Linear Algebra and its Applications 482 (2015) 149–157

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Von Neumann's trace inequality for tensors

Stéphane Chrétien *, Tianwen Wei

Laboratoire de Mathématiques, UMR 6623, Université de Franche-Comté, 16 route de Gray, 25030 Besancon, France

A R T I C L E I N F O

Article history:
Received 12 March 2015
Accepted 10 May 2015
Available online 22 May 2015
Submitted by R. Brualdi

MSC:
15A69

Keywords:
Tensors
Singular Value Decomposition (SVD)
Higher Order Singular Value Decomposition (HOSVD)
Von Neumann's trace inequality

A B S T R A C T

For two matrices in $\mathbb{R}^{n_1 \times n_2}$, von Neumann's trace inequality says that their scalar product is less than or equal to the scalar product of their singular spectrum. In this short note, we extend this result to real tensors and provide a complete study of the equality case.

1. Introduction

The goal of this paper is to generalize von Neumann's trace inequality from matrices to tensors. Consider two matrices X and Y in $\mathbb{R}^{n_1 \times n_2}$. Denote their singular spectrum, i.e. the vector of their singular values, by $\sigma(X)$ (resp. $\sigma(Y)$). The classical matrix von Neumann's inequality [6] says that

* Corresponding author.
E-mail addresses: stephane.chretien@univ-fcomte.fr (S. Chrétien), tianwen.wei@univ-fcomte.fr (T. Wei).

http://dx.doi.org/10.1016/j.laa.2015.05.015

$$\langle X, Y \rangle \leq \langle \sigma(X), \sigma(Y) \rangle,$$

and equality is achieved if and only if there exist orthogonal matrices $U \in \mathbb{R}^{n_1 \times n_1}$ and $V \in \mathbb{R}^{n_2 \times n_2}$ such that $X = U \Sigma_X V^T$ and $Y = U \Sigma_Y V^T$ are respective singular value decompositions (SVD) of X and Y. Von Neumann's trace inequality, and the characterization of the equality case in this inequality, are important in many aspects of mathematics.

For tensors, the task of generalizing von Neumann's trace inequality is rendered harder because of the necessity to appropriately define the singular values and the SVD. In this paper, we will use the SVD defined in [2], which is based on the Tucker decomposition.

Our main result is given in Theorem 3.1 below and gives a characterization of the equality case for tensors. We expect this result to be useful for the description of the subdifferential of some tensor functions as the matrix counterpart has proved for matrix functions [5]. Such functions occur naturally in computational statistics, machine learning and numerical analysis [3,4,1] due to the recent interest in sparsity promoting norms as a convex surrogate to rank penalization.

2. Main facts about tensors

Let D and $n_1, \ldots, n_D$ be positive integers. Let $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_D}$ denote a D-dimensional array of real numbers. We will also denote such arrays as tensors.

2.1. Basic notations and operations

A subtensor of $\mathcal{X}$ is a tensor obtained by fixing some of its coordinates. As an example, fixing one coordinate $i_d = k$ in $\mathcal{X}$ for some $k \in \{1, \ldots, n_d\}$ yields a tensor in $\mathbb{R}^{n_1 \times \cdots \times n_{d-1} \times n_{d+1} \times \cdots \times n_D}$. In the sequel, we will denote this subtensor of $\mathcal{X}$ by $\mathcal{X}_{i_d = k}$.

The fibers of a tensor are subtensors that have only one mode, i.e. obtained by fixing every coordinate except one. The mode-d fibers are the vectors

$$\left(\mathcal{X}_{i_1, \ldots, i_{d-1}, i_d, i_{d+1}, \ldots, i_D}\right)_{i_d = 1, \ldots, n_d}.$$

They extend the notion of columns and rows from the matrix to the tensor framework. For a matrix, the mode-1 fibers are the columns and the mode-2 fibers are the rows.

The mode-d matricization $\mathcal{X}_{(d)}$ of $\mathcal{X}$ is obtained by forming the matrix whose rows are the mode-d fibers of the tensor, arranged in a cyclic ordering; see [7] for details.

The mode-d multiplication of a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_D}$ by a matrix $U \in \mathbb{R}^{n'_d \times n_d}$, denoted by $\mathcal{X} \times_d U$, gives a tensor in $\mathbb{R}^{n_1 \times \cdots \times n'_d \times \cdots \times n_D}$. It is defined as

$$(\mathcal{X} \times_d U)_{i_1, \ldots, i_{d-1}, i'_d, i_{d+1}, \ldots, i_D} = \sum_{i_d = 1}^{n_d} \mathcal{X}_{i_1, \ldots, i_{d-1}, i_d, i_{d+1}, \ldots, i_D} U_{i'_d, i_d}.$$