
Optimal tensor-matrix and -vector contraction strategies for computing dense tensor decompositions

Oguz Kaya*¹ and Yves Robert^{2,3}

¹Laboratoire de Recherche en Informatique – Université Paris-Sud - Université Paris-Saclay – France

²Laboratoire de l'Informatique du Parallélisme – École Normale Supérieure - Lyon – France

³Innovative Computing Laboratory [Knoxville] – United States

Abstract

Tensor decompositions are able to effectively compress high dimensional data having inherent low-rank properties and are used in numerous application domains including signal processing, computer vision, quantum chemistry, electromagnetics, data analysis, and numerical data compression.

Canonical Polyadic (CP) and Tucker formulations are two prominent low-rank tensor decompositions manifesting in these applications, and these representations involve a matrix for each dimension of a tensor.

Computing these matrices is a non-linear optimization problem, which is typically linearized in one of the dimensions to update the corresponding matrix of that dimension while fixing others, and iterating in an alternating manner over all dimensions until convergence.

The linearization step constitutes the most computationally intensive part of the algorithm involving contractions with the tensor and the current state of dimension matrices in the iterative process.

An interesting problem from a combinatorial point of view arises when finding an efficient framework to minimize the computational cost of a such sequence of contractions.

We propose algorithms for finding an optimal execution scheme on a binary tree structure using dynamic programming techniques, and propose some interesting complexity results.

For a D -dimensional tensor, we are able to find the optimal strategy in $O(3^D)$ time using $O(2^D)$ space for both CP and Tucker decompositions.

Keywords: tensor decompositions, CP decomposition, Tucker decomposition, dimension trees, tensor contraction

*Speaker