Msc thesis proposal:

Efficient Implementation of Supervised Learning with the Canonical Polyadic Decomposition on GPU

 Fang Fang[[1]](#footnote-2)

**Background**

Canonical Polyadic Decomposition (CPD), also known as CANDECOMP/PARAFAC (CP), is a multilinear algebraic technique used to factorize a multiway array into a sum of component rank-1 tensors. E.g., for a $N$-order tensor $X\in R^{K\_{1}×\cdots ×K\_{N}}$ (a $N$ dimensional array), CPD expresses $X$ as

|  |  |
| --- | --- |
| $$X=\left[A\_{1},A\_{2},\cdots ,A\_{N}\right]\_{R}=\sum\_{r=1}^{R}a\_{r}^{1} ° a\_{r}^{2} °\cdots ° a\_{r}^{N},$$ | **Eq １** |

where $a\_{r}^{n}$ is a one-dimensional array, $A\_{n}=\left[a\_{1}^{n},\cdots ,a\_{R}^{n}\right]\in R^{K\_{n}×R}$ , and the rank $R$ is the minimum number of the rank-1 components needed for this decomposition.

This technique is widely applied in various fields, including chemometrics, neuroscience, and image analysis, to extract latent factors and identify hidden patterns within complex, high-dimensional data. By revealing the underlying structures of data, CPD can effectively achieve the effect of dimensionality reduction and help solve high dimensional computation problems.

Recently, in a series of research [1], [2], [3] and [4], CPD has been employed to construct a multidimensional Fourier cosine series (COS) expansion. By assuming a low rank structure as in **Eq １** for the multiway Fourier coefficients,CPD helps reduce the curse of dimensionality that has been causing the bottleneck of the classical multidimensional Fourier series expansion. As part of R&D activities at FF Quant, we have successfully applied COS empowered by CPD (COS-CPD) to solve several multi-dimensional integration problems that arise in the field of quantitative finance. The COS-CPD method has been tested to be an efficient alternative to the classical Monte Carlo method for numerical integration.

COS-CPD is essentially a supervised machine learning method that searches for the best CPD structure of the multiway Fourier coefficients, by matching the Fourier series expansion with the target function valued at a sample.

In this thesis research, we aim to develop an efficient parallel computing algorithm for COS-CPD on a graphics processing unit (GPU).

**Challenge**

The nature of our problem to solve requires the student to be able to work with tensor calculus fluently enough and implement efficient tensor calculus on GPU. E.g., the Alternate Least Squares Algorithm (ALS) is the most standard numerical optimization algorithm that is used for the CPD decomposition. It consists of several operations, e.g., MTTKRP (Matrix times Tensor Khatri Rao Product), tall and wide matrix multiplication, and the computation of the pseudo-inverse of the resultant dense matrix. In the meanwhile, there are also other state-of-art methods to solve CPD.

Note that since our aim to is to achieve dimension reduction, we will work with tensors of relatively large dimension. Thus, it is necessary to deal with the memories on GPU carefully.

Besides, our supervised machine learning approach requires to hard code optimization methods on GPU such as conjugate gradient (CG) or stochastic gradient descent (SGD) methods.

**The goal and content of this thesis**

The goal of this thesis is to have an efficient numerical algorithm and implementation for the COS-CPD method on GPU.

**Contact**

Please feel free to contact me directly if this topic is of your interest, or if you would like to learn more details: f.fang@tudelft.nl

**Reference**

1. *[Supervised Learning and Canonical Decomposition of Multivariate Functions](https://ieeexplore.ieee.org/document/9340610),* N. Kargas and N. D. Sidiropoulos, IEEE Transactions on Signal Processing, vol. 69, pp. 1097–1107, 2021, ISSN: 1053-587X. DOI: 10.1109/TSP.2021.3055000.

1. [Dimension Reduction Techniques for Multi-dimensional Numerical Integrations Based on Fourier-cosine Series Expansion](https://repository.tudelft.nl/islandora/object/uuid%3Af6ffbcd6-df14-425b-84bb-63e4e105205c). Zhimin Cheng. TU Delft, Msc thesis, 2022.

1. *[Solving Multivariate Expectations Using Dimension-reduced Fourier-cosine Series Expansion and Its Applications in Finance](http://ta.twi.tudelft.nl/users/vuik/numanal/brands_afst.pdf)*. Marnix Brands. TU Delft, Msc thesis 2023.
2. *Pricing Barrier Options Under Lévy Processes Using Dimension-reduced Cosine Expanson.* Levi Klomp. TU Delft, Msc thesis, 2023.

1. *[A Novel Option Pricing Method Based on Fourier-cosine Series Expansions](https://epubs.siam.org/doi/10.1137/080718061).* F. Fang and C. W. Oosterlee. SIAM J. Sci. Comput.,31(2):826-848, 2009.
2. [OptiCPD: Optimization for the Canonical Polyadic Decomposition Algorithm on GPUs](https://ieeexplore.ieee.org/document/10196662), Srinivasan Subramaniyan and Xiaorui Wang. 2023 IEEE International Parallel and Distributed Processing Symposium Workshops. DOI: [10.1109/IPDPSW59300.2023.00071](https://doi.org/10.1109/IPDPSW59300.2023.00071).
1. Part-time assistant professor at the Applied Mathematics Department of TU Delft; Director of of FF Quant Advisory B.V. [↑](#footnote-ref-2)