CSE 847 (Spring 2022): Machine Learning — Project
Paper Study Summary

Wei-Chien Liao (liaowei2@msu.edu)

Abstract

This paper is a summary of basic concepts of tensors, Tucker decomposition and higher order singular value decomposition (HOSVD), and variants of randomized algorithms for computing these decompositions.

Keywords: higher order singular value decomposition (HOSVD), Tucker decomposition, randomized HOSVD

1. Notation and Preliminaries

The contents in this section are mainly based on [2].

Definition 1. The **order** of a tensor is the number of dimensions, also called **ways** or **modes**.

In this paper,

- **vectors** (tensors of order 1) are denoeted by boldface lowercase letters, e.g. $\mathbf{a}$.
- **matrices** (tensors of order 2) are denoeted by boldface capital letters, e.g. $\mathbf{A}$.
- **tensors** (order ≥ 3) are denoeted by boldface Euler script letters, e.g. $\boldsymbol{\mathcal{X}}$.
- the i-th entry of a vector $\mathbf{a}$ is denoeted by a_i.
- the (i, j)-th element of a matrix $\mathbf{A}$ is denoeted by A_{ij}.
- the (i, j, k)-th element of a third-order tensor $\boldsymbol{\mathcal{X}}$ is denoeted by x_{ijk}.
- a colon ":" is used to indicate all elements of a mode. e.g. for a matrix $\mathbf{A}$,
  - $\mathbf{a}_{i:} = i$-th row of $\mathbf{A}$.
  - $\mathbf{a}_{:j} = j$-th column of $\mathbf{A}$.

Definition 2. A **fiber** is defined by fixing every index but one.

For a third-order tensor $\boldsymbol{\mathcal{X}}$,

- $\mathbf{x}_{:jk}$ = **column fibers** or **mode-1 fibers** of $\boldsymbol{\mathcal{X}}$.
- $\mathbf{x}_{i:k}$ = **row fibers** or **mode-2 fibers** of $\boldsymbol{\mathcal{X}}$.

- $\mathbf{x}_{ij:}$ = **tube fibers** or **mode-3 fibers** of $\mathcal{X}$.

Definition 3. **Slices** are two-dimensional sections of a tensor defined by fixing all but two indices.

For a third-order tensor $\mathcal{X}$,

- $\mathbf{X}_{i::}$ = **horizontal slices** of $\mathcal{X}$.
- $\mathbf{X}_{:j:}$ = **lateral slices** of $\mathcal{X}$.
- $\mathbf{X}_{::k}$ = **frontal slices** of $\mathcal{X}$.

Definition 4 (Norm of a Tensor). The **norm** of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, denoeted by $||\mathcal{X}||$, is defined as

$$||\mathcal{X}|| = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} x_{i_1 i_2 \cdots i_N}^2}. \tag{1}$$

Definition 5 (Inner Product of Tensors). The **inner product** of two same-sized tensors $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, denoeted by $\langle \mathcal{X}, \mathcal{Y} \rangle$, is defined as

$$\langle \mathcal{X}, \mathcal{Y} \rangle = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} x_{i_1 i_2 \cdots i_N} y_{i_1 i_2 \cdots i_N}}. \tag{2}$$

Thus, by the definition of norm and inner product, $\langle \mathcal{X}, \mathcal{X} \rangle = ||\mathcal{X}||^2$.

Definition 6 (Rank-one Tensors). A N-way tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is **rank one** if it can be written as the outer product of N vectors,

$$\mathcal{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)}, \tag{3}$$

for some vectors $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \cdots, \mathbf{a}^{(N)}$ and "$\circ$" denotes the vector outer product.

Definition 7. A tensor is called **cubical** if every mode is the same size. A cubical tensor is called **supersymmetric** (some literatures call this "symmetric") if its elements remain constant under any permutation of the indices.

For a 3-way tensor $\mathcal{X} \in \mathbb{R}^{I \times I \times I}$, it is supersymmetric if

$$x_{ijk} = x_{ikj} = x_{jik} = x_{jki} = x_{kij} = x_{kji} \quad \forall i, j, k = 1, \cdots I.$$

Definition 8 (Diagnoal Tensor). A tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is **diagonal** if $x_{i_1 i_2 \cdots i_N} \neq 0$ only if $i_1 = i_2 = \cdots = i_N$.

Definition 9 (Matricization). The process of reordering the elements of an N-way array into a matrix is called **matricization** This is also called **unfolding** or **flattening**.

The mode-n matricization of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is denoeted by $\mathbf{X}_{(n)}$ and arranges the mode-n fibers to be the columns of the resulting matrix.

Remark 1. It is also possible to vectorize a tensor. This process is called vectorization.

1.1. Tensor Mulitiplication

The n**-mode prodcut** of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ with a matrix $\mathbf{U} \in \mathbb{R}^{J \times I_n}$ is defined as

$$\text{eletmenwise:} \quad (\mathcal{X} \times_n \mathbf{U})_{i_1 \cdots i_{n-1} j i_{n+1} \cdots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \cdots i_N} u_{j i_n} \tag{4a}$$

$$\text{unfold tensors:} \quad \mathcal{Y} = \mathcal{X} \times_n \mathbf{U} \Leftrightarrow \mathbf{Y}_{(n)} = \mathbf{U}\mathbf{X}_{(n)} \tag{4b}$$

The result is a tensor of size $I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N$. Each mode-$n$ fiber is multiplied by the matrix $\mathbf{U}$. The following are some properties of the n-mode prodcut:

(1) For distinct modes in a series of multiplications, the order of the multiplication is irrelevant.

$$\mathcal{X} \times_m \mathbf{A} \times_n \mathbf{B} = \mathcal{X} \times_n \mathbf{B} \times_m \mathbf{A}, \quad \text{for } n \neq m. \tag{5}$$

(2) If the modes are the same

$$\mathcal{X} \times_n \mathbf{A} \times_n \mathbf{B} = \mathcal{X} \times_n (\mathbf{B}\mathbf{A}). \tag{6}$$

1.2. Some Matrix Products

- **<u>Kronecker Product</u>**

  The **Kronecker product** of two matrices $\mathbf{A} \in \mathbb{R}^{I \times J}$ and $\mathbf{B} \in \mathbb{R}^{K \times L}$, denoeted by $\mathbf{A} \otimes \mathbf{B}$, is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1J}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2J}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}\mathbf{B} & a_{I2}\mathbf{B} & \cdots & a_{IJ}\mathbf{B} \end{bmatrix} \tag{7}$$

- **<u>Khatri-Rao Product</u>**

  The **Khatri-Rao product** of two matrices $\mathbf{A} \in \mathbb{R}^{I \times K}$ and $\mathbf{B} \in \mathbb{R}^{J \times K}$, denoeted by $\mathbf{A} \odot \mathbf{B}$, is defined as

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} \mathbf{a}_1 \otimes \mathbf{b}_1 & \mathbf{a}_2 \otimes \mathbf{b}_2 & \cdots & \mathbf{a}_K \otimes \mathbf{b}_K \end{bmatrix} \tag{8}$$

- **<u>Hadamard Product</u>**

  The **Hadamard product** of two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{I \times J}$, denoeted by $\mathbf{A} * \mathbf{B}$, is defined as

$$\mathbf{A} * \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1J}b_{1J} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2J}b_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}b_{I1} & a_{I2}b_{I2} & \cdots & a_{IJ}b_{IJ} \end{bmatrix} \tag{9}$$

Some properties of these matrix prodcuts:

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD} \tag{10a}$$

$$(\mathbf{A} \otimes \mathbf{B})^\dagger = \mathbf{A}^\dagger \otimes \mathbf{B}^\dagger \tag{10b}$$

$$\mathbf{A} \odot \mathbf{B} \odot \mathbf{C} = (\mathbf{A} \odot \mathbf{B}) \odot \mathbf{C} = \mathbf{A} \odot (\mathbf{B} \otimes \mathbf{C}) \tag{10c}$$

$$(\mathbf{A} \odot \mathbf{B})^T (\mathbf{A} \odot \mathbf{B}) = (\mathbf{A}^T\mathbf{A}) * (\mathbf{B}^T\mathbf{B}) \tag{10d}$$

$$(\mathbf{A} \odot \mathbf{B})^\dagger = (\mathbf{A}^T\mathbf{A}) * (\mathbf{B}^T\mathbf{B})^\dagger (\mathbf{A} \odot \mathbf{B})^T \tag{10e}$$

2. Tucker Decomposition and HOSVD

3. Randomized Algorithms

This section summarizes some randomized algorithm for computing HOSVD discussed in [1].

References

[1] Salman Ahmadi-Asl, Stanislav Abukhovich, Maame G. Asante-Mensah, Andrzej Cichocki, Anh Huy Phan, Tohishisa Tanaka, and Ivan Oseledets. Randomized algorithms for computation of tucker decomposition and higher order SVD (HOSVD). IEEE Access, 9:28684–28706, 2021.

[2] Tamara G. Kolda and Brett W. Bader. Tensor decompositions and applications. SIAM Review, 51(3):455–500, 2009.