CSE 847 (Spring 2022): Machine Learning
Intermediate Report

Wei-Chien Liao (liaowei2@msu.edu)
Shihab Shahriar Khan (khanmd@msu.edu)

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. **Introduction and Problem Description:**
   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. **Description of the data used in the project:**
   (a) Synthetic 3rd order tensor, optionally with noise added.
   (b) COIL-100 DATA SET [1]: contains 7200 color images (100 objects under 72 different rotations).
   (c) NEIL 2 [2]: For experiments on sparse tensor. It's a 3rd order sparse tensor of shape $12092 \times 9184 \times 28818$.

5. **What We Have Done So Far:**

(a) Study the theory and mathematics of two prominent deterministic tensor decomposition algorithms called Tucker decomposition and High Order Singular Value Decomposition (HOSVD).

(b) Study the papers [3] and [4] to understand randomized tensor decomposition algorithms, and how they perform compared to their deterministic counterparts along several metrics of interest.

(c) Explore the Tensor Toolbox for MATLAB [5].

   i. Test the `HOSVD` functionality in the Toolbox.

   ii. Implement STHOSVD and HOOI algorithm.

(d) Exploration of "Tensorly" [6] library in Python.

   i. Implementation of tensor decomposition and reconstruction using lower level functions to gain a better idea of how the library and the algorithms work.

   ii. Using order-3 tensors of image and synthetic data, implemented a preliminary experiment to test the reconstruction performance of 3 algorithms: CP, Tucker and RandomizedCP. The algorithms were evaluated based on Relative Error and Running Time, following the procedure of [3]. The qualitative performance on image data was also considered.

6. **What Remains to be Done:**

   (a) Implement the randomized algorithms in [3].

   (b) Analyze the algorithms using the datasets listed in the "Description of the data" section.

   (c) Propose and evaluate potential ways further speedup can be achieved.

References

[1] Sameer A Nene, Shree K Nayar, Hiroshi Murase, et al. Columbia object image library (coil-100). 1996.

[2] Andrew Carlson, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R Hruschka, and Tom M Mitchell. Toward an architecture for never-ending language learning. In Twenty-Fourth AAAI conference on artificial intelligence, 2010.

[3] 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.

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

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

[6] Jean Kossaifi, Yannis Panagakis, Anima Anandkumar, and Maja Pantic. Tensorly: Tensor learning in python. Journal of Machine Learning Research, 20(26):1–6, 2019.