Journal of Computational Mathematics
Vol.36, No.3, 2018, 351–373.
http://www.global-sci.org/jcm
doi:10.4208/jcm.1605-m2016-0828

A TRUST-REGION-BASED ALTERNATING LEAST-SQUARES ALGORITHM FOR TENSOR DECOMPOSITIONS*

Fan Jiang and Deren Han
Jiangsu Key Laboratory of NSLSCS, School of Mathematical Sciences, Nanjing Normal University, Nanjing 210023, China
Email: 15905154902@163.com, handeren@njnu.edu.cn
Xiaofei Zhang
Information Technology Department, Chinascope, Nanjing 210023, China
Email: 879734743@qq.com

Abstract

Tensor canonical decomposition (shorted as CANDECOMP/PARAFAC or CP) decomposes a tensor as a sum of rank-one tensors, which finds numerous applications in signal processing, hypergraph analysis, data analysis, etc. Alternating least-squares (ALS) is one of the most popular numerical algorithms for solving it. While there have been lots of efforts for enhancing its efficiency, in general its convergence can not been guaranteed.

In this paper, we cooperate the ALS and the trust-region technique from optimization field to generate a trust-region-based alternating least-squares (TRALS) method for CP. Under mild assumptions, we prove that the whole iterative sequence generated by TRALS converges to a stationary point of CP. This thus provides a reasonable way to alleviate the swamps, the notorious phenomena of ALS that slow down the speed of the algorithm. Moreover, the trust region itself, in contrast to the regularization alternating least-squares (RALS) method, provides a self-adaptive way in choosing the parameter, which is essential for the efficiency of the algorithm. Our theoretical result is thus stronger than that of RALS in [26], which only proved the cluster point of the iterative sequence generated by RALS is a stationary point. In order to accelerate the new algorithm, we adopt an extrapolation scheme. We apply our algorithm to the amino acid fluorescence data decomposition from chemometrics, BCM decomposition and rank-$(L_r, L_r, 1)$ decomposition arising from signal processing, and compare it with ALS and RALS. The numerical results show that TRALS is superior to ALS and RALS, both from the number of iterations and CPU time perspectives.

Mathematics subject classification: 90C06, 90C53, 65K05
Key words: tensor decompositions, trust region method, alternating least-squares, extrapolation scheme, global convergence, regularization.

1. Introduction

The problem of decomposing a higher-order tensor into a sum of products of lower-order tensors finds more and more important applications in signal processing [3, 4], data analysis [2,32], scientific computing [5,22,24,25], biomedical engineering [1,39], machine learning [35], chemometrics [36], etc. One of the famous decompositions of higher-order tensors is CANDECOMP/PARAFAC analysis (CP, canonical polyadic decomposition), which can be dated back to the work of Hitchcock in 1927 [20,21]. However, it is out of the interests of researchers until

* Received January 5, 2017 / Revised version received March 20, 2017 / Accepted May 22, 2017 / Published online March 28, 2018 /

the study of Tucker [38], Carroll and Chang [7] and Harshman [19] in the fields of psychometries and phonetics in 1970, respectively. In signal processing, there is a generalized PARAFAC, named Block-PARAFAC [3]. The model is applied in direct-sequence code division multiple access (DS-CDMA) system, and orthogonal frequency division multiplexing (OFDM) system.

Recently, some numerical optimization algorithms were tailored for solving tensor decomposition problems [37], among which the Levenberg-Marquardt algorithm for non-linear least squares problems [15–17] attracts much attention [23]. Nevertheless, the alternating least squares (ALS) method is still the most popular one, due to its simplicity in implementation [2, 7, 19, 24, 36]. However, on one hand, ALS usually needs a large number of iterations to converge because the convergence rate of many iterations is almost null; and, on the other hand, we can not prove that the limit points of the subsequences generated by the ALS are the critical points of the least squares cost function. Many researchers thus pay a lot of attentions to enhancing the efficiency of ALS numerically by introducing skills from numerical algebra and numerical optimization. For example, the line search along the incremental direction of an old iterate to a new one is proposed by Harshman [19], and as a consequence, it needs efficient schemes in finding an 'optimal' step size. For real-valued tensors, the optimal step size can be directly calculated because it is a solution of a polynomial equation with a single argument, and the resulting algorithm is called "enhanced line search" (ELS) [10, 33]. For complex-valued tensors CP, Nion and De Lathauwer propose "enhanced line search with complex step" (ELSCS) [29,30], where a complex-valued step size factor that contains the modulus and the phase is introduced, and the optimal step size is approximated by an alternating minimization manner, i.e., find the optimal modulus for a fixed phase and then find the optimal phase for a fixed modulus. Recently, by using the classic resultant results from algebraic geometry [12], Chen, et al. [8] find that the complex-valued optimal step size can also be found by solving two single variable polynomial equations successively. Since the polynomials are with high order, they propose to solve the first polynomial equation by solving an eigenvalue problem and the algorithm is much stable. Most recently, Domanov and De Lathauwer [14] propose to reduce the CP decomposition of third-order tensors to generalized eigenvalue decomposition, which enables to use the high performance techniques from numerical algebra.

While there have been several techniques in enhancing the efficiency of ALS numerically, the results on its global convergence is in its infancy. To ensure the global convergence, Li, Kindermann, and Navasca [26] propose to use the regularization technique and name the resulting algorithm as regularized alternating least squares method, shorted as RALS method. That is, the subproblems in ALS are replaced by the new objectives which are the sum of the original least squares and the regularization terms (a quadratic term). This quadratic term, which transforms the objective function from a convex function to a strongly convex one, plays an important role. Theoretically, Li, et al. [26] prove that the limit points of the converging subsequences of the RALS are the critical points of the least squares cost function; and numerically, it speeds up ALS by successfully avoiding the swamp phenomenon, i.e., a large number of iterations with no improvements.

The RALS, though having the advantages over ALS, both from theoretical and numerical viewpoints, has some issues to be handled. Theoretically, the assertion that the limit points of the converging subsequences of the RALS are the critical points of the least squares cost function is not a satisfying result, and one wonders if the whole generated sequence converges to a critical point. Numerically or practically, the regularization parameter plays a very important role in the efficiency of the algorithm, while it is difficult to choose a proper value. To tackle