$ such that $|A| \leq |W|$, i.e., $|a_{ij}| \leq |w_{ij}|$ for each i, j.*

Lemma 3.5. *Let P be a polyhedral set in $\mathbb{R}^n$. Assume that P has at least one vertex. Then if a linear function $f : \mathbb{R}^n \to \mathbb{R}^n$ attains a maximum over P, it attains a maximum at a vertex of P.*

Lemma 3.6. *Let $a_{1i}, a_{2i}, \cdots, a_{mi} \in \mathbb{R}$ for $i = 1, 2, \cdots, n$, then*

$$\sum_{i=1}^{n} |a_{1i}a_{2i}\cdots a_{mi}| \leq \sqrt[m]{\sum_{i=1}^{n} |a_{1i}|^m} \cdot \sqrt[m]{\sum_{i=1}^{n} |a_{2i}|^m} \cdots \sqrt[m]{\sum_{i=1}^{n} |a_{mi}|^m}.$$

The equality holds if and only if $\frac{|a_{1i}|}{\sqrt[m]{\sum_{i=1}^{n} |a_{1i}|^m}} = \frac{|a_{2i}|}{\sqrt[m]{\sum_{i=1}^{n} |a_{2i}|^m}} = \cdots = \frac{|a_{mi}|}{\sqrt[m]{\sum_{i=1}^{n} |a_{mi}|^m}}$ *where* $a_{1i}, \cdots, a_{mi} \neq 0$, *and* $i = 1, 2, \cdots, n$.

Proof. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ be two vectors, $p > 1$ and $q > 1$ be numbers such that $\frac{1}{p} + \frac{1}{q} = 1$. The well-known Hölder inequality [5] reads as

$$\sum_{i=1}^{n} |x_i y_i| \leq \sqrt[p]{\sum_{i=1}^{n} |x_i|^p} \sqrt[q]{\sum_{i=1}^{n} |y_i|^q},$$

the equality holds if and only if the vectors $[|x_i|^p]$ and $[|y_i|^q]$ are linearly dependent. Let $\mathbf{y}$ be with $y_i = a_{2i} \ldots a_{mi}$ and $\mathbf{x}$ with $x_i = a_{1i}$ for all $i \in \{1, \ldots, n\}$, and $p := m$ and $q := \frac{m}{m-1}$.

Then, from the Hölder inequality, it can be deduced that

$$\begin{aligned}\sum_{i=1}^{n}|a_{1i}a_{2i}\cdots a_{mi}| &= \sum_{i=1}^{n}|x_iy_i| \\ &\le \|\mathbf{x}\|_p\|\mathbf{y}\|_q \\ &= \sqrt[m]{\sum_{i=1}^{n}|x_i|^m}\ \sqrt[\frac{m}{m-1}]{\sum_{i=1}^{n}|y_i|^{\frac{m}{m-1}}} \\ &= \sqrt[m]{\sum_{i=1}^{n}|a_{1i}|^m}\ \sqrt[\frac{m}{m-1}]{\sum_{i=1}^{n}|a_{2i}\dots a_{mi}|^{\frac{m}{m-1}}}. \end{aligned} \tag{3.1}$$

The following derivation is in the same way as deduced above,

$$\begin{aligned}\sum_{i=1}^{n}|a_{2i}\dots a_{mi}|^{\frac{m}{m-1}} &= \sum_{i=1}^{n}|a_{2i}|^{\frac{m}{m-1}}\dots|a_{mi}|^{\frac{m}{m-1}} \\ &\le \sqrt[m-1]{\sum_{i=1}^{n}(|a_{2i}|^{\frac{m}{m-1}})^{m-1}}\ \sqrt[\frac{m-1}{m-2}]{\sum_{i=1}^{n}\left(|a_{3i}|^{\frac{m}{m-1}}\dots|a_{mi}|^{\frac{m}{m-1}}\right)^{\frac{m-1}{m-2}}} \\ &= \sqrt[m-1]{\sum_{i=1}^{n}|a_{2i}|^m}\ \sqrt[\frac{m-1}{m-2}]{\sum_{i=1}^{n}|a_{3i}\dots a_{mi}|^{\frac{m}{m-2}}}. \end{aligned} \tag{3.2}$$

Thus, together with (3.1), it can be deduced that

$$\sum_{i=1}^{n}|a_{1i}a_{2i}\cdots a_{mi}| \le \sqrt[m]{\sum_{i=1}^{n}|a_{1i}|^m}\ \sqrt[m]{\sum_{i=1}^{n}|a_{2i}|^m}\ \sqrt[\frac{m}{m-2}]{\sum_{i=1}^{n}|a_{3i}\dots a_{mi}|^{\frac{m}{m-2}}}.$$

Continuously applying inductive steps to $\sum_{i=1}^{n}|a_{3i}\dots a_{mi}|^{\frac{m}{m-2}}$, we can get the final inequality.

Let's prove the equality. By the Hölder inequality, the equality holds if and only if $[|x_i|^p]$ and $[|y_i|^q]$ are linearly dependent. Thus the equality holds in (3.1) if and only if

$$|a_{1i}|^m = k_1|a_{2i}\cdots a_{mi}|^{\frac{m}{m-1}},$$

which implies $|a_{1i}| = k_1|a_{2i}\cdots a_{mi}|^{\frac{1}{m-1}}$. Similarly, the equality holds in (3.2) if and only if

$$|a_{2i}| = k_2|a_{3i}\cdots a_{mi}|^{\frac{1}{m-2}}.$$

Inductively,

$$|a_{(m-1)i}| = k_{m-1}|a_{mi}|.$$

For this formula, by reverse derivation, we deduce that $|a_{1i}|,\cdots,|a_{mi}|$ for $i=1,\cdots,n$ are mutually linear dependent. This completes the proof. □

3.1 The case for matrices

In order to explore the angles between two tensors, we firstly consider the content between two matrices. Given two s-dimensional subspaces G and H in $\mathbb{R}^n$. Given two orthonormal matrices $O_G, O_H \in \mathbb{R}^{n\times s}$, which the columns be orthogonal bases in subspaces G and H respectively. Typically, the *angles* between subspaces G and H are defined as the ranges of given matrices O_G and O_H. If $Y^\mathsf{T}(O_G^\mathsf{T} O_H)Z = diag(\varrho_1, \cdots, \varrho_s)$ is the SVD of $O_G^\mathsf{T} O_H$, then angles denoted as ϑ_k between O_G and O_H defined in the following

$$\begin{aligned} O_G Y &= [\mathbf{g}_1, \cdots, \mathbf{g}_s], \\ O_H Z &= [\mathbf{h}_1, \cdots, \mathbf{h}_s], \\ \cos\vartheta_i &= \varrho_i \text{ for } i = 1, \cdots, s. \end{aligned}$$

Given two matrices A, $B \in \mathbb{R}^{n\times r}$, then the singular value decomposition of A and B are $A = U\Sigma_A V^\mathsf{T}$ and $B = P\Sigma_B Q^\mathsf{T}$, respectively. Here $U = [\mathbf{u}_1, \cdots, \mathbf{u}_r]$, $V = [\mathbf{v}_1, \cdots, \mathbf{v}_r]$, $P = [\mathbf{p}_1, \cdots, \mathbf{p}_r]$, $Q = [\mathbf{q}_1, \cdots, \mathbf{q}_r]$, $\Sigma_A = diag(\sigma_1(A), \cdots, \sigma_r(A))$, and $\Sigma_B = diag(\sigma_1(B), \cdots, \sigma_r(B))$. Define

$$\cos\theta_i = \|P^\mathsf{T}\mathbf{u}_i\| \text{ for } i = 1, \cdots, r,$$

$$\cos\eta_i = \|Q^\mathsf{T}\mathbf{v}_i\| \text{ for } i = 1, \cdots, r,$$

$$\cos\phi_i = \|U^\mathsf{T}\mathbf{p}_i\| \text{ for } i = 1, \cdots, r,$$

$$\cos\varphi_i = \|V^\mathsf{T}\mathbf{q}_i\| \text{ for } i = 1, \cdots, r.$$

Denote

$$\begin{aligned} \cos\hat{\theta} &= \max\big\{\cos\theta_i, i = 1, \cdots, r\big\}, \\ \cos\hat{\eta} &= \max\big\{\cos\eta_i, i = 1, \cdots, r\big\}, \\ \cos\hat{\phi} &= \max\big\{\cos\phi_i, i = 1, \cdots, r\big\}, \\ \cos\hat{\varphi} &= \max\big\{\cos\varphi_i, i = 1, \cdots, r\big\}. \end{aligned} \tag{3.3}$$

As discussed above, considering the angles between two matrices, we obtain the following singular value inequality related to the angle between two matrices.

Proposition 3.7. *Let $A, B \in \mathbb{R}^{n\times p}$, $r = \min\{n, p\}$, and its ordered singular values be $\sigma_1(A) \geq \cdots \geq \sigma_r(A)$ and $\sigma_1(B) \geq \cdots \geq \sigma_r(B)$ respectively. Then*

$$\langle A, B\rangle \leq \max\{\cos\hat{\theta}\cdot\cos\hat{\eta}, \cos\hat{\phi}\cdot\cos\hat{\varphi}\} \sum_{i=1}^{r} \sigma_i(A)\sigma_i(B),$$

where $\hat{\theta}, \hat{\eta}, \hat{\phi}, \hat{\varphi}$ are defined as (3.3)*.*

Proof. Compute

$$\langle A, B\rangle = \sum_{i,j=1}^{r} \sigma_i(A)\sigma_j(B)\langle\mathbf{u}_i, \mathbf{p}_j\rangle\langle\mathbf{v}_i, \mathbf{q}_j\rangle = \sum_{i,j=1}^{r} \sigma_i(A)\sigma_j(B)M_{ij}, \tag{3.4}$$

where $M = (U^\mathsf{T} P) * (V^\mathsf{T} Q) = [M_{ij}]$. The row sum of M for each $i \in \{1, \cdots, r\}$ is as follows,

$$\begin{aligned}
\sum_{j=1}^{r} |M_{ij}| &= \sum_{j=1}^{r} |\langle \mathbf{u}_i, \mathbf{p}_j \rangle||\langle \mathbf{v}_i, \mathbf{q}_j \rangle| \\
&\leq \sqrt{\sum_{j=1}^{r} \langle \mathbf{u}_i, \mathbf{p}_j \rangle^2} \sqrt{\sum_{j=1}^{r} \langle \mathbf{v}_i, \mathbf{q}_j \rangle^2} \\
&= \sqrt{\sum_{j=1}^{r} (P^\mathsf{T} \mathbf{u}_i)_j^2} \sqrt{\sum_{j=1}^{r} (Q^\mathsf{T} \mathbf{v}_i)_j^2} \\
&= \|P^\mathsf{T} \mathbf{u}_i\| \|Q^\mathsf{T} \mathbf{v}_i\| \\
&= \cos\theta_i \cdot \cos\eta_i \\
&\leq \cos\hat{\theta} \cdot \cos\hat{\eta},
\end{aligned}$$

and its column sum for each $j \in \{1, \cdots, r\}$ is as follows,

$$\begin{aligned}
\sum_{i=1}^{r} |M_{ij}| &= \sum_{i=1}^{r} |\langle \mathbf{u}_i, \mathbf{p}_j \rangle||\langle \mathbf{v}_i, \mathbf{q}_j \rangle| \\
&\leq \sqrt{\sum_{i=1}^{r} \langle \mathbf{u}_i, \mathbf{p}_j \rangle^2} \sqrt{\sum_{i=1}^{r} \langle \mathbf{v}_i, \mathbf{q}_j \rangle^2} \\
&= \sqrt{\sum_{i=1}^{r} (U^\mathsf{T} \mathbf{p}_j)_i^2} \sqrt{\sum_{i=1}^{r} (V^\mathsf{T} \mathbf{q}_j)_i^2} \\
&= \|U^\mathsf{T} \mathbf{p}_j\| \|V^\mathsf{T} \mathbf{q}_j\| \\
&= \cos\phi_j \cdot \cos\varphi_j \\
&\leq \cos\hat{\phi} \cdot \cos\hat{\varphi},
\end{aligned}$$

Denote $\hat{M} = \min\big\{\frac{1}{\cos\hat{\theta} \cdot \cos\hat{\eta}}, \frac{1}{\cos\hat{\phi} \cdot \cos\hat{\varphi}}\big\} M$. Then, $\hat{M}$ is both row and column absolutely doubly substochastic. Therefore, $\hat{M}$ is absolutely doubly substochastic. Combining (3.4), it can be obtained that,

$$\begin{aligned}
\langle A, B \rangle &= \sum_{i,j=1}^{r} \sigma_i(A) \sigma_j(B) M_{ij} \\
&= \max\big\{\cos\hat{\theta} \cdot \cos\hat{\eta}, \cos\hat{\phi} \cdot \cos\hat{\varphi}\big\} \sum_{i,j=1}^{r} \sigma_i(A) \sigma_j(B) \hat{M}_{ij}.
\end{aligned}$$

Since $\hat{M}$ is absolutely doubly substochastic, by Lemma 3.4 there exists an absolutely doubly stochastic matrix $C = [C_{ij}] \in \mathbb{R}^{r \times r}$ such that

$$\langle A, B \rangle \leq \max\big\{\cos\hat{\theta} \cdot \cos\hat{\eta}, \cos\hat{\phi} \cdot \cos\hat{\varphi}\big\} \sum_{i,j=1}^{r} \sigma_i(A) \sigma_j(B) C_{ij},$$

Define a linear function $f(C) := \sum_{i,j=1}^{r} \sigma_i \gamma_j C_{ij}$ on the set of absolutely doubly substochastic matrices, by Lemma 3.5, it attains its maximum at a vertex. If π is a permutation such that

$C_{ij} = 1$ if and only if $j = \pi(i)$, then

$$\begin{aligned}
\langle A, B\rangle &\leq \max\big\{\cos\hat{\theta}\cdot\cos\hat{\eta}, \cos\hat{\phi}\cdot\cos\hat{\varphi}\big\}\sum_{i,j=1}^{r}\sigma_i(A)\sigma_j(B)C_{ij} \\
&= \max\big\{\cos\hat{\theta}\cdot\cos\hat{\eta}, \cos\hat{\phi}\cdot\cos\hat{\varphi}\big\}\sum_{i=1}^{r}\sigma_i(A)\sigma_{\pi(i)}(B) \qquad (3.5)\\
&\leq \max\big\{\cos\hat{\theta}\cdot\cos\hat{\eta}, \cos\hat{\phi}\cdot\cos\hat{\varphi}\big\}\sum_{i=1}^{r}\sigma_i(A)\sigma_i(B).
\end{aligned}$$

□

$\boxed{3.2}$ The case for CODT

Analogously, we could have the above ideas to define the angle between two CODTs. Given two completely orthogonally decomposable tensors $\mathcal{A}$ and $\mathcal{B}$ with size $I_1\times\cdots\times I_m$, then

$$\mathcal{A} = \Lambda_{\mathcal{A}}\cdot(U_1,\cdots,U_m),$$

and

$$\mathcal{B} = \Lambda_{\mathcal{B}}\cdot(V_1,\cdots,V_m),$$

where $U_k = [\mathbf{u}_1^{(k)},\cdots,\mathbf{u}_r^{(k)}]$ and $V_k = [\mathbf{v}_1^{(k)},\cdots,\mathbf{v}_r^{(k)}]$ for $k = 1,\cdots,m$. We define the angles of $\mathcal{A}$ and $\mathcal{B}$ as follows,

$$\begin{aligned}
\cos\theta_j^{(k)} &= \|V_j^{\mathsf{T}}\mathbf{u}_j^{(k)}\|, \text{ for } k = 1,\cdots,m \text{ and } j = 1,\cdots,r,\\
\cos\delta_j^{(k)} &= \|U_j^{\mathsf{T}}\mathbf{v}_j^{(k)}\|, \text{ for } k = 1,\cdots,m \text{ and } j = 1,\cdots,r,
\end{aligned} \qquad (3.6)$$

and

$$\begin{aligned}
\cos\theta &= \max\{\cos\theta_j^{(k)} \mid k = 1,\cdots,m \text{ and } j = 1,\cdots,r\},\\
\cos\delta &= \max\{\cos\delta_j^{(k)} \mid k = 1,\cdots,m \text{ and } j = 1,\cdots,r\},\\
\cos\eta &= \max\{\cos\theta_j^{(k)}, \cos\delta_j^{(k)} \mid k = 1,\cdots,m \text{ and } j = 1,\cdots,r\}.
\end{aligned} \qquad (3.7)$$

Considering the angle of two CODTs, we discuss the relationship between two CODTs and their SVDs based on the classic Von Neumann's inequality and obtain the following Von Neumann-type inequality.

Theorem 3.8 (Von Neumann-type inequality)**.** *Let $\mathcal{A},\mathcal{B}\in T(I_1,\cdots,I_m)$ be completely orthogonally decomposable tensors, $r = \min\{I_1,\cdots,I_m\}$ and the ordered singular values of $\mathcal{A}$ and $\mathcal{B}$ be $\sigma_1(\mathcal{A})\geq\cdots\geq\sigma_r(\mathcal{A})$ and $\sigma_1(\mathcal{B})\geq\cdots\geq\sigma_r(\mathcal{B})$, respectively. Then the following inequality holds*

$$\langle\mathcal{A},\mathcal{B}\rangle \leq \max\{\cos^m\theta, \cos^m\delta\}\cdot\sum_{i=1}^{r}\sigma_i(\mathcal{A})\sigma_i(\mathcal{B}),$$

where θ and δ are defined as (3.7). Moreover, the above equality holds if $\theta_j^{(k)} = 0$ and $\delta_j^{(k)} = 0$ for all $k\in\{1,\cdots,m\}$, $j\in\{1,\cdots,r\}$.

Proof. Since $\mathcal{A}$ and $\mathcal{B}$ are completely orthogonally decomposable tensors, they can be expressed by the definition 2.1 as follows

$$\mathcal{A} = \sum_{i=1}^{r} \sigma_i(\mathcal{A}) \mathbf{u}_i^{(1)} \otimes \cdots \otimes \mathbf{u}_i^{(m)},$$

and

$$\mathcal{B} = \sum_{i=1}^{r} \sigma_i(\mathcal{B}) \mathbf{v}_i^{(1)} \otimes \cdots \otimes \mathbf{v}_i^{(m)}.$$

Denote $U_k = [\mathbf{u}_1^{(k)}, \cdots, \mathbf{u}_r^{(k)}]$ and $V_k = [\mathbf{v}_1^{(k)}, \cdots, \mathbf{v}_r^{(k)}]$, then $U_k, V_k \in St(r, I_k)$ for $k = 1, \cdots, m$. Then

$$\begin{aligned}
\langle \mathcal{A}, \mathcal{B} \rangle &= \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) \langle \mathbf{u}_i^{(1)} \otimes \cdots \otimes \mathbf{u}_i^{(m)}, \mathbf{v}_i^{(1)} \otimes \cdots \otimes \mathbf{v}_i^{(m)} \rangle \\
&= \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) \langle \mathbf{u}_i^{(1)}, \mathbf{v}_j^{(1)} \rangle \cdots \langle \mathbf{u}_i^{(m)}, \mathbf{v}_j^{(m)} \rangle \\
&= \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) ({U_1}^{\mathsf{T}} V_1)_{ij} \cdots ({U_m}^{\mathsf{T}} V_m)_{ij} \\
&= \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) W_{ij}.
\end{aligned}$$

Denote $W = [W_{ij}] = [({U_1}^{\mathsf{T}} V_1)_{ij} \cdots ({U_m}^{\mathsf{T}} V_m)_{ij}]$ and $W_k = {U_k}^{\mathsf{T}} V_k$ for all $k \in \{1, \cdots, m\}$, then W is an absolutely doubly substochastic matrix. Actually, matrices U_k and V_k can be extended to orthogonal matrix $\tilde{U}$ and $\tilde{V}$ by adding some column $U_k^{'}$ and $V_k^{'}$, which can be denoted as $\tilde{U}_k = [U_k \ \ U_k^{'}]$ and $\tilde{V}_k = [V_k \ \ V_k^{'}]$ respectively. Then,

$$\tilde{U}_k^{\mathsf{T}} \tilde{V}_k = \begin{bmatrix} U_k^{\mathsf{T}} V_k & U_k^{\mathsf{T}} V_k^{'} \\ {U_k^{'}}^{\mathsf{T}} V_k & {U_k^{'}}^{\mathsf{T}} V_k^{'} \end{bmatrix} \quad \text{for } k \in \{1, \cdots, m\},$$

which is a orthogonal matrix. Natually, as a submatrix of the orthogonal matrix $\tilde{U}_k^{\mathsf{T}} \tilde{V}_k$, $W_k = U_k^{\mathsf{T}} V_k$ is absolutely doubly substochastic. Thus, by Lemma 3.3, W is absolutely doubly substochastic.

Combining Lemma 3.6, the row sum of W for each $i \in \{1, \cdots, r\}$ is as follows,

$$\begin{aligned}
\sum_{j=1}^{r} |W_{ij}| &= \sum_{j=1}^{r} |\langle \mathbf{u}_i^{(1)}, \mathbf{v}_j^{(1)} \rangle| \cdots |\langle \mathbf{u}_i^{(m)}, \mathbf{v}_j^{(m)} \rangle| \\
&\leq \sqrt[m]{\sum_{j=1}^{r} \langle \mathbf{u}_i^{(1)}, \mathbf{v}_j^{(1)} \rangle^m} \cdots \sqrt[m]{\sum_{j=1}^{r} \langle \mathbf{u}_i^{(m)}, \mathbf{v}_j^{(m)} \rangle^m} \\
&\leq \sqrt{\sum_{j=1}^{r} (V_1^{\mathsf{T}} \mathbf{u}_i^{(1)})_j^2} \cdots \sqrt{\sum_{j=1}^{r} (V_m^{\mathsf{T}} \mathbf{u}_i^{(m)})_j^2} \\
&= \|V_1^{\mathsf{T}} \mathbf{u}_i^{(1)}\| \cdots \|V_m^{\mathsf{T}} \mathbf{u}_i^{(m)}\| \\
&= \cos \theta_i^{(1)} \cdots \cos \theta_i^{(m)} \\
&\leq \cos^m \theta,
\end{aligned}$$

and its column sum of W for each $j \in \{1, \cdots, r\}$ is as follows,

$$\begin{aligned}
\sum_{i=1}^{r} |W_{ij}| &= \sum_{i=1}^{r} |\langle \mathbf{u}_i^{(1)}, \mathbf{v}_j^{(1)} \rangle| \cdots |\langle \mathbf{u}_i^{(m)}, \mathbf{v}_j^{(m)} \rangle| \\
&\leq \sqrt[m]{\sum_{i=1}^{r} \langle \mathbf{u}_i^{(1)}, \mathbf{v}_j^{(1)} \rangle^m} \cdots \sqrt[m]{\sum_{i=1}^{r} \langle \mathbf{u}_i^{(m)}, \mathbf{v}_j^{(m)} \rangle^m} \\
&\leq \sqrt{\sum_{i=1}^{r} (U_1^\mathsf{T} \mathbf{v}_j^{(1)})_i^2} \cdots \sqrt{\sum_{i=1}^{r} (U_m^\mathsf{T} \mathbf{v}_j^{(m)})_i^2} \\
&= \|U_1^\mathsf{T} \mathbf{v}_j^{(1)}\| \cdots \|U_m^\mathsf{T} \mathbf{v}_j^{(m)}\| \\
&= \cos \delta_j^{(1)} \cdots \cos \delta_j^{(m)} \\
&\leq \cos^m \delta.
\end{aligned}$$

Denote $\tilde{W} = \min\{\frac{1}{\cos^m \theta}, \frac{1}{\cos^m \delta}\} W$, it is obvious that $\tilde{W}$ is row absolutely doubly substochastic and column absolutely doubly substochastic, respectively. Therefore, $\tilde{W}$ is absolutely doubly substochastic. Furthermore, compute

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) W_{ij} = \max\{\cos^m \theta, \cos^m \delta\} \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) \tilde{W}_{ij}.$$

It is known that $\tilde{W}$ is absolutely doubly substochastic, by Lemma 3.4, there exists an absolutely doubly stochastic matrix $\tilde{C} = [\tilde{c}_{ij}] \in \mathbb{R}^{r \times r}$ such that $|\tilde{W}| \leq |\tilde{C}|$, then

$$\begin{aligned}
\langle \mathcal{A}, \mathcal{B} \rangle &\leq \max\{\cos^m \theta, \cos^m \delta\} \cdot \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) \big|\tilde{c}_{ij}\big| \\
&\leq \max\{\cos^m \theta, \cos^m \delta\} \cdot \max_{S \in T} \Big\{ \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) |s_{ij}| \Big\}
\end{aligned}$$

where $S = [s_{ij}]$ is an absolutely doubly-stochastic matrix, T denotes the set of all absolutely doubly-stochastic matrices. The function $f(S) = \sum\limits_{i,j=1}^{m} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) s_{ij}$ is a linear (and therefore convex) function on a polyhedral, then it attains its maximum at a vertex, i.e., a permutation matrix $P = [p_{ij}]$. If π is the permutation of $\{1, \ldots, m\}$ such that $p_{ij} = 1$ if and only if $j = \pi(i)$, then

$$\begin{aligned}
\langle \mathcal{A}, \mathcal{B} \rangle &\leq \max\{\cos^m \theta, \cos^m \delta\} \cdot \sum_{i,j=1}^{r} \sigma_i(\mathcal{A}) \sigma_j(\mathcal{B}) p_{ij} \\
&= \max\{\cos^m \theta, \cos^m \delta\} \cdot \sum_{i=1}^{r} \sigma_i(\mathcal{A}) \sigma_{\pi(i)}(\mathcal{B}) \\
&\leq \max\{\cos^m \theta, \cos^m \delta\} \cdot \sum_{i=1}^{r} \sigma_i(\mathcal{A}) \sigma_i(\mathcal{B}).
\end{aligned}$$

Furthermore, if $\theta_j^{(k)} = 0$ and $\delta_j^{(k)} = 0$ for $k \in \{1, \cdots, m\}$ and $j \in \{1, \cdots, r\}$, then the following equality holds,

$$\begin{aligned}\langle \mathcal{A}, \mathcal{B}\rangle &= \sum_{i,j=1}^{r} \sigma_i(\mathcal{A})\sigma_j(\mathcal{B})\langle \mathbf{u}_i^{(1)} \otimes \cdots \otimes \mathbf{u}_i^{(m)}, \mathbf{v}_j^{(1)} \otimes \cdots \otimes \mathbf{v}_j^{(m)}\rangle \\ &= \sum_{i,j=1}^{r} \sigma_i(\mathcal{A})\sigma_j(\mathcal{B})\langle \mathbf{u}_i^{(1)}, \mathbf{v}_j^{(1)}\rangle \cdots \langle \mathbf{u}_i^{(m)}, \mathbf{v}_j^{(m)}\rangle \\ &= \sum_{i=1}^{r} \sigma_i(\mathcal{A})\sigma_i(\mathcal{B}).\end{aligned}$$

This completes the proof. □

Corollary 3.9. *Let $\mathcal{A}, \mathcal{B} \in T(I_1, \cdots, I_m)$ be completely orthogonally decomposable tensors, $r = \min\{I_1, \cdots, I_m\}$, and its ordered singular values be $\sigma_1(\mathcal{A}) \geq \cdots \geq \sigma_r(\mathcal{A})$ and $\sigma_1(\mathcal{B}) \geq \cdots \geq \sigma_r(\mathcal{B})$ respectively. Then*

$$\|\mathcal{A} - \mathcal{B}\|^2 \geq \sum_{i=1}^{r} (\sigma_i(\mathcal{A}) - \sigma_i(\mathcal{B}))^2,$$

with the equality holds if and only if $\langle \mathcal{A}, \mathcal{B}\rangle = \sum_{i=1}^{r} \sigma_i(\mathcal{A})\sigma_i(\mathcal{B})$.

Proof. By Theorem 3.8,

$$\begin{aligned}\|\mathcal{A} - \mathcal{B}\|^2 &= \langle \mathcal{A} - \mathcal{B}, \mathcal{A} - \mathcal{B}\rangle \\ &= \sum_{i=1}^{r} \sigma_i^2(\mathcal{A}) - 2\langle \mathcal{A}, \mathcal{B}\rangle + \sum_{i=1}^{r} \sigma_i^2(\mathcal{B}) \\ &\geq \sum_{i=1}^{r} \sigma_i^2(\mathcal{A}) - 2\max\{\cos^m \theta, \cos^m \delta\} \cdot \sum_{i=1}^{r} \sigma_i(\mathcal{A})\sigma_i(\mathcal{B}) + \sum_{i=1}^{r} \sigma_i^2(\mathcal{B}) \\ &\geq \sum_{i=1}^{r} (\sigma_i(\mathcal{A}) - \sigma_i(\mathcal{B}))^2.\end{aligned}$$

Obviously, the equality holds if and only if $\langle \mathcal{A}, \mathcal{B}\rangle = \sum_{i=1}^{r} \sigma_i(\mathcal{A})\sigma_i(\mathcal{B})$. □

$\boxed{3.3}$ The case for SCODT

Having completed the generation from the Von Neumann's trace inequality to the content of CODT, let's consider the Fan's inequality to the content of tensor by using similar method. In the following, we denote $U = [\mathbf{u}_1, \cdots, \mathbf{u}_r]$ and $V = [\mathbf{v}_1, \cdots, \mathbf{v}_r]$, it can be seen that $U, V \in St(r, n)$. Define the angles between U and V as follows,

$$\begin{aligned}\cos \alpha_i &= \|V^\mathsf{T} \mathbf{u}_i\|, \text{ for } i = 1, \cdots, r, \\ \cos \beta_i &= \|U^\mathsf{T} \mathbf{v}_i\|, \text{ for } i = 1, \cdots, r, \\ \cos \alpha &= \max\{\cos \alpha_i \mid i = 1, \cdots, r\} \\ \cos \beta &= \max\{\cos \beta_i \mid i = 1, \cdots, r\}\end{aligned} \tag{3.8}$$

Lemma 3.10 (Abel's equality)**.** *Given numbers $a_1, \cdots, a_n$ and $b_1, \cdots, b_n$ arbitrarily, the following equality holds*

$$\sum_{i=1}^{n} a_i b_i = \sum_{i=1}^{n-1} \left[(a_i - a_{i+1}) \sum_{j=1}^{i} b_j\right] + a_n \sum_{j=1}^{n} b_j.$$

Lemma 3.11. *Let $\mathcal{B} = (b_{i_1 \cdots i_m}) \in S_{m,n}$ be a decomposable tensor with its eigenvalues $\lambda_1(\mathcal{B}) \geq \cdots \geq \lambda_n(\mathcal{B}) \geq 0$. Let $\mathbf{u}^{(1)}, \cdots, \mathbf{u}^{(s)}$ be a set of orthogonal vectors in $\mathbb{R}^n$. Then, the following inequality holds,*

$$\sum_{j=1}^{s} \mathcal{B}(\mathbf{u}^{(j)})^{\otimes m} \leq \sum_{i=1}^{s} \lambda_i(\mathcal{B}).$$

Proof. Without loss of generality, suppose that $\mathcal{B} \in S_{m,n}$ be a diagonal tensor. It is obvious that the diagonal elements of $\mathcal{B}$ are $b_{i \cdots i} = \lambda_i(\mathcal{B})$ for $i \in \{1, \cdots, m\}$, and

$$\mathcal{B}(\mathbf{u}^{(j)})^{\otimes m} = \sum_{i=1}^{n} \lambda_i(\mathcal{B})(\mathbf{u}_i^{(j)})^m,$$

then

$$\sum_{j=1}^{s} \mathcal{B}(\mathbf{u}^{(j)})^{\otimes m} = \sum_{j=1}^{n} \lambda_i(\mathcal{B})(\sum_{i=1}^{s} \mathbf{u}_i^{(j)})^m \leq \sum_{j=1}^{n} \lambda_i(\mathcal{B})(\sum_{i=1}^{s} \mathbf{u}_i^{(j)})^2 = \sum_{j=1}^{s} (\mathbf{u}^{(j)})^{\mathsf{T}} B \mathbf{u}^{(j)} = \sum_{i=1}^{s} \lambda_i(\mathcal{B}).$$

This completes the proof. □

Similar to the previous derivation, it is not difficult to derive the following result for SCODT.

Theorem 3.12 (Fan-type inequality)**.** *Let $\mathcal{A}, \mathcal{B} \in S_{m,n}$ be completely orthogonally decomposable symmetrical tensors, with the eigenvalues $\lambda_1(\mathcal{B}) \geq \cdots \geq \lambda_n(\mathcal{B}) \geq 0$ of $\mathcal{B}$, then the following inequality holds,*

$$\langle \mathcal{A}, \mathcal{B} \rangle \leq \min \left\{ \max\{\cos^n \alpha, \cos^n \beta\} \cdot \sum_{i=1}^{r} |\lambda_i(\mathcal{A})\lambda_i(\mathcal{B})|, \ \sum_{i=1}^{r} \lambda_i(\mathcal{A})\lambda_i(\mathcal{B}) \right\}.$$

Proof. **Part I**. Without loss of generality, suppose $\lambda_i(\mathcal{A})\lambda_i(B) \geq 0$ for $i = 1, \cdots, n$. By definition of completely orthogonally decomposable symmetrical tensors, $\mathcal{A}$ and $\mathcal{B}$ can be expressed by

$$\mathcal{A} = \sum_{i=1}^{r} \lambda_i(\mathcal{A}) \mathbf{u}_i \otimes \cdots \otimes \mathbf{u}_i,$$

and

$$\mathcal{B} = \sum_{j=1}^{r} \lambda_j(\mathcal{B}) \mathbf{v}_j \otimes \cdots \otimes \mathbf{v}_j,$$

where $U = [\mathbf{u}_1, \cdots, \mathbf{u}_r]$ and $V = [\mathbf{v}_1, \cdots, \mathbf{v}_r]$ are matrices in $Str(r, n)$. Then, the inner product of $\mathcal{A}$ and $\mathcal{B}$ expressed as

$$
\begin{aligned}
\langle \mathcal{A}, \mathcal{B} \rangle &= \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B})\langle \mathbf{u}_i \otimes \cdots \otimes \mathbf{u}_i, \mathbf{v}_j \otimes \cdots \otimes \mathbf{v}_j \rangle \\
&= \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B})\langle \mathbf{u}_i, \mathbf{v}_j \rangle \cdots \langle \mathbf{u}_i, \mathbf{v}_j \rangle \\
&= \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B})(U^\mathsf{T} V)_{ij} \cdots (U^\mathsf{T} V)_{ij}.
\end{aligned}
$$

Denote $J = [(U^\mathsf{T} V)_{ij} \cdots (U^\mathsf{T} V)_{ij}]$. It can be deduced that J is an absolutely doubly substochastic matrix. Actually, U and V can be extended to orthogonal matrix $\hat{U}$ and $\hat{V}$ by adding some column $U^{'}$ and $V^{'}$, which can be denoted as $\hat{U} = [U \quad U^{'}]$ and $\hat{V} = [V \quad V^{'}]$ respectively. Then,

$$
\hat{U}^\mathsf{T} \hat{V} = \begin{bmatrix} U^\mathsf{T} V & U^\mathsf{T} V^{'} \\ U^{'\mathsf{T}} V & U^{'\mathsf{T}} V^{'} \end{bmatrix},
$$

which is an orthogonal matrix. Natually, as a submatrix of the orthogonal matrix $\hat{U}^\mathsf{T}\hat{V}$, $U^\mathsf{T} V$ is absolutely doubly substochastic, and so is J. Consider the row sum of J for each $i \in \{1, \cdots, r\}$ as follows,

$$
\begin{aligned}
\sum_{j=1}^{r} |J_{ij}| &= \sum_{j=1}^{r} |\langle \mathbf{u}_i, \mathbf{v}_j \rangle| \cdots |\langle \mathbf{u}_i, \mathbf{v}_j \rangle| \\
&\leq \sqrt[n]{\sum_{j=1}^{r} \langle \mathbf{u}_i, \mathbf{v}_j \rangle^n} \cdots \sqrt[n]{\sum_{j=1}^{r} \langle \mathbf{u}_i, \mathbf{v}_j \rangle^n} \\
&\leq \sqrt{\sum_{j=1}^{r} (V^\mathsf{T} \mathbf{u}_i)_j^2} \cdots \sqrt{\sum_{j=1}^{r} (V^\mathsf{T} \mathbf{u}_i)_j^2} \\
&= \|V^\mathsf{T} \mathbf{u}_i\| \cdots \|V^\mathsf{T} \mathbf{u}_i\| \\
&= \cos^n \alpha_i \\
&\leq \cos^n \alpha,
\end{aligned}
$$

and its column sum of J for each $j \in \{1, \cdots, r\}$ satisfy $\sum_{i=1}^{r} |J_{ij}| \leq \cos^n \beta$. Denote $\tilde{J} = \min\{\frac{1}{\cos^n \alpha}, \frac{1}{\cos^n \beta}\} J$, it is obvious that $\tilde{J}$ is absolutely doubly substochastic. Then there exists an absolutely doubly stochastic matrix $\tilde{C} = [\tilde{c}_{ij}] \in \mathbb{R}^{r\times r}$ such that $|\tilde{J}| \leq |\tilde{C}|$, and

$$\begin{aligned}\langle \mathcal{A}, \mathcal{B}\rangle &= \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B}) J_{ij} \\ &= \max\{\cos^n \alpha, \cos^n \beta\} \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B}) \tilde{J}_{ij} \\ &\leq \max\{\cos^n \alpha, \cos^n \beta\} \cdot \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B}) \big|\tilde{c}_{ij}\big| \\ &\leq \max\{\cos^n \alpha, \cos^n \beta\} \cdot \max_{S\in T} \Big\{ \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B}) |s_{ij}| \Big\},\end{aligned}$$

where $S = [s_{ij}] \in \mathbb{R}^{r\times r}$ is a doubly stochastic matrix, T denotes the set of all the absolutely doubly stochastic matrices with size $r \times r$. The function $f(S) := \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B}) s_{ij}$ is a linear (and therefore convex) function on a polyhedral, then it attains its maximum at a vertex, for instance, at the permutation matrix $P = [p_{ij}]$. If π is the permutation of $\{1, \ldots, r\}$ such that $p_{ij} = 1$ if and only if $j = \pi(i)$, then

$$\begin{aligned}\langle \mathcal{A}, \mathcal{B}\rangle &\leq \max\{\cos^n \alpha, \cos^n \beta\} \cdot \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B}) p_{ij} \\ &= \max\{\cos^n \alpha, \cos^n \beta\} \cdot \sum_{i=1}^{r} \lambda_i(\mathcal{A})\lambda_{\pi(i)}(\mathcal{B}) \\ &\leq \max\{\cos^n \alpha, \cos^n \beta\} \cdot \sum_{i=1}^{r} \lambda_i(\mathcal{A})\lambda_i(\mathcal{B}).\end{aligned}$$

When $\alpha_i = 0$ and $\beta_j = 0$ for $i, j = 1, \cdots, r$, then the following equality holds,

$$\langle \mathcal{A}, \mathcal{B}\rangle = \sum_{i,j=1}^{r} \lambda_i(\mathcal{A})\lambda_j(\mathcal{B}) \langle \mathbf{u}_i \otimes \cdots \otimes \mathbf{u}_i, \mathbf{v}_j \otimes \cdots \otimes \mathbf{v}_j \rangle = \sum_{i=1}^{r} \lambda_i(\mathcal{A})\lambda_i(\mathcal{B}).$$

This completes the proof of the following inequality,

$$\langle \mathcal{A}, \mathcal{B}\rangle \leq \max\{\cos^n \alpha, \cos^n \beta\} \cdot \sum_{i=1}^{r} \lambda_i(\mathcal{A})\lambda_i(\mathcal{B}).$$

Part II. Let's consider this problem from another perspective. Compute

$$\langle \mathcal{A}, \mathcal{B}\rangle = \langle \Lambda \cdot (U, \cdots, U), \mathcal{B}\rangle = \langle \Lambda, \mathcal{B} \cdot (U^{\mathsf{T}}, \cdots, U^{\mathsf{T}})\rangle.$$

Denote $\mathcal{C} := \mathcal{B} \cdot (U^{\mathsf{T}}, \cdots, U^{\mathsf{T}})$, it is obvious that $\mathcal{C} = (c_{i_1 \cdots i_m})$ is a m-th order r-dimensional decomposable tensor. Then,

$$\langle \mathcal{A}, \mathcal{B}\rangle = \langle \Lambda, \mathcal{C}\rangle = \sum_{i=1}^{r} \lambda_i(\mathcal{A}) c_{i\cdots i}.$$

By Lemma 3.10 and Lemma 3.11, we deduce that

$$\begin{aligned}\langle \mathcal{A}, \mathcal{B} \rangle &= \sum_{i=1}^{r-1} \left[(\lambda_i(\mathcal{A}) - \lambda_{i+1}(\mathcal{A})) \sum_{j=1}^{i} c_{j \cdots j} \right] + \lambda_r(\mathcal{A}) \sum_{j=1}^{r} c_{j \cdots j} \\ &\leq \sum_{i=1}^{r-1} \left[(\lambda_i(\mathcal{A}) - \lambda_{i+1}(\mathcal{A})) \sum_{j=1}^{i} \lambda_j(\mathcal{B}) \right] + \lambda_r(\mathcal{A}) \sum_{j=1}^{r} \lambda_j(\mathcal{B}) \\ &= \sum_{i=1}^{r} \lambda_i(\mathcal{A}) \lambda_i(\mathcal{B}).\end{aligned}$$

The above equality holds if and only if $c_{j \cdots j}$ with $j = 1, 2, \cdots, r$ are the eigenvalue of B. Thus

$$\langle \mathcal{A}, \mathcal{B} \rangle \leq \sum_{i=1}^{r} \lambda_i(\mathcal{A}) \lambda_i(\mathcal{B}).$$

Combing part I and part II, we deduce that

$$\langle \mathcal{A}, \mathcal{B} \rangle \leq \min \left\{ \max\{\cos^n \alpha, \cos^n \beta\} \cdot \sum_{i=1}^{r} \lambda_i(\mathcal{A}) \lambda_i(\mathcal{B}), \ \sum_{i=1}^{r} \lambda_i(\mathcal{A}) \lambda_i(\mathcal{B}) \right\}.$$

This completes the proof. □

Moreover, we deduce the following proposition for SCODT.

Proposition 3.13. *Let $\mathcal{A}, \mathcal{B} \in T_{m,n}$ be completely orthogonally decomposable symmetrical tensors, then the following inequality holds,*

$$\langle \mathcal{A}, \mathcal{B} \rangle \geq \sum_{i=1}^{n} \lambda_i(\mathcal{A}) \lambda_{n+1-i}(\mathcal{B}).$$

Proof. By Theorem 3.12,

$$\langle \mathcal{A}, -\mathcal{B} \rangle \leq \sum_{i=1}^{r} \lambda_i(\mathcal{A}) \lambda_i(-\mathcal{B}) = \sum_{i=1}^{n} \lambda_i(\mathcal{A})(-\lambda_{n+1-i}(\mathcal{B})) = -\sum_{i=1}^{n} \lambda_i(\mathcal{A}) \lambda_{n+1-i}(\mathcal{B}),$$

which implies that

$$\langle \mathcal{A}, \mathcal{B} \rangle \geq \sum_{i=1}^{n} \lambda_i(\mathcal{A}) \lambda_{n+1-i}(\mathcal{B}).$$

□

$\boxed{4}$ Spectral Function for Asymmetric CODT

As an important topic in matrix theory, the theory of unitarily invariant matrix norms plays an important role in matrix spectral decomposition. In the following, a spectral decomposition property of matrices is generated to the content of tensor.

For any tensor $\mathcal{A} \in T(I_1, \cdots, I_m)$, we denote σ as the singular value map $\sigma : \mathcal{A} \to \mathbb{R}^r$, where r is the rank of $\mathcal{A}$. A real valued function f on $\mathbb{R}^n$ is called a *pre-norm*, if it is continuous, and satisfies

(i) positivity: $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$, and $f(\mathbf{x}) = 0$ if and only if $\mathbf{x} = \mathbf{0}$;

(ii) homogeneity: $f(\alpha\mathbf{x}) = |\alpha| f(\mathbf{x})$ for all real α and $\mathbf{x} \in \mathbb{R}^n$.

It is a norm if it further satisfies the triangle inequality. For a pre-norm f, its *dual norm*, denoted by f^D, is defined as

$$f^D(\mathbf{y}) = \max_{f(\mathbf{x})=1} \langle \mathbf{x}, \mathbf{y} \rangle. \tag{4.1}$$

Lemma 4.1 ([16])**.** *A pre-norm f is a norm if and only if $f = f^{DD}$.*

We call a function $g : \mathbb{R}^r \to \mathbb{R}^+$ *symmetric gauge* if it is a norm satisfies

(i) symmetry: $g(\mathbf{x}) = g(P\mathbf{x})$ for any permutation matrix $P \in \mathbb{R}^{r \times r}$;

(ii) absoluteness: $g(\mathbf{x}) = g(|\mathbf{x}|)$ for any vector $\mathbf{x} \in \mathbb{R}^r$.

For any tensor $\mathcal{Z} \in T(I_1, \cdots, I_m)$, define

$$\phi(\mathcal{Z}) := g(\sigma(\mathcal{Z})). \tag{4.2}$$

Obviously, $\phi(\cdot)$ is a unitarily invariant norm. Thus it can be seen that each unitarily invariant norm can be defined by a symmetric guage function. The following theorem proves that the unitarily invariant norm for each tensor can determine a symmetric guage function. Therefore, there is a one-to-one correspondence between unitarily invariant norm and symmetric guage function in tensor space.

Theorem 4.2. *Unitarily invariant norm in completely orthogonally decomposable tensors space is exactly the composite function of $g \circ \sigma$, where g is a symmetric gauge function.*

Proof. First, we can deduce that

$$(g \circ \sigma)^D = g^D \circ \sigma. \tag{4.3}$$

Actually, using Theorem 3.8, for any tensor $\mathcal{B} \in CODT$, there exists a tensor $\mathcal{A} \in CODT$ such that

$$(g \circ \sigma)^D(\mathcal{B}) = \max_{g(\sigma(\mathcal{A}))=1} \langle \mathcal{A}, \mathcal{B} \rangle \leq \{\max\{\cos^n \theta, \cos^n \delta\} \cdot \sum_{i=1}^{r} \sigma_i(\mathcal{A})\sigma_i(\mathcal{B}) \mid g(\sigma(\mathcal{A})) = 1\},$$

the equality holds when the angles between $\mathcal{A}$ and $\mathcal{B}$ satisfy $\theta_j^{(k)} = 0$ and $\delta_j^{(k)} = 0$ for $k \in \{1, \cdots, n\}$ and $j \in \{1, \cdots, r\}$. Moreover,

$$(g^D \circ \sigma)(\mathcal{B}) = g^D(\sigma(\mathcal{B})) = \max_{g(\sigma(\mathcal{A}))=1} \langle \sigma(\mathcal{A}), \sigma(\mathcal{B}) \rangle.$$

Thus it proves that (4.3). Furthermore, using Lemma 4.1 and formula (4.3), we derive

$$(g \circ \sigma)^{DD} = (g^D \circ \sigma)^D = g^{DD} \circ \sigma = g \circ \sigma.$$

According to Lemma 4.1 again, it deduce that $g \circ \sigma$ is a norm. This completes the proof. $\square$

$\boxed{5}$ Spectral Function for SCODT

In this section, spectral property for SCODT will be discussed. We denote λ as the eigenvalue map $\lambda : \mathcal{S} \to \mathbb{R}^n$ for each $\mathcal{S} \in SCODT(m,n)$, where $\mathcal{S} \in SCODT(m,n)$ means that $\mathcal{S}$ is a m-th order n-dimensional tensor in SCODT. Let $f : R \to [-\infty, +\infty]$ be a *extended-real* function. If f is a convex function or if a minimum point of f is being sought, then f is called *proper* if

$$f(x) > -\infty, \text{ for every } x \in \mathrm{dom}(f)$$

and if there also exists some point x_0 in its domain such that $f(x_0) < +\infty$, wherein $\mathrm{dom}(f)$ denote the domain of f:

$$\mathrm{dom}(f) = \{f(x) < +\infty \mid x \in \mathbb{R}\}.$$

Let f be a proper convex function on $\mathbb{R}^n$, the *Fenchel conjugate* is defined as

$$f^*(y) = \sup\{\langle x, y\rangle - f(x) \mid x \in \mathbb{R}\}.$$

Lemma 5.1. *Suppose the extended-real function f is proper. Then f is closed and convex if and only if $f = f^{**}$. In this case, f^* is also proper.*

In the following, with Theorem 3.12, it is not difficult to prove that there is a one-to-one relationship between invariant norm for completely orthogonally decomposable symmetrical tensors and eigenvalue function.

Theorem 5.2. *Unitarily invariant norm in completely orthogonally decomposable symmetrical tensors space is exactly the composite function of $h \circ \lambda$, where h is symmetric and convex.*

Proof. First, we should prove that

$$(h \circ \lambda)^* = h^* \circ \lambda. \tag{5.1}$$

Actually, for any $\mathcal{S} \in SCODT(m,n)$,

$$(h \circ \lambda)^*(\mathcal{S}) = \sup\{\langle \mathcal{X}, \mathcal{S}\rangle - h \circ \lambda(\mathcal{X}) \mid \mathcal{X} \in SCODT(m,n)\},$$

and

$$(h^* \circ \lambda)(\mathcal{S}) = \sup\{\langle \lambda(\mathcal{X}), \lambda(\mathcal{S})\rangle - h \circ \lambda(\mathcal{X}) \mid \lambda(\mathcal{X}) \in SCODT(m,n)\}.$$

By Theorem 3.12, it can be deduce that

$$(h \circ \lambda)^*(\mathcal{S}) \le (h^* \circ \lambda)(\mathcal{S}),$$

and the above equality is established if the angles between $\mathcal{S}$ and $\mathcal{X}$ are equal to zero. This proves (5.1). Furthermore, if h is proper, closed and convex, with using Lemma 5.1 and the equality (5.1), then we derive

$$(h \circ \lambda)^{**} = (h^* \circ \lambda)^* = h^{**} \circ \lambda = h \circ \lambda.$$

According to Lemma 5.1 again, it deduce that $h \circ \lambda$ is closed and convex. This completes the proof.

$\square$

$\boxed{6}$ Conclusion

In this paper, we considered generalizations of the important Von Neumann's trace inequality and Fan's inequality for matrices to completely orthogonally decomposable tensors and completely orthogonally decomposable symmetrical tensors respectively. Especially, the angle between two CODTs is defined and considered in it. Meanwhile, spectral properties for CODT and SCODT are studied as well.

References

[1] A. Anandkumar, R. Ge, D. Hsu, S. M. Kakade and M. T. Telgarsky, decompositions for latent variable models, *The Journal of Machine Learning Research* 15 (2014) 2773–2832.

[2] J. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples*, Springer Science and Business Media, 2000.

[3] S. Chrétien and T. Wei, Von Neumann's trace inequality for tensors, *Linear Algebra and its Applications* 482 (2015) 149–157.

[4] R. Ge, F. Huang, C. Jin and Y. Yuan, Escaping from saddle points-online stochastic gradient for tensor decomposition, *Journal of Machine Learning Research* 40 (2015) 1–46.

[5] G.H. Golub and C.F. Van Loan, *Matrix Computation*, Cambridge University Press, 1989.

[6] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[7] S. Hu, An inexact augmented Lagrangian method for computing strongly orthogonal decompositions of tensors, *Computational Optimization and Applications* 75 (2020) 701–737.

[8] S. Hu and G. Li, Convergence rate analysis for the higher order power method in best rank one approximations of tensors, *Numerische Mathematik* 140 (2018) 993–1031.

[9] S. Hu, D. Sun and K. Toh, Best nonnegative rank-one approximations of tensors, *SIAM Journal on Matrix Analysis and Applications* 40 (2020) 1527–1554.

[10] T.G. Kolda, Orthogonal tensor decompositions, *SIAM Journal on Matrix Analysis and Applications* 23 (2001) 243–255.

[11] T.G. Kolda, A counterexample to the possibility of an extension of the Eckart-Young low rank approximation theorem for the orthogonal rank tensor decomposition, *SIAM Journal on Matrix Analysis and Applications* 24 (2002) 762–767.

[12] P.M. Kroonenberg, *Three-Mode Principal Component Analysis: Theory and Applications* DSWO Press, Leiden, The Netherlands, 1983.

[13] J.M. Landsberg, *Tensors: Geometry and Applications*, AMS, Providence, RI., 2012.

[14] L.D. Lathauwer, Signal processing based on multilinear algebra, Leuven: Katholieke Universiteit Leuven, 1997.

[15] L.D. Lathauwer and B.D. Moor andJ.A. Vandewalle, Multilinear singular value decomposition, *SIAM Journal on Matrix Analysis and Applications* 21 (2000) 1253–1278.

[16] J.V. Neumann, Some matrix-inequalities and metrization of matric space, *Tomsk Univ. Rev.* 1 (1937) 286–300.

Manuscript received 29 June 2021
revised 19 October 2021
accepted for publication 27 September 2021

Jie Wang
Department of Mathematics, School of Science
China Jiliang University, Hangzhou, 310018, China
E-mail address: jiewang@cjlu.edu.cn