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.

1. Notation and Preliminaries

The contents in this section are mainly based on [1].

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 $\boldsymbol{\mathcal{X}}$.

Preprint submitted to Elsevier

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

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

2. Tucker Decomposition and HOSVD

3. Randomized Algorithms

This section summarizes some randomized algorithm for computing HOSVD discussed in [2].

1. **Project Title:**
   A Comparison of Various Randomized Higher Order Singular Value Decomposition (HOSVD) Algorithms
2. **Team Members:** Wei-Chien Liao, Shihab Shahriar Khan

3. **Description of the Problem:**
   Many applications in data sciences require processing high-order tensor data. To deal with large tensor data, dimensionality reduction techniques play an important role among many other types of algorithms. However, performing dimension reduction operations like Tucker decomposition and High Order Singular Value Decomposition (HOSVD) with deterministic algorithms are not efficient for handling large tensor data. This inefficiency can be handled by randomized algorithms. This type of algorithms accelerate classical decompostions by reducing computational complexity of deterministic methods and communications among different level of memory hierarchy. This project aims to study, implement and compare many variants of randomized algorithms, and test them with different datasets from applications such as handwritten digit classification, computer vision or signal processing to evaluate their performances.

4. **Preliminary Plan (Milestones):**
   (a) Test Tensor Toolbox for MATLAB
   (b) Study the paper [2] in the paper list.
   (c) Implement, analyze and compare the algorithms in [2]

5. **Paper List:**

[2] 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

[3] Linjian Ma and Edgar Solomonik. Fast and accurate randomized algorithms for low-rank tensor decompositions. In A. Beygelzimer, Y. Dauphin, P. Liang, and J. Wortman Vaughan, editors, Advances in Neural Information Processing Systems, 2021

[4] Tamara G. Kolda et al. Brett, W Bader. Tensor toolbox for matlab, version 3.2.1, www.tensortoolbox.org, 4 2021

[5] Rachel Minster, Arvind K. Saibaba, and Misha E. Kilmer. Randomized algorithms for low-rank tensor decompositions in the tucker format. SIAM Journal on Mathematics of Data Science, 2(1):189–215, 2020

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